Citation:

J.P. van der Stok, On the analysis of frequency curves according to a general method, in:
KNAW, Proceedings, 10 II, 1907-1908, Amsterdam, 1908, pp. 799-817

Geophysics. — *"On the analysis of frequency curves according to a general method."* By Dr. J. P. VAN DER STOK.

§ 1. In working out meteorological data statistically (climatology), frequencies of all descriptions are found. No doubt the majority are between indefinite limits as most other frequencies of different origin, but it also happens that the limits are sharply defined as in the case of observations upon the degree of cloudiness, where they lie between 0 and 10.

An intermediate form is found in the frequencies of rain showers arranged according to duration or quantity; on the one hand they are rigidly limited by the zero value, on the other hand the heavy showers are without definite limits, so that the curve gradually approaches the axis of abscissae.

The elaboration of wind-observations requires the treatment of frequencies in two dimensions, and produces curves, which differ in character from other frequency curves according to the nature of their origin.

The development in series according to the formula of BRUNS[1]) and CHARLIER, appears to be the method indicated for frequencies with indefinite limits; but the deduction of this formula is based upon a generalisation in the use of definite integrals as already pointed out by BESSEL and therefore not quite free from premises, which may be applicable to the theory of probability but have no connection with the problem in question which may be defined as the analysis of an arbitrary function between given limits. Besides, this method of deduction can hardly be applied in the case of definite limitation.

The formulae of PEARSON, as also those of CHARLIER, are entirely based upon the premises of the theory of probability and, as they are not given in series form, they only contain a definite number of constants which, in some cases, is too limited to allow a complete characterisation of the curve, particularly in the working out of frequencies of the cloudiness, as will be shown in an example in another communication.

Besides, the constants, which partly appear in exponential form,

---

[1]) H. BRUNS. Wahrscheinlichkeitsrechnung und Kollektivmasslehre, Berlin, 1906.
Idem. Beiträge zur Quotenrechnung. Kon. Sächs. Gesellsch. d. Wiss. Bnd. 58. Leipzig, 1906.

C. V. L. CHARLIER. Researches into the theory of probability. Meddel. Lunds astr. observ. Ser. II. n⁰. 4. 1906.

Idem. Ueber das Fehlergesetz. Ark. for Matem. Astron. och. Fys. Bnd. 2. n⁰. 8, 1905.

give no clear indication of the part they play in the construction of the curve, and it is not well possible to describe their function in a simple manner either verbally or graphically.

The object of this communication is to propose a general and simple method by which a curve may be found, which being integrated between certain limits, defined by the distribution of the data, will give the sums characteristic of this distribution, and that for frequencies of different kinds, as far as this is possible owing to the elements of uncertainty proceeding from the imperfection of the data which, of course, always remain.

This curve, representing the law which the phenomenon follows, should be called the frequency-curve; the curve of the aggregate values, obtained by grouping the original data within definite limits, may then be called the curve of distribution according to BRUNS. Its form depends upon the degree of condensation of the original data (Abrundung after BRUNS), but approximates more to that of the frequency curve as the condensation becomes less extensive and consequently the number of observations is greater.

Such a development of an arbitrary function can evidently be made in an infinite number of ways; it is therefore necessary to postulate some general principles.

The following premises apply to the method of development selected:

1. That the development takes place according to polynomia of an ascending degree.

2. that for the determination of the constants, the calculation of means of different orders is used, in relation to an origin favourably selected according to the requirements of the various cases.

The expression "moments" which is frequently employed, has been avoided as an unnecessary analogy with mechanical problems.

§ 2. DEVELOPMENT BETWEEN DEFINITE LIMITS.

a. No given values of the function at the limits.

The polynomia, the degree of which is indicated by a suffix, are represented by Q_n, and the series by:

$$u = A_0 Q_0 + A_1 Q_1 + A_2 Q_2 + \dots \text{etc.} \qquad (1)$$

The simplest form which can be given to the polynomia is:

$$Q_n = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots a_n$$

In this case the most practical choice for the origin of coordinates is evidently the mean between the limits as then, on integrating between the limits, all odd terms vanish; hence a separation between even and odd polynomia becomes necessary, and the general expression is:

$$Q_n = x^n + a_2 x^{n-2} + a_4 x^{n-4} + \dots a_n \qquad n \text{ even}$$
$$= x^n + a_1 x^{n-2} + a_3 x^{n-4} + \dots a_{n-2} \qquad n \text{ odd}$$

A simplification of the formulae can then be obtained by altering the scale value in such a way that the limits become ± 1, which is always possible; for the sake of convenience these limits have been omitted in the following expressions.

The means of different order are indicated by:

$$\mu_n = \int u x^n dx .$$

In order to enable us to calculate from the infinite series (1) the A-coeff. in a finite form, the unique and sufficient condition is that the a-coeff. be determined so that the condition:

$$\int Q_n x^m dx = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

is satisfied for all values of $m < n$ as then all integrals beyond the $m + 1^{\text{th}}$ term vanish and, at the same time, the a-coeff. are entirely fixed, but for an arbitrary constant factor.

If this operation has been performed, it is at once evident from (2) that:

$$\int Q_m Q_n dx = 0$$

for all values of m different from n and, further, that:

$$A_n = \alpha \int u Q_n dx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where:

$$\alpha^{-1} = \int Q_n Q_n dx = \int Q_n x^n dx .$$

The $n/2$ (n even) or $n-1/2$ (n odd) constants of the polynomium Q_n are calculated from the $n/2$ or $n-1/2$ equations:

$$\left.\begin{array}{l} \int Q_n dx = 0 \\ \int Q_n x^2 dx = 0 \\ \dots\dots \\ \int Q_n x^{n-2} dx = 0 \end{array}\right\} (n \text{ even}) \qquad \left.\begin{array}{l} \int Q_n x dx = 0 \\ \int Q_n x^3 dx = 0 \\ \dots\dots \\ \int Q_n x^{n-2} dx = 0 \end{array}\right\} (n \text{ odd})$$

or, for n even, from :

$$\frac{1}{n+1}+\frac{a_2}{n-1}+\frac{a_4}{n-3}+\dots\frac{a_n}{1}=0$$

$$\frac{1}{n+3}+\frac{a_2}{n+1}+\frac{a_4}{n-1}+\dots\frac{a_n}{3}=0$$

$$\dots\dots\dots\dots\dots$$

$$\frac{1}{2n-1}+\frac{a_2}{2n-3}+\frac{a_4}{2n-5}+\dots\frac{a_2}{n-1}=0$$

for n odd, from :

$$\frac{1}{n+2}+\frac{a_1}{n}+\frac{a_3}{n-2}+\dots\frac{a_{n-2}}{3}=0$$

$$\frac{1}{n+4}+\frac{a_1}{n+2}+\frac{a_3}{n}+\dots\frac{a_{n-2}}{5}=0$$

$$\dots\dots\dots\dots\dots$$

$$\frac{1}{2n-1}+\frac{a_1}{2n-3}+\frac{a_3}{2n-5}+\dots\frac{a_{n-2}}{n}=0$$

On eliminating successively from these equations $a_2, a_4 \dots$ or $a_1, a_3 \dots$ we find for the general expression of the polynomium :

$$Q_n = x^n - \frac{n(n-1)}{2\,.\,(2n-1)}\,x^{n-2} + \frac{n(n-1)\,(n-2)\,(n-3)}{2\,.\,4\,.\,(2n-1)\,(2n-3)}\,x^{n-4} - \text{etc.} \qquad (4)$$

i. e., but for a constant factor, that of zonal harmonics, which we shall call P-functions.

This might have been expected as the condition (2), from which (4) arises, holds good also for the P-functions.

The Q-functions may, therefore, be considered as generalized P-functions, the latter presenting a special case of the former; if we write (2) :

$$k_n \int Q_n x^m dx = 0\ ,$$

then :

$$k_n Q_n = P \quad \dots\dots\dots \quad (5)$$

if k_n be defined so that :

$$k_n\, Q_n = 1 \text{ for } x = 1.$$

The use of this constant no doubt offers advantages in treating problems relating to the potential theory, but for our purpose it would be of no importance and, in practice, entail superfluous work; some expressions certainly take a simpler form by its use, but what is thereby gained on the one hand is largely lost on the other as in calculating uQ_n in (3), we have to deal with the unnecessary factor k_n.

However the relation (5), where:

$$k_n = \frac{(2n)!}{2^n \cdot n!\, n!}$$

so that:

$$Q_n = \frac{2^n \cdot n!\, n!}{(2n)!} P_n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

is useful in deriving from the well known properties of the P-functions those of the Q-functions.

They satisfy LEGENDRE's equation as well as the zonal harmonics:

$$(x^2 - 1) \frac{d^2 Q_n}{dx^2} + 2x \frac{dQ_n}{dx} - n(n+1) Q_n = 0.$$

The recurrent formula becomes:

$$Q_{n+1} - xQ_n + \frac{n^2}{(2n+1)(2n-1)} Q_{n-1} = 0,$$

and

$$Q_n = \frac{n!}{(2n)!} \frac{d^n (x^2-1)^n}{dx^n} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6a)$$

Hence, we find:

$$\alpha^{-1} = \int Q_n\, Q_n\, dx = \frac{1}{k^2} \int P_n\, P_n\, dx = \frac{2}{k^2\,(2n+1)} = \frac{2^{2n+1}\, n!\, n!\, n!\, n!}{(2n+1)!\,(2n)!}$$

and for A_n:

$$A_n = \alpha \left[\mu_n - \frac{n(n-1)}{2.(2n-1)} \mu_{n-2} + \frac{n(n-1)(n-2)(n-3)}{2 \cdot 4 \cdot (2n-1)(2n-3)} \mu_{n-4} - \text{ect.} \right] \quad (7)$$

b. Given $u = 0$ *for* $x = \pm 1$.

The case discussed sub *a*, where nothing is supposed to be known concerning the function to be developed, will seldom occur in practice and, as all adaptation is due to the accomodating power of the A-constants the application would, in such a case, necessitate the calculation of many terms and, therefore, hardly be profitable.

Now, in dealing with observations of the degree of cloudiness, the case presents itself, that a curve has to be found, which is characterized by the limiting values mentioned above.

The observations of serene sky (cloudiness zero) and of an entirely overcast sky (cloudiness ten) ought to be considered separately from the other observations as they constitute climatological factors of peculiar importance for the description of the climate (principally in northerly latitudes). Moreover they are to be regarded rather as discrete quantities, which do not show any continuous transition to a cloudiness resp. of degree 1 or 9.

The other degrees of cloudiness may then be regarded as observations of continuous quantities subject to the above mentioned conditions.

In this case we may easily cause all terms of the series (1) to suit these conditions by simply multiplying the series by a factor that vanishes for $x = \pm 1$ e. g. $x^2 - 1$, and then applying to the new functions, which we shall call R, the same reasonings as sub a.

The degree of the polynomia is then increased by two, so that we have to start with R_2.

The general expression becomes:

$$R_{n+2} = (x^2-1)\, R'_n = (x^2-1)\, [x^n + a_2\, x^{n-2} + \ldots a_n]\ ,\ n \text{ even}$$

$$= (x^2-1)\, [x^n + a_1\, x^{n-2} + \ldots\ a_{n-2}],\ n \text{ odd}.$$

The result of this operation is evidently that the surface enclosed by the curve, as determined by the first term of the series, is not represented by a rectangle of base 2 and height 0.5 as in the case of the Q-functions, but by a parabola of base 2 and height 0.75, which makes again the surface equal to unity.

By alternately asymmetrical and symmetrical deformations the shape of this parabola is then altered by means of the next terms in such a manner as to make it approach more and more to the frequency curve corresponding to the given data.

It may be noticed here that in the case of fixed limits, there is no reason to choose for the origin of coordinates the point corresponding to the arithmetical mean; for logical and practical reasons the point intermediate between the limits is then indicated.

The condition, which has to be satisfied by the a-coeff. of the R-function, and by which they are fully determined, is now that:

$$\int R_{n+2}\, x^m\, dx = \int R'_n\, x^m\, (x^2 - 1)\, dx = 0, \quad m < n$$

or

$$\int x^{m+2}\, R'_n\, dx = \int x^m\, R'_n\, dx. \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The a-coeff. are calculated from the equations:

$$\left.\begin{array}{l}
\dfrac{1}{(n+3)(n+1)}+\dfrac{a_2}{(n+1)(n-1)}+\dfrac{a_4}{(n-1)(n-3)}+\cdots\dfrac{a_n}{3.1}=0 \\
\dfrac{1}{(n+5)(n+3)}+\dfrac{a_2}{(n+3)(n+1)}+\dfrac{a_4}{(n+1)(n-1)}+\cdots\dfrac{a_n}{5.3}=0 \\
\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\
\dfrac{1}{(2n+1)(2n-1)}+\dfrac{a_2}{(2n-1)(2n-3)}+\dfrac{a_4}{(2n-3)(2n-5)}+\cdots\dfrac{a_n}{(n+1)(n-1)}=0
\end{array}\right\}\ n \text{ even}$$

and

$$\left.\begin{array}{l}
\dfrac{1}{(n+4)(n+2)}+\dfrac{a_1}{(n+2)(n)}+\dfrac{a_3}{(n)(n-2)}+\cdots\dfrac{a_{n-2}}{5.3}=0 \\
\dfrac{1}{(n+6)(n+4)}+\dfrac{a_1}{(n+4)(n+2)}+\dfrac{a_3}{(n+2)(n)}+\cdots\dfrac{a_{n-2}}{7.5}=0 \\
\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\
\dfrac{1}{(2n+1)(2n-1)}+\dfrac{a_1}{(2n-1)(2n-3)}+\dfrac{a_3}{(2n-3)(2n-5)}+\cdots\dfrac{a_{n-2}}{(n+2)(n)}=0
\end{array}\right\}\ n \text{ odd}$$

By successive elimination of $a_2, a_4 \ldots a_1, a_3 \ldots$ we find from these equations for the general form of the R functions:

$$R_{n+2}=x^{n+2}-\frac{(n+2)(n+1)}{2\,.\,(2n+1)}x^n+\frac{(n+2)(n+1)(n)(n-1)}{2\,.\,4\,.\,(2n+1)(2n-1)}x^{n-2}-\text{etc.} \quad (9)$$

and from this expression by dividing it by x^2-1:

$$R'_n=x^n-\frac{n(n-1)}{2.(2n+1)}x^{n-2}+\frac{n(n-1)(n-2)(n-3)}{2\,.\,4\,.\,(2n+1)(2n-1)}x^{n-4}-\text{etc.} \quad . \quad (10)$$

The recurrent formula for both R and R' is:

$$R'_{n+1}-x\,R'_n+\frac{n(n+2)}{(2n+3)(2n+1)}R'_{n-1}=0$$

and the functions are solutions of the diff. equations

$$(x^2-1)\frac{d^2R_{n+2}}{dx^2}-(n+2)(n+1)R_{n+2}=0$$

$$(x^2-1)\frac{d^2R'_n}{dx^2}+4x\frac{dR'_n}{dx}-(n+3)\,n\,R'_n=0.$$

On comparing the expression for R'_n with that for Q_n it is readily seen that the R' functions may be found by differentiation of the Q_{n+1}-function, so that:

$$R'=\frac{1}{n+1}\cdot\frac{dQ_{n+1}}{dx} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

This might have been expected as the value:

$$R' = k_n \frac{dQ_{n+1}}{dx}$$

satisfies the condition (8)

$$\int x^{m+2} \frac{dQ_{n+1}}{dx}\, dx = \int x^m \frac{dQ_{n+1}}{dx}\, dx, \quad m < n$$

which is easily proved by partial integration.

Therefore the series discussed here:

$$u = \Sigma A_n R_{n+2} \qquad n = 0 . 1 . 2 \ . \ . \ .$$

might also (but for a constant factor) be written thus:

$$u = (x^2 - 1) \Sigma A_n \frac{dQ_{n+1}}{dx} \qquad n = 0 . 1 . 2 \ . \ . \ .$$

The calculation of the A-constants is based upon the evident property of the R functions that:

$$\int R_{n+2} R'_m\, dx = 0, \qquad m \text{ different from } n$$

hence

$$A_n = \beta \int u R'_n\, dx$$

where:

$$\beta^{-1} = \int R_{n+2} R'_n\, dx = \int R_{n+2}\, x^n\, dx = \int x^n (x^2 - 1) R'_n\, dx = 0$$

or, by (11)

$$\beta^{-1} = \frac{1}{n+1} \int x^n (x^2 - 1) \frac{dQ_{n+1}}{dx}\, dx$$

From the diff. equation of the R-function follows:

$$\frac{d}{dx}\left[(x^2 - 1) \frac{dQ_{n+1}}{dv}\right] = (n + 2)(n + 1)\, Q_{n+1}$$

thence:

$$\beta^{-1} = \frac{n + 2}{n + 1} \int x^{n+1} Q_{n+1}\, dx$$

or by (8):

$$\beta^{-1} = - \frac{2^{2n+1} (n + 2)!\, n!\, n!\, n!}{(2n + 3)(2n + 1)!\,(2n + 1)!}$$

and A_n is calculated by the expression:

$$A_n = \beta \left[\mu^n - \frac{n(n-1)}{2.(2n+1)} \mu^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2 . 4 . (2n+1)(2n-1)} \mu^{n-4} - \dots \text{etc.}\right] \quad (12)$$

The negative sign of β is due to our having chosen as general factor $x^2 - 1$, a quantity which, by the definition of the limits, is always negative.

As well as the Q-functions, the R-functions might be multiplied by an arbitrary, constant factor, such that any peculiar development becomes possible or also with a view of simplifying some expressions. In our case e. g. k_n might be chosen so that $\beta = 1$; practically however this would hardly afford any advantage.

c. Given $u = \frac{u_1}{u_0}$ *for* $x = \pm 1$.

As has been remarked above, in working out observations of cloudiness the case presents itself that the frequencies for the extreme limits vanish; if, however, we have to deal, not with the original observations, but with average values as, e. g. daily means, the frequencies of serene and overcast sky, although still of peculiar interest for the knowledge of the climate, cannot be regarded as discrete values because, owing to the operation of taking the means, a continuous transition of these extreme values into the intermediate values must be assumed.

In this case, when the curves assume peculiar forms quite different from the well known curves generally met with, we can take care that the conditions for the extreme limits are bound to the first term of the series whilst all other terms remain as they are in the case discussed sub b.

Now the first term must contain three constants, two for the extreme values and one for the fixing of the area.

In the expression

$$a_0 + b_0 x + c_0 x^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

the constants must satisfy the three conditions

$$u_1 = a_0 + b_0 + c_0$$
$$u_0 = a_0 - b_0 + c_0$$
$$2a_0 + \frac{2c_0}{3} = 1$$

hence:

$$4a_0 = 3 - (u_1 + u_0)$$
$$2b_0 = u_1 - u_0$$
$$4c_0 = 3(u_1 + u_0) - 3.$$

The reasoning as well as the application then remain the same as sub b; again

$$\int R_{n+2} R'_m \, dx = 0, \qquad m \text{ different from } n$$

with the exception however of the first term of the series which now assumes the form (13). In calculating A_n we have therefore to apply a correction to the expression for A_n which is easily found by remarking that:

$$(n+1)\int x^m R'_n\, dx = \int x^m \frac{dQ_{n+1}}{dx}\, dx$$
$$= \left(x^m Q_{n+1}\right)_{-1}^{+1} - m\int x^{m-1} Q_{n+1}\, dx\,.$$

For $m < n+2$ the last integral vanishes and, R_2 being of the second degree, we have to consider this case only.
We have, therefore:

$$(n+1)\int x^m R'_n\, dx = \left(x^m Q_{n+1}\right)_{-1}^{+1},\ m < 3$$

By (6) we find:

$$\left(Q_{n+1}\right)_{-1}^{+1} = \frac{2}{k_{n+1}} = \frac{2^n(n+2)!\,n!}{(2n+1)!}\ (n \text{ even})$$

whilst for n odd the expression vanishes.
Hence also:

$$\left(x^m Q_{n+1}\right)_{-1}^{+1} = \frac{2}{k_{n+1}} \qquad \left(m+n \text{ even}\right)$$

and equal to zero for $m+n$ odd; in calculating the constant A_n we have, therefore, only to apply a correction such that, instead of (12), now is used, for n odd:

$$A_n = \beta\int uR'_n\, dx - \frac{2^{n+1} b_0\, n!\ n!}{(2n+1)!} = \beta\int uR'_n\, dx - \frac{2^n(u_1-u_0)n!\,n!}{(2n+1)!} \tag{14}$$

and for n even:

$$A_n = \beta\int uR'_n\, dx - \frac{2^{n+1}(a_0+c_0)\,n!\,n!}{(2n+1)!} = \beta\int uR'_n\, dx - \frac{2^n(u_1+u_0)n!n!}{(2n+1)!} \tag{15}$$

This example of adaptation, of which many variants might be given, will suffice to demonstrate the applicability of the method to special cases.

§ 3. Development between definite limits on the one side and indefinite limits on the other.

a. No given value for the limit.

As has been noticed above, frequencies of duration and quantities of rainshowers lie between the asymmetrical limits: zero for the smallest and ∞ for the largest values.

Frequencies of this kind, therefore, offer an example of a transition between the case of fixed limits and infinite limits on both sides. As here there exists no symmetry in the limits, the zero-point cannot be chosen so that, on integrating, the odd functions vanish, hence a separation between even and odd functions would have no sense, and we are obliged to employ complete polynomia of ascending degree.

Here, as in the case discussed in § 2, there is no advantage in making the origin of coordinates coincide with the arithmetical mean and, from a logical as well as a practical standpoint, the zero-limit is indicated.

In order to develop the function between the limits ∞ and zero, the only thing to do is to multiply the series of polynomia with a suitable factor e. g. e^{-x}, so that the equation of the frequency curve becomes:

$$u = e^{-x}(A_0 S_0 + A_1 S_1 + \ldots \text{ etc.})$$
$$= A_0 \psi_0 + A_1 \psi_1 + \ldots \text{ etc.}$$

where:

$$S_n = x^n + a_1 x + a_2 x^{n-2} + \ldots a_n.$$

The conditions to be satisfied by the a coeff. are then:

$$\int_0^\infty e^{-x} S_n \, dx = 0 \;, \int_0^\infty e^{-x} x S_n \, dx = 0 \ldots \int_0^\infty e^{-x} x^{n-1} S_n \, dx = 0$$

and as:

$$\int_0^\infty e^{-x} x^n \, dx = n!,$$

the general conditional equations are:

$$n! + (n-1)!\,a_1 + (n-2)!\,a_2 + \ldots 1!\,a_{n-1} + 0!\,a_n = 0$$
$$(n+1)! + n!\,a_1 + (n-1)!\,a_2 + \ldots 2!\,a_{n-1} + 1!\,a_n = 0$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$(2n-1)! + (2n-2)!\,a_1 + (2n-3)!\,a_3 + \ldots n!\,a_{n-1} + (n-1)!\,a_n = 0$$

Hence we find for the general expression:

$$S_n = x^n - \frac{n^2}{1!} x^{n-1} + \frac{n^2(n-1)^2}{2!} x^{n-2} - \ldots (-1)^n n! \quad . \quad (16)$$

The method of calculating A_n is the same as in the former cases as here too:

$$\int_0^\infty e^{-x} S_m S_n \, dx = \int_0^\infty \psi_m S_n \, dx = 0 \, , \qquad m < n$$

and

$$A_n = \gamma \int_0^\infty u \, S_n \, dx$$

where:

$$\gamma^{-1} = \int_0^\infty e^{-x} S_n \, S_n \, dx = \int_0^\infty \psi_n \, x^n \, dx$$

but:

$$\int_0^\infty \psi_n \, S_n \, dx = - \left(\psi_n \, S_n\right)_0^\infty + 2 \int_0^\infty \psi_n \frac{dS_n}{dx} \, dx$$

or, as the last integral vanishes according to the conditions:

$$\gamma^{-1} = - \left(\psi_n \, S_n\right)_0^\infty = n! \, n!$$

because, by (16), only the last term has to be taken into account.

The expression for A_n then becomes:

$$A_n = \frac{\mu_n}{n! \, n!} - \frac{n}{1!} \frac{\mu_{n-1}}{n! \, (n-1)!} + \frac{n(n-1)}{2!} \frac{\mu_{n-2}}{n! \, (n-2)!} - \ldots \frac{(-1)^n}{n!} \quad . \quad (17)$$

by which the problem is solved.

The application to special cases will be simplified by a brief summary of the relations existing between the different quantities introduced which are analogous to those holding for zonal harmonics.

We remark that for S_n and ψ_n we can also write

$$S_n = (-1)^n \left(\frac{d}{dx} - 1\right)^{(n)} x^n \quad , \quad \psi_n = (-1)^n \frac{d^n}{dx^n} \left(e^{-x} \, x^n\right) \quad . \quad (18)$$

hence:

$$S_n = - \, n S_{n-1} + \frac{x}{n} \frac{dS_n}{dx} \quad \text{and} \quad S_n = (x-n) \, S_{n-1} - x \frac{dS_{n-1}}{dx}$$

from which the recurrent formula:

$$S_{n+1} + (2n + 1 - x) \, S_n + n^2 \, S_{n-1} = 0 \, , \quad . \quad . \quad . \quad (19)$$

can be derived, wherein for S_n as well ψ_n may be written.

Further the functions satisfy the diff. equ.:

$$x \frac{d^2 S_n}{dx^2} + (1-x) \frac{dS_n}{dx} + nS_n = 0$$

$$x \frac{d^2 \psi_n}{dx^2} + (1+x) \frac{d\psi_n}{dx} + (n+1) \, \psi_n = 0.$$

b. Given $u = 0$ for $x = 0$.

In the same manner as the Q-series has been made to suit the zero-condition of the function at the limits, the ψ-series can be made fit for the case that the function assumes the zero value for the lowest limit by multiplication with x. This case presents itself e.g. for frequencies of wind-velocity, the curve of which originates at the zero-point as absolute calms do not occur.

By this operation the degree of the polynomia is increased by one and we can write down at once the new T function from (16) by multiplication with x and, at the same time, substituting $n+1$ for n except in the binomial factors which remain the same.

The condition for the determination of the a-coëff. is now:

$$\int_0^\infty e^{-x}\, x^m\, T_{n+1}\, dx = 0, \qquad m < n$$

and the general expression:

$$T_{n+1} = x^{n+1} - \frac{n}{1!}\cdot\frac{(n+1)!}{n!}\, x^n + \frac{n(n-1)}{2!}\,\frac{(n+1)!}{(n-1)!}\, x^{n-1} \ldots (-1)^n\,(n+1)!\, x \quad (20)$$

From this evidently:

$$S_n = \frac{1}{n+1}\cdot\frac{dT_{n+1}}{dx} \quad . \; . \; . \; . \; . \; . \; . \; . \quad (21)$$

a similar relation as is shown by (11) between the Q and R functions.

Hence, if we put:

$$T_{n+1} = x T'_n$$

$$A_n = \gamma' \int_0^\infty u\, T'_n\, dx$$

where:

$$\gamma'^{-1} = \int_0^\infty e^{-x}\, T_{n+1}\, T'_n\, dx = \int_0^\infty e^{-x}\, x^n\, T_{n+1}\, dx =$$

$$\int_0^\infty e^{-x}\, x^n\, \frac{dT_{n+1}}{dx} = (n+1)\int_0^\infty e^{-x}\, x^n\, S_n\, dx = (n+1)!\, n!$$

so that:

$$A_n = \frac{\mu_n}{0!\,(n+1)!\,n!} - \frac{\mu_{n-1}}{1!\,n!\,(n-1)!} + \frac{\mu_{n-2}}{2!\,(n-1)!\,(n-2)!} - \ldots \frac{(-1)^n}{n!}. \quad (22)$$

If we call the series discussed here, the ψ'_{n+1} series, so that:

$$\psi'_{n+1} = e^{-x}\, T_{n+1} = e^{-x}\, x\, T'_n$$

we find the following relations:

$$\psi'_{n+1} = (-1)^n \frac{d^n}{dx^n}(e^{-x} x^{n+1}) = (-1)^{n+1} x \frac{d^{n+1}}{dx^{n+1}}(e^{-x} x^n) = -x \frac{d\psi_n}{dx}$$

$$x \frac{d^2 T'_n}{dx^2} + (2 - x) \frac{dT'_n}{dx} + nT'_n = 0$$

$$x \frac{d^2 T_{n+1}}{dx^2} - x \frac{dT_{n+1}}{dx} + (n+1)\, T_{n+1} = 0$$

$$x \frac{d^2 \psi'_{n+1}}{dx^2} + x \frac{d\psi'_{n+1}}{dx} + (n+1)\, \psi'_{n+1} = 0.$$

In exactly the same manner as the R-series could be expressed in diff. quot. of the Q series:

$$u_R = (x^2 - 1) \Sigma A_n \frac{dQ_{n+1}}{dx},$$

so the ψ' series might be expressed in diff. quot. of the ψ series:

$$u_{\psi'} = - x \Sigma A_n \cdot \frac{d\psi_n}{dx}.$$

In dealing with this kind of frequency curves an alteration of the scale value offers great advantages as well as in the case of fixed limits.

In the case discussed in § 2 it was possible by this artifice to simplify the limits; here such an alteration has no influence upon the limits which remain 0 and ∞ if we write hx for x, but we are able by this means to accomodate the first term of the series, by which the area is determined, according to the form of the curve, so that the task of the A-coefficients is lightened.

By the factor h, which by its nature is a positive quantity, no complication in the calculation of the constants is introduced: the series is now:

$$u = e^{-hx} [A_0 S_0(hx) + A_1 S_1(hx) + \ldots \text{ etc.}]. \quad . \quad . \quad (23)$$

and, because:

$$\gamma^{-1} = \int_0^\infty e^{-hx} S_n(hx)\, S_n(hx)\, dx = \frac{1}{h} \int_0^\infty e^{-t} S_n(t)\, S_n(t)\, dt$$

$$A_n = h \left[\frac{h^n \mu_n}{n!n!} - \frac{n}{1!} \cdot \frac{h^{n-1} \mu_{n-1}}{n!(n-1)!} + \ldots \frac{(-1)^n}{n!} \right] \quad . \quad . \quad (24)$$

We might also omit the coeff. h in (24) and write (23):

$$u = he^{-hx} [A_0 S_0(hx) + A_1 S_1(hx) + \ldots \text{ etc.}] \quad . \quad . \quad (23^a)$$

The scale value h may, of course, be chosen quite arbitrarily; it is however desirable to do this in accordance with the nature of the curve and, therefore, to calculate it methodically from the given data.

This can be done by suppressing one of the A-constants in (23a) so that the mean value corresponding with this constant can be made use of to define h.

If then we put:

$$A_1 = 0$$

we find, as $A_0 = 1$,

$$\mu_1 = A_0 h \int_0^\infty e^{-hx} \, x \, dx = \frac{1}{h}.$$

§ 4. Development between the indefinite limits $\pm \infty$.

By reasons of symmetry, in this case it is logical to take e^{-x^2} for the factor by which the limits are determined and when, for the same reason, the arithmetical mean is chosen as the origin, the polynomium can, as in the case of fixed limits, be separated into even and odd functions, because then, on integrating between the limits, the odd functions vanish.

The series becomes then:

$$u = e^{-x^2} [A_0 U_0 + A_2 U_2 + A_3 U_3 + \ldots \text{ etc.}$$
$$= A_0 \varphi_0 + A_2 \varphi_2 + A_3 \varphi_3 + \ldots \text{ etc.}$$

as, by the choice of the origin, the A_1-term has to be omitted. The conditional equation for the determination of the a constants is:

$$\int_{-\infty}^{\infty} x^m \varphi_n \, dx = 0 \qquad m < n \quad \ldots \ldots \quad (25)$$

or, generally. for n even:

$$[(n-1)(n-3)\ldots 1] + 2a_2 [(n-3)(n-5)\ldots 1] + 2^2 a_4 [(n-5)(n-7)\ldots 1] +$$
$$+ \ldots + 2^{n/2} a_n = 0$$

$$[(n+1)(n-1)\ldots 1] + 2a_2 [(n-1)(n-3) \,.\, 1] + 2^2 a_4 [(n-3)(n-5)\ldots 1] +$$
$$+ \ldots + 2^{n-2/2} a_n = 0$$

. . . .

$$[(2n-3)(2n-5)\ldots 1] + 2a_2 [(2n-5)(2n-7)\ldots 1] +$$
$$+ 2^2 a_4 [(2n-7)(2n-9)\ldots 1] + \ldots + 2a_n [(n-3)(n-5)\ldots 1] = 0$$

and, for n odd ·

55

$$[n(n-2)..\,1] + 2a_1\,[(n-2)\,(2-4)...1] + 2^2a_3\,[(n-4)\,(n-6)\,..1] + \\ + 2^{n-1/2}a_{n-2} = 0$$

$$[(n+2)\,(n)..\,1] + 2a_1\,[(n)\,(n-2)\,..1] + 2^2a_3\,[(n-2)\,(n-4) = ...1] + \\ + 2^{n-3/2}a_{n-2} = 0$$

. . . .

$$[(2n-3)\,(2n-5)...1] + 2a_1\,[(2n-5)\,(2n-7)\,..1] + \\ + 2^2a_3\,[(2n-7)\,(2n-9)...1] + 2a_{n-2}\,[(n-2)\,(n-4).\,.1] = 0.$$

From this it follows that:

$$U_n = x^n - \frac{n(n-1)}{2^2\ \ 1!}x^{n-2} + \frac{n(n-1)\,(n-2)\,(n-3)}{2^4\,.\,2!}x^{n-4} - \ldots \text{ etc.} \quad (26)$$

from which we derive that U_n and φ_n are solutions of the equations:

$$\frac{d^2U_n}{dx^2} - 2x\frac{dU_n}{dx} + 2n\,U_n = 0$$

$$\frac{\partial^2\varphi_n}{dx^2} + 2x\frac{d\varphi_n}{dx} + 2(n+1)\,\varphi_n = 0$$

and the recurrent formula becomes:

$$2\,U_{n+1} - 2x\,U_n + n\,U_{n-1} = 0$$

The A-coefficients are determined in the same manner as in all former cases:

$$\int_{-\infty}^{+\infty} e^{-x^2}\,U_m\,U_n\,dx = 0$$

for all values of m different from n so that:

$$A_n = \delta\int_{-\infty}^{+\infty} u\,U_n\,dx$$

where:

$$\delta^{-1} = \int_{-\infty}^{+\infty} e^{-x^2}\,U_n\,U_n\,dx = \int_{-\infty}^{\infty}\varphi_n\,x^n\,dx = \frac{n!}{2^n}\sqrt{\pi}$$

and

$$A_n = \frac{2^n}{\sqrt{\pi}}\left[\frac{\mu_n}{n!} - \frac{\mu_{n-2}}{2^2\,.\,1!\,(n-2)!} + \frac{\mu_{n-4}}{2^4\,.\,2!\,(n-4)!} - \text{etc.}\right]. \quad (27)$$

From the numerical values calculated by (26) and from the diff. equation it appears that, but for an arbitrary constant factor, the φ_n-functions are equal to the derivatives of the n^{th} order of φ_0 or e^{-x^2}; we might, therefore, write:

$$\varphi_n = k_n\frac{d^n\varphi_0}{dx^n},$$

this might have been expected as this value satisfies the condition (25), which can be easily proved by successive partial integration. If we put: $k_n = 1$,

$$\varphi_n = \frac{d^n e^{-x^2}}{dx^n} = (-2)^n\, U_n\, e^{-x^2}$$

and the expression for the A-coefficients becomes equal to that given by Bruns.

Therefore $\frac{d^n \varphi_0}{dx^n}$ may be substituted for φ_n for the same reasons as, instead of the Q-functions, zonal harmonics might be employed; in practice however no labour is saved by this substitution as then the polynomia are charged with superfluous coefficients. After what has been said in § 3 about a change of the scale value, it will be sufficient to remark that in this case also the great advantage which can be derived from the introduction of a scale factor is the adaptation by means of the first term of the series to the shape of the curve, the surface remaining equal to unity.

The equation of the curve then becomes:

$$u = e^{-h^2x^2} [A_0 U_0 (hx) + A_2 U_2 (hx) + etc. \quad . \quad . \qquad (28)$$

and:

$$A_n = \frac{2^n h}{\sqrt{\pi}} \left[\frac{h^n \mu_n}{n!} - \frac{h^{n-2} \mu_{n-2}}{2^2 . 1!(n-2)!} + etc.) \right] . \quad . \quad . \quad . \qquad (29)$$

The choice of the scale factor is of course quite arbitrary, but, in order to determine it in accordance with the nature of the curve, it is desirable to put $A_2 = 0$, then the average of the second order can be used for the definition of h and it is easily seen that:

$$\mu_2 = \frac{1}{2h^2} .$$

The coeff. of (29) in so far as they are independent of n may further be omitted and written before (28), then the equation of the curve becomes:

$$u = \frac{h}{\sqrt{\pi}}\, e^{-h^2x^2} [A_0 U_0 + A_3 U_3 + A_4 U_4 + etc.$$

If we take into consideration only the first term in the development, we find the exponential law in its simplest form as $A_0 = 1$.

$$u = \frac{h}{\sqrt{\pi}}\, e^{-h^2x^2} .$$

55*

§ 5. Indefinite limits, two variables.

The treatment of wind observations now offers no difficulties as, in calculating the means of different-order, the two variables (projections upon two axes arbitrarily chosen) can always be separated and the method remains in all other respects quite the same. Only, instead of one mean of each order, we can now dispose of $p+1$ means of order p.

If by V_n be denoted the same function of y as U_n is of x, the equ. of the curve assumes, as $U_0 = V_0 = 1$, the form :

$$u(x,y) = e^{-x^2-y^2}[A_0 + A_{1.0}U_1 + A_{0.1}V_1 + A_{2.0}U_2 + A_{1.1}U_1V_1 + A_{0.2}V_2 + A_{3.0}U_3 + A_{2.1}U_2V_1 + A_{1.2}U_1V_2 + A_{0.3}V_3 + etc.] \quad . \quad (30)$$

The general expression for the polynomia is :

$$U_n\,V_m$$

and as, evidently :

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-x^2-y^2}(U_nV_m)(U_pV_q)\,dxdy = 0$$

for all values of p different from n and of q different from m, we find for the A-coeff. :

$$A_{nm} = \varepsilon\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-x^2-y^2}u(U_nV_m)\,dxdy$$

where :

$$\varepsilon^{-1} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-x^2-y^2}(U_nV_m)^2\,dxdy = \frac{n!m!}{2^{n+m}}\pi \quad . \quad . \quad (31)$$

From the considerations of § 4 it follows that the function :

$$\Phi_{nm} = e^{-x^2-y^2}\,U_n V_m$$

may as well be given the form :

$$\Phi_{nm} = k_{nm}\frac{d^{n+m}}{dx^n dy^m}\Phi_0 = k_{nm}\frac{d^{n+m}}{dx^n dy^m}e^{-x^2-y^2}$$

as this satisfies the premised condition; then the series (30) assumes the form of a sum of diff. quot. like the series of Bruns and

$$\Phi_{nm} = (-2)^{n+m}\,U_nV_m\,e^{-x^2-y^2}$$

in accordance to which (31) has to be modified. If it is possible to remove the origin of coordinates to the arithmetical mean by a correction of the projections for their average value, then the terms with the coeff. $A_{1.0}$ and $A_{0.1}$ vanish from (30):

If we wish to alter the scale values according to the nature of the data, we have to write everywhere, hx and $h'y$, instead of x and y, whence

$$\varepsilon^{-1} = \frac{n!m!}{2^{n+m}} \frac{\pi}{hh'}.$$

The scale factors h and h' can then be determined by putting

$$A_{2.0} \text{ en } A_{0.2} = 0$$

and the two unmixed means of the second order can be disposed of for the determination of these constants:

$$\mu_2(x) = \frac{1}{2h^2} \text{ en } \mu_2(y) = \frac{1}{2h'^2}$$

If, further, we make the axes rotate about the origin so that they coincide with the principal axes of inertia, then also $A_{1.1}$ has to be put equal to zero and the corresponding mean

$$\mu_2\,(x, y)$$

enables us to calculate the direction of the principal axes.

The series (30) then becomes:

$$\begin{aligned} u = \varepsilon^{-x^2-y^2}[A_0 + A_{3.0}U_3 + A_{2.1}U_2V_1 + A_{1.2}U_1V_2 + A_{0.3}V_3 + \\ + A_{4.0}U_4 + A_{3.1}U_3V_1 + A_{2.2}U_2V_2 + A_{1.3}U_1V_3 + \\ + A_{0.4}V_4 + \text{enz.} \end{aligned}$$

where all terms except the first represent the deviations from the normal exponential law, the terms of odd degree being a measure of the different kinds of skewness, the terms of even degree of the different kinds of symmetrical deviations.

Chemistry. — *"Equilibria in quaternary systems."* By Prof. F. A. H. SCHREINEMAKERS.

Let us first take the system with the components: *water, ethyl alcohol, methyl alcohol* and *ammonium nitrate*; we then have at the ordinary temperature one solid substance and three solvents which are miscible in all proportions so that the resulting equilibria are very simple. The equilibria occurring in this system at 30° have been investigated and are represented in the usual manner in Fig. 1; the angular points W, M, A and Z of the tetrahedron indicate the components: water, methyl alcohol, ethyl alcohol and the salt: ammonium nitrate.

The curve wa situated on the side plane WAZ represents the solutions consisting of water and ethyl alcohol and saturated with solid salt; the curve wm represents the solutions of water and