IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

From doubly stochastic representations of *K* distributions to random walks and back again: an optics tale

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2009 J. Phys. A: Math. Theor. 42 225007

(http://iopscience.iop.org/1751-8121/42/22/225007)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.154
The article was downloaded on 03/06/2010 at 07:51

Please note that terms and conditions apply.

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **42** (2009) 225007 (13pp) doi:10.1088/1751-8113/42/22/225007

From doubly stochastic representations of *K* distributions to random walks and back again: an optics tale

O E French

Applied Mathematics Division, School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK

E-mail: oliver.french@nottingham.ac.uk

Received 8 January 2009, in final form 1 April 2009
Published 15 May 2009
Online at stacks.iop.org/JPhysA/42/225007

Abstract

A random walk model with a negative binomially fluctuating number of steps is considered in the case where the mean of the number fluctuations, $\bar{N}$, is finite. The asymptotic behaviour of the resultant statistics in the large $\bar{N}$ limit is derived and shown to give the *K* distribution. The equivalence of this model to the hitherto unrelated doubly stochastic representation of the *K* distribution is also demonstrated. The convergence to the *K* distribution of the probability density function generated by a random walk with a finite mean number of steps is examined along with the moments, and the non-Gaussian statistics are shown to be a direct result of discreteness and bunching effects.

PACS numbers: 46.65.+g, 02.50.−r

1. Introduction

The best known and most widely used statistical model for continuous variables is the Gaussian, or normal, distribution. Its ubiquity derives largely from the central limit theorem (CLT), which states that the sum of infinitely many identical independent random variables of finite mean and variance is normally distributed. However, a real physical system will contain only a finite number of such variables, and the CLT therefore represents an idealization of reality. Further appeal of the Gaussian distribution lies in its simple analytical form and the relative tractability of problems into which it is incorporated.

There has been a good deal of study into the convergence of random processes to the normal, e.g. [1], and in many cases the Gaussian provides a good approximation to fluctuations occurring in finite size systems. Nevertheless, there are many systems of interest, within the physical sciences and beyond, whose statistics display non-Gaussian behaviour even in the large system size limit, and this has prompted the development of numerous models for a variety of non-Gaussian statistics.

 © 2009 IOP Publishing Ltd Printed in the UK

A particular class of non-Gaussian statistics that has found applications in several disparate areas is the K distribution: a two-parameter class of distribution with probability density function

$$P(x) = \frac{b^2}{2\Gamma(\alpha)}\left(\frac{b^2 x}{4}\right)^{\frac{\alpha-1}{2}} K_{\alpha-1}(b\sqrt{x}), \tag{1}$$

where $\Gamma(z)$ is the gamma function and $K_{\alpha-1}(z)$ is the modified Bessel function [24] that gives the distribution its name. The K distribution has mean

$$\langle x\rangle = \frac{4\alpha}{b^2} \tag{2}$$

and normalized moments

$$x^{[r]} = \frac{\langle x^r\rangle}{\langle x\rangle^r} = \frac{r!\Gamma(r+\alpha)}{\alpha^r\Gamma(\alpha)}. \tag{3}$$

The shape parameter α describes the degree to which the K distribution is non-Gaussian, and in the limit $\alpha \to \infty$ the K distribution reverts to an exponential distribution: this being the distribution obtained for the modulus squared of a quantity described by a two-dimensional Gaussian random walk.

A special case of the distribution was first observed in entomology in the 1950s, where *trichostrongylus retortaeformis* larvae were found to be distributed spatially according to a K_0 distribution arising as a result of a two-dimensional random walk with an exponentially distributed stopping time [2]. This result was generalized in 1975 [3] to allow for gamma-distributed stopping times whereupon the K distribution was obtained and used to describe the distribution of the distance between the birthplaces of mates; so-called matrimonial distances.

The above story is repeated in optics where the K_0 distribution was used to describe doubly scattered laser light [4] before the K distribution itself was seen to be a good description of the intensity fluctuations in light passing through turbulent media [5, 6], and also those arising in the non-Rayleigh echo of microwaves from rough sea surfaces [7, 8].

In addition to the applications mentioned above, the K distribution continues to be of relevance in fields such as synthetic aperture radar [9], ultrasound [10], non-destructive materials evaluation [11], acoustics [12] and satellite imaging [13].

There are two heuristic approaches to the generation of K statistics: the doubly stochastic representation and the random walk model. In the first of these, the K distribution is obtained by 'smearing' the mean of an exponential distribution according to a gamma distribution. The latter approach involves allowing the number of steps in a two-dimensional random walk to vary according to the discrete negative binomial distribution, with the K distribution resulting in the limit of large mean step number fluctuations. Whilst it can be noted that the gamma distribution appearing in the doubly stochastic representation is the high-density limit of the negative binomial distribution used in the random walk model, there have been no attempts to date to show the conditions for the equivalence of the two approaches.

Furthermore, just as the central limit theorem requires an infinite number of fluctuating quantities in order to obtain exactly the Gaussian distribution, the random walk model requires the *mean* of the fluctuating number of steps, $\bar{N}$, to be infinite in order to retrieve the K distribution. This is clearly not physical and so the question arises: how large does $\bar{N}$ need to be in order that the statistics generated are a good fit to the K distribution?

The following section of this paper will give a brief summary of the doubly stochastic representation for the K distribution. Section 3 will consider the random walk model for the K distribution, demonstrating the emergence of Gaussian statistics for the fixed length random walk en-route. Section 4 will use a particular form for the individual step-size distribution to demonstrate in closed form the equivalence of the random walk and doubly stochastic approaches, before section 5 investigates the rate of convergence to the K distribution for a random walk with a fluctuating number of steps with finite mean. Section 6 will summarize the conclusions and identify ways in which this work can be continued forward.

2. Doubly stochastic representation

In general, a doubly stochastic distribution is the result of 'smearing' a parameter of one distribution over another [14]. Many complex distributions may be expressed in this way, and an extensive range of such distributions and their doubly stochastic representations are given in [15]. Of interest to this work, however, is the particular example of the K distribution which may be obtained by smearing the mean of an exponential distribution according to a gamma variate with mean $\bar{N}$ and shape parameter α:

$$\begin{aligned} P_K(x) &= \frac{\alpha}{\bar{N}\Gamma(\alpha)} \int_0^\infty \left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1} \exp\left\{-\frac{\alpha w}{\bar{N}}\right\} \frac{1}{w} \exp\left\{-\frac{x}{w}\right\} \mathrm{d}w \\ &= \frac{2\alpha}{\bar{N}\Gamma(\alpha)} \left(\frac{\alpha x}{\bar{N}}\right)^{\frac{\alpha-1}{2}} K_{\alpha-1}\left(2\sqrt{\frac{\alpha x}{\bar{N}}}\right). \end{aligned} \tag{4}$$

The integration variable w appears as the mean of the exponential distribution, being the last term in the integrand, and is integrated over the gamma distribution which comprises the prefactor.

The above method of obtaining the K distribution is valid for all $\alpha > 0$ and $\bar{N} > 0$. It is a mathematical ansatz, and does not derive from a consideration of a particular physical model, nor any more general physical reasoning. However, the exponential distribution appearing in the integral is known to be a facet of Gaussian intensity fluctuations and this representation hints that some composite of many Gaussian fluctuations may be responsible for the K distribution.

It is also worth noting that the above representation bears a close resemblance to the 'superstatistics' used in the description of non-equilibrium statistical mechanics [16]. Furthermore, there is a deeper connection between the K distribution and non-extensive statistics: the characteristic function of K distributed fluctuations has the same mathematical structure (to within a normalizing constant) as the power-law distribution of the Tsallis entropy. The consequences of this connection have been explored in the appendix of [17].

Although still heuristic, another method for obtaining K statistics—the random walk model—offers some insight into their physical origin. It is this model that shall be considered next.

3. Random walk model

The random walk model with fluctuating number of steps for the generation of K statistics was borne out of the study of the statistical properties of electromagnetic radiation scattered from systems containing a discrete number of scatterers. It is a heuristic model, rather than an exact solution of Maxwell's equations, and has provided considerable insight into the scattering from systems of particles, the rough sea-surface scattering of microwaves and optical propagation through turbulent media.

The electric field at the observation point is described as the superposition of contributions from a (fluctuating) number of discrete scattering centres, and the number fluctuations may be interpreted as the movement of these centres into and out of the area of illumination. Whilst this model is intuitively suitable for scattering by systems of discrete particles, it is perhaps less obvious that it should also apply to scattering by continuous media, such as rough surfaces. Nevertheless, there is a large body of experimental evidence that points to this being the case [6, 18, 19] and the question becomes one of interpreting the physical nature of the discrete scattering centres in these instances.

The coherent addition of fields from a finite number of scatterers, N, may be modelled via the process

$$\xi(t) = \sum_{k=1}^{N} a_k(t) \exp\{\mathrm{i}\phi_k(t)\} = A(t) \exp\{\mathrm{i}\Phi(t)\}, \tag{5}$$

where $\{a_k\}$ and $\{\phi_k\}$ are random variables corresponding to the amplitude and phase from each field contribution, respectively. This is equivalent to a two-dimensional random walk in complex space with a finite number of steps, N. $\{a_k\}$ are assumed to be statistically similar and independent from each other, i.e. $\langle a_k a_l\rangle = \langle a^2\rangle\delta_{kl}$, where δ_{kl} is the Kronecker symbol and the subscript is now dropped from a's. $\{\phi_k\}$ are each taken to be uniformly distributed over $[0, 2\pi)$ and in optical terms, this corresponds to the so-called strong scattering assumption in which the scatterers are assumed to be non-interacting and separated by distances much greater than the wavelength of the incident radiation. Under this condition the characteristic function, $C_N(\mathbf{u})$, is given by

$$C_N(\mathbf{u}) = C_N(u) \equiv \langle \mathrm{e}^{\,\mathrm{i}\xi\cdot\mathbf{u}}\rangle = \langle J_0(au)\rangle_a^N, \tag{6}$$

where $u = |\mathbf{u}|$ and $J_0(z)$ is the Bessel function of the first kind. The uniform angular fluctuations have been averaged over and the subscript a now indicates averaging over the step size alone. Renormalizing this step size by a factor $N^{1/2}$ and expanding the Bessel function gives for large N

$$C_{N\to\infty}(u) = \lim_{N\to\infty}\left(1 - \frac{u^2\langle a^2\rangle}{4N} + O(N^{-2})\right)^N = \exp\left\{-\frac{u^2\langle a^2\rangle}{4}\right\}, \tag{7}$$

where the averaging over a has been performed term by term on the expansion. The above characteristic function may be inverted using the inverse Fourier transform [17]

$$P_N(A) = A\int_0^\infty u J_0(uA) C_N(u)\,\mathrm{d}u \tag{8}$$

and gives rise to the Rayleigh distribution for the random walk amplitude

$$P_N(A) = \frac{2A}{\langle a^2\rangle}\exp\left\{-\frac{A^2}{\langle a^2\rangle}\right\}. \tag{9}$$

This corresponds to the amplitude fluctuations of the two-dimensional Gaussian distribution, in accordance with the CLT. Gaussian statistics are therefore derived from the two-dimensional random walk model where the (fixed) number of steps becomes large. In terms of the intensity $I = A^2$, an exponential distribution is obtained

$$P_N(I) = \frac{1}{\langle I\rangle}\exp\left\{-\frac{I}{\langle I\rangle}\right\}. \tag{10}$$

The number of steps in the walk is now allowed to vary, and is drawn from the negative binomial distribution

$$p_N = \frac{(N+\alpha-1)!}{N!(\alpha-1)!}\frac{(\bar{N}/\alpha)^N}{(1+\bar{N}/\alpha)^{N+\alpha}}, \qquad \alpha, N, \bar{N} > 0. \tag{11}$$

This distribution arises in many diverse hierarchical and multi-scale systems [17–23] and is characterized by its mean, $\bar{N}$, and the cluster parameter α. For small values of α, samples of the negative binomial distribution have a tendency to cluster together, and this clustering becomes less pronounced as α increases in size, with Poisson statistics being retrieved in the limit $\bar{N} \to \infty$.

The characteristic function for a random walk with fluctuating number of steps may be obtained by averaging over these number fluctuations

$$C_{\bar{N}}(\mathbf{u}) = \sum_{N=0}^{\infty} p_N C_N(\mathbf{u}) \tag{12}$$

and for the negative binomial distribution this yields

$$C_{\bar{N}}(\mathbf{u}) = \left[1 + \frac{\bar{N}}{\alpha}(1 - \langle J_0(au)\rangle_a)\right]^{-\alpha}. \tag{13}$$

This time the walk is scaled by the factor $\bar{N}^{1/2}$ and the mean number of steps is taken to infinity. This scaling ensures that the second moment of the amplitude fluctuations is well defined in the limit $\bar{N} \to \infty$, where the characteristic function then takes the form

$$C_{\bar{N}\to\infty}(\mathbf{u}) = \left(1 + \frac{u^2\langle a^2\rangle}{4\alpha}\right)^{-\alpha}, \tag{14}$$

which is clearly different from the previous form (7). The corresponding probability density function is again achieved by the Fourier inversion of equation (8) and results in the amplitude distribution

$$P_{\bar{N}\to\infty}(A) = \frac{2\gamma}{\Gamma(\alpha)}\left(\frac{\gamma A}{2}\right)^{\alpha} K_{\alpha-1}(\gamma A), \qquad \gamma^2 = \frac{4\alpha}{\langle a^2\rangle}, \tag{15}$$

the corresponding intensity distribution is given by

$$P_{\bar{N}\to\infty}(I) = \frac{\gamma^2}{2\Gamma(\alpha)}\left(\frac{\gamma^2 I}{4}\right)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\gamma\sqrt{I}), \tag{16}$$

which is in precisely the form of the K distribution given in equation (1) with $b = \gamma$.

The above random walk derivation of the K distribution is obtained from taking an ensemble average over the number of steps in the walk. If the number of steps is deterministic or fluctuates with a distribution whose variance is less than or equal to that of a Poisson distribution, then the resultant is Gaussian in the limit as the average number of steps becomes infinite. If the distribution has standard deviation greater than its mean, as is the case for the negative binomial distribution used above, then the limit distribution for the resultant in the same limit is non-Gaussian. Distributions with standard deviation in excess of the mean are super-Poissonian and describe clustering. These clusters introduce an effective correlation that leads to the violation of the central limit theorem.

4. Equivalence of random walk and doubly stochastic approaches

The limiting form of the random walk model given above assumes that higher moments of the step-size distribution scale with the mean number of steps in such a way that that all terms in the expansion of the characteristic function (13) vanish in the limit $\bar{N} \to \infty$, with the exception of the term containing $\langle a^2\rangle$. This circumvents the problem of exactly *how* the higher moments scale with the mean number of steps, and requires only that they vanish in the limit.

If we wish to consider the rate at which a random walk with a finite mean number of steps converges upon the limiting result given previously, it is necessary to know more than their behaviour in the $\bar{N} \to \infty$ limit: the precise nature of the scaling of higher moments is required explicitly, and this will vary from one step-size distribution to another.

By taking a particular form of the step-size distribution, $p(a)$, the scaling behaviour of all the moments is defined and expressions for convergence of the distribution and the moments may be obtained in closed form. Here, the step size is taken to be distributed according to the Rayleigh distribution

$$p(a) = \frac{2a}{\langle a^2\rangle} \exp\left\{-\frac{a^2}{\langle a^2\rangle}\right\}, \tag{17}$$

which has moments

$$\langle a^r\rangle = \frac{r}{2}\Gamma\left(\frac{r}{2}\right)\langle a^2\rangle^{r/2}. \tag{18}$$

The averaging of the Bessel function required in equation (13) may be performed directly to give

$$\langle J_0(au)\rangle_a = \exp\left\{-\frac{u^2\langle a^2\rangle}{4}\right\}. \tag{19}$$

The algebraic form of equation (13), with the help of the identity that defines the gamma function [24]

$$\Gamma(z) = k^z \int_0^\infty t^{z-1}\,\mathrm{e}^{-kt}, \tag{20}$$

enables the characteristic function to be written in the following form:

$$C_{\bar{N}}(\mathbf{u}) = \frac{1}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1} \exp\left\{-\left(1+\frac{\bar{N}}{\alpha}\right)t\right\} \exp\left\{\frac{\bar{N}t}{\alpha}\exp\left\{-\frac{u^2\langle a^2\rangle}{4}\right\}\right\}\mathrm{d}t. \tag{21}$$

The corresponding probability distribution is obtained by using equation (8)

$$P_{\bar{N}}(A) = A\int_0^\infty u J_0(Au) C_{\bar{N}}(\mathbf{u})\,\mathrm{d}u. \tag{22}$$

The walk is again scaled by the square root of the mean number of steps through making the transformation

$$u \to \left(\frac{1}{\bar{N}}\right)^{1/2} u, \qquad A \to \bar{N}^{1/2}A, \tag{23}$$

which has the effect of normalizing $\langle A^2\rangle = \langle a^2\rangle$, and the probability distribution is then given by, after exchanging the order of integration,

$$P_{\bar{N}}(A) = \frac{A}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1}\exp\left\{-\left(1+\frac{\bar{N}}{\alpha}\right)t\right\}\int_0^\infty u J_0(Au)\exp\left\{\frac{\bar{N}t}{\alpha}\exp\left\{-\frac{u^2\langle a^2\rangle}{4\bar{N}}\right\}\right\}\mathrm{d}u\,\mathrm{d}t. \tag{24}$$

Now making the substitution $w = \bar{N}t/\alpha$ and expressing the 'exponential of an exponential' term as an infinite series yields

$$\begin{aligned} P_{\bar{N}}(A) &= \frac{\alpha A}{\bar{N}\Gamma(\alpha)}\int_0^\infty \left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1}\exp\left\{-\left(1+\frac{\alpha}{\bar{N}}\right)w\right\} \\ &\quad\times \sum_{k=0}^{\infty}\frac{w^k}{k!}\int_0^\infty u J_0(Au)\exp\left\{-\frac{ku^2\langle a^2\rangle}{4\bar{N}}\right\}\mathrm{d}u\,\mathrm{d}w \\ &= \frac{2\alpha\delta(A)}{\bar{N}\Gamma(\alpha)}\int_0^\infty\left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1}\exp\left\{-\left(1+\frac{\alpha}{\bar{N}}\right)w\right\}\mathrm{d}w \\ &\quad+ \frac{\alpha A}{\bar{N}\Gamma(\alpha)}\int_0^\infty\left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1}\exp\left\{-\left(1+\frac{\alpha}{\bar{N}}\right)w\right\}\sum_{k=1}^{\infty}\frac{w^k}{k!}\frac{2\bar{N}}{k\langle a^2\rangle}\exp\left\{-\frac{\bar{N}A^2}{k\langle a^2\rangle}\right\}\mathrm{d}w, \end{aligned} \tag{25}$$

where we have used the relations

$$\int_0^\infty u J_0(Au)\exp\left\{-\frac{ku^2\langle a^2\rangle}{4\bar{N}}\right\}\mathrm{d}u = \frac{2\bar{N}}{k\langle a^2\rangle}\exp\left\{-\frac{\bar{N}A^2}{k\langle a^2\rangle}\right\}, \qquad k \geqslant 1$$

$$\int_0^\infty u J_0(Au)\,\mathrm{d}u = \frac{2\delta(A)}{A},$$

and $\delta(x)$ is the Dirac delta function. Transforming into $I = A^2$ and again using equation (20), it is found that

$$P_{\bar{N}}(I) = 2\delta(I)\left(1+\frac{\bar{N}}{\alpha}\right)^{-\alpha} + \frac{\alpha}{\langle a^2\rangle\Gamma(\alpha)}\int_0^\infty \left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1}\exp\left\{-\frac{\alpha w}{\bar{N}}\right\} \times \sum_{k=1}^{\infty}\frac{w^k}{kk!}\exp\left\{-\frac{\bar{N}I}{k\langle a^2\rangle}-w\right\}\mathrm{d}w. \tag{26}$$

The first term comprises the product of a delta function with p_0, the probability that there are no steps in the random walk. This therefore corresponds to the probability of zero intensity arising as a result of no steps being taken. The relative magnitude of this term decreases with increasing $\bar{N}$, vanishing from the random walk in the limit $\bar{N} \to \infty$. It should be noted that there is another contribution to the probability of zero intensity: that from the second term in the equation, which is the probability that zero intensity arises from walks with nonzero numbers of steps, i.e. that the walk returns to the origin.

It is now convenient to write the summation term in equation (26) as

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{w^k}{kk!}\exp\left\{-\frac{\bar{N}I}{k\langle a^2\rangle}-w\right\} &= -\frac{\langle a^2\rangle}{\bar{N}}\frac{\mathrm{d}}{\mathrm{d}I}\sum_{k=1}^{\infty}\frac{w^k}{k!}\exp\left\{-\frac{\bar{N}I}{k\langle a^2\rangle}-w\right\}\\ &\equiv -\frac{\langle a^2\rangle}{\bar{N}}\frac{\mathrm{d}f_\infty(I,w)}{\mathrm{d}I},\end{aligned} \tag{27}$$

where

$$f_\infty(I,w) \equiv \lim_{\bar{N}\to\infty}\sum_{k=1}^{\infty}\frac{w^k}{k!}\exp\left\{-\frac{\bar{N}I}{k\langle a^2\rangle}-w\right\}.$$

Replacing the summation with an integral over k, noting that $k! = \Gamma(k+1)$ and then making the substitution $x = k/\bar{N}$ yields

$$\begin{aligned} f_\infty(I,w) &\approx \lim_{\bar{N}\to\infty}\int_{1/\bar{N}}^{\infty}\frac{\bar{N}w^{\bar{N}x}}{\Gamma(\bar{N}x+1)}\exp\left\{-\frac{I}{\langle a^2\rangle x}-w\right\}\mathrm{d}x\\ &= \lim_{\bar{N}\to\infty}\int_{1/\bar{N}}^{\infty}\frac{\bar{N}w^{\bar{N}x}}{\sqrt{2\pi}\,\mathrm{e}^{-\bar{N}x}(\bar{N}x)^{\bar{N}x+1/2}}\exp\left\{-\frac{I}{\langle a^2\rangle x}-w\right\}\mathrm{d}x\\ &= \lim_{\bar{N}\to\infty}\int_{1/\bar{N}}^{\infty}\left(\frac{\bar{N}}{2\pi x}\right)^{\frac{1}{2}}\exp\left\{\bar{N}x\left(1-\ln\frac{\bar{N}x}{w}\right)-w\right\}\exp\left\{-\frac{I}{\langle a^2\rangle x}\right\}\mathrm{d}x,\end{aligned}$$

where Stirling's approximation [24] has been used for the asymptotic behaviour of the gamma function with large argument. For large $\bar{N}$, the integrand will be dominated by the region around its single turning point, which occurs at $x = w/\bar{N}$ and is a maximum. Expanding the argument of the exponential around this point to second order gives

$$f_\infty(I,w) = \lim_{\bar{N}\to\infty}\int_{1/\bar{N}}^{\infty}\left(\frac{\bar{N}}{2\pi x}\right)^{\frac{1}{2}}\exp\left\{-\frac{\bar{N}^2}{2w}\left(x-\frac{w}{\bar{N}}\right)^2\right\}\exp\left\{-\frac{I}{\langle a^2\rangle x}\right\}\mathrm{d}x.$$

Making a final substitution of $y = (\bar{N}/w)^{1/2}x$ and rearranging yields

$$f_\infty(I,w) = \lim_{\bar{N}\to\infty} \int_{1/\bar{N}}^{\infty} \left(\frac{\bar{N}}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{\bar{N}}{2}\left(y - \sqrt{\frac{w}{\bar{N}}}\right)^2\right\}$$

$$\times \left(\frac{\sqrt{w/\bar{N}}}{y}\right)^{\frac{1}{2}} \exp\left\{-\frac{I}{\langle a^2\rangle y\sqrt{w/\bar{N}}}\right\} \mathrm{d}y$$

$$= \int_0^\infty \delta\left(y - \sqrt{\frac{w}{\bar{N}}}\right)\left(\frac{\sqrt{w/\bar{N}}}{y}\right)^{\frac{1}{2}} \exp\left\{-\frac{I}{\langle a^2\rangle y\sqrt{w/\bar{N}}}\right\} \mathrm{d}y$$

$$= \exp\left\{-\frac{\bar{N}I}{w\langle a^2\rangle}\right\} \tag{28}$$

since

$$\lim_{\bar{N}\to\infty} \left(\frac{\bar{N}}{2\pi}\right)^{\frac{1}{2}} \exp\left\{\frac{\bar{N}}{2}\left(y - \sqrt{\frac{w}{\bar{N}}}\right)^2\right\} = \delta\left(y - \sqrt{\frac{w}{\bar{N}}}\right) \tag{29}$$

is simply the Gaussian definition of the Dirac delta function. Therefore

$$\lim_{\bar{N}\to\infty} \sum_{k=1}^{\infty} \frac{w^k}{kk!} \exp\left\{-\frac{\bar{N}I}{k\langle a^2\rangle}\right\} = -\frac{\langle a^2\rangle}{\bar{N}} \frac{\mathrm{d}f_\infty(I,w)}{\mathrm{d}I} = \frac{1}{w}\exp\left\{-\frac{\bar{N}I}{w\langle a^2\rangle}\right\}, \tag{30}$$

and reinserting this into equation (26) yields

$$\lim_{\bar{N}\to\infty} P_{\bar{N}}(I) = \frac{\alpha}{\bar{N}\Gamma(\alpha)} \int_0^\infty \left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1} \exp\left\{-\frac{\alpha w}{\bar{N}}\right\} \frac{\bar{N}}{w\langle a^2\rangle} \exp\left\{-\frac{\bar{N}I}{w\langle a^2\rangle}\right\} \mathrm{d}w, \tag{31}$$

since the δ-function term vanishes in the limit. Taking the doubly stochastic representation, given in equation (4), with $x = I$ and then applying the scaling $I \to \bar{N}I/\langle a^2\rangle$ results in the following:

$$P_K(I) = \frac{\alpha}{\bar{N}\Gamma(\alpha)} \int_0^\infty \left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1} \exp\left\{-\frac{\alpha w}{\bar{N}}\right\} \frac{\bar{N}}{w\langle a^2\rangle} \exp\left\{-\frac{\bar{N}I}{w\langle a^2\rangle}\right\} \mathrm{d}w, \tag{32}$$

which is precisely that obtained from the limit of the random walk model.

It has therefore been demonstrated that the doubly stochastic representation of the K distribution emerges naturally from the random walk model. As alluded to earlier, this results from the continuum limit of the negative binomial distribution being the gamma distribution, which manifests itself mathematically through the gamma transformation of equation (20).

5. Convergence of moments

The above equivalence occurs only in the strict limit $\bar{N} \to \infty$ but such a limit is not the one that corresponds to a real physical system, where the mean of the number fluctuations will necessarily be finite. Nevertheless, the intensity fluctuations of a finite system are found to be a good approximation to the K distribution in a number of systems [5–7], and it would be useful to know over what range this approximation is valid.

This is investigated through a consideration of the normalized intensity moments

$$I^{[r]} = \frac{\langle I^r\rangle}{\langle I\rangle^r} \tag{33}$$

for the random walk with finite $\bar{N}$. It is possible to rewrite equation (26) as

$$P_{\bar{N}}(I) = \frac{\alpha}{\bar{N}\Gamma(\alpha)} \int_0^\infty \left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1} \exp\left\{-\left(1+\frac{\alpha}{\bar{N}}\right)w\right\} \sum_{k=1}^{\infty} \frac{w^k}{k!} \frac{\bar{N}}{k\langle a^2\rangle} \exp\left\{-\frac{\bar{N}I}{k\langle a^2\rangle}\right\} \mathrm{d}w, \quad (34)$$

where the delta function term has been omitted since it does not contribute to the moments. The I dependence is now in the form of an exponential distribution and averaging over the intensity then yields

$$\begin{aligned}\langle I^r\rangle &= \frac{\alpha}{\bar{N}\Gamma(\alpha)} \int_0^\infty \left(\frac{\alpha w}{\bar{N}}\right)^{\alpha-1} \exp\left\{-\left(1+\frac{\alpha}{\bar{N}}\right)w\right\} \sum_{k=1}^{\infty} \frac{w^k r!}{k!} \left(\frac{k\langle a^2\rangle}{\bar{N}}\right)^n \mathrm{d}w \\ &= \frac{r!}{\Gamma(\alpha)} \left(\frac{\langle a^2\rangle}{\bar{N}}\right)^r \left(1+\frac{\bar{N}}{\alpha}\right)^{-\alpha} \sum_{k=1}^{\infty} \frac{k^r\Gamma(k+\alpha)}{k!} \left(1+\frac{\alpha}{\bar{N}}\right)^{-k},\end{aligned} \quad (35)$$

where equation (20) has been used to perform the integration over w. The summation may be performed by noting that it is always possible to write

$$k^n = \sum_{j=1}^{n} \frac{c_j k!}{(k-j)!}, \quad (36)$$

where c_j are constants that may easily be deduced for a given n. The normalized moments are then given by, recalling that for the rescaled walk $\langle I\rangle = \langle a^2\rangle$,

$$I^{[r]} = \frac{r!}{\Gamma(\alpha)\bar{N}^r} \left(1+\frac{\bar{N}}{\alpha}\right)^{-\alpha} \sum_{j=1}^{r} c_j S_j, \quad (37)$$

where

$$\begin{aligned}S_j &= \sum_{k=1}^{\infty} \frac{\Gamma(k+\alpha)}{(k-j)!} \left(1+\frac{\alpha}{\bar{N}}\right)^{-k} \\ &= \Gamma(\alpha+j) \left(1+\frac{\bar{N}}{\alpha}\right)^{\alpha} \left(\frac{\bar{N}}{\alpha}\right)^{j}.\end{aligned} \quad (38)$$

The above may be used to calculate the following:

$$I^{[1]} = 1, \quad (39)$$

$$I^{[2]} = 2\left[\left(1+\frac{1}{\alpha}\right) + \frac{1}{\bar{N}}\right], \quad (40)$$

$$I^{[3]} = 6\left[\left(1+\frac{1}{\alpha}\right)\left(1+\frac{2}{\alpha}\right) + \frac{3}{\bar{N}}\left(1+\frac{1}{\alpha}\right) + \frac{1}{\bar{N}^2}\right], \quad (41)$$

$$I^{[4]} = 24\left[\left(1+\frac{1}{\alpha}\right)\left(1+\frac{2}{\alpha}\right)\left(1+\frac{3}{\alpha}\right) + \frac{6}{\bar{N}}\left(1+\frac{1}{\alpha}\right)\left(1+\frac{2}{\alpha}\right) + \frac{7}{\bar{N}^2}\left(1+\frac{1}{\alpha}\right) + \frac{1}{\bar{N}^3}\right], \quad (42)$$

which are all of the form

$$I^{[r]} = I_K^{[r]} + O(\bar{N}^{-1}), \quad (43)$$

where $I_K^{[r]}$ is the corresponding moment of the K distribution given by

$$I_K^{[r]} = \frac{r!\Gamma(r+\alpha)}{\alpha^r\Gamma(\alpha)} = r!\prod_{j=0}^{r-1}\left(1+\frac{j}{\alpha}\right). \quad (44)$$

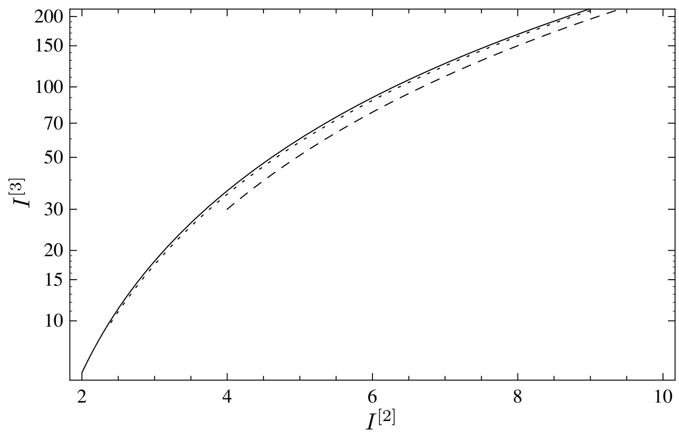

Figure 1. The third moment of the FMRW plotted against the second on a log-linear scale for $\alpha = 1/2$. The dashed line is calculated for $\bar{N} = 1$ whilst the dotted line is for $\bar{N} = 5$. The solid line is the corresponding curve for the K distribution.

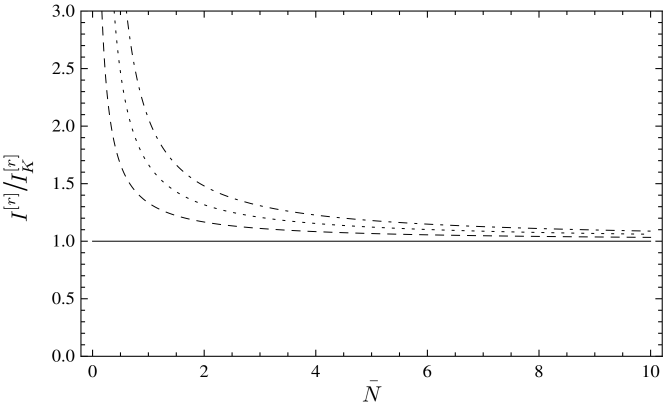

Figure 2. The plot of the ratio $I^{[r]}/I_K^{[r]}$ against $\bar{N}$ for $\alpha = 1/2$. The first, second, third and fourth moments are given by the solid, dashed, dotted and dot-dashed lines, respectively (from bottom to top).

Figure 1 shows the plot of the third moment against the second moment of the finite mean random walk (FMRW) on a log-linear scale. It can be seen that for $\bar{N} = 5$ the curve is practically indistinguishable from that of the K distribution, and that even for $\bar{N} = 1$ there is reasonable agreement. The starting point of the plot in each case is determined by the minimum value of $I^{[2]}$ that may be achieved for a given $\bar{N}$, which may be deduced from equation (40). This clearly shows the rapid convergence with $\bar{N}$ of the FMRW to K statistics.

Figure 2 plots the ratio $I^{[r]}\big/I_K^{[r]}$ against $\bar{N}$, for $\alpha = 0.5$ and r from 1 to 4. The mean of the FMRW is the same as for the K distribution for all $\bar{N}$, but the higher moments of the FMRW are always greater than the corresponding K distribution moments, indicated by the ratio always exceeding unity. There is again rapid convergence to the K distribution moments with increasing $\bar{N}$ in both cases, with the higher moments being slower to converge than the lower ones.

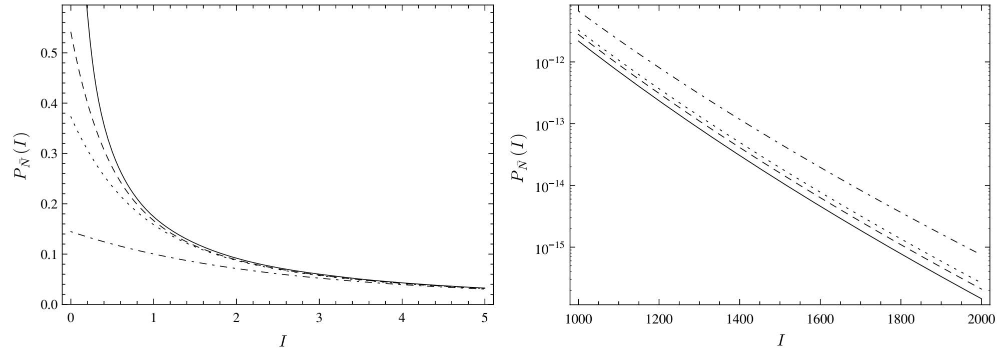

Figure 3. Plot of the probability density function, $P_{\bar{N}}(I)$, against intensity, *I*, for $\alpha = 1/2$. The left hand figure is plotted on linear axes for small values of *I*, whilst the right is plotted on a log-linear scale for larger values of *I*. In each case, the solid line is the *K* distribution ($\bar{N} \to \infty$) and the dashed, dotted and dot-dashed lines are the FMRW model with $\bar{N} = 10$, 6 and 2 respectively.

This behaviour can be understood by considering the effect of finite $\bar{N}$ on the form of the probability density function given in equation (26). As already mentioned, the δ-function term corresponds to an additional weight at $I = 0$ as a result of there being no steps in the walk, and this is non-vanishing for finite $\bar{N}$. This weight does not contribute directly to the moments, but merely serves to satisfy the conservation of probability. However, the fact that a finite proportion of the probability is embodied in this δ-function term will cause a decrease in the moments as a result of the associated redistribution of the probability density.

This effect is opposed for nonzero values of the intensity by a shifting of the distribution weight away from the 'front end' and into the 'tail'. This can be seen from figure 3, which is a plot of the probability density function, $P_{\bar{N}}(I)$, against the intensity, *I*, for $\alpha = 0.5$ and varying values of $\bar{N}$. Also shown is the *K* distribution for the same value of α. The infinite series in equation (26) has been truncated at the value $k = k_{\max}$, which causes a small error in the tail, but for a given *I* this error can be made vanishingly small by taking $k_{\max}$ to be sufficiently large and does not affect the conclusions drawn here. It can be seen that the value of the probability density function for the FMRW is less than that of the *K* distribution for smaller values of *I*, but is greater for larger values.

In the limit $\alpha \to \infty$, the number fluctuations are Poisson in nature and yet the normalized second moment, equation (40), still differs from the Gaussian value of 2 for finite $\bar{N}$. Since in this limit the bunching properties of the negative binomial are no longer at work, it can only be the discreteness of the number fluctuations that are giving rise to this deviation. In this limit, and for large yet finite $\bar{N}$, it is possible to write

$$I^{[2]} = 2\left(1 + \frac{1}{\bar{N}}\right), \tag{45}$$

$$I^{[3]} = 6\left(1 + \frac{1}{\bar{N}}\right)\left(1 + \frac{2}{\bar{N}}\right) + O(\bar{N}^{-2}), \tag{46}$$

$$I^{[4]} = 24\left(1 + \frac{1}{\bar{N}}\right)\left(1 + \frac{2}{\bar{N}}\right)\left(1 + \frac{3}{\bar{N}}\right) + O(\bar{N}^{-2}), \tag{47}$$

and the moments take a similar form to the *K* distribution moments given in equation (44) with $\bar{N}$ replacing α. This suggests that the effect upon the fluctuations of finite $\bar{N}$ in the $\alpha \to \infty$ limit is similar to that of the bunching of the negative binomial distribution. This is intuitive in the sense that for finite $\bar{N}$ the number fluctuations, even for $\alpha \to \infty$, are bunched around certain values: the integer values of N. It would appear therefore that it is the bunching effect that gives rise to the non-Gaussian statistics observed in many systems.

6. Conclusions and further work

In this paper, it has been demonstrated that two heuristic models for the generation of K statistics that are only superficially related at first glance are in fact equivalent. This equivalence derives from the emergence of the gamma distribution in the high-density limit of the negative binomial distribution, the latter describing the number fluctuations in the random walk model.

When the mean of the number fluctuations in the random walk model is finite, the moments display a rapid convergence to K statistics with increasing $\bar{N}$, although this rate of convergence is slower for higher moments. The effect of finite $\bar{N}$ is to introduce discreteness into the variation in step size, and this has an analogous effect to the bunching inherent in negative binomial statistics.

Although the results obtained here were obtained for a particular form of the step-size distribution, there will be no qualitative difference were another form to be used. The effect of the step-size distribution is manifested in the coefficients of the inverse powers of $\bar{N}$ in equations (40)–(42), which in the most general case contain expressions of the form $\langle a^{2j}\rangle/\langle a^2\rangle^j$ [25]. For the Rayleigh distribution, these factors reduce simply to $j!$ but other distributions will give slight modifications to this. For example, if the step-size distribution was itself taken to be K distributed, then

$$\frac{\langle a^{2j}\rangle}{\langle a^2\rangle^j} = \frac{j!\Gamma(\alpha+j)}{\alpha^j\Gamma(\alpha)}, \tag{48}$$

resulting in different numerical factors, but no essentially new results.

Furthermore, by taking the step-size distribution to Rayleigh, it has been possible to better distil the effect of discreteness upon the generation of non-Gaussian statistics. The Rayleigh distribution is that which arises in two-dimensional random walks and results in Gaussian statistics. It does therefore not, in itself, contribute to the onset of the non-Gaussian fluctuations observed in the FMRW model. These are instead due to a combination of the bunching and intrinsic discreteness of the negative binomial distribution.

It would be interesting to investigate in a similar vein the convergence of random walks which display a directional bias, deriving from a non-uniform angular distribution of ϕ_i's, or which do not satisfy the assumption that a_i's and ϕ_i's are independent. Both of these situations arise in scattering problems where the strong scattering assumption is invalid [19], and have been shown in the limit to give the generalized K distribution [26].

Fluctuations that display power-law behaviour on both an inner and an outer scale have been shown to arise in the large mean fluctuation number limit of random walks where the step size is drawn from a power-law distribution and the number of steps fluctuates according to a negative binomial distribution [17]. It would be interesting also to investigate the effect of finite mean step size upon a random walk whose step-size distribution exhibits power-law behaviour over a finite range only. This would impose a third scale which, depending on the size of $\bar{N}$, would either suppress the outer scale induced by the power-law fluctuations or have the effect of limiting the outer power-law behaviour to a finite region, rather than over the infinite region seen in [17].

Furthermore, whilst the number fluctuations considered here are negative binomial in nature, there are other distributions that will give a nonzero variance in the high-density limit and give rise to non-Gaussian statistics other than the K distribution. Indeed, some work has already been performed on random walks in the large mean limit for which the number fluctuations are governed by a power law, and singular behaviour was observed in the resulting distribution. The effect of finite mean upon such a system might well yield interesting results and modify this singular behaviour. Such power-law number fluctuations have been used in the resolution of a paradox in the consideration of sandpiles: the apparent infinite displacement, or flight, of grains within the pile was revealed to be instead due to a power-law distributed number of subflights, each of finite length. Large though the systems described by sandpile dynamics are, it would be informative to consider the effects of finite mean upon the resulting distribution.

Through the investigation of these further examples, it will be possible to gain additional insight into the role played by the intrinsic discreteness of finite systems in generating non-Gaussian fluctuations.

Acknowledgments

This work was funded by the Engineering and Physical Sciences Research Council. The author would like to thank K I Hopcraft and E Jakeman for valuable discussions on this work.

References

[1] Gnedenko B V and Kolmogorov A N *Limit Distributions for Sums of Independent Random Variables* (Reading, MA: Addison-Wesley)
[2] Broadbent S R and Kendall D G 1953 *Biometrics* **9** 460–6
[3] Yasuda N 1975 *Theor. Popul. Biol.* **7** 156–67
[4] Bertolotti M, Crosignani B and DiPorto P 1970 *J. Phys. A: Math. Gen.* **3** L37–L38
[5] Jakeman E and McWhirter J G 1977 *J. Phys. A: Math. Gen.* **3** 1122
[6] Parry G and Pusey P N 1979 *J. Opt. Soc. Am.* **69** 796–8
[7] Jakeman E and Pusey P N 1976 *IEEE Trans. Antennas Propag.* **24** 806–14
[8] Ward K D, Tough RJA and Watts S 2006 *Sea Clutter: Scattering, the K Distribution and Radar Performance* (London: IET)
[9] Chung P J *et al* 2005 *IEEE Trans. Signal Process.* **53** 397
[10] Clifford L *et al* 1993 *Ultrasound Med. Biol.* **19** 487–95
Molthen R C *et al* 1995 *Ultrasound Med. Biol.* **21** 161–70
Molthen R C *et al* 1998 *Ultrasound Med. Biol.* **24** 93–100
[11] Keyes T K and Tucker W T 1999 *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **46** 883–7
[12] Caruthers J W 2006 *J. Acoust. Soc. Am.* **119** 3449
Johnson S F *et al* 2006 *J. Acoust. Soc. Am.* **120** 3142
[13] Doulgeris A P, Anfinsen S N and Eltoft T 2008 *Trans. Geosci. Remote Sens.* **46** 2999–3009
[14] Cox D R 1955 *J. R. Stat. Soc.* B **17** 129–64
[15] Teich M C and Diament P 1989 *J. Opt. Soc. Am.* A **6** 80–91
[16] Beck C and Cohen EGD 2003 *Physica* A **322** 267–75
Beck C, Cohen EGD and Rizzo S 2005 Special issue and directory: nonextensive statistical mechanics: new trends, new perspectives *Europhys. News* **36** 189–92
[17] Hopcraft K I, Jakeman E and Tanner RMJ 1999 *Phys. Rev.* E **60** 5327–43
[18] Jakeman E and Pusey P N 1976 *Phys. Rev. Lett.* **40** 546–50
[19] Jakeman E and Tough RJA 1988 *Adv. Phys.* **37** 471–529
[20] Smith J M, Hopcraft K I and Jakeman E 2007 *Phys. Rev.* E **77** 031112
[21] Jakeman E 1980 *J. Phys. A: Math. Gen.* **13** 31
[22] Madhava Rao N and Kaila K L 2007 *Geophys. J. Int.* **85** 283–90
[23] Bittner E *et al* 2007 *Europhys. Lett.* **78** 58002
[24] Abramowitz M and Stegun I A 1972 *Handbook of Mathematical Functions* (New York: Dover)
[25] Pusey P N 1977 *Photon Correlation Spectroscopy and Velocimetry* ed H Z Cummins and E R Pike (New York: Plenum)
[26] Jakeman E and Tough RJA 1987 *J. Opt. Soc. Am.* A **4** 1764–72