DISTRIBUTION OF THE EXTREME VALUES OF THE SUM OF n SINE WAVES PHASED AT RANDOM*

BY

S. O. RICE

Bell Telephone Laboratories, Inc., New York

1. Introduction. The statistical behavior of the sum of n sine waves phased at random has been studied in connection with a number of technical problems. These include radio wave fading and overloading in multichannel telephony.[1]

When the sine waves are of unit amplitude their sum may be written as

$$z = \sum_{m=1}^{n} \cos \varphi_m \tag{1.1}$$

where $\varphi_1, \varphi_2, \cdots, \varphi_n$ are independent random angles, each distributed uniformly over the range $-\pi$ to π. z cannot exceed n. When $n - 2 < z \leq n$ the probability density of z may be expressed as a power series in $(n - z)$, as is shown in Sec. 2.

There is a close relation between the distribution of z and the problem of the random walk in two dimensions, and the two are often treated together. Several equations connecting them are given in Sec. 3. In Sec. 4 the results of Sec. 2 are used to obtain the first few terms in a series for the distribution of the extreme values in the random walk problem.

When n is large the central portion of the distribution for z approaches a normal law. In Sec. 5 an attempt is made to obtain an approximation to the distribution over the entire range of z by interpolating between the normal law result for small z and the results of Sec. 2 which hold for extreme values of z. The work is carried out first for the random walk distribution and then translated to the z distribution. This procedure is used because the random walk distribution seems to be better suited to our method of interpolation than does the z distribution. Figure 1 is associated with the interpolation between the results given by Pearson[2] and Rayleigh[3] for the random walk and those of Sec. 4.

I wish to express my thanks for the many helpful suggestions concerning this paper which I have received from Mr. John Riordan and others.

2. Series for the probability density of z when z is near n. Let $q_n(z)$ denote the probability density of the random variable z defined by (1.1). Then, when $n - 1 <$

*Received Jan. 27, 1954.

[1]See, for instance, W. R. Bennett, Distribution of the sum of randomly phased components, Quarterly Appl. Math. **5**, 385-393 (Jan. 1948). References to earlier work will be found in this paper. Mention should also be made of papers by F. Horner, Phil. Mag. (7) **37**, 145-162 (1946), and R. D. Lord, Phil. Mag. (7) **39**, 66-71 (1948). The second paper gives our equation (3.5).

[2]Drapers' Co. Research Memoirs Biometric Series III. *Math. Contributions to the theory of evolution—XV. A mathematical theory of random migration,* Karl Pearson assisted by John Blakeman, London (1906).

[3]Rayleigh, Phil. Mag. **10**, 73 (1880) and Phil. Mag. **37**, 321-347 (1919).

$z < n + 1$,

$$q_{n+1}(z) = \int_{z-1}^{n} q_1(z - u)q_n(u)\, du$$
$$= \frac{1}{\pi}\int_{z-1}^{n} [1 - (z - u)^2]^{-1/2} q_n(u)\, du. \tag{2.1}$$

The change of variables

$$x = n + 1 - z, \qquad u = z - 1 + vx$$

carries (2.1) into

$$q_{n+1}(n + 1 - x) = \frac{(x/2)^{1/2}}{\pi}\int_0^1 [v(1 - vx/2)]^{-1/2} q_n[n - (1 - v)x]\, dv \tag{2.2}$$

which holds when $0 < x < 2$. When we place the assumed expansion

$$q_n(n - x) = A_n x^{n/2-1}\left[1 + \sum_{k=1}^{\infty} a_{nk}(x/4)^k/n(n + 2) \cdots (n + 2k - 2)\right] \tag{2.3}$$

in (2.2) and use

$$q_1(z) = q_1(1 - x) = \pi^{-1}[2x - x^2]^{-1/2} \tag{2.4}$$

we obtain

$$A_n = (2\pi)^{-n/2}/\Gamma(n/2) \tag{2.5}$$

and a set of recurrence relations, the lth of which is

$$a_{n+1,l} = \sum_{k=0}^{l} a_{n,l-k}\alpha_k^2/k!, \qquad l = 0, 1, 2, \cdots, \tag{2.6}$$

where $a_{n0} = 1$, $\alpha_0 = 1$ and

$$\alpha_k = 1\cdot 3\cdot 5 \cdots (2k - 1).$$

The series in (2.3) converges for $| x | < 2$ as may be seen by substituting $q_n(n - x) = x^{-1+n/2}f_n(x)$, $f_1(x) = \pi^{-1}(2 - x)^{-1/2}$ in (2.2). An integral is obtained which may be used to show in succession that $f_2(x)$, $f_3(x)$, $\cdots$ $f_n(x)$ are analytic functions of x inside the circle $| x | = R$, $R < 2$, in the complex x-plane.

Equations (2.6) and $a_{1k} = \alpha_k^2/k!$ lead to

$$\sum_{k=0}^{\infty} a_{nk}s^k = \left[\sum_{j=0}^{\infty} \alpha_j^2 s^j/j!\right]^n. \tag{2.7}$$

This is merely a formal result because the series on the right does not converge.

It is interesting to note that (2.7) fits in with some heuristic manipulation of the integral

$$q_n(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-izt}[J_0(t)]^n\, dt. \tag{2.8}$$

Thus if we raise the asymptotic expression

$$J_0(t) \sim \sum_{m=0}^{\infty} \frac{\pi^{-1/2}\alpha_m^2}{m!4^m}\left[\frac{e^{it}}{(i2t)^{m+1/2}} + \frac{e^{-it}}{(-i2t)^{m+1/2}}\right] \tag{2.9}$$

(where $-\pi < \arg t < \pi$ and $\arg(-i) = -\pi/2$) to the nth power we obtain the sum of terms of the form $\exp i(n - 2l)t$ times a series in $1/t$ with $l = 0, 1, \cdots n$. Let the t-plane be cut along the negative real t axis and let the limits of integration in (2.8) be $-i \pm \infty$ instead of $\pm\infty$. When we substitute (2.9) in (2.8), assume $n - 2 < z < n$, and use

$$\int_{-i-\infty}^{-i+\infty} \frac{e^{ixt}}{(it)^{\nu}}\, dt = \frac{2\pi x^{\nu-1}}{\Gamma(\nu)}, \qquad x > 0$$

$$0, \qquad x < 0$$

only the terms multiplied by $\exp[i(n - z)t]$ (corresponding to $l = 0$, $x = n - z$) contribute to the value of the integral. Furthermore, they lead to the same series in x for $q_n(n - x)$ as does (2.7). If this procedure could be justified and generalized it might lead to an expression for (2.8) which would supplement the one obtained by the method used by W. R. Bennett.[1]

The coefficients a_{nk} may be expressed in terms of Bell's Y polynomials.* Thus, multiplying both sides of (2.7) by $t^n/n!$ and summing n from 0 to ∞ shows that a_{nk} is the coefficient of $t^n s^k/n!$ in

$$[\exp t]\left[\exp \sum_{j=1}^{\infty} \left(t\alpha_j^2 \frac{s^j}{j!}\right)\right] = e^t \sum_{k=0}^{\infty} \frac{s^k}{k!} Y_k(t\alpha_1^2, t\alpha_2^2, \cdots, t\alpha_k^2)$$

where the $Y_k(y_1, y_2 \cdots y_k)$ are Bell's polynomials:

$$Y_0 = 1, \qquad Y_1 = y_1, \qquad Y_2 = y_2 + y_1^2, \qquad Y_3 = y_3 + 3y_2y_1 + y_1^3, \cdots.$$

Consequently,

$$a_{nk} = \frac{n!}{k!} Y_k(t\alpha_1^2, t\alpha_2^2, \cdots, t\alpha_k^2), \tag{2.10}$$

where, after writing the right hand side as a polynomial in t, t^m is replaced by $1/(n - m)!$. We obtain in this way

$$q_n(n - x) = \frac{(x/2\pi)^{n/2-1}}{2\pi\Gamma(\frac{1}{2}n)}\left[1 + \frac{x}{4} + \frac{(n + 8)x^2}{32(n + 2)} + \frac{(n^2 + 24n + 200)x^3}{384(n + 2)(n + 4)} + \cdots\right], \tag{2.11}$$

where the series converges when $0 \leq x < 2$.

The probability $\Psi(E)$ that $|z| \geq E$ is given by

$$\Psi_n(E) = 2\int_E^n q_n(z)\, dz = 2\int_0^{n-E} q_n(n - x)\, dx \tag{2.12}$$

and when $n - 2 < E \leq n$ this leads to

$$\Psi_n(E) = \frac{2}{\Gamma[\frac{1}{2}(n + 2)]}\left[\frac{n - E}{2\pi}\right]^{n/2}\left[1 + \frac{n(n - E)}{4(n + 2)} + \frac{n(n + 8)(n - E)^2}{32(n + 2)(n + 4)} + \cdots\right]. \tag{2.13}$$

---

*Actually what is used is the slightly more general version of these polynomials given by John Riordan, *Derivatives of composite functions*, Bull. Amer. Math. Soc. **52**, 664-667 (1946).

3. Relation between z and the problem of the random walk. The random variable z, defined as the sum of n cosines by (1.1), may be regarded as the projection of the resultant r (of n unit random vectors) on the x-axis. Hence, we may write $z = r \cos \theta$ where θ is a random angle distributed uniformly over the interval $(0, 2\pi)$. The probability $p_n(r)\, dr$ that the length of the resultant lies between r and $r + dr$ is given by the random walk distribution when the n elementary linear walks are of unit length each. We shall use $\Phi_n(r)$ to denote the probability that the resultant equals or exceeds r. Then the connection between z and r leads to the following relations between the probability functions:

$$\Psi_n(E) = \frac{2}{\pi} \int_E^n p_n(r) \operatorname{arc\,cos} (E/r)\, dr, \tag{3.1}$$

$$\Psi_n(E) = \frac{2E}{\pi} \int_E^n r^{-1}(r^2 - E^2)^{-1/2} \Phi_n(r)\, dr, \tag{3.2}$$

$$\Phi_n(r) = -r \frac{d}{dr} \int_r^n (E^2 - r^2)^{-1/2} \Psi_n(E)\, dE, \tag{3.3}$$

$$\Phi_n(r) = r^2 \int_r^n E^{-1}(E^2 - r^2)^{-1/2}[E^{-1}\Psi_n(E) - \Psi_n'(E)]\, dE, \tag{3.4}$$

$$q_n(z) = \frac{1}{\pi} \int_z^n (r^2 - z^2)^{-1/2} p_n(r)\, dr, \tag{3.5}$$

$$p_n(r) = -2r \int_r^n (z^2 - r^2)^{-1/2}[dq_n(z)/dz]\, dz. \tag{3.6}$$

In these equations E, r and z are assumed to be less than n and $p_1(r)$ is to be interpreted as an impulse function. In (3.6) n must exceed two but this causes no difficulty since it is known that

$$p_2(r) = (2/\pi)(4 - r^2)^{-1/2}, \qquad -2 < r < 2. \tag{3.7}$$

In going from (3.1) to (3.2) we have integrated by parts. Setting $n^2 - r^2 = \xi$, $n^2 - E^2 = x$ in (3.2) converts it to a special case of Abel's integral equation* whose solution gives (3.3). The remaining equations are obtained by the same kind of analysis.

4. Random walk distributions when r is near n. When $n > 2$ and $n - 2 < r < n$, substitution of the expression (2.11) for $q_n(z)$ in the integral (3.6) for the probability density $p_n(r)$ of r gives

$$p_n(r) = \frac{n^{1/2}}{2\pi\Gamma[\frac{1}{2}(n-1)]} \left(\frac{n-r}{2\pi}\right)^{(n-3)/2} \left[1 + \frac{(n-1)(n-r)}{4n} + \frac{(n^2 + 4n - 9)(n-r)^2}{32n^2} + \cdots \right]. \tag{4.1}$$

When $n = 2$ the method fails, but it is not difficult to show from (3.7) that (4.1) also holds in this case. Expression (4.1) may also be obtained from the recurrence relation for $p_n(r)$ by a method similar to that used in Sec. 2 to obtain $q_n(n - x)$. Pearson[2] has

*See, for example, Whittaker and Watson, *Modern analysis*, 4th ed., Cambridge, 1927, p. 229.

given, essentially, the leading term in (4.1) and has given one or two more terms for $n = 3, 4, 5, 6$.

Integrating (4.1) termwise gives

$$\Phi_n(r) = \int_r^n p_n(\rho)\, d\rho = \frac{n^{1/2}}{\Gamma[\frac{1}{2}(n+1)]}\left[\frac{n-r}{2\pi}\right]^{(n-1)/2}\left[1 + \frac{(n-1)^2(n-r)}{4n(n+1)} + \frac{(n-1)(n^2+4n-9)}{32n^2(n+3)}(n-r)^2 + \cdots\right] \tag{4.2}$$

which holds for $n - 2 < r < n$.

5. Approximations for $\Phi_n(r)$ and $\Psi_n(E)$. Numerical values of the various probability densities and distributions have been given by Pearson[2], Slack[4], Bennett[1] and others for values of n up to 10 (and somewhat beyond in certain cases.) The values of $\Phi_n(r)$ and $\Psi_n(E)$ given by Bennett were computed from series which converge for all values of r and E between zero and n. In this section we shall consider a method of estimating values of $\Phi_n(r)$ and $\Psi_n(E)$ which involves only a small amount of calculation but which, of course, lacks the accuracy of the computations mentioned above.

When n is large, but E and r of moderate size, it is known that

$$\Phi_n(r) \approx \exp(-r^2/n), \tag{5.1}$$

$$\Psi_n(E) \approx 1 - \operatorname{erf}(E/n^{1/2}), \tag{5.2}$$

$$\operatorname{erf}(x) = 2\pi^{-1/2}\int_0^x \exp(-x^2)\, dx.$$

Our method of estimation is roughly equivalent to interpolating between values given by these formulas and those given by the formulas of Sec. 2 and 4. We shall first work with the random walk distribution.

Equation (5.1) suggests that we introduce a function y of r and n defined by

$$\Phi_n(r) = e^{-ny} \tag{5.3}$$

or

$$y = -\frac{1}{n}\log \Phi_n(r). \tag{5.4}$$

Comparison of (5.1) and (5.3) shows that, for small values of r/n, $y \approx (r/n)^2$ and hence in this case y depends "much more" on r/n than on n. The same is true when r/n is nearly unity since Eq. (4.2) for $\Phi_n(r)$ leads to (assuming n large so that the series may be approximated by $\exp[(n-r)/4]$)

$$y \approx -\frac{n-1}{2n}\log\left(1 - \frac{r}{n}\right) - \frac{1}{4}\left(1 - \frac{r}{n}\right) + \frac{1}{2}\log\frac{\pi}{e}, \tag{5.5}$$

and this again tends to become a function of r/n only as $n \to \infty$.

In order to test the dependence of y on r/n, Bennett's values of $\Phi_n(r)$ for $n = 6$ and 10 were used to compute y from (5.4). The results are plotted as the "exact" values of y

[4]Margaret Slack, *The probability distributions of sinusoidal oscillations combined in random phase*, J.I.E.E., Pt. III, **93**, 76-86 (1946).

(indicated by the small triangles and circles) shown in Fig. 1. It is seen that the two sets of values tend to follow a common curve.

The dashed curves in Fig. 1 were computed from the approximation (5.5) for $n = 6, 10$ and ∞. Although (5.5) is valid only for $r/n \approx 1$, it yields values of y which are fairly

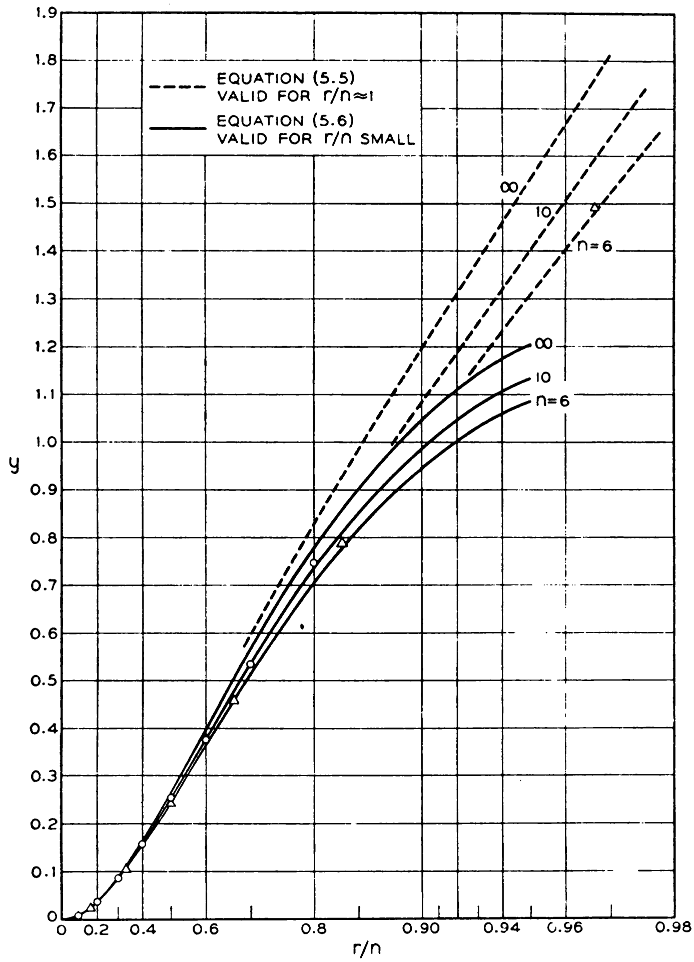

FIG. 1. The triangles ($n = 6$) and the circles ($n = 10$) show exact values of $y = (-1/n) \log_e \Phi_n(r)$. $\Phi_n(r)$ is the probability that the resultant of n random two-dimensional unit vectors is longer than r. The curves show the approximations (5.5) and (5.6) for $n = 6$, 10, and ∞.

close to the exact values even for values of r/n as small as 0.4. The solid curves were computed from

$$y \approx \left(\frac{r}{n}\right)^2 + \frac{1}{4}\left(\frac{r}{n}\right)^4 + \frac{5}{36}\left(\frac{r}{n}\right)^6 - \frac{1}{n}\left[\frac{1}{2}\left(\frac{r}{n}\right)^2 + \frac{1}{3}\left(\frac{r}{n}\right)^4\right] + \frac{1}{12n^2}\left(\frac{r}{n}\right)^2 \tag{5.6}$$

which holds only for small values of r/n and large values of n. Expression (5.6) is obtained from

$$\Phi_n(r) \approx e^{-x}\left[1 - \frac{f_1}{2n} - \frac{2f_2}{3n^2} + \frac{(6n - 11)f_3}{8n^3} + \cdots\right], \tag{5.7}$$

where $x = r^2/n$ and

$$f_m = -x\,{}_1F_1(-m; 2; x) = -x + \frac{m}{1!2!}x^2 - \frac{m(m-1)}{2!3!}x^3 + \cdots,$$

${}_1F_1(\ )$ being a confluent hypergeometric function. Expression (5.7) may be obtained by integrating a result given by Pearson[2] and, later, by Rayleigh[3]. Pearson's formula suggests that the next two terms in (5.7) are $+ (50n - 57)f_4/15n^4 - (270n^2 - 2125n + 1892)f_5/144n^5$.

Thus, a rough idea of how $\Phi_n(r)$ behaves for all values of r/n and a particular value of n may be obtained by (i) computing approximations to y from (5.5) and (5.6),* (ii) plotting them on semi-log paper as shown in Fig. 1, (iii) joining the two portions by a smooth curve, and (iv) using these approximate values of y to compute $\Phi_n(r)$ from (5.3).

When we turn to $\Psi_n(E)$, the procedure used in dealing with $\Phi_n(r)$ suggests that a new function y' of E and n be defined by

$$\Psi_n(E) = 1 - \operatorname{erf}\,[(ny')^{1/2}]. \tag{5.8}$$

However, it is found that the expression for y' when $E/n \approx 1$ does not have the simplicity of its analogue (5.5). Instead of following this line of thought further, we note that the analogy between y and y' suggests that y' should not differ greatly from the function obtained by replacing r by E in y. That the difference is small may be verified by comparing Bennett's exact values of $\Psi_{10}(E)$ with approximate ones obtained from (5.8). In using (5.8), the values of y' are taken to be those of y as computed from (5.4) and Bennett's exact values of $\Phi_{10}(r)$. The exact and approximate values are shown in the following table for three values of E.

E	exact $\Psi_{10}(E)$	approx. $\Psi_{10}(E)$
1	.6604	.6624
5	.02337	.02440
8	.00009	.00011

Thus, once we have obtained approximate values of y from curves of the type shown in Fig. 1, approximate values of $\Phi_n(r)$ and $\Psi_n(E)$ may be obtained readily from (5.3) and (5.8) (with y in place of y').

*Instead of (5.6) one may use (5.7) and (5.4). This gives more accurate values of y at the cost of more computation.