FACTS AND FIGURES PERTAINING TO THE BIVARIATE RAYLEIGH DISTRIBUTION

by

J.A. Battjes
Dept. of Civil Engineering
Delft University of Technology
The Netherlands

1969

SUMMARY

There is some evidence that the heights and the periods squared of wind waves each are Rayleigh-distributed. If this is the case, it may be surmised that their joint distribution is also of the Rayleigh type. The bivariate Rayleigh distribution, which is known in the field of statistical communications theory, may thus be applicable to wind wave problems. However, in the field of coastal and ocean engineering, this distribution does not seem to be known. In this note attention is drawn to its existence, and its main properties are listed.

LIST OF CONTENTS

	page
1. INTRODUCTION	1
2. THE BIVARIATE RAYLEIGH DISTRIBUTION	3
2.1. Background	3
2.2. The probability density function	3
2.3. The cumulative probability function	4
2.4. The marginal probability density functions	4
2.5. The moments	5
2.6. The correlation coefficient	6
2.7. Case of zero correlation	8
2.8. Case of 100% correlation	8
2.9. Regression lines	9
LIST OF REFERENCES	11

1. INTRODUCTION

In the fields of harbour-, coastal- and ocean-engineering, knowledge of the statistical distributions of heights and periods of wind waves is often of great importance. Many efforts, both theoretical and experimental, have been made to determine such distributions.

Empirical wave height distributions were first given by Putz (1952). At about the same time, Longuet-Higgins (1952) pointed out that theoretically the Rayleigh distribution should be applicable to the heights of wind waves with a narrow spectrum, provided the waves be not too steep. It has since been established empirically that the Rayleigh distribution applies to wind wave heights with a fair degree of accuracy, even if the waves are steep and do not have a narrow spectrum. In the latter case, the wave height may be defined as the difference of the maximum (crest) height and minimum (trough) height between two successive zero up- or downcrossings.

The analytical determination of the distribution of wave periods (intervals between successive zero up- or down crossings) is difficult. So far only approximate solutions have been obtained (Longuet-Higgins, 1962). Empirical distribution functions have been first proposed by Putz (1952). Bretschneider (1959) noted that the squares of the periods of wind waves had a Rayleigh distribution with about the same degree of accuracy as the wave heights. This conclusion seems to have also been reached by Russian authors (Bretschneider, 1965).

The distributions mentioned so far pertain to the heights (H) and periods (T) separately. The problem of the joint distribution has not been solved by theoretical means. Bretschneider (1959) approached this problem empirically. Having observed that the marginal distributions of H and T^2 are both of the Rayleigh type, he assumed that this would also be the case for the joint distribution.[1] However, he was unable to elaborate this assumption because the bivariate Rayleigh distribution was unknown to him. To quote:

> "The joint distribution of wave heights and lengths (or wave heights and periods) in general is difficult to describe completely for all conditions of correlation. ... The fact that both marginal distributions are of the same type is of some help. The bivariate asymptotic problem of joint distribution for the Rayleigh type (or a modified Rayleigh type) distribution has yet to be solved".

It was unknown to Bretschneider that this problem, which also arises in the field of communications theory, had already been solved by Uhlenbeck (1943) and Rice (1944, 1945).

The fact that the theoretical bivariate Rayleigh distribution is known seems to have remained unnoticed in the fields of coastal- and ocean-engineering. Apparently, there is a communications gap between researchers dealing with wind waves and those working in the field of statistical communications theory. It is the purpose of this note to contribute to bridging that gap. One way of doing this would be to merely mention a few references. However, it is deemed more useful to reproduce the main results, because not all of these are available in any single publication known to the author. Also, some corrections and additions will be made.[2)]

It is emphasized that this note in what follows deals solely with the theoretical bivariate Rayleigh distribution. At this time no theoretical or empirical evidence is offered to prove or to disprove its possible applicability to wind waves.

---

1) This assumption cannot mathematically be justified. In fact, Fréchet (1951, 1956) has shown that corresponding to each given pair of marginal distribution functions, an infinite number of bivariate distribution functions can exist.

2) Some minor corrections appear in paragraph 2.5. The additions consist of the material presented in the paragraphs 2.6. through 2.9.

2. THE BIVARIATE RAYLEIGH DISTRIBUTION

2.1. Background

The envelope of a narrow-band random noise signal has a Rayleigh distribution. Similarly, the values $X(t)$ and $Y = X(t+\tau)$ of such an envelope at two different times separated by a certain time lag τ have a joint Rayleigh distribution. The derivation of this distribution seems to have been given for the first time by Uhlenbeck (1943) and again, independently, by Rice (1944-1945). The method employed is essentially similar to the one used in deriving the one-dimensional Rayleigh distribution. The derivation will not be reproduced here.

2.2. The probability density function (p.d.f.)

The joint p.d.f. $f(x,y)$ of two stochastic variables X and Y is defined by

$$f(x,y)\,dx\,dy = Prob\left[x < X \leq x+dx \text{ and } y < Y \leq y+dy\right] \qquad (1)$$

X and Y are said to have a joint Rayleigh distribution if their p.d.f. is given by

$$f(x,y) = \frac{\pi^2}{4}\frac{xy}{(\bar{X}\bar{Y})^2(1-k^2)}\, e^{-\frac{\pi}{4}\frac{(x/\bar{X})^2+(y/\bar{Y})^2}{1-k^2}}\, I_0\left(\frac{\pi k}{2(1-k^2)}\frac{xy}{\bar{X}\bar{Y}}\right) \qquad (2)$$

$$\text{for } x \geq 0 \text{ and } y \geq 0 \;;\; f(x,y) = 0 \text{ otherwise.}$$

in which a bar denotes average values and I_0 is the modified Bessel function of the first kind of order zero (Uhlenbeck, 1943; Rice, 1944-'45; Nakagami, 1964). The parameter k is defined by

$$k = \frac{\int_0^\infty S(\omega)\cos(\omega_c-\omega)\tau\,d\omega}{\int_0^\infty S(\omega)\,d\omega} \qquad (3)$$

in which $S(\omega)$ is the power spectral density of the narrow-band random noise, with center frequency ω_c. $S(\omega)$ is non-negative; therefore, $|k| \leq 1$.

The case $k=1$ occurs when $\tau=0$, i.e. when $X=Y$, or when there is 100% correlation between X and Y. For very large τ, $k \to 0$, while X and Y approach stochastic independence and, therefore, become uncorrelated. Although in these special cases k equals the coefficient of correlation between X and Y, this is not generally the case. The general relationship will be dealt with in section 2.6.

In the following, the definition of k as given by Eq.(3) will not be needed; k will merely function as a parameter which fulfills the inequality $|k| \leq 1$.

For variables which have been normalized so as to make their average value equal to one, the joint Rayleigh probability density becomes

$$f(x,y) = \frac{\pi^2}{4} \frac{xy}{1-k^2} e^{-\frac{\pi}{4}\frac{x^2+y^2}{1-k^2}} I_0\left(\frac{\pi}{2}\frac{k}{1-k^2}xy\right) \quad \text{for } \begin{cases} x \geq 0 \\ y \geq 0 \end{cases} \tag{4}$$

$$= 0 \qquad \text{otherwise}$$

This is the form which will be used subsequently.

2.3. The cumulative probability function

The author has not succeeded in expressing the cumulative probability function, or the distribution function, in a finite number of known functions. It appears to be necessary to compute it numerically if and when the need arises.

2.4. The marginal probability density functions

The marginal p.d.f. of X is defined by

$$f(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \tag{5}$$

Substitution of Eq. (4) gives rise to an integral which may be evaluated by using the following series expression for I_0 (Abramowitz and Stegun, 1965):

$$I_0(t) = \sum_{j=0}^{\infty} \frac{t^{2j}}{2^{2j}(j!)^2} \tag{6}$$

and by integrating termwise. The result is, as expected, the one-dimensional Rayleigh p.d.f. :

$$f(x) = \frac{\pi}{2} x \, e^{-\frac{\pi}{4}x^2} \qquad \text{for } x \geq 0$$
$$= 0 \qquad \text{for } x < 0 \tag{7}$$

Similarly,

$$f(y) = \frac{\pi}{2} y \, e^{-\frac{\pi}{4}y^2} \qquad \text{for } y \geq 0$$
$$= 0 \qquad \text{for } y < 0 \tag{8}$$

2.5. The moments

The moment matrix $\mu_{m,n}$ is defined by

$$\mu_{m,n} = \overline{X^m Y^n} = \int_0^\infty \int_0^\infty x^m y^n f(x,y) \, dx \, dy \tag{9}$$

Substituting Eqs. (4) and (6) and integrating termwise, it can be shown that

$$\mu_{m,n} = \left(\frac{4}{\pi}\right)^{\frac{m+n}{2}} (1-k^2)^{\frac{m+n}{2}+1} \sum_{j=0}^{\infty} \frac{\Gamma\left(j+1+\frac{m}{2}\right)\Gamma\left(j+1+\frac{n}{2}\right) k^{2j}}{(j!)^2} \tag{10}$$

Apart from a constant factor, the infinite series in the right hand member is the same as the series expression for a hypergeometric function (Magnus and Oberhettinger, 1943), so that Eq.(10) may be written as

$$\mu_{m,n} = \left(\frac{4}{\pi}\right)^{\frac{m+n}{2}} (1-k^2)^{\frac{m+n}{2}+1} \Gamma\left(\frac{m}{2}+1\right)\Gamma\left(\frac{n}{2}+1\right) {}_2F_1\left(\frac{m}{2}+1, \frac{n}{2}+1; 1; k^2\right) \tag{11}$$

(Middleton, 1960), or, after performing a known transformation of the hypergeometric function:

$$\mu_{m,n} = \left(\frac{4}{\pi}\right)^{\frac{m+n}{2}} \Gamma\left(\frac{m}{2}+1\right)\Gamma\left(\frac{n}{2}+1\right) {}_2F_1\left(-\frac{m}{2}, -\frac{n}{2}; 1; k^2\right) \tag{12}$$

(The equations 9.22 in Middleton (1960), corresponding to Eqs. (10) and (11) above, contain misprints).

By definition resp. normalization, the zeroth- and first moments are equal to one. The second moments are given by

$$\overline{X^2} = \overline{Y^2} = \frac{4}{\pi} \tag{13}$$

and

$$\overline{XY} = {}_2F_1\left(-\tfrac{1}{2}, -\tfrac{1}{2}; 1; k^2\right) \tag{14}$$

The last expression is equivalent to

$$\overline{XY} = \frac{4}{\pi} E(k) - \frac{2}{\pi}(1-k^2)\, K(k) \tag{15}$$

in which $K(k)$ and $E(k)$ are the complete elliptic integrals of the first and second kind of modulus k. Eq. (15) is due to Uhlenbeck (1943). It may be proved by using series expressions of the various transcendental functions in the right hand members of Eqs. (14) and (15), or of Eq. (10) with $m = n = 1$.

2.6. The correlation coefficient

The coefficient of correlation between X and Y is defined by

$$\rho = \frac{\overline{XY} - \bar{X}\bar{Y}}{(\overline{X^2} - \bar{X}^2)^{1/2}(\overline{Y^2} - \bar{Y}^2)^{1/2}} \tag{16}$$

Substitution of $\bar{X} = \bar{Y} = 1$ and of Eqs. (13) and (15) gives

$$\rho = \frac{E(k) - \frac{1}{2}(1-k^2)\, K(k) - \frac{\pi}{4}}{1 - \frac{\pi}{4}} \tag{17}$$

A graph of ρ *vs.* k is shown in figure 1 by the full line.

$E(k)$ and $K(k)$ are even functions of k. It follows that ρ is also an even function of k. Two special cases occur when $k = 0$ and $k = 1$:

$$E(0) = K(0) = \frac{\pi}{2}, \text{ so}$$
$$\rho = 0 \text{ if } k = 0 \tag{18}$$

$$E(1) = 1\,; \quad K(k) \to \infty \text{ as } k \to 1, \text{ but } (1-k^2)K(k) \to 0, \text{ so}$$
$$\rho = 1 \text{ if } k = 1 \tag{19}$$

This confirms the remarks made in section 2.2 following the definition of k.

Utilizing series representations for E and K in Eq. (17), or for the gamma-functions in Eq. (10) with $m = n = 1$, the following series representation for ρ may be obtained:

$$\rho = a\left[k^2 + \sum_{j=2}^{\infty}\left\{\frac{1\cdot 3\cdot 5\cdot \ldots\ldots (2j-3)}{4\cdot 6\cdot 8\cdot \ldots\ldots (2j)}\right\}^2 k^{2j}\right] \tag{20}$$

or

$$\rho = a\left(k^2 + \frac{k^4}{16} + \frac{k^6}{64} + \text{------}\right) \tag{21}$$

in which

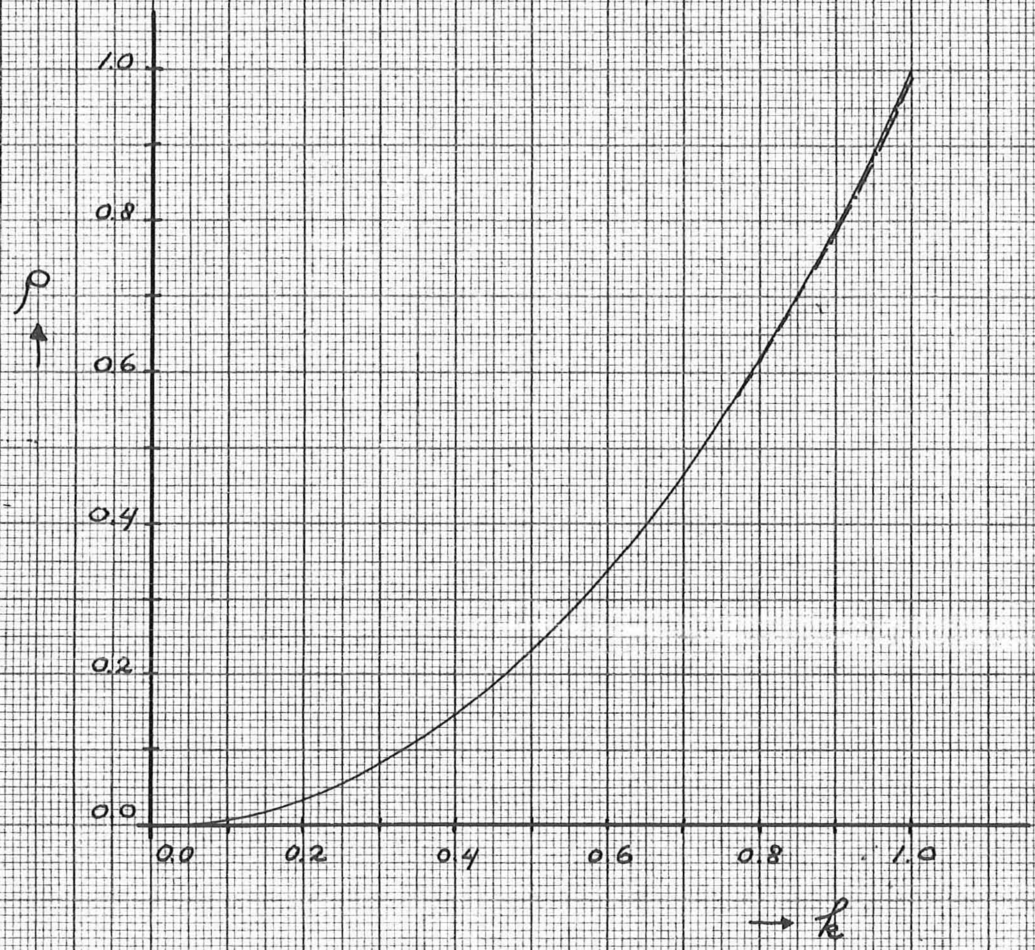
$$a = \frac{\pi}{16 - 4\pi} \tag{22}$$

It should be noted that ρ is a non-negative function of k. Thus, two stochastic variables with a joint Rayleigh distribution cannot be negatively correlated. Bretschneider (1959) already arrived at a similar, though more restricted, result. He proved that two stochastic variables could not be jointly Rayleigh distributed if $\rho = -1$.

In order to obtain an explicit expression for k as a function of ρ, the series given by Eq. (21) has been inverted, with the result

$$k^2 = \frac{\rho}{a} - \frac{1}{16}\left(\frac{\rho}{a}\right)^2 - \frac{1}{128}\left(\frac{\rho}{a}\right)^3 - \ldots\ldots \tag{23}$$

The two series given by Eqs. (21) and (23) are rapidly converging for $|k| \leq 1$, $|\rho| \leq 1$. Graphs of ρ vs. k, based on the first three terms of the series, are given in figure 1. Only for values of ρ and k close to 1 is there a visible deviation from the exact relationship. The truncation error after 3 terms is less than 0.1% for all $|k| < 0.7$ and less than 1% for all $|k| < 0.95$.



$$\rho = \frac{E(k) - \frac{1}{2}(1-k^2)K(k) - \frac{\pi}{4}}{1 - \frac{\pi}{4}}$$

$$k^2 = \frac{\rho}{a} - \frac{1}{16}\frac{\rho^2}{a^2} - \frac{1}{128}\frac{\rho^3}{a^3}$$

$$\rho = a\left(k^2 + \frac{k^4}{16} + \frac{k^6}{64}\right)$$

$$a = \frac{\pi}{16 - 4\pi}$$

Figure 1

2.7. Case of zero correlation

The stochastic variables X and Y are said to be uncorrelated if the coefficient of correlation ρ, given by Eq. (16), is zero. If $\rho = 0$, $k = 0$. Substitution of this value in Eq. (4) yields

$$\left.\begin{aligned} f(x,y) &= \frac{\pi^2}{4} x y \, e^{-\frac{\pi}{4}(x^2+y^2)} \quad \text{for} \begin{cases} x \geq 0 \\ y \geq 0 \end{cases} \\ &= 0 \quad \text{otherwise} \end{aligned}\right\} \text{for } k=0 \tag{24}$$

which is equal to the product of the marginal p.d. functions as given by Eqs. (7) and (8). Thus, two stochastic variables with a joint Rayleigh distribution are stochastically independent if they are uncorrelated.

2.8. Case of 100% correlation

If 100% correlation occurs then the two variables X and Y are linearly dependent. Moreover, X and Y are identically distributed. Therefore, $X = Y$ if $\rho = 1$. The two-dimensional p.d.f. actually degenerates to a one-dimensional one. The joint probability density must then be zero for all $x \neq y$ and be infinite for $x = y$. This may be shown formally by investigating the behavior of $f(x,y)$ as $k \to 1$. To this end, the following asymptotic expression for $I_0(t)$ (Abramowitz and Stegun, 1965) is substituted in Eq. (4):

$$I_0(t) = \frac{e^t}{\sqrt{2\pi t}} \left\{ 1 + \mathcal{O}(t^{-1}) \right\} \tag{25}$$

After some algebraic manipulations, the result can be written as follows:

$$f(x,y) \to \frac{\pi\sqrt{xy}}{2\sqrt{k}} \, e^{-\frac{\pi}{2}\frac{xy}{1+k}} \left[\frac{e^{-\frac{(x-y)^2}{2\sigma}}}{\sqrt{2\pi\sigma}} \right] \quad \text{as } k \to 1 \tag{26}$$

in which

$$\sigma = \frac{2}{\pi}(1 - k^2) \tag{27}$$

The expression in the brackets is formally equal to a Gaussian p.d.f. with independent variable $(x-y)$, zero mean and mean square deviation σ. As k approaches 1, σ approaches 0. The Gaussian p.d.f. then approaches Dirac's unit impulse "function" $\delta(x-y)$ (Lighthill, 1966). In fact, this may be considered as a definition of δ . Thus,

$$f(x,y) = \frac{\pi}{2}\sqrt{xy}\, e^{-\frac{\pi}{4}xy}\, \delta(x-y) \qquad \text{for } k=1 \tag{28}$$

which may also be written as

$$f(x,y) = f(x)\, \delta(x-y) \qquad \text{for } k=1 \tag{29}$$

or as

$$f(x,y) = f(y)\, \delta(x-y) \qquad \text{for } k=1 \tag{30}$$

in which $f(x)$ and $f(y)$ are the marginal p.d. functions given by Eqs. (7) and (8).

2.9. Regression lines

The regression of X on Y is given by

$$\bar{X}_y = \frac{\int_0^\infty x\, f(x,y)\, dx}{\int_0^\infty f(x,y)\, dx} \tag{31}$$

The expression which results after substitution of Eq. (4) has been evaluated using equations 11.4.28 and 13.1.27 from Abramowitz and Stegun (1965), with the result

$$\bar{X}_y = \sqrt{1-k^2}\, M(-\tfrac{1}{2}, 1, z y^2) \tag{32}$$

in which M is the confluent hypergeometric function, and z is a coefficient given by

$$z = -\frac{\pi}{4}\, \frac{k^2}{1-k^2} \tag{33}$$

It appears that in general the regression of X on Y is non-linear. Exception should be made for the two trivial cases of stochastic independence and linear dependence of X and Y.

In the case of stochastic independence, $k = 0$, substitution of which gives

$$\overline{X}_y = 1 \qquad \text{if} \quad k = 0, \tag{34}$$

as expected.

In the case of linear dependence of X and Y, $k = 1$. As $k \to 1$, $z \to -\infty$. Using asymptotic expansions of the confluent hypergeometric function, it may be shown that

$$M\left(-\tfrac{1}{2}, 1, zy^2\right) \to \frac{ky}{\sqrt{1-k^2}} \qquad \text{if} \quad k \to 1 \tag{35}$$

Therefore,

$$\overline{X}_y = y \qquad \text{if} \quad k = 1, \tag{36}$$

as expected.

The regression of Y on X is obtained by interchanging x and y in the preceding expressions, because the bivariate Rayleigh distribution is symmetric with respect to x and y.

References

Abramowitz, M. and Stegun, I.A.: "Handbook of Mathematical Functions", Dover Publications, Inc., New York, 1046 pages, 1965.

Bretschneider, C.L.: "Wave Variability and Wave Spectra for Wind-generated Gravity Waves", Techn. Mem. 118, B.E.B., Dept. of the Army, 192 pages, 1959.

Bretschneider, C.L.: "Generation of Waves by Wind; State of the Art", NESCO Report SN-134-6, 96 pages, 1965.

Lighthill, M.J.: "Fourier Analysis and Generalised Functions", Cambridge Univ. Press, 79 pages, 1966.

Longuet-Higgins, M.S. (1952): "On the Statistical Distribution of the Heights of Sea Waves", Journal of Marine Research, vol. 11, No. 3, p. 245 - 266.

Longuet-Higgins, M.S.: "The distribution of intervals between zeros of a stationary random function", Phil. Trans. Roy. Soc., vol. 254, No. 1047, p. 557-599, 1962.

Magnus, W. and Oberhettinger, F.: "Formeln und Sätze für die speziellen Funktionen der mathematischen Physik", Berlin, 172 p., 1943.

Middleton, D.: "Statistical Communication Theory", Mc. Graw-Hill Book Company, Inc., New York, 1140 pages, 1960

Nakagami, M.: "On the Intensity Distribution..and its Application to Signal Statistics", Radio Science, vol. 68D, No. 9, p.995-1003, sept. 1964.

Putz, R.A.: "Statistical Distribution for Ocean Waves," Trans. A.G.U., Vol. 33, No. 5, p. 685-692, 1952.

Rice, S.O.: "Mathematical Analysis of Random Noise", The Bell System Techn. Jour., Vol. 23, p. 282-332, 1944; Vol. 24, p. 46-156, 1945.

Uhlenbeck, G.E.: "Theory of Random Processes", MIT Radiation Lab. Rept. 454, 1943.

Fréchet, M.: "Sur les tableaux de corrélation dont les marges sont données", Annales de l'Université de Lyon, Section A, p. 53-77, 1951.

Fréchet, M.: "Sur les tableaux de corrélation dont les marges sont données", Comptes Rendus de l'Académie des Sciences, Tome 242, p. 2426-2428, 1956.