Report 2013

AD 633745

DAVID TAYLOR MODEL BASIN
WASHINGTON, D.C.

DEPARTMENT OF THE NAVY

CLEARINGHOUSE
FOR FEDERAL SCIENTIFIC AND
TECHNICAL INFORMATION

Hardcopy	Microfiche		
$2.60	$0.50	29 pp	as

ARCHIVE COPY

HYDROMECHANICS

AERODYNAMICS

STRUCTURAL MECHANICS

APPLIED MATHEMATICS

ACOUSTICS AND VIBRATION

THE GRAM-CHARLIER APPROXIMATION OF THE NORMAL LAW AND THE STATISTICAL DESCRIPTION OF A HOMOGENEOUS TURBULENT FLOW NEAR STATISTICAL EQUILIBRIUM

Code 1

by

Joseph Kampé de Fériet
University of Lille, France

Distribution of this document is unlimited.

APPLIED MATHEMATICS LABORATORY
RESEARCH AND DEVELOPMENT REPORT

DDC
RECEIVED
APR 13 1966
C

March 1966 Report 2013

THE GRAM-CHARLIER APPROXIMATION OF THE NORMAL LAW AND THE STATISTICAL DESCRIPTION OF A HOMOGENEOUS TURBULENT FLOW NEAR STATISTICAL EQUILIBRIUM

by

Joseph Kampé de Fériet
University of Lille, France

Distribution of this document is unlimited.

March 1966 Report 2013

FOREWORD

This report was prepared by Professor J. Kampé de Fériet during his visit at the Applied Mathematics Laboratory, David Taylor Model Basin in 1963. It was completed by the author while he was a visiting professor at The Catholic University of America in 1964-1965.

We wish to express to Professor Kampé de Fériet our gratitude for his contributions to the studies of stochastic processes which are pursued in this laboratory.

Francois N. Frenkiel

Applied Mathematics Laboratory

TABLE OF CONTENTS

	Page
ABSTRACT	1
INTRODUCTION	1
ONE-DIMENSIONAL GRAM-CHARLIER SERIES	6
GRAM-CHARLIER APPROXIMATION OF THE PROBABILITY DENSITY FOR ONE RANDOM VARIABLE	9
JOINT PROBABILITY DENSITY OF TWO RANDOM VARIABLES	14
HERMITE POLYNOMIALS OF TWO VARIABLES	15
GRAM-CHARLIER APPROXIMATION OF A TWO-DIMENSIONAL PROBABILITY DENSITY	19
REFERENCES	23

BLANK PAGE

ABSTRACT

The probability density distribution is discussed for two correlated random variables based on an approximation to a normal (Gaussian) law using Hermite polynomials of two variables.

INTRODUCTION

In many physical theories the following problem must be solved:

Given a large number of experimental values of p quantities $x_1, \ldots x_p$, which are considered as random variables having a <u>known</u> joint probability density

$$f(x_1, \ldots x_p, \theta_1, \ldots \theta_q)$$

how can an estimate $\hat{\theta}_1, \ldots \hat{\theta}_q$ be made of the q unknown parameters $\theta_1, \ldots \theta_q$ such that the surface

$$z = f(x_1, \ldots x_q, \hat{\theta}_1, \ldots \hat{\theta}_q)$$

fits the "cloud" of the experimental points as well as possible?

This problem has been solved by several well-known methods during the last few decades; each of these solutions is based on a particular definition of the measure of the discrepancy between the surface and the experimental cloud. There has been much discussion about the best choice for the measure of the discrepancy among the different schools. Nevertheless, it can be said, that in many physical problems estimates can now be made following one or the other of these methods which lead to results of great significance for theoretical physics.

But the situation in fluid dynamics is very different if we want to study a turbulent flow statistically. For instance, taking the most simple case of a homogeneous turbulence, let us suppose that a great number of measurements have been made of velocity components (U_1, U_2, U_3) and (V_1, V_2, V_3) at two different points with a constant interval of time between the two measurements. No theory has yet been established giving the mathematical expression for the joint probability density

$$f(U_1, U_2, U_3, V_1, V_2, V_3, \theta_1, \ldots \theta_q)$$

which these six random variables must follow, and thus we are confronted by a much more difficult situation. Before making any estimate of the unknown parameters $\theta_1, \ldots \theta_q$, we must first guess which particular function must be used for the joint probability density.

For this problem, statistical mechanics is able only to give a few hints, which are, however, of great value. In the classical case of a mechanical system with a finite number of degrees of freedom, the probability law, according to J. W. Gibbs, is given by the canonical distribution in the phase space. It has been shown that for some continuous media[1] having a countable number of degrees of freedom (for instance a vibrating string) this result is still true. It is still the normal law which maximizes the entropy (as defined in information theory), and thus generalizes the canonical distribution to the function space describing the phases of the continuous medium.

It seems natural to take as phase space for a homogeneous turbulent flow, the function space of all vector functions $U_1 (X_1, X_2, X_3)$, $U_2 (X_1, X_2, X_3)$, $U_3 (X_1, X_2, X_3)$ which are square integrable in any bounded domain (this space is called Λ in Reference 2; the reason leading to this choice is discussed in Reference 3). The main feature of Λ is that the energy of any finite portion of the fluid is always finite. A point in Λ is defined by a countable number of coordinates (Λ can be considered as a direct sum of a countable number of Hilbert spaces). Is it then possible to extend the results proved for the vibrating string to the turbulent flow of an inviscid fluid? The answer to this question is negative. The canonical distribution of Gibbs and its extension to some continuous media is essentially based on the hypothesis that the systems are <u>conservative</u>; the canonical distribution and its extensions correspond to a <u>statistical equilibrium</u>.

---

[1]References are listed on page 23.

If we assume, as we always do in research on turbulence, that the fluid is viscous, then there is a constant dissipation of energy. If through the maximum entropy principle, we associate normal law and statistical equilibrium, we cannot expect that the joint probability density of the random variables would be the normal probability density.

The logical consequence of the preceding remarks is that not being able to take advantage of the results linked with the statistical equilibrium, we must start as near as possible to this case, i.e., as a first step we suppose that our turbulent flow is near the statistical equilibrium. In mathematical terms we must make the hypothesis that the joint probability density is given by the first terms of a series

$$f_o\,(x_1, \ldots x_p, \theta_1, \ldots \theta_q) + f_1\,(x_1, \ldots x_p, \theta_1, \ldots \theta_q) + \ldots$$

where f_o is the normal probability density and the subsequent terms f_1, f_2 ... mark the difference between the actual statistical state of the system and the statistical equilibrium.

Obviously, the joint probability density can be put in the form

$$f_o\,(x_1, \ldots x_p, \theta_1, \ldots \theta_q)\; P(x_1, \ldots x_p, \theta_1, \ldots \theta_q)$$

Where

$$P = 1 + \frac{f_1}{f_o} + \frac{f_2}{f_o} + \ldots \quad (f_o > 0)$$

As in most of the approximations made in physics, it seems logical to begin by assuming that P is a polynomial in $x_1, \ldots x_p$. Because of the exponential form of f_o, the simplest way to express this polynomial seems to be to use the Gram-Charlier series, based on Hermite polynomials.

In the following pages we will summarize the essential features of the theory, giving special attention to the two-dimensional case, because most of the literature has been confined to the one-dimensional case.

For a general exposition of the theory of Hermite polynomials in one and several variables, see Reference 4; for the one-dimensional Gram-Charlier series, we refer to Cramer, pp 221-231,[5] and Kendall, pp 145-150,[6]

where further bibliographical references are given. We are indebted to Dr. Lieblein, who has pointed out to us papers, References 7 and 8, which are among the very few devoted to the two-dimensional case.

Note that the Gram-Charlier approximation of the bivariate normal law, suggested here, differs on an essential point from that used in the literature known to us.

As a rule, to represent a function of two variables f(x, y), one uses the polynomials $H_m(x)\ H_n(y)$. From a purely theoretical point of view this is perfectly correct because it is well known that if a sequence

$$\phi_o\ (x),\ \phi_1\ (x), \ldots \phi_n(x), \ldots\ldots$$

defines an orthonormal basis in the Hilbert space of functions of one variable f(x) (which are square integrable), the sequence of the product

$$\phi_o\ (x)\ \phi_o\ (y), \ldots\ \ \phi_m(x)\ \phi_n(y), \ldots$$

constitutes an orthonormal basis in the Hilbert space of functions of two variables f(x,y).

This assumption leads to the following representation of the bivariate probability density

$$P(x,y) = \frac{1}{2\pi\ \sigma\ \tau}\ e^{-\left[\frac{(x-m)^2}{2\sigma^2} + \frac{(y-n)^2}{2\tau^2}\right]} \left[\sum_{o,\ o}^{+\infty,\ +\infty} A_{j,k}\ H_j\left(\frac{x-m}{\sigma}\right)\ H_k\left(\frac{y-n}{\tau}\right)\right]$$

Correct from the point of view of the representation of a function of two variables by a series, this development has a major disadvantage. If we assume

$$A_{o,\ o} = 1$$

and all

$$A_{j,k} = 0 \quad (j \geq 1,\ k \geq 1)$$

the function

$$P_o\ (x,y) = \frac{1}{2\ \pi\ \sigma\ \tau}\ e^{-\frac{(x-m)^2}{2\ \sigma^2} - \frac{(y-n)^2}{2\ \tau^2}}$$

<u>does not represent the bivariate normal law in the general case</u>; it only fits if the two normal random variables x and y are independent.

Consequently, a limited development

$$P(x,y) = \frac{1}{2\pi\ \sigma\ \tau}\ e^{-\frac{(x-m)^2}{2\sigma^2} - \frac{(y-n)^2}{2\ \tau^2}} \left[\sum_{j+k=0}^{j+k=2n} A_{j,k}\ H_j\ \frac{(x-m)}{\sigma}\ H_k\ \frac{(y-n)}{\tau}\right]$$

is a possible approximation of the real probability density, but it is not, strictly speaking, its Gram-Charlier approximation.

To generalize the one-dimensional Gram-Charlier approximation to bivariate probability density, we must start from the exponential

$$\frac{1}{2\pi\ \sigma\ \tau} \exp\left[-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma^2} - 2r\frac{xy}{\sigma\tau} + \frac{y^2}{\tau^2}\right)\right]$$

and use polynomials in (x,y) deduced from this general normal law exactly as the H_n in one variable is deduced from the one-dimensional law. These polynomials will reduce to products $H_j\ \frac{(x-m)}{\tau}\ H_k\ \frac{(y-n)}{\tau}$ if, and only if, $r = 0$.

Fortunately, these polynomials are known. They were discovered by Hermite in 1864 (Reference 4, p 363). It is precisely in the case of two variables that he has made his most original contribution. It is now known that in the one-variable case, Tschebyscheff was his predecessor by 4 years. But in the two-variable case Hermite introduced two adjunct sequences

$$G_{m,n}\ (x,y) \quad \text{and} \quad H_{m,n}\ (x,y)$$

of polynomials, connected with two positive definite quadratic forms

$$\phi\ (x,y) = ax^2 + 2\ bxy + cy^2$$

$$a \geq o,\ b \geq 0 \quad \Delta = ac - b^2 \leq o$$

and its adjunct

$$\psi\ (\xi,\ \eta) = \frac{c}{\Delta}\xi^2 - \frac{2b}{\Delta}\ \xi\eta + \frac{a}{\Delta}\eta^2$$

These two sequences of polynomials are biorthogonal.

The remarkable properties of the polynomials G and H very easily lead to the true extension of the Gram-Charlier approximation for a bivariate density.

Not only are the formulas for the computation of the coefficients much simpler, but the meaning of the development is also changed. If we suppose all the coefficients $A_{j,k} = 0$ except $A_{o,o} = 1$, then the unique term left coincides with the general bivariate law with an arbitrary correlation coefficient r.

The main contribution here is thus to substitute for the usual series a new Gram-Charlier series based on the polynomials $G_{m,n}(x,y)$ and $H_{m,n}(x,y)$ introduced by Ch. Hermite.

The computation of the sufficient condition for the polynomial $P_4(x)$ to be positive was made by Mr. M. Pine.

ONE-DIMENSIONAL GRAM-CHARLIER SERIES

The Hermite polynomials in one variable x are defined by the generating function

$$e^{hx - \frac{h^2}{2}} = \sum_{o}^{+\infty} \frac{h^n}{n!} H_n\ (x) \qquad [1]$$

from where we deduce

$$H_n\ (x) = (-1)^n\ e^{\frac{x^2}{2}} \frac{d^n}{dx^n}\left(e^{\frac{-x^2}{2}}\right) \qquad [2]$$

If n is even,

$$H_n(-x) = H_n(x)$$

and if n is odd,

$$H_n(-x) = -H_n(x)$$

We have

$$H_o(x) = 1$$

$$H_1(x) = x$$

$$H_2(x) = x^2 - 1$$

$$H_3(x) = x^3 - 3x$$

$$H_4(x) = x^4 - 6x^2 + 3$$

$$H_5(x) = x^5 - 10x^3 + 15x$$

$$H_6(x) = x^6 - 15x^4 + 45x^2 - 15$$

$$H_7(x) = x^7 - 21x^5 + 105x^3 - 105x$$

$$H_8(x) = x^8 - 28x^6 + 210x^4 - 420x^2 + 105$$

$$H_9(x) = x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x$$

$$H_{10}(x) - x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945$$

The Hermite polynomials satisfy the following equations

$$\frac{dHn}{dx} = nH_{n-1} \qquad \frac{d^k H_n}{dx^k} = n(n-1)\ldots(n-k+1)\, H_{n-k} \qquad [3]$$

$$H_n(x) - x\, H_{n-1}(x) + (n-1)\, H_{n-2}(x) = 0 \qquad [3']$$

$$\frac{d^2 Hn}{d\,x^2} - x\,\frac{d\,Hn}{dx} + n\,H_n = 0 \qquad [3'']$$

The polynomial $H_n(x)$ has n real roots.

We have the pair of conjugate Fourier transforms

$$e^{\frac{-x^2}{2}} H_n(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ikx - \frac{k^2}{2}} (ik)^n \, dk \qquad [4]$$

(x real)

$$(ik)^n e^{-\frac{k^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx - \frac{x^2}{2}} H_n(x) \, dx \qquad [5]$$

(k real)

The Hermite polynomials have the fundamental orthogonality property

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} H_m(x) H_n(x) \, dx = 2\pi n! \, \delta_{m,n} \qquad [6]$$

which results, for the Fourier coefficients A_n, in

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left[\sum_{o}^{+\infty} A_n H_n(x) \right] \qquad [7]$$

and

$$A_n = \frac{1}{n!} \int_{-\infty}^{+\infty} f(x) H_n(x) \, dx \qquad [8]$$

The right-hand side of Equation [7] is known as the Gram-Charlier series of f(x); we will find in Reference 4, pp 351-355, sufficient conditions for the validity of the representation of a given function f(x) by this series.

GRAM-CHARLIER APPROXIMATION OF THE PROBABILITY DENSITY FOR ONE RANDOM VARIABLE

Let us consider a random variable X having moments of all orders. We put

$$\bar{X} = m \qquad [9]$$

$$\overline{(X - m)^2} = \sigma^2 \qquad [10]$$

and consider the moment centered at expectation

$$\overline{(X - m)^k} = M_k, \quad k = 3,4,\ldots. \qquad [11]$$

Let us assume that the probability density of x is given by:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma^2}} P_{2n}(x) \qquad [12]$$

where $P_{2n}(x)$ is a polynomial of degree 2n (if we take a polynomial of odd degree 2n + 1, we will have the consequence that for $x \to -\infty$ or $x \to +\infty$, depending on the coefficient of x^{2n+1}, the probability density P (x) will have very large negative values, which is absurd). It is always possible to express any polynomial as a sum of Hermite polynomials. Thus we can write

$$P(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma^2}} \left[\sum_{o}^{2n} A_j H_j \left(\frac{x-m}{\sigma}\right)\right] \qquad [13]$$

Let us note that P(x) is an entire function of order 2 of x.

From the orthogonal property, Equation [6], we get the remarkably simple expression for the coefficients

$$A_n = \frac{1}{n\,!}\,\overline{H_n\left(\frac{X - m}{\sigma}\right)} \qquad [14]$$

namely,

$$A_o = 1 \qquad [15]$$

$$A_1 = 0 \qquad [16]$$

$$A_2 = 0 \qquad [17]$$

$$A_3 = \frac{1}{3\,!}\,\frac{\mu_3}{\sigma^3} \qquad [18]$$

$$A_4 = \frac{1}{4\,!}\left(\frac{\mu_4}{\sigma^4} - 3\right) \qquad [19]$$

$$A_5 = \frac{1}{5\,!}\left(\frac{\mu_5}{\sigma^5} - 10\frac{\mu_3}{\sigma^3}\right) \qquad [20]$$

$$A_6 = \frac{1}{6!}\left(\frac{\mu_6}{\sigma^6} - 15\,\frac{\mu_4}{\sigma^4} + 15\right) \qquad [21]$$

We can have equivalent expressions in terms of the cumulants K_n in place of the centered moments μ_n. Let us remember that $\phi(t)$ being the characteristic function

$$\phi(t) = e^{\overline{i\,t\,X}}$$

the cumulants K_n are defined by the expression

$$\log \phi(t) = \sum_{1}^{+\infty} K_n \frac{(it)^n}{n!} \qquad [22]$$

the series being convergent in the circle $|t| < \delta$, and where $\phi(t) > 0$. We find

$$K_1 = 0$$
$$K_2 = \sigma^2$$
$$K_3 = \mu_3$$
$$K_4 = \mu_4 - 3\,\sigma^2$$

Thus, with this notation, we obtain the equivalent expressions

$$A_o = 1$$
$$A_1 = 0$$
$$A_2 = 0$$
$$A_3 = K_3\ 6\ \sigma^2$$
$$A_4 = K_4\ 24\ \sigma^4$$

The expression of the characteristic function $\phi(t)$ corresponding to the probability density $P(x)$ defined by Equation [13] can easily be computed. From Equation [5] we obtain

$$(i\sigma t)^n\, e^{itm - \frac{\sigma^2 t^2}{2}} = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{+\infty} e^{it\,x - \frac{(x-m)^2}{2\,\sigma^2}}\, H_n\left(\frac{x-m}{\sigma}\right)dx \qquad [23]$$

thus

$$\phi(t) = e^{imt - \frac{\sigma^2 t^2}{2}}\left[\sum_{0}^{2n} A_j\,(i\,\sigma\,t)^j\right] \qquad [24]$$

The simplest way to compute the moments μ_n in terms of the A_j (i.e., the inverse formula of Equation [14]) is to compute the coefficients of t in the development of the entire function of t defined by [24].

If we assume that all the A_n for $n \geq 1$ are equal to zero, the probability density $P(x)$ is simply the normal law.

The first step, to use the Gram-Charlier approximation, would be to try

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \left[1 + A_3 H_3 \left(\frac{x-m}{\sigma}\right) + A_4 H_4 \left(\frac{x-m}{\sigma}\right)\right] \qquad [25]$$

the coefficients A_3 and A_4 being computed from the known moments μ_3 and μ_4 by formulas [18] and [19]. In order that [24] be the <u>exact</u> expression of the probability density of x, all the coefficients A_5, A_6, A_7,...computed from the given μ_n by [20], [21]...must be <u>zero</u>.

Let us note that the characteristic function corresponding to [25] has the expression

$$\phi(t) = e^{imt - \frac{\sigma^2 t^2}{2}} \left[1 + A_3 (i\,\sigma t)^3 + A_4 (1\,\sigma t)^4\right] \qquad [26]$$

Thus $\phi(t)$ is an entire function of order 2.

All the moments μ_3, μ_4, μ_5, can easily be expressed in terms of A_3 and A_4, by computing (from Equation [25]) the development of $\phi(t)$ in a power series

$$\phi(t) = 1 - imt - \frac{\sigma^2 t^2}{2} + \mu_3 \frac{(it)^3}{3!} + \mu_4 \frac{(it)^4}{4!} + \ldots$$

One way to test the accuracy of the Gram-Charlier approximation, limited to a polynomial of order 4, would be as follows:

a) compute A_3 and A_4 from the known values of μ_3 and μ_4 by formulas [18] and [19];

b) from these values of A_3 and A_4 compute the moments $\hat{\mu}_5$, $\hat{\mu}_6$, $\hat{\mu}_7$ . . .

$$\hat{\mu}_5 = 10\,\sigma^2\,\mu_3$$

$$\hat{\mu}_6 = 15\ \sigma^2\ \mu_4 - 30\ \sigma^6$$

and then compare these values with the known values μ_5, μ_6, μ_7.

We could consider the approximation as being good if the differences

$$|\hat{\mu}_5 - \mu_5|,\ |\hat{\mu}_6 - \mu_6|,\ |\hat{\mu}_7 - \mu_7|$$

are all small.

It is interesting to note that in that case the skewness of P(x) is produced by A_3 only; if $A_3 = 0$, the function P(x) is even in (x-m).

As we have already pointed out, in order that P(x) might be a probability density, the polynomial $P_{2n}(x)$ must satisfy the condition

$$P_{2n}(x) \geq 0 \qquad [27]$$

for all real x; this condition seems to have been neglected very often. Let us observe that

$$P_4(x) = 1 + A_3\ H_3\left(\frac{x-m}{\sigma}\right) + A_4\ H_4\left(\frac{x-m}{\sigma}\right)$$

is a continuous function of A_3, A_4. For $A_3 = A_4 = 0$ we have

$$P_4(x) = 1$$

and the condition is satisfied. Thus, if

$$|A_3| < \beta \quad 0 < A_4 < \alpha$$

we are sure that Equation [27] is satisfied.

Putting

$$a = 3\frac{A_3}{A_4} \quad b = 3 + \frac{1}{A_4}$$

we find the following bounds

$$b \geq 9, \quad a \leq \frac{\sqrt{-48 + \sqrt{5766 - 2840}}}{9} \qquad [28]$$

JOINT PROBABILITY DENSITY OF TWO RANDOM VARIABLES

Let (X,Y) be a pair of random variables; suppose we know the <u>moments</u>

$$\overline{X^j Y^k} = m_{j,k} \quad \text{for} \quad 0 \leq (j + k) \leq j \qquad [29]$$

We write m and n for $m_{1,o}$ and $m_{o,1}$
(the expectations of X and Y)

$$\bar{X} = m \quad \bar{Y} = n \qquad [30]$$

From $m_{j,k}$ we can compute the <u>moments centered at expectation</u>

$$\overline{(X-m)^j \quad (Y-n)^k} = \mu_{j,k} \qquad [31]$$

We write σ^2 and τ^2 for $\mu_{2,o}$, $\mu_{o,2}$ (the variances of X and Y) and $r\,\sigma\,\tau$ for $\mu_{1,1}$ (r being the correlation of X and Y). We always suppose

$$\begin{aligned} &\sigma > 0 \quad \tau > 0 \\ &\overline{(X-m)^2} = \sigma^2 \\ &\overline{(Y-m)^2} = \tau^2 \\ &\overline{(X-m) \quad (Y-n)} = r\,\sigma\,\tau \quad |r| \leq 1 \end{aligned} \qquad [32]$$

If (X,Y) follows the normal law,

$$\text{Prob}\left[x < X < x + dx,\ y < Y < y + dy\right] = p\ (x,y)\ dxdy \qquad [33]$$

$$p\ (x,y) = \frac{\sqrt{\Delta}}{2\pi} \exp\left[-\frac{1}{2}\phi\ (x-m,\ y-n)\right]$$

where

$$\phi\ (x,y) = \frac{1}{(1 - r^2)}\left[\frac{x^2}{\sigma^2} - 2\ r\frac{xy}{\sigma t} + \frac{y^2}{\tau^2}\right] \qquad [34]$$

$$\Delta = \frac{1}{\sigma^2\ \tau^2\ (1 - r^2)}$$

To represent the given $m_{j,k}$ (or $\mu_{j,k}$) as well as possible, we will try a probability law of the type

$$p\ (x,y) = \frac{\sqrt{\Delta}}{2\pi} \exp\left[\frac{1}{2}\phi\ (x - m,\ y - n)\right] P\ (x,y) \qquad [35]$$

where P (x,y) is a suitably chosen polynomial. We will give an expression for P (x,y) in terms of the Hermite polynomials of two variables.

HERMITE POLYNOMIALS OF TWO VARIABLES

Consider a positive definite quadratic form

$$\phi\ (x,y) = ax^2 + 2\ bxy + cy^2 \qquad [36]$$

$$a > 0 \quad c > 0 \quad \Delta = ac - b^2 > 0$$

and its adjunct

$$\psi\ (\xi\ \eta) = \frac{c}{\Delta}\xi^2 - 2\frac{b}{\Delta}\xi\ \eta + \frac{a}{\Delta}\eta^2 \qquad [37]$$

Putting

$$\xi = ax + by \quad \eta = bx + cy \tag{38}$$

we have the identity

$$\psi(\xi, \eta) = \phi(x, y) \tag{39}$$

Following Reference 4, pp 363-387, we define the Hermite polynomials

$H_{m,n}(x,y)$ and $G_{m,n}(x,y)$ by

$$\exp\left[h(ax + by) + k(bx + cy) - \frac{1}{2}\phi(h,k)\right] = \sum_{0}^{+\infty} \frac{h^m}{m!}\frac{k^n}{n!} H_{mn}(x,y) \tag{40}$$

$$\exp\left[hx + ky - \frac{1}{2}\psi(h,k)\right] = \sum_{0}^{+\infty} \frac{h^m}{m!}\frac{k^n}{n!} G_{m,n}(x,y) \tag{41}$$

Let us recall some useful properties of those polynomials

$$H_{m,n}(x,y) = (-1)^{m+n} e^{\frac{1}{2}\phi(x,y)} \frac{\partial^{m+n}}{\partial x^m \partial y^n}\left[e^{-\frac{1}{2}\phi(x,y)}\right] \tag{42}$$

$$G_{m,n}(x,y) = (-1)^{m,n} e^{\frac{1}{2}\psi(\xi,\eta)} \frac{\partial^{m,n}}{\partial \xi^m \partial \eta^n}\left[e^{-\frac{1}{2}\psi(\xi\ \eta)}\right] \tag{43}$$

The polynomials $H_{m,n}$ and $G_{m,n}$ verify the following conditions:

$$\begin{cases} \dfrac{\partial}{\partial x} H_{m,n}(x,y) = am\, H_{m-1,n} + bn\, H_{m,n-1} \\[2ex] \dfrac{\partial}{\partial y}, H_{m,n}(x,y) = bm\, H_{m-1,n} + cn\, H_{m,n-1} \end{cases} \tag{44}$$

$$\begin{cases} \dfrac{\partial}{\partial x} G_{m,n}(x,y) = m\, G_{m-1,\,n} \\[2ex] \dfrac{\partial}{\partial y} G_{m,n}(x,y) = n\, G_{m,\,n-1} \end{cases} \qquad [45]$$

$$H_{m,n}(x,y) - \xi H_{m-1,\,n}(x,y) + a(m-1) H_{m-2,\,n}(x,y) \qquad [46]$$

$$+ bn\, H_{m-1,\,n-1}(x,y) = 0$$

$$H_{m,n}(x,y) - \eta H_{m,n-1}(x,y) + bm\, H_{m-1,\,n-1}(x,y)$$

$$+ c(n-1) H_{m,n-2}(x,y) = 0$$

$$G_{m,n}(x,y) - x\, G_{m-1,\,n}(x,y) + \frac{c}{\Delta}(m-1) G_{m-2,\,n}(x,y) \qquad [47]$$

$$- \frac{b}{\Delta} n\, G_{m-1,\,n-1}(x,y) = 0$$

$$G_{m,n}(x,y) - y\, G_{m,n-1}(x,y) - \frac{b}{\Delta} m\, G_{m-1,\,n-1}(x,y)$$

$$+ \frac{a}{\Delta}(n-1) G_{m,n-2}(x,y) = 0$$

The first polynomials $H_{m,n}$ have the following expressions:

$$H_{o,o}(x,y) = 1 \qquad [48]$$

$$H_{1,o}(x,y) = \xi$$

$$H_{o,1}(x,y) = \eta$$

$$H_{2,o}(x,y) = \xi^2 - a$$

$$H_{1,1}(x,y) = \xi\,\eta - b$$

$$H_{o,2}(x,y) = \eta^2 - c$$

$$H_{3,o}(x,y) = \xi^3 - 3\,a\,\xi$$

$$H_{2,1}(x,y) = \xi^2\,\eta - 2b\,\xi - a\,\eta$$

$$H_{1,2}(x,y) = \xi\eta^2 - 2b\,\eta - c\,\xi$$

$$H_{o,3}(x,y) = \eta^3 - 3\,c\,\eta$$

$$H_{4,o}(x,y) = \xi^4 - 6\,a\,\xi^2 + 3\,a^2$$

$$H_{3,1}(x,y) = \xi^3\eta - 3\,b\,\xi^2 - 3\,a\,\xi\,\eta + 3\,a\,b$$

$$H_{2,2}(x,y) = \xi^2\,\eta^2 - c\,\xi^2 - 2\,b\,\xi\,\eta - a\,\eta^2 + ac + 2\,b^2$$

$$H_{1,3}(x,y) = \xi\,\eta^3 - 3\,b\,\eta^2 - 3\,c\,\xi\,\eta + 3\,c\,b$$

$$H_{o,4}(x,y) = \eta^4 - 6\,c\,\eta^2 + 3\,c^2$$

The first polynomials $G_{m,n}$ have the following expressions:

$$G_{o,o}(x,y) = 1 \qquad [49]$$

$$G_{1,o}(x,y) = x$$

$$G_{o,1}(x,y) = y$$

$$G_{2,o}(x,y) = x^2 - \frac{c}{\Delta}$$

$$G_{1,1}(x,y) = x\,y + \frac{b}{\Delta}$$

$$G_{o,2}(x,y) = y^2 - \frac{a}{\Delta}$$

$$G_{3,o}(x,y) = x^3 - 3\,\frac{c}{\Delta}\,x$$

$$G_{2,1}(x,y) = x^2\,y + 2\,\frac{b}{\Delta}\,x - \frac{c}{\Delta}\,y$$

$$G_{1,2}(x,y) = xy^2 + 2\,\frac{b}{\Delta}\,y - \frac{a}{\Delta}\,x$$

$$G_{o,3}(x,y) = y^3 - 3\frac{a}{\Delta}y$$

$$G_{4,o}(x,y) = x^4 - 6\frac{c}{\Delta}x^2 + 3\frac{c^2}{\Delta^2}$$

$$G_{3,1}(x,y) = x^3 y - \frac{3c}{\Delta}xy + \frac{3b}{\Delta}x^2 - \frac{3bc}{\Delta^2}$$

$$G_{2,2}(x,y) = x^2 y^2 - \frac{a}{\Delta}x^2 - \frac{c}{\Delta}y^2 + \frac{4b}{\Delta}xy + \frac{ac}{\Delta^2} + 2\frac{b^2}{\Delta^2}$$

$$G_{1,3}(x,y) = xy^3 - \frac{3a}{\Delta}xy + \frac{3b}{\Delta}y^2 - \frac{3ab}{\Delta^2}$$

$$G_{o,4}(x,y) = y^4 - 6\frac{a}{\Delta}y^2 + 3\frac{a^2}{\Delta^2}$$

The two sequences of polynomials $H_{m,n}$ and $G_{m,n}$ have the orthogonality property

$$\frac{\sqrt{\Delta}}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\phi(x,y)} H_{m,n}(x,y)\, G_{p,q}(x,y)\, dxdy = \delta_{m,p}\,\delta_{n,q}\, m!\, n! \qquad [50]$$

This is the most original contribution of Hermite: the $H_{m,n}$ or the $G_{m,n}$ are not orthogonal to themselves but they are biorthogonal, i.e., they are orthogonal to each other in this very special way.

Let us note that the polynomials $H_{m,n}(x,y)$ and $G_{m,n}(x,y)$ degenerate in a product $H_m(x)\, H_n(y)$ if, and only if, the coefficient r is equal to zero.

GRAM-CHARLIER APPROXIMATION OF A TWO-DIMENSIONAL PROBABILITY DENSITY

Let us now suppose that in the probability law [35] the polynomial P (x,y) is the sum of a finite number of Hermite polynomials

$$H_{m,n}\left(\frac{x-m}{\sigma}, \frac{y-n}{\tau}\right)$$

In the quadratic forms [36] and [37] we will assume

$$a = \frac{1}{1 - r^2}, \quad b = \frac{r}{1 - r^2}, \quad c = \frac{1}{1 - r^2} \qquad [51]$$

thus

$$\Delta = ac - b^2 = \frac{1}{1 - r^2} \quad |r| \leq 1 \qquad [52]$$

and [36] becomes

$$\phi\ (x,y) = \frac{1}{1 - r^2}\ (x^2 - 2\ r\ x\ y + y^2) \qquad [53]$$

the adjunct form [37] being

$$\psi\ (\xi,\ \eta) = \xi^2 + 2\ r\ \xi\ \eta + \eta^2 \qquad [54]$$

the variables $(\xi,\ \eta)$ and (x,y) being connected by

$$\xi = \frac{1}{1 - r^2}\ (x - r\ y) \quad \eta = \frac{1}{1 - r^2}\ (- r\ x + y) \qquad [55]$$

Now it is obvious that, when (x,y) follows the normal law, we have

$$p_o\ (x,y) = \frac{1}{2\ \pi\ \sigma\ \tau \sqrt{1 - r^2}} \exp\left[- \frac{1}{2}\ \phi\left(\frac{x - m}{\sigma}, \frac{y - n}{\tau}\right)\right] \qquad [56]$$

where the quadratic form ϕ is defined by [53].

Let us point out again that $p_o\ (x,y)$ is the probability for a pair of random variables X, Y which are not necessarily independent; it applies even if the correlation r of X, Y is not zero. Now the principle of our method is to use the two-dimensional Gram-Charlier series:

$$p\ (x,y) = p_o\ (x,y)\left[\sum_{j + k = 0}^{j + k = 4} A_{j,k}\ H_{j,k}\left(\frac{x - m}{\sigma}\right), \left(\frac{y - n}{\tau}\right)\right] \qquad [57]$$

where $H_{j,k}\ (x,y)$ are the Hermite polynomials [48] a, b, c, ξ, η having the values [51] and [55].

Because of the orthogonality property [50], we obviously have the fundamental result

$$A_{j,k} = \frac{1}{j!\ k!}\ \overline{G_{j,k}\left(\frac{x-m}{\sigma}\right)\left(\frac{y-n}{\tau}\right)} \qquad [58]$$

Equations [49] give the values of the coefficients $A_{j,k}$ in terms of the moments $\mu_{j,k}$ (defined by [31]); we have

$$\begin{aligned}
&A_{o,o} = 1 \qquad A_{1,o} = 0 \qquad A_{o,1} = 0 \\
&A_{2,o} = 0 \qquad A_{1,1} = 0 \qquad A_{o,2} = 0 \\
&A_{3,o} = \frac{\mu_{3,o}}{6\ \sigma^3} \qquad A_{2,1} = \frac{\mu_{2,1}}{2\ \sigma^2\ \tau} \qquad A_{1,2} = \frac{\mu_{1,2}}{2\ \sigma\ \tau^2}
\end{aligned} \qquad [59]$$

$$A_{o,3} = \frac{\mu_{o,3}}{6\ \tau^3},$$

$$A_{4,o} = \frac{1}{24}\left(\frac{\mu_{4,o}}{\sigma^4} - 3\right),$$

$$A_{3,1} = \frac{1}{6}\left(\frac{\mu_{3,1}}{\sigma^3\ \tau} - 3\ r\right),$$

$$A_{2,2} = \frac{1}{4}\left(\frac{\mu_{2,2}}{\sigma^2\ \tau^2} - 1 - 2\ r^2\right),$$

$$A_{1,3} = \frac{1}{6}\left(\frac{\mu_{1,3}}{\sigma\ \tau^3} - 3\ r\right),$$

$$A_{o,4} = \frac{1}{24}\left(\frac{\mu_{o,4}}{\tau^4} - 3\right) .$$

Of course, as for the one-dimensional case, to test the fitness of the approximation, we must compute $A_{j,k}$ for $j + k = 5$ and see whether they are small.

$$A_{5,o} = \frac{1}{120}\left[\frac{\mu_{5,o}}{\sigma^5} - 10\,\frac{\mu_{3,o}}{\sigma^3}\right]$$

etc.

REFERENCES

1. J. Kampé de Fériet, in Proc. Symposia App. Math. Vol. 13, pp 165-198 (American Mathematical Society, 1962).

2. G. Birkhoff and J. Kampé de Fériet, Journal of Math. and Mech., Vol. 7, pp 663-704 (1958), Part C.

3. G. Birkhoff and J. Kampé de Fériet, Journal of Math. and Mech., Vol. 11, pp 319-340 (1962).

4. P. Appell and J. Kampé de Fériet, Fonctions hypergeometriques et hyperspheriques, Polynomes d Hermite. (Gauthiers-Villars, Paris, 1926)

5. H. Cramer, Mathematical Methods of Statistics (Princeton University Press, 1946.

6. M. G. Kendall, The advanced theory of statistics, Vol. I, (Ch. Griffin, London, 1948).

7. M. G. Kendall, Biometrika, Vol. 36, pp 177-190 (1949).

8. J. S. Pretorius, Biometrika, Vol. 22, pp 109-221 (1930).

BLANK PAGE

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY *(Corporate author)* David Taylor Model Basin, Washington D.C. 20007	2a REPORT SECURITY CLASSIFICATION Unclassified 2b GROUP

3 REPORT TITLE

THE GRAM-CHARLIER APPROXIMATION OF THE NORMAL LAW AND THE STATISTICAL DESCRIPTION OF A HOMOGENEOUS TURBULENT FLOW NEAR STATISTICAL EQUILIBRIUM

4 DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Final

5 AUTHOR(S) *(Last name, first name, initial)*

Kampé de Fériet, Joseph

6 REPORT DATE March 1966	7a TOTAL NO. OF PAGES 28	7b NO. OF REFS 8
8a CONTRACT OR GRANT NO. b. PROJECT NO. c. d	9a ORIGINATOR'S REPORT NUMBER(S) 2013 9b OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*	

10 AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY David Taylor Model Basin Washington, D.C. 20007

13 ABSTRACT

The probability density distribution is discussed for two correlated random variables based on an approximation to a normal (Gaussian) law using Hermite polynomials of two variables.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Gram-Charlier Distribution						
Turbulence						
Probability						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

Unclassified

Security Classification