

ON POLYNOMIAL EXPANSIONS OF SECOND-ORDER DISTRIBUTIONS*

E. WONG† AND J. B. THOMAS‡

Summary. Under certain general conditions the problem of finding the principal solution to the Fokker-Planck equation can be reduced to an eigenvalue problem of the Sturm-Liouville type. In particular, when the eigenfunctions are orthogonal polynomials, the resulting second-order probability distribution functions are of a well-known class. The conditions which give rise to polynomial solutions are investigated in this paper.

Introduction. Barrett and Lampard [1, 2] have investigated a class of two-dimensional probability density functions $p(x, y)$, which admit expansions of the form

$$(1) \quad p(x, y) = p_1(x)p_2(y) \sum_{n=0}^{\infty} A_n \varphi_n(x) \theta_n(y),$$

where $\varphi_n(x)$ and $\theta_n(y)$ are polynomials of degree n , orthonormal with respect to the first order densities $p_1(x)$ and $p_2(y)$, respectively. Since (1) represents an expansion of a function of two variables in a single sum, the expansion is said to be diagonal. If $p(x, y)$ is symmetric in the variables x and y , as it is in all the examples shown by Barrett and Lampard, (1) becomes

$$(2) \quad p(x, y) = p(x)p(y) \sum_{n=0}^{\infty} A_n \varphi_n(x) \varphi_n(y).$$

It was shown that two important examples of such two-dimensional densities are those of the Gaussian and chi-square distributions with Hermite and Laguerre polynomials as the corresponding orthogonal polynomials.

In part, the diagonal polynomial expansions of the Gaussian and chi-square two-dimensional density functions can be understood as consequences of the fact that the Fokker-Planck equations for the Gaussian and chi-square Markoff processes reduce to the differential equations for Hermite and Laguerre polynomials. In this note, the necessary and sufficient conditions under which the Fokker-Planck equation reduces to an eigenvalue problem with polynomial solutions are derived. It is further shown that these conditions restrict the solutions to three distinct classes. In addition to the Gaussian and chi-square cases, there exists a class of

* Received by the editors April 25, 1961 and in final revised form January 24, 1962.

† International Business Machines Corporation, Research Center, P. O. Box 218, Yorktown Heights, New York.

‡ Princeton University, Princeton, New Jersey.

stationary Markoff processes, for which the two-dimensional densities admit diagonal expansions in terms of polynomials, the polynomials being Jacobi polynomials. This class of two-dimensional densities represents a natural extension of some familiar first order densities, and may be useful as models in analysis.

The eigenvalue problem. It is well known [3] that the principal solution $p(x_0|x; t)$ of the Fokker-Planck equation,

$$(3) \quad \frac{1}{2} \frac{\partial^2}{\partial x^2} [B(x)p(x_0|x; t)] - \frac{\partial}{\partial x} [A(x)p(x_0|x; t)] = \frac{\partial p(x_0|x; t)}{\partial t}, \quad t \geq 0,$$

represents the conditional probability density function of a one-dimensional stationary Markoff process. If it is assumed that an equilibrium density function $p(x)$ exists such that

$$(4) \quad p(x) = \lim_{t \rightarrow \infty} p(x_0|x; t),$$

and that

$$(5) \quad \lim_{t \rightarrow \infty} \frac{\partial p(x_0|x; t)}{\partial t} = 0,$$

then, (3) reduces, in the limit, to

$$(6) \quad \frac{1}{2} \frac{d^2}{dx^2} (B(x)p(x)) - \frac{d}{dx} (A(x)p(x)) = 0.$$

Equation (6) can be integrated once to yield

$$(7) \quad \frac{1}{2} \frac{d}{dx} (B(x)p(x)) - A(x)p(x) = \text{const.}$$

Sufficient conditions for the constant in (7) to be zero have been investigated by Andronov, Pontryagin and Witt [4]. Here, we postulate the existence of a principal solution $p(x_0|x; t)$ of (3) which satisfies the conditions of (4) and (5) and the following:

$$(8) \quad \frac{1}{2} \frac{d}{dx} (B(x)p(x)) - A(x)p(x) = 0.$$

As a consequence [5], the equilibrium density $p(x)$ satisfies the differential equation

$$(9) \quad \frac{dp(x)}{dx} = \frac{1}{B(x)} \left(2A(x) - \frac{dB(x)}{dx} \right) p(x).$$

With the usual substitution of $(f(t)p(x)\varphi(x))$ for $p(x_0|x; t)$, (3) separates into

$$(10) \quad \frac{df(t)}{dt} = -\lambda f(t)$$

and

$$(11) \quad \frac{1}{2} \frac{d^2}{dx^2} (B(x)p(x)\varphi(x)) - \frac{d}{dx} (A(x)p(x)\varphi(x)) = -\lambda p(x)\varphi(x).$$

With the use of (8), (11) becomes

$$(12) \quad \frac{1}{2} \frac{d}{dx} \left(B(x)p(x) \frac{d\varphi(x)}{dx} \right) + \lambda p(x)\varphi(x) = 0,$$

which is of the Sturm-Liouville form. Assume the eigenvalues to be discrete, this being the case of interest at present. Then it follows from (12) that the eigenfunctions $\varphi(x)$ are orthogonal with respect to $p(x)$ and can be normalized so that

$$(13) \quad \int_{x_1}^{x_2} p(x)\varphi_m(x)\varphi_n(x) dx = \delta_{mn},$$

where x_1 and x_2 are the limits of the range of interest.

Equation (10) is easily solved and yields

$$(14) \quad f_n(t) = k_n e^{-\lambda_n t}.$$

Thus, under the assumption of discrete eigenvalues, $p(x_0|x; t)$ can be written as

$$(15) \quad p(x_0|x; t) = p(x) \sum_{n=0}^{\infty} G_n(x_0) e^{-\lambda_n t} \varphi_n(x).$$

The functions $G_n(x_0)$ can be obtained from the requirement that $p(x_0|x; t)$ be the principal solution, i.e.,

$$(16) \quad p(x_0|x; 0) = \delta(x - x_0).$$

Therefore (15) becomes

$$(17) \quad p(x) \sum_{n=0}^{\infty} G_n(x_0) \varphi_n(x) = \delta(x - x_0).$$

After multiplying both sides of (17) by $\varphi_m(x)$ and integrating over the range x_1 to x_2 , $G_n(x)$ is found to be

$$(18) \quad G_n(x) = \varphi_n(x_0),$$

where the orthonormality condition (13) has been used.

Thus (15) becomes

$$(19) \quad p(x_0|x; t) = p(x) \sum_{n=0}^{\infty} e^{-\lambda_n t} \varphi_n(x_0) \varphi_n(x).$$

The joint probability density function $p(x_0, x; t)$ is given by

$$(20) \quad p(x_0, x_0; t) = p(x_0)p(x) \sum_{n=0}^{\infty} e^{-\lambda_n t} \varphi_n(x_0)\varphi_n(x).$$

When the $\varphi_n(x)$ are orthogonal polynomials, (20) shows that $p(x_0, x; t)$ is precisely of the form studied by Barrett and Lampard. Now the problem is to determine the conditions under which the eigenfunctions $\varphi_n(x)$ form a complete orthonormal set of polynomials.

Conditions for polynomial solutions. Consider the integral defined by

$$(21) \quad I = \int_{x_1}^{x_2} \frac{d}{dx} \left(B(x)p(x) \frac{d\varphi_n(x)}{dx} \right) G_m(x) dx,$$

where $G_m(x)$ is an arbitrary polynomial of degree $m < n$. Since $\varphi_n(x)$ is a solution of (12), this integral becomes

$$(22) \quad I = -2\lambda_n \int_{x_1}^{x_2} p(x)\varphi_n(x)G_m(x) dx.$$

Now if $\varphi_n(x)$ is a polynomial of degree n , the arbitrary polynomial $G_m(x)$ can be expanded in terms of the eigenfunctions $\varphi_k(x)$ as

$$(23) \quad G_m(x) = \sum_{k=0}^m a_k \varphi_k(x).$$

Substitution of (23) into (22) yields

$$(24) \quad I = -2\lambda_n \sum_{k=0}^m a_k \int_{x_1}^{x_2} p(x)\varphi_n(x)\varphi_k(x) dx.$$

The right hand side of (24) is zero from the orthogonality condition (13) and the fact that m is less than n .

If the right-hand side of (21) is integrated by parts twice, the integral I can be expressed as

$$(25) \quad I = \left\{ B(x)p(x) \left(G_m(x) \frac{d\varphi_n(x)}{dx} - \varphi_n(x) \frac{dG_m(x)}{dx} \right) \right\} \Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} \varphi_n(x) \frac{d}{dx} \left(B(x)p(x) \frac{dG_m(x)}{dx} \right) dx.$$

Let (25) be rewritten as

$$(26) \quad I = I_1 + I_2.$$

where I_2 is given by

$$(27) \quad I_2 = \int_{x_1}^{x_2} \varphi_n(x) \frac{d}{dx} \left(B(x)p(x) \frac{dG_m(x)}{dx} \right) dx.$$

With the use of (23) and (12), I_2 can be rewritten as

$$(28) \quad I_2 = -2 \sum_{k=0}^m a_k \lambda_k \int_{x_1}^{x_2} p(x) \varphi_n(x) \varphi_k(x) dx.$$

From the orthogonality condition and the fact that m is less than n , it is seen that I_2 is again zero. Thus it follows immediately that I_1 also vanishes.

The fact that I_1 and I_2 vanish separately is expressed by the following equations:

$$(29) \quad \left\{ B(x)p(x) \left(G_m(x) \frac{d\varphi_n(x)}{dx} - \varphi_n(x) \frac{dG_m(x)}{dx} \right) \right\} \Big|_{x_1}^{x_2} = 0, \quad m < n,$$

$$(30) \quad \int_{x_1}^{x_2} \varphi_n(x) \frac{d}{dx} \left(B(x)p(x) \frac{dG_m(x)}{dx} \right) dx = 0, \quad m < n.$$

Equation (29) is satisfied for arbitrary polynomial $G_m(x)$ if and only if

$$(31) \quad \begin{aligned} B(x_1)p(x_1)\varphi_n'(x_1) &= B(x_2)p(x_2)\varphi_n'(x_2) \\ &= B(x_1)p(x_1)\varphi_n(x_1) = B(x_2)p(x_2)\varphi_n(x_2) = 0 \end{aligned}$$

where prime denotes derivative with respect to x . Equation (31) is satisfied if either $B(x_1)p(x_1) = 0$ or if $\varphi_n'(x_1) = \varphi_n(x_1) = 0$, and if either $B(x_2)p(x_2) = 0$ or $\varphi_n'(x_2) = \varphi_n(x_2) = 0$. The condition that φ_n and its derivative both vanish at either end point forces $\varphi_n(x)$ to be zero. Therefore $B(x)p(x)$ must vanish at both end points, i.e.,

$$(32) \quad B(x_1)p(x_1) = B(x_2)p(x_2) = 0.$$

Equation (30) can be written, with the use of (8), as

$$(33) \quad \int_{x_1}^{x_2} p(x)\varphi_n(x) \left(B(x) \frac{d^2 G_m(x)}{dx^2} + A(x) \frac{dG_m(x)}{dx} \right) dx = 0.$$

Equation (33) is to be satisfied for all polynomials $G_m(x)$ of degree less than n . In particular, for $G_m(x) = x$, (33) becomes

$$(34) \quad \int_{x_1}^{x_2} p(x)A(x)\varphi_n(x) dx = 0, \quad n = 2, 3, \dots$$

As a consequence of (34), $A(x)$ must be a polynomial of degree not more than one, i.e.,

$$(35) \quad A(x) = ax + b.$$

Similarly, for $G_m(x) = x^2$, it is found that $B(x)$ must be a polynomial of degree not more than two, i.e.,

$$(36) \quad B(x) = cx^2 + dx + e.$$

Thus, (35) and (36) are necessary conditions in order for (33) to be satisfied. That (35) and (36) are also sufficient conditions is obvious, since now (33) can be written as

$$(37) \quad \int_{x_1}^{x_2} p(x)\varphi_n(x)G_m(x) dx = 0, \quad m < n,$$

where $G_m(x)$ is a polynomial of degree m . In addition, in order for the polynomials to be normalizable, $p(x)$ must be such that all the moments are finite, i.e.,

$$(38) \quad \int_{x_1}^{x_2} x^n p(x) dx < \infty, \quad n = 0, 1, \dots, n < \infty.$$

Equations (32), (35), (36) and (38) represent a set of necessary conditions for (12) to yield as eigenfunctions a complete orthonormal set of polynomials. The sufficiency of these conditions remains to be proved, since $\varphi_n(x)$ has been assumed to be a polynomial of degree n in (23) through (38).

With the use of (35) and (36), (12) can be rewritten as

$$(39) \quad \frac{1}{2}(cx^2 + dx + e) \frac{d^2\varphi_n(x)}{dx^2} + (ax + b) \frac{d\varphi_n(x)}{dx} + \lambda_n \varphi_n(x) = 0.$$

Now let $\varphi_n(x)$ be given by

$$(40) \quad \varphi_n(x) = \sum_{k=0}^n d_k^{(n)} x^k,$$

and substitute (40) into (39). Equating the coefficient of each power of x to zero results in the following set of equations:

$$(41) \quad \left(\frac{1}{2}n(n-1) + na + \lambda_n\right)\alpha_n^{(n)} = 0,$$

$$(42) \quad \left(\frac{1}{2}(n-1)(n-2)c + (n + (n-1)a + \lambda_n)\alpha_{n-1}^{(n)} + \left(\frac{1}{2}n(n-1)d + nb\right)\alpha_n^{(n)}\right) = 0,$$

and

$$(43) \quad \left(\frac{1}{2}k(k-1)c + ka + \lambda_n\right)\alpha_k^{(n)} + \left(\frac{1}{2}k(k+1)d + (k+1)b\right)\alpha_{k+1}^{(n)} + \left(\frac{1}{2}(k+1)(k+2)e\right)\alpha_{k+2}^{(n)} = 0, \quad k = 0, 1, 2, \dots, n-2.$$

Equation (41) yields the solution for λ_n . From (42) and (43) the coefficients $\alpha_k^{(n)}$ for values of k up to $n-1$ can be expressed in terms of $\alpha_n^{(n)}$. The coefficient $\alpha_n^{(n)}$ in turn can be found using the normalization integral (13), the convergence of which is assured by (38).

Summarizing, necessary and sufficient conditions for (12) to yield as eigenfunctions a complete orthonormal set of polynomials are as follows:

$$(32) \quad B(x_1)p(x_1) = B(x_2)p(x_2) = 0,$$

$$(35) \quad A(x) = ax + b,$$

$$(36) \quad B(x) = cx^2 + dx + e,$$

and

$$(38) \quad \int_{x_1}^{x_2} x^n p(x) dx < \infty, \quad n = 0, 1, \dots, n < \infty.$$

It is of interest to note that (35) and (36) together with (9) imply that $p(x)$ belongs to the Pearson system of distributions.

Construction of a class of two-dimensional distributions. The conditions (32), (35), (36), and (38) derived in the previous section restrict the density function $p(x)$ to be one of three forms [7]. Without loss of generality these forms can be taken as

$$(44) \quad p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty \leq x \leq \infty,$$

$$(45) \quad p(x) = \frac{1}{\Gamma(\alpha + 1)} x^\alpha e^{-x}, \quad 0 \leq x \leq \infty, \alpha > -1,$$

and

$$(46) \quad p(x) = \frac{1}{2^{\alpha+\beta+1}} \frac{\Gamma(\alpha + \beta + 2)}{\Gamma(\alpha + 1)\Gamma(\beta + 1)} (1 - x)^\alpha (1 + x)^\beta, \quad -1 \leq x \leq 1, \alpha, \beta > -1.$$

Equation (44) represents the familiar Gaussian case with Hermite polynomials as the corresponding polynomials. For special values of α ; namely, $\alpha = n - \frac{1}{2}$, $n = 0, 1, 2, \dots$, (45) represents the density function of the chi-square distribution. The corresponding polynomials are the Laguerre polynomials $L_n^\alpha(x)$.

Equation (46) represents the density function for the Pearson type I distribution [6, 8]. For special values of α and β (46) reduces to density functions of well-known distributions. For example, the case of $\alpha = \beta = 0$ yields the uniform density function

$$(47) \quad p(x) = \frac{1}{2}, \quad -1 \leq x \leq 1;$$

and the case of $\alpha = \beta = \frac{1}{2}$ yields

$$(48) \quad p(x) = \frac{1}{\pi} \frac{1}{\sqrt{1 - x^2}}, \quad -1 \leq x \leq 1,$$

the density function for the distribution of a sine wave with unit amplitude and random phase [9]. The results thus far show that it is possible to construct a class of second-order density functions as diagonal expansions in

terms of orthogonal polynomials based on distributions of Pearson type I (see also [10, 11]).

The polynomials orthonormalized with respect to the density function $p(x)$ of (46) are the Jacobi polynomials [12]

$$\begin{aligned}
 \varphi_n(x) &= \frac{(-1)^n}{2^n} \\
 (49) \quad &\sqrt{\frac{(2n + \alpha + \beta + 1)\Gamma(n + \alpha + \beta + 1)\Gamma(\alpha + 1)\Gamma(\beta + 1)}{\Gamma(\alpha + \beta + 2)\Gamma(n + \alpha + 1)\Gamma(n + \beta + 1)n!}} \\
 &\quad \times (1 - x)^{-\alpha}(1 + x)^{-\beta} \frac{d^n}{dx^n} [(1 - x)^{n+\alpha}(1 + x)^{n+\beta}].
 \end{aligned}$$

It should be noted that (49) represents a rather wide class of polynomials, which includes Legendre and Chebyshev and other Gegenbauer-type polynomials as special cases.

The eigenvalues λ_n can be determined from (12). To proceed, let (12) be rewritten as

$$(50) \quad \frac{1}{2}B(x) \frac{d^2\varphi_n(x)}{dx^2} + A(x) \frac{d\varphi_n(x)}{dx} + \lambda_n \varphi_n(x) = 0.$$

The functions $A(x)$ and $B(x)$ satisfy (9) and (35) and (36). For the $p(x)$ of (46), these functions are

$$(51) \quad A(x) = \gamma(\beta - \alpha) - \gamma(\alpha + \beta + 2)x,$$

and

$$(52) \quad B(x) = 2\gamma(1 - x^2),$$

where γ is an arbitrary positive constant. If (51) and (52) are substituted into (50) and the coefficient of the x^n term is set equal to zero, it is found that

$$(53) \quad -\gamma n(n - 1) - \gamma(\alpha + \beta + 2)n + \lambda_n = 0,$$

or

$$(54) \quad \lambda_n = \gamma n(n + \alpha + \beta + 1).$$

The construction of the class of two-dimensional density functions based on the type I Pearson distributions is completed by the use of (46), (49) and (54) in (20). These density functions have the form

$$(55) \quad p(x_0, x; t) = p(x_0)p(x) \sum_{n=0}^{\infty} e^{-\gamma n(n+\alpha+\beta+1)t} \varphi_n(x_0)\varphi_n(x),$$

TABLE 1

Conditions for Polynomial Solutions	Results		
	Density	Range	Polynomials
(1) $B(x_1)p(x_1) = B(x_2)p(x_2) = 0$	$k_1 e^{-1/2 x^2}$	$-\infty, \infty$	Hermite
(2) $A(x) = ax + b$	$k_2 x^\alpha e^{-x}$	$0, \infty$	Laguerre
(3) $B(x) = cx^2 + dx + e$	$k_3(1-x)^\alpha(1+x)^\beta$	$-1, 1$	Jacobi
(4) $\int_{x_1}^{x_2} x^n p(x) dx < \infty, n < \infty$			

with $p(x)$ and $\varphi_n(x)$ given by (46) and (49). As was noted by Barrett and Lampard, the normalized covariance function $\rho(t)$ defined by

$$(56) \quad \rho(t) \triangleq \left\langle \left(\frac{x - \mu}{\sigma} \right) \left(\frac{x_0 - \mu}{\sigma} \right) \right\rangle,$$

where μ and σ^2 are the mean and variance respectively, can be expressed as

$$(57) \quad \rho(t) = \langle \varphi_1(x) \varphi_1(x_0) \rangle.$$

Equation (55) shows that, for the class of Markoff processes, with joint density expressed by (55), $\rho(t)$ is given by

$$(58) \quad \rho(t) = e^{-\gamma(\alpha+\beta+2)t}.$$

Thus (55) can be generalized to the form

$$(59) \quad p(x_0, x) = p(x_0)p(x) \sum_{n=0}^{\infty} (\rho)^{[n(\alpha+\beta+1)]/[1+\alpha+\beta+2]} \varphi_n(x_0) \varphi_n(x).$$

As in the cases of Gaussian and chi-square distributions, the second-order distribution represented by (59) is uniquely determined by the first-order distribution and the covariance functions ρ .

Conclusion. The principal results of this paper are summarized in Table 1.

The results in this paper find applications in a number of areas. Among these are cross correlation of outputs of nonlinear devices [13, 14, 15] and optimization of nonlinear networks [16]. In addition, the solutions of the Fokker-Planck equation are of general interest in problems involving noise and other stochastic processes (see, for example, [17]). Specific applications are found in the investigation of dynamic systems [4], the level crossings of random processes [18], and the distribution of functionals [19].

REFERENCES

1. J. F. BARRETT AND D. G. LAMPARD, *An expansion for some second-order probability distributions and its application to noise problems*, IRE Transactions on Information Theory, IT-1 (1955), pp. 10-15.

2. K. S. MILLER, R. I. BERNSTEIN AND L. E. BLUMENSON, *Generalized Rayleigh process*, Quart. Appl. Math., 16 (1958), pp. 137-145. See, in particular, reference to private communication from D. G. Lampard.
3. M. C. WANG AND G. E. UHLENBECK, *On the theory of the Brownian motion II*, Rev. Mod. Phys., 17 (1945), pp. 323-342.
4. A. A. ANDRONOV, L. S. PONTRYAGIN, AND A. A. WITT, *On the statistical investigation of dynamic systems*, J. Exp. and Th. Phys. (U. S. S. R.), 3(1933), p. 165.
5. W. FELLER, *An Introduction to Probability Theory and Its Application*, John Wiley and Sons, New York, 1957, Chapter XIV.
6. H. CRAMÉR, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946, pp. 248-249.
7. D. JACKSON, *Fourier Series and Orthogonal Polynomials*, The Mathematical Association of America, pp. 142-148; 1941.
8. W. P. ELDERTON, *Frequency Curves and Correlation*, 3rd ed., Cambridge University Press, Cambridge, England, 1938.
9. S. D. RICE, *Mathematical analysis of random noise*, Bell System Tech. J., 24 (1945), p. 236.
10. E. WONG, *Vector stochastic processes in problems of communication theory*, Ph.D. Dissertation, Department of Electrical Engineering, Princeton University, 1959.
11. S. KARLIN AND J. MCGREGOR, *Classical diffusion processes and total positivity*, J. Math. Anal. Appl., 1 (1960), pp. 163-183.
12. A. ERDÉLYI, et al., *Higher Transcendental Functions*, vol. 2, McGraw-Hill, New York, 1955.
13. J. J. BUSSGANG, *Cross-correlation function of amplitude-distorted Gaussian signals*, Res. Lab. of Electronics, M.I.T., Cambridge, Mass., Tech. Rep. 216, 1952.
14. J. L. BROWN, JR., *On a cross-correlation property for stationary random processes*, IRE Trans. on Information Theory, IT-3 (1957), pp. 28-31.
15. R. LEIPNIK, *The effect of instantaneous nonlinear devices on cross-correlation*, IRE Trans. on Information Theory, IT-4 (1958), pp. 73-76.
16. L. A. ZADEH, *Classification of nonlinear filters*, IRE WESCON Conv. Rec., 1957.
17. D. MIDDLETON, *An Introduction to Statistical Communication Theory*, McGraw-Hill, New York, 1960, Chapter X.
18. S. O. RICE, *Distribution of the durations of fades in radio transmission: Gaussian noise model*, Bell System Tech. J., 37 (1958), pp. 581-635.
19. D. A. DARLING AND A. J. F. SIEGERT, *A systematic approach to a class of problems in the theory of noise and other random phenomena—Part I*, IRE Trans. on Information Theory, IT-3 (1957), pp. 32-37.