IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Unimodality of photocount distribution for optical noise field

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

1996 J. Phys. A: Math. Gen. 29 7105

(http://iopscience.iop.org/0305-4470/29/22/012)

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

Download details:
IP Address: 171.66.16.70
The article was downloaded on 02/06/2010 at 04:04

Please note that terms and conditions apply.

J. Phys. A: Math. Gen. **29** (1996) 7105–7111. Printed in the UK

Unimodality of photocount distribution for optical noise field

Yuri P Virchenko
Single Crystal Institute, Kharkov, 310001, Ukraine

Received 1 April 1996, in final form 23 July 1996

Abstract. The photocounts statistics for the optical radiation registered by a detector, is considered. The 'one-top' property of the Mandel probability distribution of random photocounts is proved provided the stochastic optical field is Gaussian. The noise low-intensity electromagnetic field has the Gaussian property due to the central limit theorem.

1. Introduction

In the theory of low-intensity optical-field detection, the probability distributions $P(n)$ of a random number n of registered photons is studied. If the photons have a stochastic origin, then the theory gives the following simple formula [1–3]:

$$P(n) = \frac{1}{n!}\langle[\Omega(a)]^n \exp(-\Omega(a))\rangle \tag{1}$$

where

$$\Omega(a) = \int_0^T |a(t)|^2 \, \mathrm{d}t. \tag{2}$$

Here, T is the registration time, $a(t)$ is a complex function of time—the coherent photon amplitude at the time t. In the general case, the amplitude $a(t)$ is a random function (i.e. a random process from a mathematical view-point [4]). Therefore, the symbol $\langle . \rangle$ in (1) denotes the expectation connected with the probability distribution for the random function $a(t)$. The distribution character is mainly defined by the influence of random perturbations when the electromagnetic field spreads in medium before detection. It may be diverse in different physical situations. Therefore, in fact the formula (1) is very complicated despite its illusory simplicity, and it may describe qualitatively diverse distributions $P(n)$. In particular, the distributions $P(n)$ may have a different number of maxima (or, similarly, different quantities of the numbers n for which the difference $(P(n+1) - P(n))$ changes its sign). This characteristic is a very important qualitative property of probabilistic distribution. The presence of several maxima points to some physical phenomena taking place in a system described by probabilistic distribution. For instance, noise-induced phase transitions in a physical system are described by means of a probabilistic distribution having several maxima [5]. In contrast, if the distribution used, has only one maximum, then it apparently corresponds to a trivial physical situation when the described random value is almost definite and the maximum points to a stable state of the system. In the theory of continuous probabilistic distributions, such a distribution is called unimodal [6]. The notion

0305-4470/96/227105+07$19.50 © 1996 IOP Publishing Ltd

of unimodality may be applied to discrete probabilistic distributions. In this case, it will play the analogous role, i.e. a unimodal discrete distribution describes a random physical value having a stable, almost definite meaning. In particular, the probabilistic distributions (1) describing photocounts of electromagnetic radiation not connected with information transmission, is apparently unimodal. It is natural to set the problem of mathematical justification of this physical conjecture in the case when the optical field is purely noise. To formulate this problem exactly, it is necessary to specify mathematically the probabilistic properties of noise. In this paper, we suppose that electromagnetic noise has a sufficiently small intensity, therefore, we can neglect nonlinear effects when it is spread in a medium. We also suppose that the detected noise is a sum of a large number of random small summands at each instant. According to the central limit theorem, we may consider that such a noise is a Gaussian random process [4].

In the present paper we solve the posed problem in the case when the process $a(t)$ is Gaussian and it has zero expectation. Earlier, we have solved this problem for a particular case [7] contained in the theorem proved below. As for the photocount distribution, the unimodality problem is an important theoretical one in view of the above mentioned reasons. But, of course, the obtained result does not solve this problem entirely.

Besides the general theoretical importance, the property of unimodality has an applied significance. For example, note the information transmission theory and, in particular, laser communication theory [8, 9]. From the view-point of these theories, the unimodality of photocount distributions is in order for the application of the ideal receiver theory [9, 10]. In this case, it is necessary to guarantee the uniqueness of the intersection point between the photocount distributions of a purely noise optical field and an optical field transmitting a signal. The unique intersection point is the threshold level of the ideal receiver, the existence of which allows us to define the probabilities of mistaken recognition (signal registration or signal loss).

In section 2, we define the known Karuhnen–Loève expansion [11–13] of an arbitrary Gaussian process to study the optical noise under consideration, and also we give the well known physical example for which our theorem is applicable. In section 3, we present a mathematical definition of unimodal discrete probabilistic distributions. Further, in section 4, we study the general form of characteristic functions of photocount distributions in the Gaussian case. Finally section 5 is devoted to the proof of our main result.

2. Gaussian amplitudes

Let $a(t)$ be a Gaussian process having zero average, $\langle a(t)\rangle = 0$. Based on the Karuhnen–Loève theorem [11–13], we can state that for each fixed time interval $[0, T]$ the amplitude $a(t)$ may be presented as a superposition of some non-random mutually orthogonal modes $\psi_k(t), k = 1, 2, 3, \ldots$ with random complex coefficients α_k which are Gaussian and statistically independent,

$$a(t) = \sum_k \alpha_k \psi_k(t). \tag{3}$$

Furthermore, $\Re[\alpha_k]$ and $\Im[\alpha_k]$ are statistically independent and identically distributed. Thus, the functions $\psi_k(t)$ and the coefficients α_k possess the following properties

$$\int_0^T \psi_k(t)\psi_m^*(t)\,\mathrm{d}t = \delta_{km}$$
$$\langle \alpha_k\alpha_m^*\rangle = \delta_{km} \qquad \langle \alpha_k\alpha_m\rangle = 0 \qquad k, m = 1, 2, 3, \ldots \tag{4}$$

and there is the non-random number $\lambda_k > 0$ for each random variable α_k that the distribution density $f_k(\alpha)$ of α_k on a complex plane α is determined by the formula

$$f_k(\alpha) = (\lambda_k/\pi)\exp(-\lambda_k|\alpha|^2). \tag{5}$$

As mentioned in section 1, the optical field always has the Gaussian property and, consequently, the amplitude $a(t)$ has the properties (3)–(5), if it contains only a noise component with sufficiently small intensity, due to which we can neglect nonlinear effects when evolving the field. It takes place due to both the linearity of the field evolution equations and the Gaussian property of a noise electromagnetic radiation source modelled by white noise in time with a spatially distributed intensity. Since white noise is a Gaussian random process, and random solutions of the linear field equations inherit the Gaussian property of the source, then the set of solutions forms a random Gaussian process having zero average.

The following particular example of noise electromagnetic field plays an important role in quantum optics. It illustrates the above conclusion on Gaussian property of a low-intensity noise optical field. Let $a(t)$ be an amplitude of one-mode optical noise with the Lorentz spectrum having the width $\nu > 0$ [9, 14]. The functions $a(t)$ are the trajectories of the complex stationary Ornstein–Uhlenbeck process $a(t)$ [4]. These trajectories are subjected to the Langevin equation

$$\dot{a}(t) + \nu a(t) = \varphi(t) \tag{6}$$

where the complex white noise

$$\varphi(t) = \varphi_1(t) + \mathrm{i}\varphi_2(t)$$
$$\langle\varphi_1(t)\varphi_2(t')\rangle = 0 \qquad \langle\varphi_j(t)\varphi_j(t')\rangle = \sigma\delta(t-t') \qquad j = 1, 2$$

generates the optical noise $a(t)$ according to the above-mentioned mechanism. Here $\varphi_j(t)$, $j = 1, 2$, are the statistically independent real white noises having the identical intensities equal to σ. On the basis of the Langevin equations, they generate the stationary real Ornstein–Uhlenbeck processes $a_1(t), a_2(t)$ which compose the complex process $a(t)$, $a(t) = a_1(t) + \mathrm{i}a_2(t)$. The process $a(t)$ is Gaussian due to the Gaussian property of the white noise $\varphi(t)$ and to the linearity of (6).

In accordance with the Karuhnen–Loève theorem, the numbers λ_k (the functions $\psi_k(t)$) in this example, are the characteristic numbers (the eigen-functions) of the integral operator with the kernel coinciding with the correlation function $\langle a(t)a^*(t')\rangle = (\sigma/\nu)\exp(-\nu|t-t'|)$ [9, 15],

$$\psi_k(t) = \lambda_k \int_0^T \langle a(t)a^*(t')\rangle\psi_k(t')\,\mathrm{d}t' \qquad k = 1, 2, 3, \ldots.$$

3. Discrete unimodal distributions

By analogy with the theory of continuous probabilistic distributions theory, let us introduce the notion of discrete distributions unimodality.

Definition. The probability distribution $P(n)$ of an integer random value $n \in \mathbb{Z}$ is called a unimodal one, if the function $(P(n+1) - P(n))$ has no more than one change of sign, i.e. there exists the number m for which the following inequalities are valid,

$$P(n) \geqslant P(n-1) \quad \text{if } n \leqslant m \qquad \text{and} \qquad P(n) \geqslant P(n+1) \quad \text{if } n \geqslant m.$$

The number m is called the top of the distribution.

The significance of the introduced notion, is connected with the fact that the unimodality property characterizes the probabilistic distributions of the random photocounts $n \in \mathbb{Z}$ for which any physical phenomena in principle (for example, nonlinear effects or noise-induced phase transitions) are absent.

The introduced notion possesses the property of completeness which is analogous to that of the unimodal class in continuous distributions theory [6].

Statement. Let $P_N(n)$, $N = 1, 2, 3, \ldots$ be a sequence of unimodal distributions in the above sense. Let this sequence converge to the distribution $P(n)$. Then the distribution $P(n)$ is also unimodal.

The proof of this statement follows immediately by passing to the limit $N \to \infty$ in the inequalities

$$P_N(n) \geqslant P_N(n-1) \qquad \text{if } n \leqslant m_N \qquad \text{and} \qquad P_N(n) \geqslant P_N(n+1) \qquad \text{if } n \geqslant m_N$$

for each fixed n even if the tops m_N do not converge to any value of m (for mathematical details, see [6]).

4. Characteristic function $\Omega(a)$

As a rule, the distribution (1) is not calculated explicitly, i.e. it is impossible to fulfil explicitly averaging over the probability distribution of the random process $a(t)$. The distribution $P(n)$ determined by (1), (2) is very complicated even in the considered case, when the amplitude $a(t)$ is presented in the form described in section 2. As a rule, in this case averaging is not fulfilled explicitly due to the fact that the numbers λ_k are not known explicitly. But despite this, there exists an analytic expression of the characteristic function

$$F(\lambda) = \langle \exp(-\mathrm{i}\lambda\Omega(a)) \rangle \qquad \lambda > 0$$

of the random value $\Omega(a)$ determined by (2), which is valid for any complex Gaussian amplitude $a(t)$. It is just the distinctive circumstance that permits us to prove the unimodality of the distribution $P(n)$ in the case under consideration. The expression takes the form:

$$F(\lambda) = \prod_{k=1}^{\infty} (1 + \mathrm{i}\lambda/\lambda_k)^{-1} \qquad \lambda_k > 0. \tag{7}$$

The formula (7) is obtained directly by substituting the expansion (3) in $\Omega(a)$, using (4) and by averaging the function $\exp(-\mathrm{i}\lambda\Omega(a))$, using the densities (5), i.e.

$$\Omega(a) = \int_0^T |a(t)|^2 \,\mathrm{d}t = \sum_{k=1}^{\infty} |\alpha_k|^2$$

$$\langle \exp(-\mathrm{i}\lambda\Omega(a)) \rangle = \prod_{k=1}^{\infty} \int \exp(-\mathrm{i}\lambda|\alpha|^2) f_k(\alpha) \,\mathrm{d}^2\alpha.$$

In the last formula, the integration in each factor is fulfilled on the complex plane of α.

The probability distribution $P(n)$ is expressed in terms of $F(\lambda)$ in the following way

$$P(n) = (\mathrm{i}^n/n!) \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} F(\lambda)|_{\lambda=-\mathrm{i}}.$$

5. Unimodality theorem

Now we can prove our main result.

Theorem. *The probability distribution $P(n)$ is unimodal, if the complex random amplitude $a(t)$ is a complex Gaussian process with zero expectation.*

Proof. (A) The first step of our proof is its reduction to a simpler problem. We introduce the random values Ω_N determined by the characteristic functions

$$F_N(\lambda) = \langle \exp(-\mathrm{i}\lambda\Omega_N)\rangle = \prod_{k=1}^{N}(1+\mathrm{i}\lambda/\lambda_k)^{-1}. \tag{8}$$

The functions $F_N(\lambda)$ are characteristic in fact, because they are presented by the product of the characteristic functions of exponential distributions [6],

$$\begin{aligned} &(1+\mathrm{i}\lambda/\lambda_k)^{-1} = \int_0^\infty \exp(-\mathrm{i}\lambda s) g_k(s)\,\mathrm{d}s \\ &g_k(s) = \lambda_k \exp(-\lambda_k s) \qquad s > 0;\; k = 1, 2, \ldots. \end{aligned} \tag{9}$$

The probability distributions of the random values Ω_N weakly converge to the probability distribution of $\Omega(a)$, since $F_N(\lambda) \to F(\lambda)$ as $N \to \infty$ [6]. Therefore, the probability distributions

$$P_N(n) = \frac{1}{n!}\langle \Omega_N^n \exp(-\Omega_N)\rangle = (\mathrm{i}^n/n!)\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}F_N(\lambda)|_{\lambda=-\mathrm{i}} \tag{10}$$

converge to $P(n)$ as $N \to \infty$. Using the statement, we conclude that it is sufficient to prove the unimodality of the distributions $P_N(n)$, $N = 1, 2, 3, \ldots$.

(B) The next step is the obtaining of the recursion relation between the distributions $P_N(n)$. From (8), (9), we find that

$$F_N(\lambda) = \int_0^\infty \mathrm{e}^{-\mathrm{i}\lambda s}(g_1 * \ldots * g_N)(s)\,\mathrm{d}s \tag{11}$$

where $\lambda > 0$ and the symbol $*$ denotes the convolution of the densities,

$$\begin{aligned} (g_m * g_k)(s) &= \int_0^s g_m(s-t) g_k(t)\,\mathrm{d}t \\ &= \lambda_m\lambda_k \exp(-\lambda_m s)\int_0^s \exp[(\lambda_m-\lambda_k)t]\,\mathrm{d}t \qquad m, k = 1, \ldots, N. \end{aligned} \tag{12}$$

Equation (11) is valid, since convolution of densities passes to the product of the corresponding characteristic functions. From (10), (11) it follows that

$$P_N(n) = (1/n!)\int_0^\infty t^n \mathrm{e}^{-t}(g_1 * \ldots * g_N)(t)\,\mathrm{d}t.$$

Then using (12) and transposing the integration order, we have

$$\begin{aligned} P_N(n;\lambda_1\ldots\lambda_N) = (\lambda_1\ldots\lambda_N/n!)&\int_0^\infty \exp[(\lambda_2-\lambda_1)s_1]\,\mathrm{d}s_1\ldots \\ \ldots\int_{s_{N-1}}^\infty \exp[(\lambda_N-\lambda_{N-1})s_{N-1}]\,\mathrm{d}s_{N-1}&\int_{s_{N-1}}^\infty s_N^n \exp[-(1+\lambda_N)s_N]\,\mathrm{d}s_N. \end{aligned} \tag{13}$$

where we renamed $P_N(n) \Rightarrow P_N(n; \lambda_1 \dots \lambda_N)$, showing the dependence of distributions $P_N(n)$ on variables $\lambda_1, \dots, \lambda_N$. Further, we get the recursion relation, having calculated explicitly the internal integral in (13)

$$P_N(n; \lambda_1 \dots \lambda_N) = (\lambda_N/(1+\lambda_N)) \sum_{k=0}^{n} (1+\lambda_N)^{k-n} P_{N-1}(k; \lambda_1, \dots, \lambda_{N-1}). \tag{14}$$

Here, without loss of generality, we suppose that the values $\lambda_1, \dots, \lambda_N$ are different. Equation (14) permits us to calculate recurrently the distributions $P_N(n)$, using the explicit expression of $P_1(n; \lambda_1)$

$$P_1(n; \lambda_1) = \lambda_1/(1+\lambda_1)^{1+n}. \tag{15}$$

(C) At this step, we prove a statement which is the key for the proof. Hence, we introduce the convolution operation for the discrete probability distributions $p(n), q(n), n = 0, 1, 2, 3, \dots$. The operation which we shall denote by the symbol $\circ$, assigns the new probability distribution $r(n)$ to each pair p, q. It is calculated by the formula

$$r(n) = (p \circ q)(n) = \sum_{k=0}^{n} p(n-k)q(k) = \sum_{k=0}^{n} p(k)q(n-k)$$

for $n = 0, 1, 2, \dots$.

Lemma. *Let $q(n)$ be a unimodal probability distribution, i.e. the difference $[q(n+1)-q(n)]$ has no more than one change of sign. Then the probability distribution $(p \circ q)(n)$ is also unimodal, if $p(n)$ is the geometric distribution $p(n) = (1-z)z^n$, $0 < z < 1$.*

Proof of Lemma. The difference

$$(1-z)^{-1}(r(n+1) - r(n)) = z^{n+1}q(0) + \sum_{k=0}^{n} z^k[q(n+1-k) - q(n-k)]$$

$$= z^n \left(zq(0) + \sum_{k=0}^{n} z^{-k}[q(k+1) - q(k)] \right)$$

has the same number of sign changes as the difference

$$(p(n))^{-1}[r(n+1) - r(n)] = zq(0 + R(n)$$

where

$$R(n) = \sum_{k=0}^{n} z^{-k}[q(k+1) - q(k)].$$

The function $R(n)$ is unimodal, since the difference

$$R(n+1) - R(n) = z^{-(n+1)}[q(n+2) - q(n+1)]$$

has no more than one change of sign. Therefore, there does not exist more than one value m for which $R(m) > -zq(0)$ and $R(m+1) \leqslant -zq(0)$.

(D) If we introduce the geometrical distributions

$$p_l(n) = \lambda_l(1+\lambda_l)^{-(1+n)} \qquad \lambda_l > 0, l = 1, 2, \dots, N$$

then the recursion relation (14) is interpreted as

$$P_N(n; \lambda_1, \dots, \lambda_N) = (p_N \circ P_{N-1}(.; \lambda_1, \dots, \lambda_{N-1}))(n).$$

Further we repeat this recurrence and use the explicit expression (15) for $P_1(n; \lambda_1)$, and as a result, we have

$$P_N(n) = (p_1 \circ p_2 \circ \dots \circ p_N)(n).$$

The proof of the unimodality of $P_N(n)$ for each $N = 1, 2, \ldots$ is obtained now by means of induction on number N, using the statement of the lemma, taking into account the fact that $P_1(n)$ monotonously decreases and, consequently, is unimodal.

The proof of the theorem follows from the statements A and D. □

6. Conclusion

In this note, we have shown that the photocount distributions (1), (2) of noise optical field is characterized by unimodality. This property does not depend on the field intensity, registration time and on-field generating mechanism, provided that the noise is of a Gaussian type. The revealed unimodality is important from a theoretical view-point and for the applicability of the theory of the ideal receiver to optical signal recognition. Usually, the problem of photocount distribution unimodality is not discussed in optical signal transmission theory. But it is implied silently that such a unimodality takes place. However, physical situations exist in which the unimodality of photocount distribution is absent either from the theoretical view-point [16] or from the experimental one [17, 18]. In this note, we have ascertained only the sufficient conditions which guarantee the presence of unimodality. Finally, it should be noted that photocount distribution unimodality in the presence of a signal is also important for optical communication theory together with the unimodality of purely noise optical field photocounts.

Acknowledgment

This work was supported by the International Science Foundation.

References

[1] Glauber R I 1963 *Phys. Rev. Lett.* **10** 84
[2] Mandel L and Wolf E 1965 *Rev. Mod. Phys.* **37** 231
[3] Lax M 1968 *Fluctuation and Coherence Phenomena in Classical and Quantum Physics* (New York: Gordon and Breach)
[4] Gardiner C W 1985 *Handbook of Stochastic Methods* 2nd edn (Berlin: Springer)
[5] Horsthemke W and Lefever R 1984 *Noise-induced Transitions* (Berlin: Springer)
[6] Lukacs E 1970 *Characteristic Functions* 2nd edn (London: Griffin)
[7] Virchenko Yu P and Mazmanishvili A A 1995 Photocounts distribution unimodality *Prob. Inform. Trans.* **31** 90
[8] Pratt W K 1970 *Laser Communication Systems* (New York: Wiley)
[9] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[10] Price R 1962 Error probabilities for adaptive multichannel reception of binary signals IRE *Trans. Inform. Theor.* **IT-8**
[11] Loève M 1946 Sur les fonctions aleatoires stationnaires du second ordre *Rev. Sci. Instrum.* **84** 195
[12] Loève M 1955 *Probability Theory* (New York: Academic)
[13] Karhunen K 1947 Über linearen Methoden in der Wahrscheinlichkeit sre-chung *Ann. Acad. Sci. Fennicae.* A **1**
[14] Klauder J R and Sudarshan E C G 1968 *Fundamentals of Quantum Optics* (New York: Benjamin)
[15] Siegert A J F 1957 A systematic approach to a class of problems in the theory of noise and other random phenomena—part II, examples *Trans. IRE* **IT-3** 38
[16] Perina J and Krepelka J 1992 Quantum statistics of stimulated Raman and hyper-Raman scattering of light with pump depletion *J. Mod. Opt.* **39** 1029
[17] Perina J and Krepelka J 1992 Forward and back-wave mixing of non-classical light pump depletion *J. Mod. Opt.* **39** 2405
[18] Perina J 1992 *Quantum Statistics of Linear an Nonlinear Optical Phenomena* (Dordrecht: Kluwer)