On the statistics of K-distributed noise

1980 J. Phys. A: Math. Gen. 13 31

(http://iopscience.iop.org/0305-4470/13/1/006)

J. Phys. A: Math. Gen., **13** (1980) 31–48. Printed in Great Britain

On the statistics of K-distributed noise

E Jakeman
Royal Signals and Radar Establishment, Malvern, Worcestershire, UK

Received 25 January 1979, in final form 18 May 1979

Abstract. When the number of steps in a random walk varies, the distribution of the resultant vector components in the limit of large mean step number may be non-Gaussian. In this paper the statistics and temporal correlation properties of one class of such non-Gaussian limit distributions are derived and some of its potential applications are reviewed briefly.

1. Introduction

The class of modified Bessel function or K distributions

$$p(A)=\frac{2b}{\Gamma(\alpha)}\left(\frac{bA}{2}\right)^{\alpha}K_{\alpha-1}(bA) \qquad \alpha>0 \tag{1}$$

has recently been found to provide an excellent model for the amplitude statistics of the scattered radiation in a wide variety of experiments involving scattering from turbulent media (Jakeman and Pusey 1978). It has been suggested that this may be because they are limit distributions in a certain type of random walk problem. It is well known that if the number of steps in a random walk is increased without limit then the components of the resultant vector have a Gaussian distribution (the central limit theorem). It is less well known that if the number of steps is itself a statistical variable then this result does not necessarily hold. Although there is a vast literature devoted to the central limit theorem and the conditions governing convergence to Gaussian statistics, little consideration appears to have been given to the possibility of non-Gaussian limit distributions arising from number fluctuations. Equation (1) defines such a class of distributions, generated by negative binomial number fluctuations. In this paper a discussion of the consequences of the negative binomial assumption will be given. Some higher-order joint statistics of the associated limit distributions will be derived and the properties of two related distributions will also be discussed. Number fluctuations in the random walk problem will be treated here as a subject of interest in its own right and no attempt will be made in the present work to justify the basic model for the number fluctuation distribution. However, comparisons with experimental data will be given to demonstrate the relevance and potential range of applications of limit distributions related to the class (1).

The negative binomial distributions

$$P_N=\binom{N+\alpha-1}{N}\frac{(\bar{N}/\alpha)^N}{(1+\bar{N}/\alpha)^{N+\alpha}}, \qquad \alpha>0, \tag{2}$$

0305-4470/80/010031+18\$01.00

form a two-parameter class, each member of which is characterised by its mean $\bar{N}$ and normalised variance $\alpha^{-1}+\bar{N}^{-1}$. The most familiar member of the class is the geometric distribution, when $\alpha=1$, which, for example, describes the photon statistics of thermal (incoherent) light. The parameter α appearing in (2) can in fact be regarded as a measure of the degree of bunching in a time sequence of events in which N counts the number of events in a fixed small sample time. These distributions are often introduced as models for variable mean Poisson processes but they assume a more important and fundamental role in queueing theory and more particularly in population statistics (e.g. Bartlett 1966) where they occur as equilibrium distributions in the birth–death–immigration process. Various theoretical treatments of this latter type of process exist in the literature and in § 2 a rate-equation approach will be used to derive a number of results relevant to the random walk problem.

In § 3 limit distributions associated with negative binomial number fluctuations will be derived and discussed and results on some useful related distributions will be considered. In § 4 comparisons with presently available experimental data will be given. A general discussion and summary are given in § 5.

2. Number fluctuations

In this section the theory of negative binomial number fluctuations will be developed using a rate-equation approach to solve the birth–death–immigration problem.

The transitions between adjacent population levels in a population with birth rate λ, death rate μ and immigration rate ν are illustrated in figure 1. Immigration is somewhat analogous to spontaneous emission, being independent of population, and with the applications of § 4 in mind it is, perhaps, more appropriate to call this contribution to the rate equation 'spontaneous creation'. It will be assumed throughout that births, deaths and spontaneous creations occur as uncorrelated random events. The rate equation for the process may thus be written

$$\frac{dP_N}{dt}=\mu(N+1)P_{N+1}-[(\lambda+\mu)N+\nu]P_N+[\lambda(N-1)+\nu]P_{N-1} \tag{3}$$

where $P_N(t)$ is the probability of finding a population of N individuals at time t. A

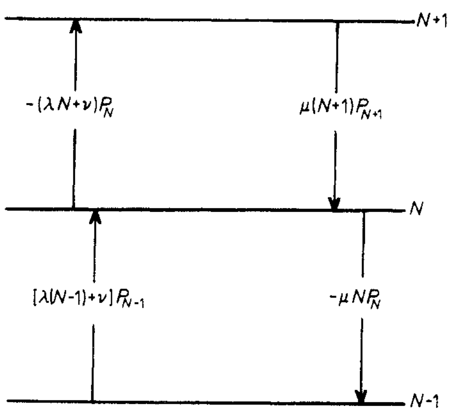

Figure 1. Population transitions leading to negative binomial number fluctuations.

partial differential equation for the generating function

$$Q(z,t)=\langle(1-z)^N\rangle=\sum_{N=0}^{\infty}(1-z)^N P_N(t) \tag{4}$$

may be deduced from (3) without difficulty:

$$\frac{\partial Q}{\partial t}=z[-\mu+\lambda(1-z)]\frac{\partial Q}{\partial z}-\nu z Q. \tag{5}$$

If it is assumed that M individuals are present initially then equation (5) must be solved subject to the boundary conditions

$$Q(0,t)=1, \qquad Q(z;0)=(1-z)^M. \tag{6}$$

The first of these conditions is just an expression of the unit normalisation of $P_N(t)$ whilst the second results from the initial condition $P_N(0)=\delta_{NM}$. The transient solution for the population statistics is (e.g. Bartlett 1966)

$$Q(z,t)=\left(\frac{\lambda-\mu}{\lambda-\mu+\lambda(\theta-1)z}\right)^{\nu/\lambda}\left(\frac{\lambda-\mu+(\mu\theta-\lambda)z}{\lambda-\mu+\lambda(\theta-1)z}\right)^M \tag{7a}$$

where

$$\theta(t)=\exp[(\lambda-\mu)t]. \tag{7b}$$

An equilibrium distribution for large times exists if the death rate μ is greater than the birth rate λ. Setting $\theta=0$ in equation (7) gives

$$Q(z,\infty)=(1+\bar{N}z/\alpha)^{-\alpha} \tag{8a}$$

where

$$\bar{N}=\nu/(\mu-\lambda), \qquad \alpha=\nu/\lambda. \tag{8b}$$

This is the generating function for the class of negative binomial distributions defined by equation (2) of the last section. The normalised higher-factorial moments may be evaluated from (8) by repeatedly differentiating with respect to z and then setting $z=0$:

$$\langle N(N-1)\dots(N-r+1)\rangle/\bar{N}^r=\prod_{n=0}^{r-1}\left(1+\frac{n}{\alpha}\right). \tag{9}$$

In the equilibrium regime all the higher-order temporal coherence properties of the population number fluctuations can in principle be calculated from equation (3) via the transient solution (7). For example, the joint distribution $P_{MN}(0,t)$ of finding M individuals present at time zero and N at time t may be expressed in the form

$$P_{MN}(0,t)=P_M P_{(M)N}(t) \tag{10}$$

where P_M is the equilibrium probability of finding M individuals present, given by equations (2) and (8), and $P_{(M)N}(t)$ is the probability of finding N individuals present at time t conditional on the presence of M at time zero as given by the transient solution (7). The joint generating function

$$Q(z,0;z',t)=\langle(1-z)^M(1-z')^N\rangle=\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}(1-z)^M(1-z')^N P_M P_{(M)N}(t) \tag{11}$$

can be evaluated analytically and expressed in terms of θ, $\bar{N}$ and , (equations (7) and (8)) as

$$Q(z, 0; z', t) = [(1+\bar{N}z/\alpha)(1+\bar{N}z'/c\ ) - zz'\vartheta(1+\bar{N}/\alpha)\bar{N}/\alpha]^{-}\ , \tag{12}$$

and this result readily yields the bilinear moment (correlation) function of the number fluctuations

$$\frac{\langle N(0)N(t)\rangle}{\bar{N}^2} = 1 + \left(\frac{1}{\alpha} + \frac{1}{\bar{N}}\right)\theta(t). \tag{13}$$

3. Limit distributions

The results of the last section will now be used to evaluate the limit distributions generated by a random walk with variable step number in the limit of large mean step number. In order to make the calculations explicit and to facilitate comparison with experimental data, only the two-dimensional walk will be considered. No additional complications appear to arise in the case of other dimensionalities.

3.1. Single-fold statistics

Consider the N-step two-dimensional random walk

$$\mathscr{E}(t) = \sum_{j=1}^{N} a_j(t) \exp(\mathrm{i}\phi_j(t)) = A(t) \exp(\mathrm{i}\Phi(t)) \tag{14}$$

where $\{a_j(t)\}$ and $\{\phi_j(t)\}$ are statistically independent sets of random variables describing the length and orientation, respectively, of the steps as a function of time. It will be assumed that different members of each set are statistically identical but independent, and that the $\{\phi_j(t)\}$ are uniformly distributed over 2π radians. The phase $\Phi(t)$ of the resultant vector will then also be uniformly distributed, and assuming first that N is fixed, its characteristic function can be expressed in terms of the zeroth-order Bessel function

$$C_N(\boldsymbol{u}) = \langle \exp[\mathrm{i}(u_1\mathscr{E}_1 + u_2\mathscr{E}_2)]\rangle = \langle J_0(ua)\rangle^N = \langle J_0(uA)\rangle \tag{15}$$

where the subscripts indicate real and imaginary parts and $u = |\boldsymbol{u}| = (u_1^2 + u_2^2)^{1/2}$. In order to derive the asymptotic distribution for the case of large N it is convenient to renormalise the step length a by a factor $\sqrt{N}$ and examine the limit of the expression $\langle J_0(ua/\sqrt{N})\rangle^N$ as N tends to infinity. It is not difficult to show that, *whatever the step-length distribution*,

$$\lim_{N\to\infty} C_N(u) = \exp(-\tfrac{1}{4}u^2\langle a^2\rangle) \tag{16}$$

corresponding to the Rayleigh distribution

$$p(A) = 2(A/\langle A^2\rangle)\exp(-A^2/\langle A^2\rangle). \tag{17}$$

Defining the intensity by

$$I = A^2, \tag{18}$$

equation (17) is easily converted into the intensity fluctuation distribution

$$p(I) = (1/\langle I\rangle)\exp(-I/\langle I\rangle) \tag{19}$$

with normalised moments

$$n^{[r]} = \langle I^r\rangle/\langle I\rangle^r = r!. \tag{20}$$

This result expresses the fact that in the large-N limit the real and imaginary parts of the vector $\mathscr{E}$ are zero-mean, independent Gaussian variables as predicted by the central limit theorem. A corollary of this result is that if the step lengths $\{a_j\}$ in the random walk (14) are Rayleigh distributed then so is the resultant amplitude A (but with scaled mean square $\langle A^2\rangle = N\langle a^2\rangle$) *whatever the value of N*. The distribution (17) is thus *stable* with respect to the convolution (15).

Now suppose that N is itself a statistical variable, independent of the $\{a_j\}$ and $\{\phi_j\}$ and distributed according to equation (2). Averaging equation (15) over the fluctuations in N and renormalising the step lengths by a factor $\sqrt{\bar{N}}$ leads to

$$C_{\bar{N}}(u) = [1 + (\bar{N}/\alpha)(1 - \langle J_0(ua/\sqrt{\bar{N}})\rangle)]^{-\alpha}. \tag{21}$$

The associated limit distributions can easily be derived:

$$\lim_{\bar{N}\to\infty} C_{\bar{N}}(u) = [1 + u^2\langle a^2\rangle/4c\ ]^{-} \tag{22}$$

corresponding to the K distributions

$$p(A) = \frac{2b}{\Gamma(\alpha)}\left(\frac{bA}{2}\right)^{o} K_{\alpha-1}(bA) \tag{23}$$

with $b = 2(\alpha/\langle A^2\rangle)^{1/2}$ and normalised intensity moments (cf equation (20))

$$n^{[r]} = r!\Gamma(r+\alpha)/\alpha^r\Gamma(\alpha). \tag{24}$$

Like equation (17) the result (23) holds *whatever the properties of the individual steps*. It is interesting that if the step lengths $\{a_j\}$ appearing in equation (14) are K-distributed then so is the resulting amplitude *for any fixed N*. The class (23) is, in fact, *infinitely divisible* with respect to the convolution (15): a useful analytical property which derives in part from the infinite divisibility of the underlying negative binomial distribution. Unlike the Rayleigh distribution (17), however, K distributions are not stable, because increasing the number of K-distributed steps in the random walk (16) increases the *order* of the distribution (23) of the resultant amplitude as well as its mean square (Jakeman and Pusey 1978). The asymptotic behaviour of (23) and (24) with large α should thus be equivalent to that obtained by taking N large in the usual random walk problem. This is indeed the case, for as α tends to infinity these two formulae reduce to equations (17) and (20) predicted by the central limit theorem.

The functions (23) have rarely figured in the literature as probability distributions (Nakagami and Ota 1957, Beckmann 1967) but are commonly used as model structure (correlation) functions following the suggestion of Tatarski (1961). The normalised moments (24) provide a useful basis for comparison with experiment as they depend only on the single parameter α. The second moment

$$n^{[2]} = 2(1+\alpha^{-1}) \tag{25}$$

is a convenient measure of the deviation from Gaussian statistics against which the higher moments can be plotted. Figure 2 shows moments plotted in this way for

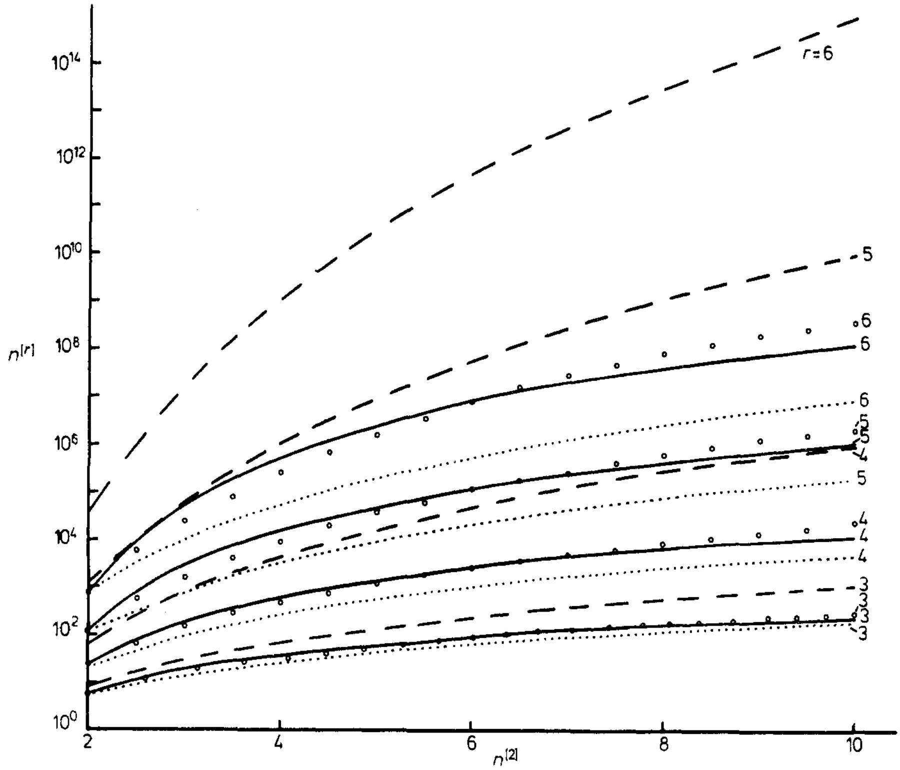

Figure 2. Higher-order normalised intensity moments as a function of the second moment for various distributions: broken curves, lognormal; open circles, Weibull; dotted curves, gamma; full curves, K.

comparison with other non-Gaussian distributions. It can be shown that the normalised moments of K distributions always lie between those of the Rayleigh and lognormal distributions with the same values of $n^{[1]}$ and $n^{[2]}$ (Jakeman and Pusey 1976). Clearly the Weibull distributions

$$p(A)=2\beta bA(bA^2)^{\beta-1}\exp[-(bA^2)^\beta] \tag{26}$$

with $b=\Gamma(1+\beta^{-1})/\langle A^2\rangle$ and normalised intensity moments

$$n^{[r]}=\Gamma(1+r/\beta)/\Gamma^r(1+\beta^{-1}) \tag{27}$$

are close to K distributions for a wide range of second moment values. Indeed they are identical for second moments of two ($\alpha\to\infty, \beta=1$) and six ($\alpha=\frac{1}{2}, \beta=\frac{1}{2}$). However, their analytical properties are less attractive than those of K distributions (in particular they are not infinitely divisible) and it is not clear how the theory of their higher-order joint distributions should be formulated. It will now be shown that in the case of K distributions these can be derived on the basis of the rate equation (3).

3.2. Correlation properties

Consider first the correlation function $\langle \mathscr{E}(t)\mathscr{E}(t')\rangle$ with $t'=t+\tau$ and $\tau>0$. Out of N steps present in the walk at the initial time t, N_s will survive to the later time t', $N-N_s$

steps will disappear and $N' - N_s$ new steps will appear, where N' is the number of steps at time t'. The correlation function of the complex amplitude can thus be expressed in the form

$$\langle \mathscr{E}(t)\mathscr{E}^*(t')\rangle = \left\langle \left(\sum_s a_s\, \mathrm{e}^{\mathrm{i}\phi_s} + \sum_d a_d\, \mathrm{e}^{-\mathrm{i}\phi_d}\right)\left(\sum_s a_s'\, \mathrm{e}^{\mathrm{i}\phi_s'} + \sum_n \mathrm{e}^{-\mathrm{i}\phi_n'}\right)\right\rangle \tag{28}$$

where s labels the members of the original population which survive to time t', d labels those which die and n the new members of the population which appear during the period τ. Since the sets $\{\phi_s\}$, $\{\phi_d\}$ and $\{\phi_n\}$ are statistically independent of each other and $\{\phi_i\}$ are uniformly distributed, (28) immediately reduces to

$$\langle \mathscr{E}(t)\mathscr{E}^*(t')\rangle = \langle N_s\rangle\langle a(t)a(t')\rangle\langle \exp[\mathrm{i}(\phi(t) - \phi(t'))]\rangle \tag{29}$$

and since the original population is depleted by a statistically random (Bernoulli) death process the normalised first-order correlation function may finally be written as

$$g^{(1)}(\tau) = \exp(-\mu\tau)\langle a(0)a(\tau)\rangle\langle \exp[\mathrm{i}(\phi(0) - \phi(\tau))]\rangle/\langle a^2\rangle. \tag{30}$$

The second-order or intensity correlation function $\langle I(t)I(t')\rangle$ can be derived using a similar approach. A straightforward calculation gives the normalised form

$$g^{(2)}(\tau) = 1 + \left(1 + \frac{1}{\alpha}\right)|g^{(1)}(\tau)|^2 + \left(\frac{1}{\alpha} + \frac{1}{\bar{N}}\right)\theta(\tau) + \frac{1}{\bar{N}}\left(\frac{\langle a^2(0)a^2(\tau)\rangle}{\langle a^2\rangle^2} - 1\right)\exp(-\mu\tau) \tag{31}$$

where $g^{(1)}(\tau)$ is defined by equation (30) and $\theta(\tau)$ by equation (7). Formulae (30) and (31) are exact consequences of the negative binomial assumption. In order to generate the correlation properties associated with the distributions (23) (i.e. with K-distributed noise), $\bar{N}$ must be increased without limit as in the derivation of equation (22) from equation (21). The first-order correlation function is unaffected by this procedure, being independent of $\bar{N}$, but the intensity correlation function becomes

$$g^{(2)}(\tau) = 1 + \left(1 + \frac{1}{\alpha}\right)|g^{(1)}(\tau)|^2 + \frac{1}{\alpha}\theta(\tau), \tag{32}$$

which is identical to the result (25) for the second moment when $\tau = 0$. The relation (32) replaces the familiar factorisation theorem for a complex Gaussian field (the Siegert relation) and reduces to it only in the limit of large α. Both the lifetime of the steps and their cross-section fluctuations enter into formula (30) for the field correlation function, but orientational (phase) fluctuations may often dominate the time dependence of this quantity. Two types of contribution to the second-order correlation function (32) can be identified: a number fluctuation term $\theta(\tau)$ and an 'interference' term proportional to $|g^{(1)}(\tau)|^2$. The time dependence of these two contributions will usually be quite different.

The higher-order joint distribution properties of the resultant of a random walk with variable step number may also be calculated by following a procedure similar to that leading to equation (28). Thus the characteristic function may be reduced as follows:

$$\begin{aligned} C_{\bar{N}}(\boldsymbol{u}, \boldsymbol{v}) &= \langle \exp[\mathrm{i}(u_1\mathscr{E}_1 + u_2\mathscr{E}_2 + v_1\mathscr{E}_1' + v_2\mathscr{E}_2')]\rangle \\ &= \langle A^{N_s}B^{N-N_s}D^{N'-N_s}\rangle \end{aligned} \tag{33a}$$

where

$$A = \langle \exp[\mathrm{i}(au_1\cos\phi + au_2\sin\phi + a'v_1\cos\phi' + a'v_2\sin\phi')]\rangle, \tag{33b}$$

$$B = \langle \exp[\mathrm{i}(au_1\cos\phi + au_2\sin\phi)]\rangle, \tag{33c}$$

$$D = \langle \exp[\mathrm{i}(a'v_1\cos\phi' + a'v_2\sin\phi')]\rangle. \tag{33d}$$

The averages in (33) are carried out over the step lengths and orientations, subscripts label real and imaginary parts and the primed and unprimed parameters refer to the later and earlier times as before. Evaluation of equation (33*a*) appears to be difficult except in certain limiting cases and warrants further investigation.

3.3. *Related distributions*

As a first example, consider the *incoherent* random walk corresponding to equation (14):

$$I(t)=\sum_{j=1}^{N} a_j^2(t). \tag{34}$$

The Laplace transform of the probability distribution of this quantity is given by

$$Q_N(s)=\langle\exp(-sI)\rangle=\langle\exp(-sa^2)\rangle^N \tag{35}$$

and the associated limit distribution can be derived in the usual way by dividing a by $\sqrt{N}$ (so that $\langle I\rangle=\langle a^2\rangle$) and taking N large:

$$\lim_{N\to\infty} Q_N(s)=\exp(-s\langle a^2\rangle), \qquad p(I)=\delta(I-\langle I\rangle). \tag{36}$$

The intensity (34) is therefore constant in this limit. On the other hand, if N varies according to the distribution (2), the average form of equation (35) after scaling a by a factor $\sqrt{\bar{N}}$ is

$$Q_{\bar{N}}(s)=[1+(\bar{N}/\alpha)(1-\langle\exp(-sa^2/\bar{N})\rangle)]^{-\alpha} \tag{37}$$

and the Laplace transform of the associated limit distribution is given by

$$\lim_{\bar{N}\to\infty} Q_{\bar{N}}(s)=(1+s\langle I\rangle/\alpha)^{-\alpha}. \tag{38}$$

Inverse Laplace transformation of this formula generates the class of gamma distributions

$$P(I)=\alpha(\alpha I/\langle I\rangle)^{\alpha-1}\exp(-\alpha I/\langle I\rangle)/\langle I\rangle\Gamma(\alpha), \tag{39}$$

which is just the continuous analogue of the discrete negative binomial class (2). These distributions are discussed extensively in the literature. When $\alpha=\frac{1}{2}n$ and $\langle I\rangle=2\sigma$ the intensity is said to be a chi-square variate with n degrees of freedom and if n is integral it can be expressed as the sum of the squares of n Gaussian variables. The class (39) has often been proposed as a model for non-Gaussian statistics and has also appeared in studies of infinite divisibility (Gnedenko and Kolomogorov 1954). The term m distribution was coined for the sub-class $n\geq 1$ in an early detailed analysis of their properties (Nakagami 1943). The normalised moments of (39) are given by

$$n^{[r]}=\langle I^r\rangle/\langle I\rangle^r=\Gamma(r+\alpha)/\alpha^r\Gamma(\alpha) \tag{40}$$

and are smaller by a factor $r!$ than those of the corresponding K distributions (figure 2). In fact, it is possible to show that a K-distributed variable can be factored into the product of two independent gamma variates (Nakagami and Ota 1957), although this simple factorisation cannot be extended to the *correlation properties* of K-distributed noise generated through the rate equation (3).

The joint statistics of a gamma variate can be derived from (34) using the rate equation (3) and the procedure of the last section. For example, the joint Laplace transform $Q(s, s') = \langle \exp[-(sI(t) + s'I(t'))] \rangle$ and corresponding distribution are given by

$$Q(s, s') = [(1 + s\langle I \rangle/\alpha)(1 + s'\langle I \rangle/\alpha) - ss'\langle I \rangle^2 \ (\tau)/ \ ^2]^{-} , \tag{41}$$

$$P(I, I') = \frac{2}{\langle I \rangle^2 \Gamma(\alpha)(1-\theta)} \left(\frac{\alpha}{\langle I \rangle} \sqrt{\frac{II'}{\theta}} \right)^{-1} \exp\left(-\frac{\alpha(I + I')}{\langle I \rangle (1-\theta)} \right) I_{\alpha-1}\left(\frac{2\alpha\sqrt{II'\theta}}{\langle I \rangle (1-\theta)} \right), \tag{42}$$

where $I_{\alpha-1}$ is a modified Bessel function of the first kind of order $\alpha - 1$. The normalised first-order intensity correlation function is

$$\langle II' \rangle / \langle I \rangle^2 = 1 + \theta(\tau)/\alpha, \tag{43}$$

which may be obtained directly from equation (32) by neglecting the interference term proportional to $|g^{(1)}(\tau)|^2$.

As a second example of distributions related to class (1), the statistics of K-distributed noise plus a constant amplitude component will be considered briefly. In the case of coherent addition of K-distributed noise with a randomly phased but constant amplitude 'step' (homodyning), the characteristic function of the limit distribution from equation (22) is (a_0 is the constant amplitude component)

$$\lim_{\bar{N} \to \infty} C_{\bar{N}}(u) = J_0(ua_0)/(1 + u^2\langle a^2 \rangle/4\alpha)^{\alpha}. \tag{44}$$

Inversion of this formula is difficult but the moments can be expressed as finite sums:

$$n^{[r]} = \frac{(r!)^2}{(1+x)^r} \sum_{n=0}^{r} \binom{n + \alpha - 1}{n} \frac{(x/\alpha)^n}{[(r-n)!]^2} \tag{45a}$$

where

$$x = \langle a^2 \rangle / a_0^2 \tag{45b}$$

is the ratio of the mean intensity of the K-distributed noise to the intensity of the constant amplitude component. Thus, for example,

$$\langle A^2 \rangle = a_0^2(1 + x), \tag{46}$$

$$n^{(2)} = [1 + 4x + 2x^2(1 + \alpha^{-1})]/(1 + x)^2. \tag{47}$$

Plots of formula (47) for the second moment against x take an interesting form if $\alpha = x/\epsilon$ with ϵ constant. Curves for various values of ϵ are shown in figure 3, and will be discussed futher in the next section. Note, however, that $n^{(2)}$ rises from unity at $x = 0$ to a maximum value of $(\epsilon^2 + 4\epsilon + 2)/(2\epsilon + 1)$ at $x = (1 + \epsilon)/\epsilon$, then decreases asymptotically to 2 as x becomes large.

4. Comparison with experiment

The class of distributions (1) was originally proposed as a model for non-Gaussian microwave sea echo (Jakeman and Pusey 1976). When a radar illuminates a large area of the sea it is usually found that the probability distribution of the envelope of the return signal can be approximated by the Rayleigh distribution (17). This is a

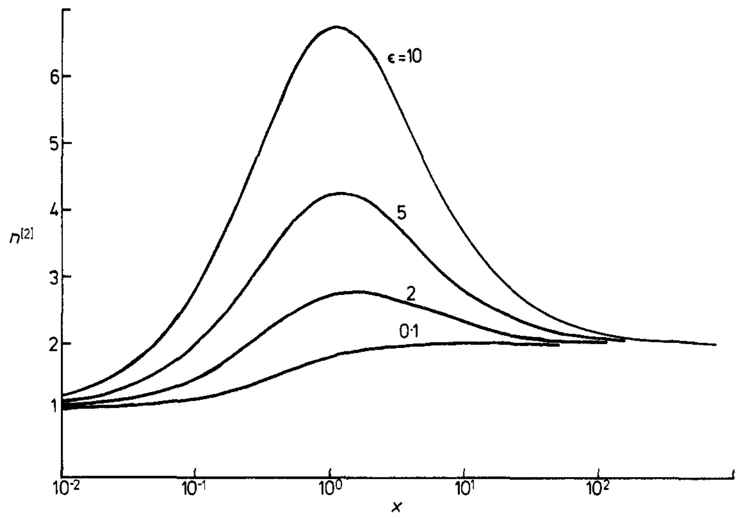

Figure 3. Second normalised intensity moment of homodyned K-distributed noise (see equation (47) of text).

consequence of the central limit theorem since the signal can be thought of as being the vector sum, equation (14), of randomly phased components from a large number of independent 'scatterers'. However, by using a narrow beam-divergence and short pulse-length, it is possible to illuminate areas of the sea of linear dimensions comparable with the longer wavelengths of the sea surface. Under these conditions large deviations from Rayleigh statistics are often found and the original idea was to represent the scattered radiation by equation (14) with N relatively small. The expression (15) for the generating function then represents a formal solution of the scattering problem exhibiting the correct dependence on N which is itself expected to be proportional to the illuminated area. For modelling purposes a simple analytical form for the amplitude statistics is required, however, and this necessitates a distribution for the $\{a_j\}$ for which the convolution (15) can be inverted. As already mentioned in § 3.1, K distributions fulfil this requirement. They also have moments which lie between those of the Rayleigh distribution and those of the lognormal distribution with the same mean and variance—a property shared with experimental sea-echo measurements. Bearing in mind the *ad hoc* assumptions of the model, the agreement between experimental data and K-distributed noise shown in figure 4 is surprisingly good. Recent analysis (Jakeman and Pusey 1977) of the area dependence of the statistics has not, however, confirmed the tenets of the original model, which clearly predicts that (for fixed N)

$$n^{[2]} = 2(1 + 1/\alpha N) \tag{48}$$

where α is the order of the K distributions characterising the $\{a_j\}$. Thus if the illuminated area is doubled the deviation from the Gaussian value of two should be halved. Such behaviour is not supported by the experimental data, which indeed show the reverse trend in some cases. This is almost certainly due to instrumental difficulties such as system noise, which affects the accuracy of the measured amplitude distribution for small amplitude fluctuations and hence the normalisation of the lower moments. However, if, as seems likely, the K-distributed nature of the amplitude fluctuations

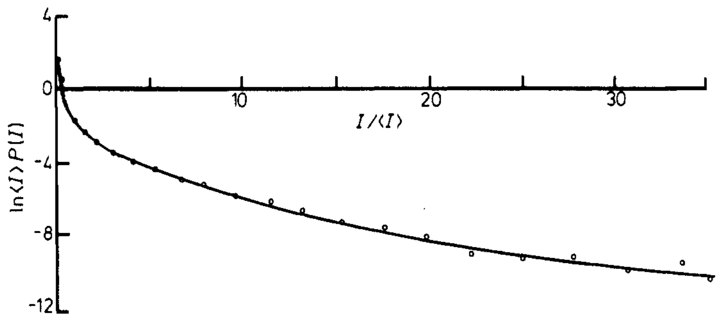

Figure 4. Experimental sea-echo data (○) compared with K-distributed noise (full curve, $\alpha = 0{\cdot}5$) (Jakeman and Pusey 1977).

derives from spatial bunching of the scatterers on the sea surface rather than from the properties of the individual scattering centres (i.e. the $\{a_j\}$) then the area dependence implied by equation (48) will only arise when the illuminated area exceeds the characteristic bunching length.

Comparison of experimental data with the higher moments of the intensity fluctuation distribution (figure 2) affords a more sensitive measure of goodness of fit than comparison with the distributions themselves shown in figure 4. Because of normalisation difficulties, mentioned earlier, large uncertainties appear in high-order moments calculated from the sea-echo data. Measurements of optical frequency scintillation using photon counting techniques, on the other hand, are completely free from the spurious noise and nonlinearities with bedevil both detection at microwave frequencies and analogue processing methods (Cummins and Pike 1974) and it is therefore data from light scattering experiments which provide the most convincing evidence in support of K-distributed noise. The first experiments of this kind, specifically designed to investigate enhanced non-Gaussian fluctuations, measured the scintillation of laser light scattered into the Fraunhofer region by a thin layer of nematic liquid crystal in turbulent motion (Pusey and Jakeman 1975). The system is discussed at length in earlier papers but, briefly, consists of a 25 μm thick layer of the nematic liquid crystal MBBA contained between two glass slides. The initially clear, aligned state is driven into turbulent motion by the application of a potential difference of 15 or 20 volts across the layer. This perturbs the direction of the local optic axis and causes strong scattering of incident light, the sample taking on the appearance of ground glass. In this 'dynamic scattering' mode the layer behaves as a deep random phase screen, introducing into an incident plane-wave path differences of the order of an optical wavelength, which vary randomly in time and with position in the illuminated area. When the latter is only the order of 10 μm or so, large non-Gaussian fluctuations in the scattered intensity occur in the Fraunhofer region. The data taken in these experiments were originally interpreted in terms of a fixed-N, facet model which gave only moderate agreement with the measured higher-order single-fold statistical properties. Comparison with K-distributed noise was made only recently and is reproduced in figure 5. Agreement is excellent even for the fifth normalised moment.

The number fluctuation model described in the preceding sections also makes predictions concerning the second-order or intensity correlation function, namely equation (32). Two complications prevent direct comparison of this formula with data on the liquid crystal system, however. Firstly, the deviation of the second normalised

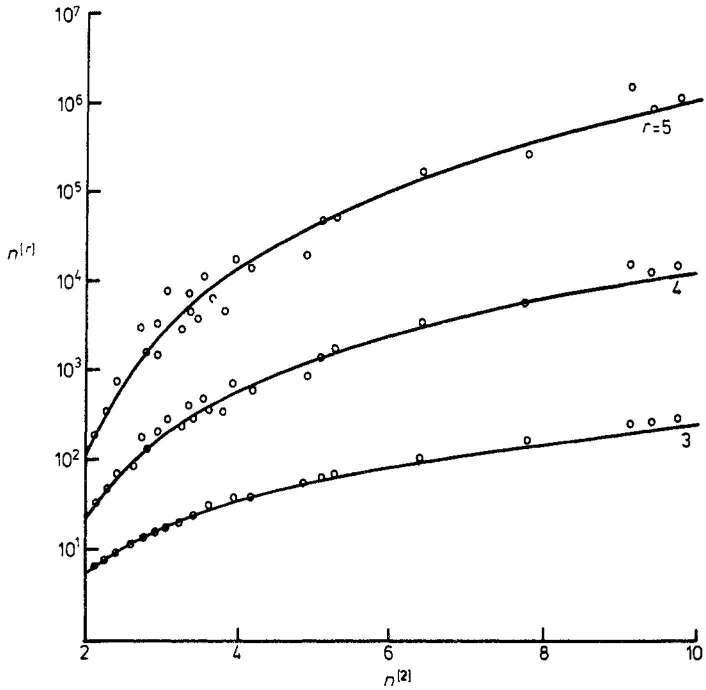

Figure 5. Liquid crystal scattering data compared with the moments of K-distributed noise (Jakeman and Pusey 1978).

intensity moment from the Gaussian value of two was found experimentally to be inversely proportional to the illuminated area, suggesting that this was still much larger than the outer scale (bunching length) of the turbulence. Secondly, although the measured correlation function exhibited decay on two time scales, the size of the two contributions was estimated by cross-correlation of the output between two spatially separated detectors (Pusey and Jakeman 1975). Although some qualitative predictions of the effect of a spatial distribution of scatterers on the statistics can be made, these are not included in equation (32). Moreover, this formula is restricted to detection at a single space point.

More recent optical scintillation experiments using photon counting techniques have concentrated on Fresnel region effects. Laboratory experiments have investigated the statistical properties of light scattered by strongly turbulent layers of air and water as a function of propagation distance beyond the scattering region (Parry *et al* 1977, Parry and Gray 1979). Outdoor experiments have included the measurement of scintillation effects which occur when a laser beam propagates through an extended region of turbulence in the atmosphere (Parry and Pusey 1979) and also measurements of the twinkling of starlight (Jakeman *et al* 1978, Parry and Walker 1979). In both the layer and extended region experiments the normalised second intensity moment varies with propagation distance in a manner rather similar to figure 3. It first increases from unity at short propagation distances (where there are only phase fluctuations) to a maximum typically greater than the Gaussian value of two at a distance possibly corresponding to the focussing length of the larger-scale fluctuations in refractive index. Thereafter a decrease towards two at large propagation distances is observed, as interference between contributions from many independent scatterers begins to dominate the statistics. K distributions seem to provide a good fit to the measured intensity statistics in the peak region and beyond. Figure 6 shows various data taken in this region from the above-mentioned systems compared with the moments of K-distributed noise. Preliminary results only are given for atmospheric propagation measurements. Data taken in a more recent comprehensive set of experiments confirm the significance of K distributions in the context of propagation through *extended* regions of turbulent media, but this will be published elsewhere (Parry and Pusey 1979). Although intensity fluctuations in the region preceding the focussing peak are not found to be K-distributed, certain definite trends were observed in the layer scattering systems. These enabled the detection position in the stellar scintillation measurements to be established as being in a sub-region immediately preceding the peak. Here the fluctuations appear to lie between lognormal and K, whilst very close to the scattering layer, the higher-order normalised intensity moments appear to increase faster than lognormal. Between these two sub-regions lies a short range of propagation paths for which lognormal statistics were observed. These observations, together with the close resemblance between the measured focussing curve and figure 3, suggest that the coherent addition of K-distributed noise and a constant amplitude component might provide a useful model for the statistics of radiation scattered by a turbulent layer over the whole range of propagation distances from the scatterer to the far field. Figure 7 compares experimental data obtained when a laser beam is scattered by a turbulent layer of air with the moments defined by equation (45). The two parameters x and α are obtained by fitting the second and third normalised intensity moments and the suitability of the model is measured by the agreement between experiment and theory

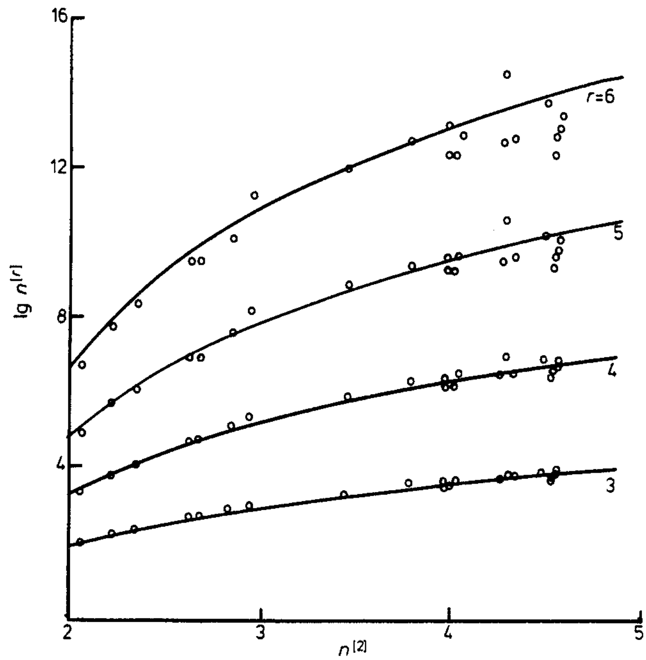

Figure 6. Data from various turbulence scattering experiments compared with the moments of K-distributed noise (Parry *et al* 1977, together with additional unpublished data).

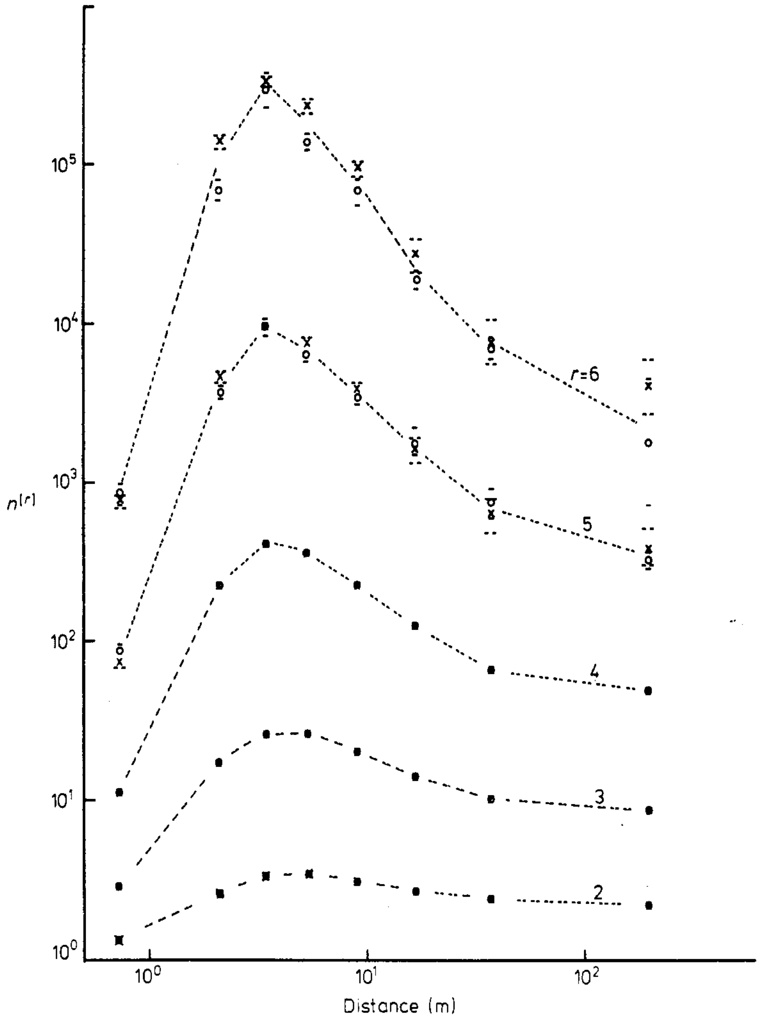

Figure 7. Normalised moments of the measured intensity fluctuation distribution of laser light scattered by a turbulent layer of air (○) plotted as a function of propagation distance: comparison with homodyned K-distributed noise (×) (Parry *et al* 1977, together with additional unpublished data; the dotted lines are drawn as an aid to data identification).

for the fourth, fifth and sixth moments. Figure 8 shows log plots of x and α as a function of propagation distance and reveals an extensive region of linear behaviour for both these quantities. Analysis of data taken in experiments on a turbulent layer of water in terms of the same statistical model gives almost identical plots, showing linear regions with the same slope. This suggests that there may be some underlying physical basis for using homodyned K distributions to describe intensity fluctuations in the Fresnel region of a turbulent scattering layer. Further evidence in support of this conjecture is provided by the knowledge that Rice statistics (homodyned Gaussian noise) is predicted theoretically for intensity fluctuations in the Fresnel region far from a random phase screen characterised by a *single* length scale (Gaussian phase correlation function). A discussion of this problem is presented elsewhere (Jakeman and McWhirter 1977).

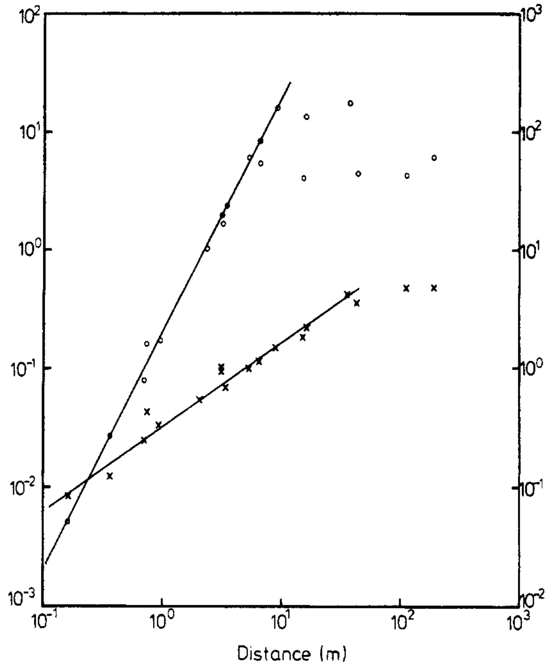

Figure 8. Parameters x (○) and α (×) used in equation (45) to obtain figure 7 and to fit some additional data on the same scattering system, plotted as a function of propagation distance.

As a final application of the results given in § 3, a comparison is shown in figure 9 of data on the intensity statistics of white light scattered by a turbulent layer of air and the gamma distributions (39). These measurements were not made under identical conditions to the laser scattering experiments so that confirmation of the ratio between the normalised moments (24) and (40) is not possible. Moreover, the bandwidth of the white light was unfortunately insufficient to eliminate interference effects completely, i.e. the scattering process was not fully incoherent. Nevertheless the data compare favourably with the normalised moments (40), which is encouraging if not conclusive. As in the case of figure 6, only data from the focussing peak in the second-moment plot and beyond are included. In order to characterise the statistics for all propagation distances in the white light experiments, the sum of a constant amplitude component plus gamma-distributed noise would seem appropriate by analogy with the coherent illumination case.

5. Discussion

The aim of this paper has been to show that the inclusion of step number fluctuations in the random walk problem can lead to new and useful classes of limit distribution. An in-depth study of the general problem has not been attempted but some simple limiting statistical properties associated with negative binomial number fluctuations have been calculated to demonstrate the interesting possibilities of this type of random walk. Quite apart from investigations of limit distributions arising from other kinds of number fluctuations, much further work is required to establish, for K-distributed noise, the full

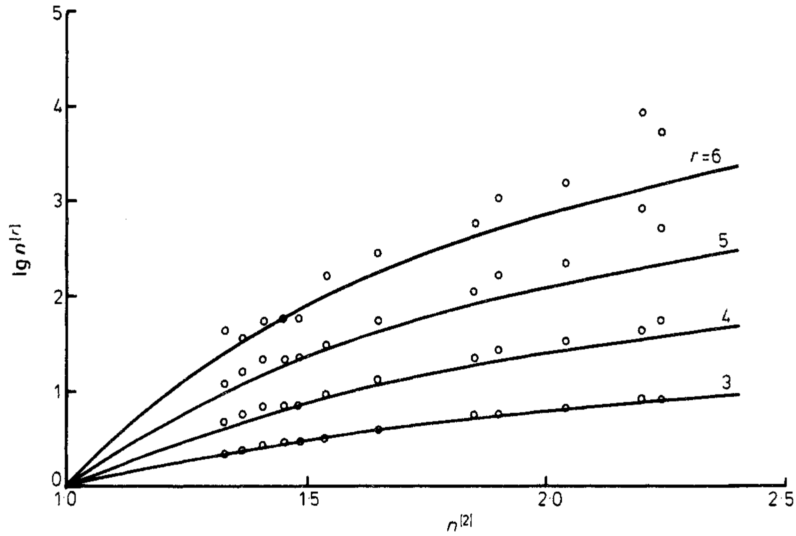

Figure 9. Data on the statistics of white light scattered by a turbulent layer of air compared with the moments of gamma-distributed noise (Parry *et al* 1977, together with additional unpublished data).

range of statistical properties (distributions of integrated intensity, level crossings, maxima, etc) presently available in the case of the more familiar Gaussian noise. The choice of model for the present study has been governed by the requirements of mathematical tractability and physical relevance, and these two points warrant some further comment.

Taking first the question of mathematical tractability, three considerations have determined the course of the work: (i) the existence of a closed form for the limit distributions in terms of familiar tabulated mathematical functions, (ii) the infinite divisibility of these limit distributions and (iii) the ability to calculate higher-order (joint) statistical properties. Closely related to (ii) is a requirement in modelling applications that the fixed-N convolution (15) should not introduce an additional parameter into the theory. This can be achieved by choosing any discrete compound Poisson distribution for the number fluctuations (Feller 1968) leading to a generalised form of equation (21):

$$C_{\bar{N}}(u) = \exp[-\bar{M}(1 - \overline{\langle J_0(ua/\sqrt{\bar{N}})\rangle^n})]. \tag{49}$$

Here the bar indicates averaging over the fluctuations in n, and $\bar{M}$ is a parameter which scales under the convolution (15), i.e. $\bar{M} \to N\bar{M}$. Formula (49) reduces to the negative binomial result (21) when n is logarithmically distributed: $p(n) = q^n/n \ln[1/(1-q)]$. However it is not easy to find other examples based on (49) which also satisfy requirement (i) above. By contrast it is easy to construct models satisfying requirement (i) if the 'no additional parameter' constraint is relaxed. Thus the uniform distribution $p(N) = (R+1)^{-1}$ for $N \leqslant R$ gives in place of (22)

$$\lim_{\bar{N}\to\infty} C_{\bar{N}}(u) = [1 - \exp(-2u^2\langle a^2\rangle)]/2u^2\langle a^2\rangle, \tag{50}$$

corresponding to the exponential-integral limit distribution

$$p(A) = (A/\langle A^2\rangle)E_1(A^2/2\langle A^2\rangle) \tag{51}$$

with normalised intensity moments

$$n^{[r]} = r!2^r/(r+1) \tag{52}$$

where

$$E_1(x) = \int_x^\infty \frac{e^{-y}}{y}\,dy. \tag{53}$$

The single parameter $\langle A^2\rangle$ characterising (51) clearly does not scale under the convolution (15), which consequently introduces N as a second parameter. Moreover, further assumptions are needed before high-order (joint) statistical properties can be calculated and requirement (iii) above satisfied.

Turning now to the question of the relevance of K-distributed noise to physical systems, it is evident that the limit distributions generated by negative binomial number fluctuations in the two-dimensional random walk problem provide an accurate model for the single-interval statistics of radiation scattered by a range of systems involving turbulent media. Although the random walk problem occurs widely throughout physics (and indeed in many other branches of science), so that K distributions may well prove useful in other fields, it is the experimental data from these turbulent scattering systems which have motivated the present work and it is the bunching of scattering centres in space and time which is held to be responsible for the particular statistics of the scattered radiation. Unfortunately, presently available experimental data do not adequately test the negative binomial model, but the bunching of scatterers in multi-scale systems is an intuitively reasonable hypothesis: both the modulation of small wavelets on the sea surface by an underlying larger-scale structure and intermittancy in turbulently mixing systems would undoubtedly lead to clustering of scattering centres. It is significant in this context that fluctuations in light scattered by specially designed single-scale surfaces are typically not K-distributed according to recent experimental evidence (Parry and Gray 1979). The population process defined by equation (3) is, of course, by no means the only model for number fluctuations (although it may be the simplest one satisfying requirements (i)–(iii) above). It does appear to be consistent, however, with a picture of turbulence in which large eddies are 'spontaneously created' and then 'give birth' to generations of smaller daughter eddies through a cascade process which terminates when the smallest eddies 'die' due to viscous dissipation. According to § 3 the size of the eddies (step length or scattering power) would not affect the limit distribution of the amplitude of scattered radiation so that the rate equation (3) would completely characterise the scattering process (apart from spatial effects mentioned in § 4) in the limit of high population levels.

To summarise, it has been shown in this paper that negative binomial fluctuations in the number of steps in a two-dimensional random walk lead to a new class of limit distributions, the K distributions, which appears to provide a good model for the single-interval statistics of radiation scattered from turbulent media. A simple population model reminiscent of the cascade description of turbulence has been used to generate higher-order joint statistical properties but no attempt has been made to justify the model through a rigorous mathematical formulation of the scattering problem. Some progresś on this front is reported elsewhere (Hoenders *et al* 1979). It has been demonstrated that every compound Poisson distribution of step number fluctuations generates a class of infinitely divisible limit distributions. The negative binomial model appears to be the simplest example of physical interest for which the random walk problem can be solved in terms of tabulated functions.

Acknowledgments

The author is indebted to Drs G Parry and P N Pusey, not only for advice and encouragement during the preparation of this work but also for making available much new and unpublished data.

References

Bartlett M S 1966 *An Introduction to Stochastic Processes* (Cambridge: Cambridge University Press) ch 3
Beckmann P 1967 *Probability in Communication Engineering* (New York: Harcourt, Brace and World) p 167
Cummins H Z and Pike E R (ed.) 1974 *Photon Correlation and Light-Beating Spectroscopy* (New York: Plenum)
Feller W 1968 *An Introduction to Probability Theory and Its Application* (New York: Wiley) vol 1 p 288
Gnedenko B V and Kolmogorov A N 1954 *Limit Distributions for Sums of Independent Random Variables* (Reading, Mass: Addison-Wesley)
Hoenders B J, Jakeman E, Baltes H P and Steinle B 1979 *Opt. Acta* **26** *to appear*
Jakeman E and McWhirter J G 1977 *J. Phys. A: Math. Gen.* **10** 1599–643
Jakeman E, Parry G, Pike E R and Pusey P N 1978 *Contemp. Phys.* **19** 127–45
Jakeman E and Pusey P N 1976 *IEEE Trans. Antennas Propag.* **AP24** 806–14
—— 1977 *Radar 77* (London: IEE) pp 105–9
—— 1978 *Phys. Rev. Lett.* **40** 546–50
Nakagami M 1943 *J. Inst. Elec. Commun. Engrs, Japan* **27** 145
Nakagami M and Ota M 1957 *Rep. Radio Wave Propag. Res. Committee, Japan*
Parry G and Gray P 1978 *Optical propagation through random phase screens*, presented at 'Optics 78', Bath, September 1978
Parry G and Pusey P N 1979 *J. Opt. Soc. Am.* **69** 796–8
Parry G, Pusey P N, Jakeman E and McWhirter J G 1977 *Opt. Commun.* **22** 195–200
Parry G and Walker J 1979 *Opt. Acta* **26** 563–74
Pusey P N and Jakeman E 1975 *J. Phys. A: Math. Gen.* **8** 392–410
Tatarski V I 1961 *Wave Propagation in a Turbulent Medium* (New York: McGraw-Hill) ch 1