On Approximation to a Certain Type of Integrals.

By **Ragnar Frisch** (Oslo).

Introduction.

Some years ago, in the course of an analysis of upper and lower limits for incomplete moments of statistical distributions I established an elementary summation formula[1] which proved rather useful for the purpose I had in view. Subsequently the formula was generalized by professor STEFFENSEN, who showed[2] that the formula in question could be looked upon as giving the first term of an expansion in a certain type of series. Professor STEFFENSEN established re-

[1] Sur les semi-invariants et moments employés dans l'étude des distribution statistiques. Skrifter utg. av Det Norske Videnskaps-Akademi i Oslo. 1926, p. 26. Formula (42 a).

[2] On the Sum or Integral of a Product of two Functions. This Journal 1927, p. 44—70.

17

currence formulae for the coefficients of the series and computed the second, third and fourth term and the corresponding remainders[1], but did not arrive at a general, explicite expression for the coefficient of the n-th term and the corresponding remainder. A year later I found these expressions accidentally while I was working on some other problem. I also discovered the real nature of the procedure in question which proved to be a certain kind of least square fitted polynomial approximation. I did not, however, at the time publish the result. Taking the question up again later I found that the whole problem could be considerably generalized. The type of generalization in question is analogous to the generalization from polynomials to arbitrary functions.

Also in this general case can the remainder be brought on a form analogous to the form which professor STEFFENSEN had considered. This transformation is rendered possible by the general mean value theorem expressed in formula (1.1). This theorem is rather interesting in itself a part from the particular use which is presently made of it. It seems to be susceptible of a rather wide application in various sorts of practical problems involving the evaluation of a remainder. The formula in question contains for instance as a very special case the mean value formula for divided differences. My demonstration of this formula was first outlined in the Comptes rendus, Paris.[2] Shortly afterwards I received a letter from professor PÓLYA who drew my attention to the fact that the essence of the mean value theorem in question is contained in his paper in the Transactions of the American Mathematical Society, Vol. 24 (1922) (published 1924) p. 323. Formula (1.1) itself, which is the formula that has to be used in numerical applications, is not given by professor PÓLYA. But it can be deduced from PÓLYA's formula in the following way.

With the notation of Section 1 of the present paper PÓLYA's formula is

[1] loc. cit., p. 58.
[2] Académie des Sciences. Séance du 22 avril 1929.

$$(0.1)\quad \operatorname{sgn}\begin{vmatrix} a_0(x_0) \cdots a_{n-1}(x_0)\, F(x_0) \\ a_0(x_1) \cdots a_{n-1}(x_1)\, F(x_1) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_0(x_n) \cdots a_{n-1}(x_n)\, F(x_n) \end{vmatrix} = \operatorname{sgn}\begin{vmatrix} a_0\ (\xi) \cdots a_{n-1}\ (\xi)\, F\ (\xi) \\ a'_0\ (\xi) \cdots a'_{n-1}(\xi)\, F'\ (\xi) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a^{(n)}_0(\xi) \cdots a^{(n)}_{n-1}(\xi)\, F^{(n)}(\xi) \end{vmatrix}$$

sgn Z designating $+1$, 0 or -1 according as Z is positive, zero or negative and $F(x)$ being a function satisfying the same condition as $f(x)$ of Section 1. Let us insert $F(x)=f(x)-\lambda a_n(x)$ in (0.1), λ being a constant to be disposed of. Since by hypothesis the determinant $\Delta_n(x)$ befined in Section 1 does not vanish in the interval considered, the determinant

$$\begin{vmatrix} a_0(x_0) \cdots a_n(x_0) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_0(x_n) \cdots a_n(x_n) \end{vmatrix}$$

must be different from zero by virtue of PÓLYA's formula. It is therefore possible to determine λ (as a function of $x_0 \ldots x_n$) in such a way that the left hand side determinant in (0.1) vanishes. We simply have to put

$$\lambda = \begin{vmatrix} a_0(x_0) \cdots a_{n-1}(x_0)\, f(x_0) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_0(x_n) \cdots a_{n-1}(x_n)\, f(x_n) \end{vmatrix} : \begin{vmatrix} a_0(x_0) \cdots a_n(x_0) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_0(x_n) \cdots a_n(x_n) \end{vmatrix}.$$

The right hand side determinant in (0.1) must consequently also vanish, which gives $\lambda = \dfrac{F_n(\xi)}{\Delta_n(\xi)}$. Equating the two expressions for λ we get (1.1). This demonstration of (1.1) was indicated to me by professor PÓLYA in his letter already referred to.

My own demonstration proceeds on different lines and contains some points which are not discussed by PÓLYA, for instance formula (1.9) which may be looked upon as a generalization of (1.1) to the case where the numerator and denominator are constructed by two different sets of functions $a_0(x)$, $a_1(x) \ldots$ and $b_0(x)$, $b_1(x) \ldots$. I have therefore found it worth while to give my own demonstration in full. This is done in Section 1.

The mean value formula (1. 1) can be connected with a generalization of the notion of difference operation, which leads amongst others to a method of decomposing a given determinant in elementary factors. This aspect of the problem is analyzed in Section 5.

The expansion of the integral of a product which is obtained by the method of the present article has an interesting application to the damping problem discussed by professor MEIDELL. This problem is considered in Section 7.

1. A General Mean Value Theorem.

Let $a_0(x)$, $a_1(x) \dots a_n(x)$ be a sequence of real functions of the real variable x, possessing derivatives up to the order n, and further such that the WRONSKIANS

$$A_k = A_k(x) = \begin{vmatrix} a_0 & a_1 & & a_k \\ a'_0 & a'_1 & \cdots & a'_k \\ \cdot & \cdot & \cdot & \cdot \\ a_0^{(k)} & a_1^{(k)} & \cdots & a_k^{(k)} \end{vmatrix} \quad \text{where} \quad a_h^{(\varkappa)} = \frac{d^\varkappa a_h}{dx^\varkappa}$$

$$(k = 0, 1 \dots n)$$

do not vanish in a certain intervall (i). Let $f(x)$ be a function possessing derivatives up to the order n in (i). Finally let $x_0 x_1 \dots x_n$ be a system of $(n+1)$ values of x in (i). There exist at least one value $x = \xi$ in (i) such that

$$\frac{\begin{vmatrix} a_0(x_0) & \cdots & a_{n-1}(x_0) & f(x_0) \\ a_0(x_1) & \cdots & a_{n-1}(x_1) & f(x_1) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_n) & \cdots & a_{n-1}(x_n) & f(x_n) \end{vmatrix}}{\begin{vmatrix} a_0(x_0) & \cdots & a_{n-1}(x_0) & a_n(x_0) \\ a_0(x_1) & \cdots & a_{n-1}(x_1) & a_n(x_1) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_n) & \cdots & a_{n-1}(x_n) & a_n(x_n) \end{vmatrix}} = \frac{\begin{vmatrix} a_0 & \cdots & a_{n-1} & f \\ a'_0 & \cdots & a'_{n-1} & f' \\ \cdot & \cdot & \cdot & \cdot \\ a_0^{(n)} & \cdots & a_{n-1}^{(n)} & f^{(n)} \end{vmatrix}_\xi}{\begin{vmatrix} a_0 & \cdots & a_{n-1} & a_n \\ a'_0 & \cdots & a'_{n-1} & a'_n \\ \cdot & \cdot & \cdot & \cdot \\ a_0^{(n)} & \cdots & a_{n-1}^{(n)} & a_n^{(n)} \end{vmatrix}_\xi} \tag{1. 1}$$

The denominator of the left hand side of (1. 1) *is different from zero provided the values $x_0 \dots x_n$ are distinct.*

In order to establish this let us put

$$\alpha_k = \alpha_k(x) = D_{(\alpha)k} \frac{a_{k+1}}{a_0} \quad \text{and} \quad \varphi_k = \varphi_k(x) = D_{(\alpha)k} \frac{f}{a_0}$$

$D_{(\alpha)k}$ designating the operation $D_{(\alpha)k} = \frac{d}{d\alpha_{k-1}} \frac{d}{d\alpha_{k-2}} \cdots \frac{d}{d\alpha_0}$ with $D_{(\alpha)0} = 1$, i. e. $\alpha_0 = \frac{a_1}{a_0}$ and $\varphi_0 = \frac{f}{a_0}$.

Further we shall consider the determinants obtained from A_k by replacing a_k by f and by a_{k+1} respectively, i. e.

$$F_k = F_k(x) = \begin{vmatrix} a_0 & a_1 & \cdots & a_{k-1} & f \\ a'_0 & a'_1 & \cdots & a'_{k-1} & f' \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_0^{(k)} & a_1^{(k)} & \cdots & a_{k-1}^{(k)} & f^{(k)} \end{vmatrix}$$

$$A_k = A_k(x) = \begin{vmatrix} a_0 & a_1 & \cdots & a_{k-1} & a_{k+1} \\ a'_0 & a'_1 & \cdots & a'_{k-1} & a'_{k+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_0^{(k)} & a_1^{(k)} & \cdots & a_{k-1}^{(k)} & a_{k+1}^{(k)} \end{vmatrix}$$

By convention we put $F_0 = f$ and $\bar{A}_0 = a_1$.

Finally we shall consider the determinants

$$\mathfrak{G}_{n-k}(y_0 y_1 \cdots y_{n-k}) = \begin{vmatrix} \left(D_{(\alpha)k} \frac{a_k}{a_0}\right)_{y_0} & \left(D_{(\alpha)k} \frac{a_{k+1}}{a_0}\right)_{y_0} & \cdots & \left(D_{(\alpha)k} \frac{a_n}{a_0}\right)_{y_0} \\ \left(D_{(\alpha)k} \frac{a_k}{a_0}\right)_{y_1} & \left(D_{(\alpha)k} \frac{a_{k+1}}{a_0}\right)_{y_1} & \cdots & \left(D_{(\alpha)k} \frac{a_n}{a_0}\right)_{y_1} \\ \cdot & \cdot & \cdot & \cdot \\ \left(D_{(\alpha)k} \frac{a_k}{a_0}\right)_{y_{n-k}} & \left(D_{(\alpha)k} \frac{a_{k+1}}{a_0}\right)_{y_{n-k}} & \cdots & \left(D_{(\alpha)k} \frac{a_n}{a_0}\right)_{y_{n-k}} \end{vmatrix} .$$

Similarly let $\mathcal{F}_{n-k}(y_0 \cdots y_{n-k})$ be the determinant obtained from $\mathscr{A}_{n-k}(y_0 \cdots y_{n-k})$ by replacing a_n by f. In particular we have

$$\mathscr{F}_0(x) = D_{(a)\,n} \frac{f}{a_0} = \varphi_n. \tag{1. 2}$$

I shall first show that we have the explicit formulae

$$\varphi_k = \frac{F_k}{A_k} \qquad (k=0, 1 \cdots n) \tag{1. 3}$$

$$\alpha_k = \frac{\overline{A}_k}{A_k} \qquad (k=0, 1 \cdots (n-1)) \tag{1. 4}$$

$$\alpha'_k = \frac{d\alpha_k}{dx} = \frac{A_{k-1}A_{k+1}}{A_k^2} \qquad (k=0, 1 \cdots (n-1)) \tag{1. 5}$$

where by convention $A_{-1}=1$.

In fact, supposing that these formulae are exact for k, we get

$$\varphi_{k+1} = \frac{d}{d\alpha_k}\varphi_k = \frac{d}{d\alpha_k}\frac{F_k}{A_k} = \frac{1}{\alpha'_k}\frac{d}{dx}\frac{F_k}{A_k} = \frac{\begin{vmatrix} A_k & F_k \\ A'_k & F'_k \end{vmatrix}}{A_{k-1}A_{k+1}}.$$

Now, if a k rowed determinant whose elements are functions of a certain variable is differentiated with respect to this variable, the result can be written as the sum of the k determinants obtained by first differentiating the elements in the first row (leaving the other elements unchanged), next differentiating the elements in the second row (leaving the other elements unchanged) etc. If we consider a determinant of the type F_k we therefore have

$$F'_k = \frac{dF_k}{dx} = \begin{vmatrix} a_0 & \cdots & a_{k-1} & f \\ a'_0 & \cdots & a'_{k-1} & f' \\ \cdot & \cdot & \cdot & \cdot \\ a_0^{(k-1)} & \cdots & a_{k-1}^{(k-1)} & f^{(k-1)} \\ a_0^{(k+1)} & \cdots & a_{k-1}^{(k+1)} & f^{(k+1)} \end{vmatrix}$$

that is, F_k is differentiated simply by raising the differential exponent of the last row one unit. Similarly for A'_k. Both F'_k and A'_k are consequently minors in F_{k+1}. By virtue of SYLVESTER's theorem on the minors in a determinant we therefore have

$$\begin{vmatrix} A_k & F_k \\ A'_k & F'_k \end{vmatrix} = A_{k-1}F_{k+1}$$

and hence $\varphi_{k+1} = \frac{F_{k+1}}{A_{k+1}}$, which is (1. 3) for $(k+1)$. Since (1. 4) is the special case $f=a_{k+1}$ of (1. 3), (1. 4) also holds good for $(k+1)$. Consequently

$$\alpha'_{k+1} = \frac{d\alpha_{k+1}}{dx} = \frac{d}{dx}\frac{\overline{A}_{k+1}}{A_{k+1}} = \frac{\begin{vmatrix} A_{k+1} & \overline{A}_{k+1} \\ A'_{k+1} & \overline{A}'_{k+1} \end{vmatrix}}{A_{k+1}^2}.$$

Since the last expression by SYLVESTER's theorem is equal to $\frac{A_k A_{k+2}}{A_{k+1}^2}$, we see that also (1. 5) holds good for $(k+1)$. Since $\varphi_0 = \frac{F_0}{A_0}$ and $\alpha'_0 = \frac{A_{-1}A_1}{A_0^2}$ the formulae (1. 3)—(1. 5) hold good generally.

By virtue of (1. 5) all the derivatives α'_k are finite and different from zero in the intervall (i). That is, all the functions α_k are monotonic in (i).

Besides the set of functions $a_0(x) \cdots a_n(x) f(x)$, let us consider another set $b_0(x) \cdots b_n(x)\, g(x)$ satisfying the same conditions as the first set. Let β_k, γ_k, B_k, $\mathscr{G}_k$ etc. be the functions corresponding to α_k, φ_k, A_k, $\mathscr{F}_k$ etc., and let $D_{(\beta)\,k}$ be the operation corresponding to $D_{(\alpha)\,k}$.

We introduce the notations

$$P_{n-k}(y_0 \cdots y_{n-k}) = \alpha'_{k-1}(y_0)\cdot\alpha'_{k-1}(y_1)\cdots\alpha'_{k-1}(y_{n-k})$$

$$Q_{n-k}(y_0 \cdots y_{n-k}) = \beta'_{k-1}(y_0)\cdot\beta'_{k-1}(y_1)\cdots\beta'_{k-1}(y_{n-k}).$$

In what follows we shall have to consider several sets of arguments $y_0 y_1 \cdots$, $y'_0 y'_1 \cdots$, $y''_0 y''_1 \cdots$ etc. For shortnes we write

$$P_{n-k} = P_{n-k}(y_0 \cdots y_{n-k}),\quad P'_{n-k-1} = P_{n-k-1}(y'_0 \cdots y'_{n-k-1}),$$

$$P''_{n-k-2} = P_{n-k-2}(y''_0 \cdots y''_{n-k-2}) \text{ etc. And similarly}$$

$$\mathscr{F}'_{n-k-1} = \mathscr{F}_{n-k-1}(y'_0 \cdots y'_{n-k-1}) \text{ etc.}$$

Assuming $y_0 < y_1 < \cdots < y_{n-k}$ we shall consider the function of y

$$(\mathscr{F}_{n-k})_{y_0=y} \cdot \mathscr{G}_{n-k} - (\mathscr{G}_{n-k})_{y_0=y} \cdot \mathscr{F}_{n-k}.$$

This function vanishes for $y=y_0$ (because the minuend and subtrahend then become equal), and for $y=y_1$ (because two rows in the determinants then become identical). There must consequently exist a value y'_0 between y_0 and y_1 such that

$$(1.6)\qquad \left(\frac{\partial \mathscr{F}_{n-k}}{\partial y_0}\right)_{y_0=y'_0} \cdot \mathscr{G}_{n-k} = \left(\frac{\partial \mathscr{G}_{n-k}}{\partial y_0}\right)_{y_0=y'_0} \cdot \mathscr{F}_{n-k}.$$

Next, consider the function of y

$$\left(\frac{\partial \mathscr{F}_{n-k}}{\partial y_0}\right)_{\substack{y_0=y'_0\\ y_1=y}} \cdot \mathscr{G}_{n-k} - \left(\frac{\partial \mathscr{G}_{n-k}}{\partial y_0}\right)_{\substack{y_0=y'_0\\ y_1=y}} \cdot \mathscr{F}_{n-k}.$$

This function vanishes for $y=y_1$ (by virtue of (1.6)) and for $y=y_2$ (because two rows become identical). There must consequently exist a value y'_1 between y_1 and y_2, such that

$$(1.7)\qquad \left(\frac{\partial^2 \mathscr{F}_{n-k}}{\partial y_0 \partial y_1}\right)_{\substack{y_0=y'_0\\ y_1=y'_1}} \cdot \mathscr{G}_{n-k} = \left(\frac{\partial^2 \mathscr{G}_{n-k}}{\partial y_0 \partial y_1}\right)_{\substack{y_0=y'_0\\ y_1=y'_1}} \cdot \mathscr{F}_{n-k}.$$

If we continue in this way we get an equation where on the left hand side each row in the determinant $\mathscr{F}_{n-k}$, with exception of the last, is differentiated with respect to its argument, and on the right hand side each row in the determinant $\mathscr{G}_{n-k}$, with exception of the last, is differentiated with respect to its

argument. At the same time, the argument in each row (with exception of the last) is replaced by another argument situated between the original argument of the row in question and the next following. The last row in the determinants $\mathscr{F}_{n-k}$ and $\mathscr{G}_{n-k}$ are left unchanged.

The quantities in the first column of $\mathscr{F}_{n-k}$ being

$$D_{(\alpha)\,k}\frac{\alpha_k}{\alpha_0} = \frac{d}{d\alpha_{k-1}}\left(D_{(\alpha),\,k-1}\frac{\alpha_k}{\alpha_0}\right) = \frac{d\alpha_{k-1}}{d\alpha_{k-1}} = 1,$$

we see that, when the differentiations $\frac{\partial}{\partial y_0}, \frac{\partial}{\partial y_1} \cdots \frac{\partial}{\partial y_{n-k-1}}$ are performed on $\mathscr{F}_{n-k}$, there are only zeros left in the first column, except the last element in the first column, which is equal to unity. The determinant in question therefore reduces to an $(n-k)$ rowed determinant, which does not contain explicitely the argument y_{n-k} (although the effect of this argument is still present because the arguments y' depend on y_{n-k}). In the $(n-k)$ rowed determinant thus obtained let us take the quantity $\alpha'_k(y'_0) \neq 0$ out of the first row, the quantity $\alpha'_k(y'_1) \neq 0$ out of the second row ... and finally the quantity $\alpha'_k(y'_{n-k-1}) \neq 0$ out of the last row. A part from the sign $(-)^{n-k}$, the resulting determinant is nothing else than

$$\mathscr{F}'_{n-k-1} = \mathscr{F}_{n-k-1}(y'_0 \cdots y'_{n-k-1}).$$

Taking the quantity $\beta'_k(y'_0)$ out of the first row, the quantity $\beta'_k(y'_1)$ out of the second row, etc. on the right hand side we get the equation

$$(1.8)\qquad P'_{n-k-1} \cdot \mathscr{F}'_{n-k-1} \cdot \mathscr{G}_{n-k} = Q'_{n-k-1} \cdot \mathscr{G}'_{n-k-1} \cdot \mathscr{F}_{n-k}.$$

Next, let us consider the function of y

$$P'_{n-k-1} \cdot (\mathscr{F}'_{n-k-1})_{y'_0=y} \cdot \mathscr{G}_{n-k} - Q'_{n-k-1} \cdot (\mathscr{G}'_{n-k-1})_{y'_0=y} \cdot \mathscr{F}_{n-k}.$$

This function vanishes for $y=y'_0$ (by virtue of (1.8)) and for $y=y'_1$ (because two rows become identical). There must consequently exist a value y''_0 between y'_0 and y'_1 such that

$$P'_{n-k-1}\cdot\left(\frac{\partial \mathcal{F}'_{n-k-1}}{\partial y'_0}\right)_{y'_0=y''_0}\cdot\mathcal{G}_{n-k}=Q'_{n-k-1}\cdot\left(\frac{\partial \mathcal{G}_{n-k-1}}{\partial y'_0}\right)_{y'_0=y''_0}\cdot\mathcal{F}_{n-k}.$$

If we continue in this way we get

$$P'_{n-k-1}\cdot P''_{n-k-2}\cdot\mathcal{F}''_{n-k-2}\cdot\mathcal{G}_{n-k}=$$
$$=Q'_{n-k-1}\cdot Q''_{n-k-2}\cdot\mathcal{G}''_{n-k-2}\cdot\mathcal{F}_{n-k}.$$

Repeating the process we get

$$P'_{n-k-1}\cdot P''_{n-k-2}\cdots P_0^{(n-k)}\cdot\mathcal{F}_0^{(n-k)}\cdot\mathcal{G}_{n-k}=$$
$$=Q'_{n-k-1}\cdot Q''_{n-k-2}\cdots Q_0^{(n-k)}\cdot\mathcal{G}_0^{(n-k)}\cdot\mathcal{F}_{n-k}.$$

In particular if we put $k=0$ and make a slight change in the notation of the arguments, we finally get by (1. 2) and (1. 3) the following formula:

$$(1.9)\quad B_n(\xi)\cdot P_0P_1\cdots$$
$$\cdots P_{n-1}\cdot\begin{vmatrix}a_0\,(\xi)\cdots a_{n-1}\,(\xi)\,f(\xi)\\ a'_0\,(\xi)\cdots a'_{n-1}\,(\xi)\,f'\,(\xi)\\ \cdots\cdots\cdots\\ a_0^{(n)}\,(\xi)\cdots a_{n-1}^{(n)}\,(\xi)\,f^{(n)}\,(\xi)\end{vmatrix}\cdot\begin{vmatrix}b_0\,(x_0)\cdots b_{n-1}(x_0)\,g\,(x_0)\\ b_0\,(x_1)\cdots b_{n-1}(x_1)\,g\,(x_1)\\ \cdots\cdots\cdots\\ b_0\,(x_n)\cdots b_{n-1}(x_n)\,g\,(x_n)\end{vmatrix}=$$

$$A_n(\xi)\cdot Q_0Q_1\cdots$$
$$\cdots Q_{n-1}\cdot\begin{vmatrix}b_0\,(\xi)\cdots b_{n-1}\,(\xi)\,g\,(\xi)\\ b'_0\,(\xi)\cdots b'_{n-1}\,(\xi)\,g'\,(\xi)\\ \cdots\cdots\cdots\\ b_0^{(n)}\,(\xi)\cdots b_{n-1}^{(n)}\,(\xi)\,g^{(n)}\,(\xi)\end{vmatrix}\cdot\begin{vmatrix}a_0\,(x_0)\cdots a_{n-1}\,(x_0)\,f\,(x_0)\\ a_0\,(x_1)\cdots a_{n-1}\,(x_1)\,f\,(x_1)\\ \cdots\cdots\cdots\\ a_0\,(x_n)\cdots a_{n-1}\,(x_n)\,f\,(x_n)\end{vmatrix}$$

where

$$P_0=\alpha'_{n-1}(\xi)$$
$$P_1=\alpha'_{n-2}(\xi_{10})\cdot\alpha'_{n-2}(\xi_{11})$$
$$P_2=\alpha'_{n-3}(\xi_{20})\cdot\alpha'_{n-3}(\xi_{21})\cdot\alpha'_{n-3}(\xi_{22})$$
$$\cdots\cdots\cdots\cdots$$

$$Q_0=\beta'_{n-1}(\xi)$$
$$Q_1=\beta'_{n-2}(\xi_{10})\cdot\beta'_{n-2}(\xi_{11})$$
$$Q_2=\beta'_{n-3}(\xi_{20})\cdot\beta'_{n-3}(\xi_{21})\cdot\beta'_{n-3}(\xi_{22})$$
$$\cdots\cdots\cdots\cdots$$

ξ being a value between ξ_{10} and ξ_{11}, ξ_{10} a value between ξ_{20} and ξ_{21}, and ξ_{11} a value between ξ_{21} and ξ_{22}, and so on, finally $\xi_{n-1,0}$, $\xi_{n-1,1}\cdots\xi_{n-1,n-1}$ a set of values separating the values $x_0x_1\cdots x_n$.

If we put $b_k(x)=x^k$, $g(x)=x^n$ and $f(x)=a_n(x)$, that is

$$\beta_k(x)=\frac{0!\;1!\cdots(k-1)!\;(k+1)!}{0!\;1!\cdots(k-1)!\;k!}\,x=(k+1)\,x,$$

and consequently $\beta'_k(x)=k+1$, we see from (1. 9) that

$$\begin{vmatrix}a_0(x_0)\cdots a_n(x_0)\\ \cdots\cdots\\ a_0(x_n)\cdots a_n(x_n)\end{vmatrix}$$

cannot vanish if the arguments $x_0\cdots x_n$ are distinct. This being so, we get (1. 1) by putting $b_k(x)=a_k(x)$ and $g(x)=a_n(x)$ in (1. 9).

Formula (1. 1) therefore holds good if the arguments are distinct, and by a limiting process we see that it must also be true if some of the arguments coincide, the numerator and the denominator of the left hand side of (1. 1) then taking on a form with one or more rows differentiated a certain number of times.

The essential point in the formula (1. 1) is that it is sufficient to consider a single value ξ. If we introduce $(n+1)$ values $\xi_0\,\xi_1\cdots\xi_n$, we can formulate the following proposition:

Let $a_0(x)\cdots a_n(x)$ and $b_0(x)\cdots b_n(x)$ be two sequences of functions satisfying the condition

$$A_0\neq 0,\ A_1\neq 0\cdots A_{n-1}\neq 0\ \text{and}\ B_0\neq 0,\ B_1\neq 0\cdots B_n\neq 0$$

in the interval (i).

If $x_0 x_1 \cdots x_n$ is a system of values in (i), there also exist a system $\xi_0 \xi_1 \cdots \xi_n$ in (i) such that

$$(1.10) \qquad \frac{\begin{vmatrix} a_0(x_0) & \cdots & a_n(x_0) \\ a_0(x_1) & \cdots & a_n(x_1) \\ \cdot & \cdot & \cdot \\ a_0(x_n) & \cdots & a_n(x_n) \end{vmatrix}}{\begin{vmatrix} b_0(x_0) & \cdots & b_n(x_0) \\ b_0(x_1) & \cdots & b_n(x_1) \\ \cdot & \cdot & \cdot \\ b_0(x_n) & \cdots & b_n(x_n) \end{vmatrix}} = \frac{\begin{vmatrix} a_0(\xi_0) & \cdots & a_n(\xi_0) \\ a'_0(\xi_1) & \cdots & a'_n(\xi_1) \\ \cdot & \cdot & \cdot \\ a_0^{(n)}(\xi_n) & \cdots & a_n^{(n)}(\xi_n) \end{vmatrix}}{\begin{vmatrix} b_0(\xi_0) & \cdots & b_n(\xi_0) \\ b'_0(\xi_1) & \cdots & b'_n(\xi_1) \\ \cdot & \cdot & \cdot \\ b_0^{(n)}(\xi_n) & \cdots & b_n^{(n)}(\xi_n) \end{vmatrix}}$$

In fact, let

$$\mathcal{A} = \mathcal{A}(x_0 x_1 \cdots x_n) = \begin{vmatrix} a_0(x_0) & \cdots & a_n(x_0) \\ a_0(x_1) & \cdots & a_n(x_1) \\ \cdot & \cdot & \cdot \\ a_0(x_n) & \cdots & a_n(x_n) \end{vmatrix}$$

and let $\mathcal{B} = \mathcal{B}(x_0 x_1 \cdots x_n)$ be the analogous determinant for the sequence $b_0(x), b_1(x) \cdots$.

Let $x_0 \cdots x_n$ be a system of $(n+1)$ distinct values in the interval (i) and consider the function of x

$$(\mathcal{A})_{x_0=x} \cdot \mathcal{B} - (\mathcal{B})_{x_0=x} \cdot \mathcal{A}.$$

This function vanishes for $x=x_0, x_1 \cdots x_n$. Its n-th derivative consequently vanishes in a point ξ_n. That is, we have

$$\left(\frac{\partial^n \mathcal{A}}{\partial x_0^n}\right)_{x_0=\xi_n} \cdot \mathcal{B} = \left(\frac{\partial^n \mathcal{B}}{\partial x_0^n}\right) \cdot \mathcal{A}.$$

Next, consider the function of x

$$\left(\frac{\partial^n \mathcal{A}}{\partial x_0^n}\right)_{\substack{x_0=\xi_n \\ x_1=x}} \cdot \mathcal{B} - \left(\frac{\partial^n \mathcal{B}}{\partial x_0^n}\right)_{\substack{x_0=\xi_n \\ x_1=x}} \cdot \mathcal{A}.$$

This function vanishes for $x=x_1, x_2 \cdots x_n$. Its $(n-1)$-th derivative consequently vanishes in a point ξ_{n-1}. If we continue

in this way, we get (1.10). We see that we may put ξ_0 in this formula equal to any of the arguments $x_0 \cdots x_n$. The formula (1.10) contains as a special case the formula of SCHWARZ.[1]

2. The Expansion of the Integral of a Product in a Series containing a Sequence of Arbitrary Functions.

Let $f(x)$ and $g(x)$ be two functions of the variable x. Our problem is to evaluate the integral

$$(2.1) \qquad \int_\alpha^\beta f(x)\, g(x)\, dx.$$

In order to do so we shall take as our point of departure TCHEBYCHEFF's idea of approximating the mean of a product by the product of the means. This amounts to putting

$$(2.2) \qquad \int_\alpha^\beta f(x)\, g(x)\, dx = \frac{1}{\beta-\alpha} \int_\alpha^\beta f(x)\, dx \cdot \int_\alpha^\beta g(x)\, dx + R.$$

The explicite expression for the remainder R in this formula is

$$(2.3) \qquad R = \frac{1}{2(\beta-\alpha)} \int_\alpha^\beta dx \int_\alpha^\beta dy\, (f(x)-f(y))\,(g(x)-g(y)).$$

The formula (2.3) is proved by simply multiplying out the product $(f(x)-f(y))(g(x)-g(y))$ and integrating each term separately. The formula (2.2) with the remainder (2.3) is the elementary summation (integration) formula referred to in the Introduction. From this formula we immediately deduce TCHEBYCHEFF's inequality for integrals.

[1] H. A. SCHWARZ: Verallgemeinerung eines analytischen Fundamentalsatzes. Ges. Mathem. Abh. 2, 1880, p. 301.

The remainder (2. 3) can also be written in the form

$$R = \int_{\alpha}^{\beta} (f(x) - f_0)(g(x) - g_0)\, dx \tag{2. 4}$$

where $f_0 = \dfrac{1}{\beta - \alpha} \displaystyle\int_{\alpha}^{\beta} f(x)\, dx$ is the mean of $f(x)$ over the interval (α, β) and similarly g_0 is the mean of $g(x)$. The formula (2. 4) can also be proved by simply multiplying out the integrand and integrating each term separately.

If we introduce the deviation factors p and q defined by

$$p^2 = \int_{\alpha}^{\beta} (f(x) - f_0)^2\, dx$$

$$q^2 = \int_{\alpha}^{\beta} (g(x) - g_0)^2\, dx$$

and the coefficient of correlation between the two functions $f(x)$ and $g(x)$ over the interval (α, β)

$$r = \frac{1}{pq} \int_{\alpha}^{\beta} (f(x) - f_0)(g(x) - g_0)\, dx$$

we can write (2. 4) in the form

$$R = pqr. \tag{2. 5}$$

This shows that Tchebycheff's inequality for integrals can be generalized into the following proposition: *The mean of a product is greater or less than the product of the means according as the two variables are positively or negatively correlated. If they are uncorrelated the mean of the product is equal to the product of the means.* Since r is a coefficient of correlation,

we always have $|r| \leqq 1$, which shows that we have the error limit

$$|R| \leqq pq \tag{2. 6}$$

and therefore a fortiori

$$|R| \leqq (\beta - \alpha) AB \tag{2. 7}$$

where A is the maximum deviation of $f(x)$ from its mean, and B the maximum deviation of $g(x)$ from its mean, the deviations being taken regardless of sign.

The formula (2. 2) can be interpreted as the formula obtained by *fitting a constant to one of the functions, say $g(x)$ by least squares over the interval* (α, β), and evaluating (2. 1) under the assumption that the function in question is equal to this constant over the domain of integration. The result is evidently independent of which one of the two functions is replaced by a least square determined constant. The procedure is therefore symmetric in the two functions f and g.

A natural generalization of this idea would be *to replace one of the two functions f and g by a least square fitted polynomial over the domain of integration.* It will presently appear that the result thus obtained is independent of which one of the two functions is replaced by a polynomial. The procedure is therefore symmetric in the two functions f and g.

A still more general procedure would be to replace one of the two functions f and g by a linear form in a set of given linearly independent functions $a_0(x), a_1(x) \cdots$, the constants of the form to be determined by least squares over the interval (α, β). This is the procedure which shall now be considered.

We introduce the numerical coefficients

$$a_{ij} = \int_{\alpha}^{\beta} a_i(x)\, a_j(x)\, dx. \tag{2. 8}$$

Since a_{ij} is the product moment (over the intervall (α, β)) of the two functions $a_i(x)$ and $a_j(x)$, the symmetric matrix

$$(2.9)\qquad (a_{ij}) = \begin{pmatrix} a_{00} & a_{01} & \cdots & a_{0k} \\ a_{10} & a_{11} & \cdots & a_{1k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k0} & a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

is a $(k+1)$ rowed GRAM-ian matrix whose determinant

$$(2.10)\qquad A^{(k+1)} = \begin{vmatrix} a_{00} & a_{01} & \cdots & a_{0k} \\ a_{10} & a_{11} & \cdots & a_{1k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k0} & a_{k1} & \cdots & a_{kk} \end{vmatrix}$$

can be expressed in the form[1]

$$(2.11)\qquad (k+1)!\,A^{(k+1)} = \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots \int_\alpha^\beta dx_k \begin{vmatrix} a_0(x_0) & \cdots & a_k(x_0) \\ a_0(x_1) & \cdots & a_k(x_1) \\ \cdot & \cdot & \cdot \\ a_0(x_k) & \cdots & a_k(x_k) \end{vmatrix}^2 .$$

Since the integrand in (2.11) is symmetric in all the variables, we can drop the left hand side prefactor $(k+1)!$ if the domain of integration is changed to $\int_\alpha^\beta dx_0 \int_{x_0}^\beta dx_1 \cdots \int_{x_{k-1}}^\beta dx_k$ that is $\alpha \leqq x_0 \leqq \cdots \leqq x_k \leqq \beta$ (evidently it does not restrict generality if we assume $\alpha < \beta$).

If the functions $a_0(x)$, $a_1(x) \cdots$ are linearly independent, the integrand of (2.11) cannot vanish identically in $x_0 \cdots x_k$. The determinant $A^{(k+1)}$ must therefore be positive, not zero.

The elements of the reciprocal of the matrix (2.9) will be denoted a^*_{ij}. I. e. a^*_{ij} is the quantity which is obtained by leaving out the i-th column and the j-th row from the determinant $A^{(k+1)}$, then dividing the k rowed determinant obtained by $A^{(k+1)}$, and finally multiplying by $(-)^{i+j}$.

[1] See for instance KOWALEWSKI; Determinantentheorie. Leipzig 1909 p. 321.

The quantities

$$(2.12)\qquad m_\varkappa = \int_\alpha^\beta a_\varkappa(x) f(x)\, dx \quad \text{and} \quad n_\varkappa = \int_\alpha^\beta a_\varkappa(x)\, g(x)\, dx$$

will be called the (general) moments of f and g respectively, taken with respect to the functions $a_\varkappa$ over the interval (α, β). For $a_\varkappa(x) = (x-\alpha)^\varkappa$, $m_\varkappa$ and $n_\varkappa$ are the ordinary power moments taken about the beginning of the interval of integration.

If the linear form

$$(2.13)\qquad G(x) = g_0 \cdot a_0(x) + \cdots + g_k \cdot a_k(x) = \Sigma_\varkappa\, g_\varkappa \cdot a_\varkappa(x)$$

where the $g_\varkappa$ are constants, is fitted to $g(x)$ by least squares over the interval (α, β) the coefficients $g_\varkappa$ are determined by the linear system

$$\int_\alpha^\beta a_i(x) \cdot (g(x) - \Sigma_\varkappa g_\varkappa \cdot a_\varkappa(x))\, dx = 0$$
$$(i = 0, 1 \cdots k)$$

that is

$$(2.14)\qquad \Sigma_\varkappa\, a_{i\varkappa} \cdot g_\varkappa = n_i \qquad (i = 0, 1 \cdots k).$$

Since $A^{(k+1)} \neq 0$ the system (2.14) always has a uniquely determined solution, namely

$$(2.15)\qquad g_i = \Sigma_\varkappa\, a^*_{i\varkappa} \cdot n_\varkappa \qquad (i = 0, 1 \cdots k).$$

If we insert $G(x)$ instead of $g(x)$ in (2.1) and perform the integration, we get

$$(2.16)\qquad P_k = \Sigma_{ij} m_i a^{*}_{ij} n_j = -\begin{vmatrix} 0 & n_0 & n_1 & \cdots & n_k \\ m_0 & a_{00} & a_{01} & \cdots & a_{0k} \\ m_1 & a_{10} & a_{11} & \cdots & a_{1k} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ m_k & a_{k0} & a_{k1} & \cdots & a_{kk} \end{vmatrix} : \begin{vmatrix} a_{00} & a_{01} & \cdots & a_{0k} \\ a_{10} & a_{11} & \cdots & a_{1k} \\ \cdot & \cdot & \cdots & \cdot \\ a_{k0} & a_{k1} & \cdots & a_{kk} \end{vmatrix}$$

$$(k = 0, 1, 2 \cdots).$$

Since the right hand side of the last formula is symmetric in the functions $f(x)$ and $g(x)$ we see that it does not matter which one of the functions $f(x)$ and $g(x)$ is replaced by a least square fitted linear form. That is, we have

$$(2.17)\qquad \int_{\alpha}^{\beta} f(x)\, G(x)\, dx = \int_{\alpha}^{\beta} F(x)\, g(x)\, dx = P_k$$

$F(x) = \Sigma_{\varkappa} f_{\varkappa} \cdot a_{\varkappa}(x)$ being the linear form which is fitted by least squares to $f(x)$ over the interval (α, β).

If the functions $a_0(x)$, $a_1(x) \cdots$ are orthogonal, i. e. if

$$\int_{\alpha}^{\beta} a_i(x)\, a_j(x)\, dx = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}$$

the expression for P_k reduces to

$$(2.18)\qquad P_k = m_0 n_0 + m_1 n_1 + \cdots + m_k n_k .$$

The expression (2. 16) respectively (2. 18) might be considered as the k-th order approximation to the integral (2. 1). That is, we may put

$$(2.19)\qquad \int_{\alpha}^{\beta} f(x)\, g(x)\, dx = P_k + R_k$$

P_k being given by (2. 16), respectively (2. 18) and R_k being the remainder after $(k+1)$ terms. The formula (2. 18) which holds good when the functions $a_0(x)$, $a_1(x) \cdots$ are orthogonal,

has the practical advantage that the coefficients of the series thus obtained, are independent of the number of terms retained (i. e. independent of k). In practice one would of course choose functions $A_0(x)$, $A_1(x) \cdots$ which are orthogonal over some standard interval, say (0, 1), and then put

$$a_i(x) = \frac{1}{\sqrt{\beta - \alpha}}\, A_i\left(\frac{x - \alpha}{\beta - \alpha}\right),$$

so that the functions $a_0(x)$, $a_1(x) \cdots$ are orthogonal over the interval (α, β).

I now proceed to an analysis of the remainder R_k.

3. The Explicite Expression for the Remainder.

Let $F(x) = \sum_{i=0}^{k} f_i \cdot a_i(x)$ and $G(x) = \sum_{j=0}^{k} g_j \cdot a_j(x)$ be the linear forms in the $(k+1)$ functions $a_0(x) \cdots a_k(x)$ (e. g. the k-th degree polynomials) which are fitted by least squares to $f(x)$ and $g(x)$ respectively over the interval (α, β). Further let $\varphi(x) = f(x) - F(x)$ and $\gamma(x) = g(x) - G(x)$ be the deviations of $f(x)$ and $g(x)$ respectively from these linear forms.

Finally let p and q be the deviation factors defined by

$$p^2 = \int_{\alpha}^{\beta} \varphi^2(x)\, dx \quad \text{and} \quad q^2 = \int_{\alpha}^{\beta} \gamma^2(x)\, dx$$

and let

$$r = \frac{1}{pq} \int_{\alpha}^{\beta} \varphi(x)\, \gamma(x)\, dx$$

be the coefficient of correlation between the deviation of $f(x)$ and the deviation of $g(x)$ over the interval (α, β).

With these notations the explicite expression for the remainder R_k of the expansion (2. 19) *is*

$$R_k = \int_\alpha^\beta \varphi(x)\,\gamma(x)\,dx \tag{3. 1}$$

which can also be written

$$R_k = pqr. \tag{3. 2}$$

In order to prove this we notice that the integral (3. 1) by virtue of (2. 17) can be written as

$$\int_\alpha^\beta f(x)\,g(x)\,dx - 2\,P_k + \int_\alpha^\beta F(x)\,G(x)\,dx. \tag{3. 3}$$

Since the linear forms $F(x)$ and $G(x)$ are least square fitted, the moments of these forms, taken with respect to the functions $a_\varkappa$, i. e.

$$M_\varkappa = \int_\alpha^\beta a_\varkappa(x)\,F(x)\,dx$$

$$N_\varkappa = \int_\alpha^\beta a_\varkappa(x)\,G(x)\,dx$$

up to the order $\varkappa = k$ are equal to the moments $m_\varkappa$ and $n_\varkappa$ of the function $f(x)$ and $g(x)$ respectively. Fitting a linear form in a set of given functions by moments (taken with respect to these functions) or by least squares, namely amounts to the same. In fact, we have

$$M_\varkappa = \int_\alpha^\beta a_\varkappa(x)\,\Sigma_i f_i\,a_i(x)\,dx = \int_\alpha^\beta a_\varkappa(x)\,\Sigma_i\,\Sigma_j\,a_i(x)\,a_{ij}^*\,m_j\,dx =$$

$$= \Sigma_j\,m_j\,\Sigma_i\,a_{ij}^* \int_\alpha^\beta a_i(x)\,a_\varkappa(x)\,dx = \Sigma_j\,m_j\,\Sigma_i\,a_{i\varkappa}\,a_{ij}^*.$$

Now, if $0 \leqq \varkappa \leqq k$, we have $\Sigma_i\,a_{i\varkappa}\,a_{ij}^* \begin{cases} 0 & (j \neq \varkappa) \\ 1 & (j = \varkappa), \end{cases}$ that is $M_\varkappa = m_\varkappa$. Similarly we get $N_\varkappa = n_\varkappa$ if $0 \leqq \varkappa \leqq k$.

If the numerical coefficients $g_\varkappa$ in the linear form $G(x)$ are expressed in terms of the moments $n_\varkappa$, the last integral of (3. 3) will therefore be

$$\int_\alpha^\beta F(x)\cdot\Sigma_i\,g_i\cdot a_i(x)\,dx = \int_\alpha^\beta F(x)\cdot\Sigma_{ij}\,a_{ij}^*\,n_j\,a_i(x)\,dx =$$

$$= \Sigma_{ij}\,M_i\,a_{ij}^*\,n_j = \Sigma_{ij}\,m_i\,a_{ij}^*\,n_j = P_k.$$

This establishes (3. 1).

The formula (3. 1) may be looked upon as a generalization of the remainder expression (2. 4). I now proceed to establish a lemma by which it is possible to transform (3. 1) in such a way as to obtain an explicite remainder expression which is a generalization of (2. 3).

Lemma: Let $\varphi(x)$ and $\gamma(x)$ be two functions, at least one of which has all its moments (taken with respect to the functions $a_0(x)$, $a_1(x)\cdots$) up to the order k equal to zero. That is, we have either

$$\int_\alpha^\beta a_\varkappa(x)\,\varphi(x)\,dx = 0 \quad \text{or} \quad \int_\alpha^\beta a_\varkappa(x)\,\gamma(x)\,dx = 0 \quad \text{or both.}$$

$$(\varkappa = 0,\ 1 \cdots k).$$

Then we have the identity

$$(\varkappa+1)!\,A^{(\varkappa)} \int_\alpha^\beta \varphi(x)\,\gamma(x)\,dx = \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots \tag{3. 4}$$

$$\cdots \int_\alpha^\beta dx_\varkappa \begin{vmatrix} a_0(x_0) & \cdots & a_{\varkappa-1}(x_0) & \varphi(x_0) \\ a_0(x_1) & \cdots & a_{\varkappa-1}(x_1) & \varphi(x_1) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_\varkappa) & \cdots & a_{\varkappa-1}(x_\varkappa) & \varphi(x_\varkappa) \end{vmatrix} \cdot \begin{vmatrix} a_0(x_0) & \cdots & a_{\varkappa-1}(x_0) & \gamma(x_0) \\ a_0(x_1) & \cdots & a_{\varkappa-1}(x_1) & \gamma(x_1) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_\varkappa) & \cdots & a_{\varkappa-1}(x_\varkappa) & \gamma(x_\varkappa) \end{vmatrix}$$

where $\varkappa$ *is any of the integers* $\varkappa = 1, 2 \cdots (k+1)$ *and* $A^{(\varkappa)}$ *is a* $\varkappa$*-rowed determinant of the form* (2. 10).

In order to prove (3. 4) we notice that we have

$$\begin{vmatrix} a_0(x_0) \cdots a_{\varkappa-1}(x_0)\,\varphi(x_0) \\ a_0(x_1) \cdots a_{\varkappa-1}(x_1)\,\varphi(x_1) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_0(x_\varkappa) \cdots a_{\varkappa-1}(x_\varkappa)\,\varphi(x_\varkappa) \end{vmatrix} = \sum_{r=0}^{\varkappa} (-)^{\varkappa+r} D_r\, \varphi(x_r)$$

where D_r denotes the minor corresponding to the element $\varphi(x_r)$ of the last row. Similarly for the determinant where the function in the last column is $\gamma(x)$. The right hand side integral in (3. 4) is therefore equal to

$$\sum_{r=0}^{\varkappa} \sum_{s=0}^{\varkappa} (-)^{r+s} \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots \int_\alpha^\beta dx_\varkappa\ \varphi(x_r)\,\gamma(x_s)\,D_r\,D_s.$$

The general term of this double sum must vanish if $r \neq s$. In fact, if $r \neq s$, the product $\gamma(x_s)\,D_r$ is independent of x_r, and D_s (considered as a function of x_r) is a linear form in $a_0(x_r) \cdots a_{\varkappa-1}(x_r)$, where $\varkappa - 1 \leqq k$. Performing the integration over x_r we see that if $\varphi(x)$ has vanishing moments (with respect to the functions $a_0(x)$, $a_1(x) \cdots$) up to the order k, the term in question must vanish. Similar argument if $\gamma(x)$ has vanishing moments up to the order k. The right hand side of (3. 4) therefore reduces to

$$\sum_{r=0}^{\varkappa} \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots \int_\alpha^\beta dx_\varkappa\ \varphi(x_r)\,\gamma(x_r)\,D_r^2.$$

Performing here first the integration over x_r and next the integration over the remaining $\varkappa$ variables we get by (2. 11)

$$\sum_{r=0}^{\varkappa} \left\{ \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots \int_\alpha^\beta dx_\varkappa \cdot D_r^2 \right\} \cdot \int_\alpha^\beta \varphi(x)\,\gamma(x)\,dx =$$
$$[\text{except } dx_r]$$

$$= \sum_{r=0}^{\varkappa} \varkappa!\, A^{(\varkappa)} \int_\alpha^\beta \varphi(x)\,\gamma(x)\,dx = (\varkappa+1)!\, A^{(\varkappa)} \int_\alpha^\beta \varphi(x)\,\gamma(x)\,dx$$

which is the left hand side of (3. 4).

Since the integrand to the right in (3. 4) is symmetric in all the variables, we may drop the left hand side prefactor $(\varkappa+1)!$ if the domain of integration is changed to $\int_\alpha^\beta dx_0 \cdot \int_{x_0}^\beta dx_1 \cdots \int_{x_{\varkappa-1}}^\beta dx_\varkappa$, that is $\alpha \leqq x_0 \leqq \cdots \leqq x_\varkappa \leqq \beta$.

Choosing different values of $\varkappa$ in (3. 4) we get different expressions for the remainder R_k. The case $\varkappa = k+1$ is of special interest. We namely have

$$(3.5)\quad \begin{vmatrix} a_0(x_0) & \cdots a_k(x_0) & \varphi(x_0) \\ a_0(x_1) & \cdots a_k(x_1) & \varphi(x_1) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ a_0(x_{k+1}) & \cdots a_k(x_{k+1}) & \varphi(x_{k+1}) \end{vmatrix} = \begin{vmatrix} a_0(x_0) & \cdots a_k(x_0) & f(x_0) \\ a_0(x_1) & \cdots a_k(x_1) & f(x_1) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ a_0(x_{k+1}) & \cdots a_k(x_{k+1}) & f(x_{k+1}) \end{vmatrix}.$$

This simply follows from the fact that, $F(x) = f(x) - \varphi(x)$ being a linear form in $a_0(x) \cdots a_k(x)$, we have

$$\begin{vmatrix} a_0(x_0) & \cdots a_k(x_0) & F(x_0) \\ a_0(x_1) & \cdots a_k(x_1) & F(x_1) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ a_0(x_{k+1}) & \cdots a_k(x_{k+1}) & F(x_{k+1}) \end{vmatrix} = 0$$

identically in $x_0 \cdots x_{k+1}$.

The formula (3. 5) also holds good, of course, if φ is replaced by γ and f by g. This gives the following explicite expression for the remainder

$$(3.6) \qquad R_k = \frac{1}{(k+2)!\,\Delta^{(k+1)}} \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots$$

$$\cdots \int_\alpha^\beta dx_{k+1} \begin{vmatrix} a_0(x_0) & \cdots & a_k(x_0) & f(x_0) \\ a_0(x_1) & \cdots & a_k(x_1) & f(x_1) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_{k+1}) & \cdots & a_k(x_{k+1}) & f(x_{k+1}) \end{vmatrix} \cdot \begin{vmatrix} a_0(x_0) & \cdots & a_k(x_0) & g(x_0) \\ a_0(x_1) & \cdots & a_k(x_1) & g(x_1) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_{k+1}) & \cdots & a_k(x_{k+1}) & g(x_{k+1}) \end{vmatrix}$$

Collecting the results, we see that if $a_0(x)$, $a_1(x) \cdots$ is a sequence of arbitrary functions we have the identity

$$\int_\alpha^\beta f(x)\, g(x)\, dx = P_k + R_k$$

where P_k is given by (2. 16) and R_k by (3. 6), the moments $m_\varkappa$ and $n_\varkappa$ and the coefficients a_{ij} being defined by (2. 12) and (2. 8) respectively, and $\Delta^{(k+1)}$ being defined by (2. 10).

4. Evaluation of the Remainder.

The evaluation of the remainder R_k may be attacked in two different ways, either by taking the formula (3. 2) or the formula (3. 6) as a point of departure. The first formula gives an interesting intuitive interpretation of the circumstances on which the closeness of the approximation depends: The approximation is all the better, the smaller the *deviation* of each function from its least square fitted linear form, and the smaller the *correlation* between these deviations. If Δ is interpreted as the maximum absolute deviation of f from its least square fitted form, and similarly for B, we see from (3. 2) that (2. 7) holds good in the general case.

In the present Section we shall be concerned with the evaluation obtained from (3. 6). The formula (3. 2), or rather the special case (2. 5), is utilized for the purpose which we have in view in Section 7.

We introduce the notation

$$(4.1) \qquad \Delta(x_0 x_1 \cdots x_\varkappa) = \begin{vmatrix} a_0(x_0) & \cdots & a_\varkappa(x_0) \\ \cdot & \cdot & \cdot \\ a_0(x_\varkappa) & \cdots & a_\varkappa(x_\varkappa) \end{vmatrix}$$

$$(4.2) \qquad F(x_0 x_1 \cdots x_\varkappa) = \begin{vmatrix} a_0(x_0) & \cdots & a_{\varkappa-1}(x_0) f(x_0) \\ \cdot & \cdot & \cdot \\ a_0(x_\varkappa) & \cdots & a_{\varkappa-1}(x_\varkappa) f(x_\varkappa) \end{vmatrix}.$$

In particular $F(x_0) = f(x_0)$. Similarly $G(x_0 x_1 \cdots x_\varkappa)$ denotes the determinant obtained by replacing $f(x)$ by $g(x)$ in (4. 2).

Further we shall consider the ratios

$$(4.3) \qquad f(x_0 x_1 \cdots x_\varkappa) = \frac{F(x_0 x_1 \cdots x_\varkappa)}{\Delta(x_0 x_1 \cdots x_\varkappa)}$$

and similarly $g(x_0 x_1 \cdots x_\varkappa)$ if f is replaced by g in (4. 3). In the next Section I shall show that the ratios $f(x_0 \cdots x_\varkappa)$ and $g(x_0 \cdots x_\varkappa)$ may be looked upon as a generalization of the notion of divided difference.

In order that we may consider the ratios $f(x_0 \cdots x_\varkappa)$ and $g(x_0 \cdots x_\varkappa)$ up to the order $\varkappa = k+1$ we introduce a new arbitrary function $a_{k+1}(x)$ satisfying the same conditions as the other functions in the set $a_0(x) \cdots a_k(x)$. This functions $a_{k+1}(x)$ might be choosen independently of the functions $a_0(x) \cdots a_k(x)$, and with the only purpose of »squizing» the remainder. This is possible because the principal term P_k does not involve $a_{k+1}(x)$.

Introducing the ratios $f(x_0 \cdots x_\varkappa)$ and $g(x_0 \cdots x_\varkappa)$ in (3. 6) we get

$$(4.4) \qquad R_k = \frac{1}{(k+2)!\,\Delta^{(k+1)}} \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots$$

$$\cdots \int_\alpha^\beta dx_{k+1} f(x_0 \cdots x_{k+1})\, g(x_0 \cdots x_{k+1})\, \Delta^2(x_0 \cdots x_{k+1})$$

$A^2(x_0 \cdots x_{k+1})$ being non negative, there exists a point $(\xi_0 \cdots \xi_{k+1})$ in the domain of integration such that

$$R_k = \frac{f(\xi_0 \cdots \xi_{k+1})\, g(\xi_0 \cdots \xi_{k+1})}{(k+2)!\ A^{(k+1)}} \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots \int_\alpha^\beta dx_{k+1}\, A^2(x_0 \cdots x_{k+1}).$$

That is, by virtue of (2. 11)

$$R_k = \frac{A^{(k+2)}}{A^{(k+1)}} f(\xi_0 \cdots \xi_{k+1})\, g(\xi_0 \cdots \xi_{k+1}). \tag{4. 5}$$

But, by (1. 1) this expression is equal to

$$R_k = \frac{A^{(k+2)}}{A^{(k+1)}} \left. \frac{\begin{vmatrix} a_0 & a_1 & \cdots a_k & f \\ a'_0 & a'_1 & \cdots a'_k & f' \\ \cdot & \cdot & \cdots & \cdot \\ a_0^{(k+1)} & a_1^{(k+1)} & \cdots a_k^{(k+1)} & f^{(k+1)} \end{vmatrix}}{\begin{vmatrix} a_0 & a_1 & \cdots a_k & a_{k+1} \\ a'_0 & a'_1 & \cdots a'_k & a'_{k+1} \\ \cdot & \cdot & \cdots & \cdot \\ a_0^{(k+1)} & a_1^{(k+1)} & \cdots a_k^{(k+1)} & a_{k+1}^{(k+1)} \end{vmatrix}} \right|_\xi \cdot \left. \frac{\begin{vmatrix} a_0 & a_1 & \cdots a_k & g \\ a'_0 & a'_1 & \cdots a'_k & g' \\ \cdot & \cdot & \cdots & \cdot \\ a_0^{(k+1)} & a_1^{(k+1)} & \cdots a_k^{(k+1)} & g^{(k+1)} \end{vmatrix}}{\begin{vmatrix} a_0 & a_1 & \cdots a_k & a_{k+1} \\ a'_0 & a'_1 & \cdots a'_k & a'_{k+1} \\ \cdot & \cdot & \cdots & \cdot \\ a_0^{(k+1)} & a_1^{(k+1)} & \cdots a_k^{(k+1)} & a_{k+1}^{(k+1)} \end{vmatrix}} \right|_\eta \tag{4. 6}$$

ξ and η being two values in the interval (α, β). Since $A^{(k+1)}$ and $A^{(k+2)}$ are positive, and the two determinants in the denominator of (4. 6) are of the same sign, according to our assumptions regarding the set $a_0(x) \cdots a_{k+1}(x)$, we see that, if none of the two functions of x

$$\begin{vmatrix} a_0 & \cdots a_k & f \\ \cdot & \cdots & \cdot \\ a_0^{(k+1)} & \cdots a_k^{(k+1)} & f^{(k+1)} \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} a_0 & \cdots a_k & g \\ \cdot & \cdots & \cdot \\ a_0^{(k+1)} & \cdots a_k^{(k+1)} & g^{(k+1)} \end{vmatrix}$$

change sign in the interval (α, β), the remainder is positive or negative according as the two functions are of the same sign or of opposite sign. For $a_\varkappa(x) = x^\varkappa$, the two functions in question are $f^{(k+1)}(x)$ and $g^{(k+1)}(x)$ (a part from a positive, constant factor). For $k=0$ this criterion involves the monotony of f and g, that is it gives TCHEBYCHEFF's inequality for integrals. For $k=1$ the criterion involves the convexity of f and g, and so on.

The remainder expression (4. 4) can also be written

$$R_k = \frac{A^{(k+2)}}{A^{(k+1)}} H_{k+1} \tag{4. 7}$$

where the quantity

$$H_{k+1} = \frac{1}{(k+2)!\ A^{(k+2)}} \int_\alpha^\beta dx_0 \int_\alpha^\beta dx_1 \cdots$$

$$\cdots \int_\alpha^\beta dx_{k+1} f(x_0 \cdots x_{k+1})\, g(x_0 \cdots x_{k+1})\, A^2(x_0 \cdots x_{k+1})$$

may be looked upon as the average value of the product of the (general) $(k+1)$-th order difference of $f(x)$ and the (general) $(k+1)$-th order difference of $g(x)$; the average being taken over the interval (α, β) and with the positive weights $A^2(x_0 \cdots x_{k+1})$.

5. A Generalization of the Notion of Difference Operation and a General Interpolation Formula of Newton's Type.

Let a sequence of functions $a_0(x)$, $a_1(x) \cdots$ satisfying the condition of Section 1, be given. Let $A(x_0 \cdots x_\varkappa)$ and $F(x_0 \cdots x_\varkappa)$ be the determinants defined by (4. 1) and (4. 2) and let $\bar{A}(x_0 \cdots x_\varkappa)$ be the determinant obtained from $A(x_0 \cdots x_\varkappa)$ by replacing $a_\varkappa(x)$ by $a_{\varkappa+1}(x)$. By (1. 9) and the assumptions regarding the set $a_0(x)$, $a_1(x) \cdots$ all the ratios

$$\alpha(x_0 \cdots x_\varkappa) = \frac{\bar{A}(x_0 \cdots x_\varkappa)}{A(x_0 \cdots x_\varkappa)}, \text{ in particular } \alpha(x_0) = \frac{a_1(x_0)}{a_0(x_0)} \tag{5. 1}$$

are finite. The quantities $\alpha(x_0 \cdots x_\varkappa)$ are evidently symmetric in all the arguments.

We shall find an expression for the difference

$$\alpha(x_0\cdots[x_i]\cdots x_\varkappa)-\alpha(x_0\cdots[x_j]\cdots x_\varkappa),$$

where [ ] designates exclusion of the argument in question. Considering the minors in the determinant $A(x_0\cdots x_\varkappa)$ we have by SYLVESTER's theorem

$$\bar{A}(x_0\cdots[x_i]\cdots x_\varkappa)\cdot A(x_0\cdots[x_j]\cdots x_\varkappa)-$$
$$-\bar{A}(x_0\cdots[x_j]\cdots x_\varkappa)\cdot A(x_0\cdots[x_i]\cdots x_\varkappa)=$$
$$=A(x_0\cdots[x_i]\cdots[x_j]\cdots x_\varkappa)\cdot A(x_0\cdots x_\varkappa).$$

We have here assumed $i<j$, which is expressed by the fact that $[x_i]$ is written before $[x_j]$ in the string of arguments of the first quantity A on the right hand side of the last equation. If $i>j$ a minus should be added on one side of the equation. By virtue of the last equation we get

$$(5.2)\quad \alpha(x_0\cdots[x_i]\cdots x_\varkappa)-\alpha(x_0\cdots[x_j]\cdots x_\varkappa)=$$
$$=\frac{A(x_0\cdots[x_i]\cdots[x_j]\cdots x_\varkappa)\cdot A(x_0\cdots x_\varkappa)}{A(x_0\cdots[x_i]\cdots x_\varkappa)\cdot A(x_0\cdots[x_j]\cdots x_\varkappa)}.$$

This equation shows that the difference (5.2) can not vanish provided the arguments $x_0\cdots x_\varkappa$ are distinct and the functions $a_0(x)$, $a_1(x)\cdots$ satisfy the conditions specified in Section 1. We assume of course that $i\neq j$, otherwise the first determinant in the right hand side numerator of (5.2) would have no meaning.

Jf (5.2) is written out for the first values of $\varkappa$, we get

$$-(\alpha(x_0)-\alpha(x_1))=\frac{1\cdot A(x_0x_1)}{A(x_0)\cdot A(x_1)}$$

$$-(\alpha(x_0x_1)-\alpha(x_1x_2))=\frac{A(x_1)\cdot A(x_0x_1x_2)}{A(x_0x_1)\cdot A(x_1x_2)}$$

$$-(\alpha(x_0x_1x_2)-\alpha(x_1x_2x_3))=\frac{A(x_1x_2)\cdot A(x_0x_1x_2x_3)}{A(x_0x_1x_2)\cdot A(x_1x_2x_3)}$$

etc.

We shall now take the quantities α as a base for the definition of a general difference operation.

Let $f(x)$ be a given function. The zero order difference of $f(x)$ for the argument x_0 is defined as $\frac{f(x_0)}{a_0(x_0)}$. It should be noticed that the zero order difference is not simply the function f itself as in the case of ordinary divided differences.

The first order difference is defined as

$$f(x_0x_1)=\frac{\frac{f(x_0)}{a_0(x_0)}-\frac{f(x_1)}{a_0(x_1)}}{\alpha(x_0)-\alpha(x_1)}.$$

The second order difference is defined as

$$f(x_0x_1x_2)=\frac{f(x_0x_1)-f(x_1x_2)}{\alpha(x_0x_1)-\alpha(x_1x_2)}.$$

Generally the $\varkappa$-th order difference is defined by

$$(5.3)\qquad f(x_0\cdots x_\varkappa)=\frac{f(x_0\cdots[x_i]\cdots x_\varkappa)-f(x_0\cdots[x_j]\cdots x_\varkappa)}{\alpha(x_0\cdots[x_i]\cdots x_\varkappa)-\alpha(x_0\cdots[x_j]\cdots x_\varkappa)}.$$

[ ] designating that the argument in question is omitted.

The general differences defined by the recurrence operation (5.3) are uniquely determined. And the explicite expression for the $\varkappa$-th order general difference has the following simple form

$$(5.4)\qquad f(x_0\cdots x_\varkappa)=\frac{F(x_0\cdots x_\varkappa)}{A(x_0\cdots x_\varkappa)}=\frac{\begin{vmatrix} a_0(x_0) & \cdots & a_{\varkappa-1}(x_0) & f(x_0)\\ a_0(x_1) & \cdots & a_{\varkappa-1}(x_1) & f(x_1)\\ \cdot & \cdot & \cdot & \cdot\\ a_0(x_\varkappa) & \cdots & a_{\varkappa-1}(x_\varkappa) & f(x_\varkappa)\end{vmatrix}}{\begin{vmatrix} a_0(x_0) & \cdots & a_{\varkappa-1}(x_0) & a_\varkappa(x_0)\\ a_0(x_1) & \cdots & a_{\varkappa-1}(x_1) & a_\varkappa(x_1)\\ \cdot & \cdot & \cdot & \cdot\\ a_0(x_\varkappa) & \cdots & a_{\varkappa-1}(x_\varkappa) & a_\varkappa(x_\varkappa)\end{vmatrix}}$$

which shows that the definition (5. 3) is independent of the choice of the arguments x_i and x_j which are omitted alternately.

The formula (5. 4) may be proved by complete induction. In fact, suppose that the formula is exact for $(\varkappa-1)$. By the same argument as in the case of (5. 2), we get

$$(5.5)\qquad f(x_0\cdots[x_i]\cdots x_\varkappa)-f(x_0\cdots[x_j]\cdots x_\varkappa)=$$
$$=\frac{\Delta(x_0\cdots[x_i]\cdots[x_j]\cdots x_\varkappa)\cdot F(x_0\cdots x_\varkappa)}{\Delta(x_0\cdots[x_i]\cdots x_\varkappa)\cdot\Delta(x_0\cdots[x_j]\cdots x_\varkappa)}.$$

Taking the ratio between (5. 5) and (5. 2) we get (5. 4) for $\varkappa$. Since (5. 4) obviously holds good in the case $\varkappa=0$, it holds generally.

If we insert $f(x)=a_\varkappa(x)$ in (5. 4) we get $a_\varkappa(x_0\cdots x_\varkappa)=1$, and if we insert $f(x)=a_\lambda(x)$ where $\lambda<\varkappa$ we get $a_\varkappa(x_0\cdots x_\varkappa)=0$. We therefore have the proposition: *The $\varkappa$-th order difference of $a_\varkappa(x)$, (that is, of the $\varkappa$-th function in the set which defines the difference operation) is unity and the higher order differences of $a_\varkappa(x)$ are zero.*

This property will be referred to as the *finiteness* of the general difference operation. This property not only illustrates the character of the general difference operation. It also throws an interesting light upon the ordinary difference operation. The fact that the $\varkappa$-th order divided difference of a $\varkappa$-th degree polynomial (with unity coefficient of $x^\varkappa$) is equal to 1, and the higher differences zero, is in reality not characteristic for a polynomial as such. It is rather an expression for the fact that the ordinary divided difference operation can be looked upon as an operation *defined by the notion of polynomials of increasing degree.* The ordinary divided difference in Newton's sense is namely the special case $a_\varkappa(x)=p_\varkappa(x)$ of (5. 4) where $p_\varkappa(x)$ is an *arbitrary* polynomial of degree $\varkappa$ with unity coefficient of $x^\varkappa$. For this reason the ordinary divided differences might be called *polynomial* differences as distinguished from the general differences (5. 4).

In the case where the functions $a_\varkappa(x)$ are polynomials

of degree $\varkappa$, with unity coefficient of $x^\varkappa$, the determinant $\Delta(x_0\cdots x_\varkappa)$ reduces to the VANDERMONDE, determinant[1]

$$(5.6)\qquad \Delta(x_0\cdots x_\varkappa)=\begin{vmatrix}1 & x_0 & x_0^2 & \cdots & x_0^\varkappa\\ 1 & x_1 & x_1^2 & \cdots & x_1^\varkappa\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 1 & x_\varkappa & x_\varkappa^2 & \cdots & x_\varkappa^\varkappa\end{vmatrix}$$

which can be written in the form

$$(5.7)\qquad \Delta(x_0\cdots x_\varkappa)=\prod_{j>i}(x_j-x_i)$$

where j runs through the values $j=1,2\cdots\varkappa$ and i through the values $i=0,1\cdots(j-1)$.

The expression (5. 7) can be looked upon as the decomposition of the determinant (5. 6) in elementary factors. Such a decomposition in elementary factors is also possible in the general case (4. 1). We namely have

$$(5.8)\qquad \begin{vmatrix}a_0(x_0) & a_1(x_0) & \cdots & a_\varkappa(x_0)\\ a_0(x_1) & a_1(x_1) & \cdots & a_\varkappa(x_1)\\ \cdot & \cdot & \cdot & \cdot\\ a_0(x_\varkappa) & a_1(x_\varkappa) & \cdots & a_\varkappa(x_\varkappa)\end{vmatrix}=$$
$$= a_0(x_0)\cdot a_0(x_1)\cdots a_0(x_\varkappa)$$
$$(-)^\varkappa(\alpha(x_0)-\alpha(x_1))\cdot(\alpha(x_1)-\alpha(x_2))\cdots$$
$$\cdots(\alpha(x_{\varkappa-1})-\alpha(x_\varkappa))$$
$$(-)^{\varkappa-1}(\alpha(x_0x_1)-\alpha(x_1x_2))\cdot(\alpha(x_1x_2)-\alpha(x_2x_3))\cdots$$
$$\cdots(\alpha(x_{\varkappa-2}x_{\varkappa-1})-\alpha(x_{\varkappa-1}x_\varkappa))$$
$$\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot$$
$$(-)^1(\alpha(x_0x_1\cdots x_{\varkappa-1})-\alpha(x_1x_2\cdots x_\varkappa)).$$

[1] $p_\varkappa(x)$, $(\varkappa=0,1\cdots k)$ being an arbitrary $\varkappa$-th degree polynomial with unity coefficient of $x^\varkappa$, we evidently have $\begin{vmatrix}p_0(x_0) & \cdots & p_k(x_0)\\ \cdot & \cdot & \cdot\\ p_0(x_k) & \cdots & p_k(x_k)\end{vmatrix}=\begin{vmatrix}1 & x_0 & \cdots & x_0^k\\ \cdot & \cdot & \cdot & \cdot\\ 1 & x_k & \cdots & x_k^k\end{vmatrix}$. This is readily proved by a suitable procedure of column subtractions performed on the left hand side determinant.

In fact, in the left hand side determinant of (5. 8) take the factor $a_0(x_0)$ out of the first row, the factor $a_0(x_1)$ out of the second row etc. This being done, subtract the next to last row from the last, the third to last from the next to last ... and finally the first row from the second. In the determinant thus obtained take the factor $-(\alpha(x_0)-\alpha(x_1))$ out of the second row, the factor $-(\alpha(x_1)-\alpha(x_2))$ out of the third row, an so on. The determinant thus obtained is

$$\begin{vmatrix} \frac{a_0(x_0)}{a_0(x_0)} & \frac{a_1(x_0)}{a_0(x_0)} & \cdots & \frac{a_\varkappa(x_0)}{a_0(x_0)} \\ a_0(x_0x_1) & a_1(x_0x_1) & \cdots & a_\varkappa(x_0x_1) \\ a_0(x_1x_2) & a_1(x_1x_2) & \cdots & a_\varkappa(x_1x_2) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_{\varkappa-1}x_\varkappa) & a_1(x_{\varkappa-1}x_\varkappa) & \cdots & a_\varkappa(x_{\varkappa-1}x_\varkappa) \end{vmatrix}.$$

In this determinant subtract the next to last row from the last, the third to last row from the next to last ... and finally the second row from the third. Then take the factor $-(\alpha(x_0x_1)-\alpha(x_1x_2))$ out of the third row, the factor $-(\alpha(x_1x_2)-$ $-\alpha(x_2x_3))$ out of the fourth row etc. If we continue in this way, we finally obtain the determinant

$$(5.9)\qquad \begin{vmatrix} \frac{a_0(x_0)}{a_0(x_0)} & \frac{a_1(x_0)}{a_0(x_0)} & \cdots & \frac{a_\varkappa(x_0)}{a_0(x_0)} \\ a_0(x_0x_1) & a_1(x_0x_1) & \cdots & a_\varkappa(x_0x_1) \\ a_0(x_0x_1x_2) & a_1(x_0x_1x_2) & \cdots & a_\varkappa(x_0x_1x_2) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_0x_1\cdots x_\varkappa) & a_1(x_0x_1\cdots x_\varkappa) & \cdots & a_\varkappa(x_0x_1\cdots x_\varkappa) \end{vmatrix}.$$

But in this determinant all the elements below the principal diagonal are zero, and all the elements in the principal diagonal are equal to unity, by virtue of the finiteness of the general difference operation. This establishes (5. 8). We could also prove (5. 8) by introducing the expression (5. 2) for each difference in the right hand side of (5. 8) and multiply together all these expressions.

In the polynomial case we have

$$(x_i-x_0)(x_i-x_1)\cdots(x_i-x_{i-1})(x_{i+1}-x_i)(x_{i+2}-x_i)\cdots(x_\varkappa-x_i)\cdot$$
$$\cdot\Delta(x_0\cdots[x_i]\cdots x_\varkappa)=\Delta(x_0\cdots x_\varkappa)$$

and consequently by (5. 2)

$$(5.10)\qquad \alpha(x_0\cdots[x_i]\cdots x_\varkappa)-\alpha(x_0\cdots[x_j]\cdots x_\varkappa)=x_j-x_i.$$

Introducing this in (5. 8) wee see that (5. 7) is the polynomial case of (5. 8).

The formula (5. 10) is also readily verified by noticing that we have in the polynomial case

$$\alpha(x_0\cdots x_\varkappa)=x_0+\cdots+x_\varkappa.$$

In fact, consider the polynomial

$$p(x)=\begin{vmatrix} 1 & \cdots & x_0^{\varkappa+1} \\ \cdot & \cdot & \cdot \\ 1 & \cdots & x_\varkappa^{\varkappa+1} \\ 1 & \cdots & x^{\varkappa+1} \end{vmatrix} : \begin{vmatrix} 1 & \cdots & x_0^\varkappa \\ \cdot & \cdot & \cdot \\ 1 & \cdots & x_\varkappa^\varkappa \end{vmatrix}.$$

On the one side we have $p(x)=x^{\varkappa+1}-(x_0+\cdots+x_\varkappa)\cdot x^\varkappa+\cdots$, and on the other side $p(x)=x^{\varkappa+1}-\alpha(x_0\cdots x_\varkappa)\cdot x^\varkappa+\cdots$. Hence $\alpha(x_0\cdots x_\varkappa)=x_0+\cdots+x_\varkappa$.

Newtons well known interpolation formula

$$f(x)=f(x_0)+(x-x_0)f(x_0x_1)+(x-x_0)(x-x_1)f(x_0x_1x_2)+\cdots+$$
$$+(x-x_0)\cdots(x-x_{k-1})f(x_0x_1\cdots x_k)+R_k$$

where

$$R_k=(x-x_0)(x-x_1)\cdots(x-x_k)f(x_0x_1\cdots x_kx)$$

follows immediately from the very notion of divided difference. In fact, according to the definition of the divided difference we have

$$\begin{vmatrix} 1 & x_0^1 & \cdots & x_0^k & f(x_0) \\ 1 & x_1^1 & \cdots & x_1^k & f(x_1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_k^1 & \cdots & x_k^k & f(x_k) \\ 1 & x^1 & \cdots & x^k & f(x) \end{vmatrix} = \begin{vmatrix} 1 & x_0^1 & \cdots & x_0^{k+1} \\ \cdot & \cdot & \cdot & \cdot \\ 1 & x_k^1 & \cdots & x_k^{k+1} \\ 1 & x^1 & \cdots & x^{k+1} \end{vmatrix} \cdot f(x_0 x_1 \cdots x_k x).$$

By retaining the term containing $f(x)$ on the left hand side of the last equation and transposing the rest of the determinant to the right hand side, we get immediately

$$(5.11) \qquad f(x) = -\begin{vmatrix} 1 & x_0^1 & \cdots & x_0^k & f(x_0) \\ 1 & x_1^1 & \cdots & x_1^k & f(x_1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_k^1 & \cdots & x_k^k & f(x_k) \\ 1 & x^1 & \cdots & x^k & 0 \end{vmatrix} : \begin{vmatrix} 1 & x_0^1 & \cdots & x_0^k \\ \cdot & \cdot & \cdot & \cdot \\ 1 & x_k^1 & \cdots & x_k^k \end{vmatrix} + R_k$$

R_k having the expression given on the preceding page.

This is nothing else than LAGRANGE's interpolation formula with remainder. And carrying out on the denominator and the first $(k+1)$ rows of the numerator of (5.11) the same manipulation of the rows as we did in the case of (5.9), we get NEWTON's formula.

This method can be applied not only in the case of polynomial differences but also in the general case. This leads to the following interpolation formula of LAGRANGE type

$$(5.12) \quad f(x) =$$

$$-\begin{vmatrix} a_0(x_0) & a_1(x_0) & \cdots & a_k(x_0) & f(x_0) \\ a_0(x_1) & a_1(x_1) & \cdots & a_k(x_1) & f(x_1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_0(x_k) & a_1(x_k) & \cdots & a_k(x_k) & f(x_k) \\ a_0(x) & a_1(x) & \cdots & a_k(x) & 0 \end{vmatrix} : \begin{vmatrix} a_0(x_0) & a_1(x_0) & \cdots & a_k(x_0) \\ a_0(x_1) & a_1(x_1) & \cdots & a_k(x_1) \\ \cdot & \cdot & \cdot & \cdot \\ a_0(x_k) & a_1(x_k) & \cdots & a_k(x_k) \end{vmatrix} + R_k$$

where

$$(5.13) \qquad R_k = \frac{A(x_0 \cdots x_k x)}{A(x_0 \cdots x_k)} f(x_0 \cdots x_k x) =$$

$$= \frac{A(x_0 \cdots x_k x)}{A(x_0 \cdots x_k)} \cdot \frac{\begin{vmatrix} a_0 & \cdots & a_k & f \\ a'_0 & \cdots & a'_k & f' \\ \cdot & \cdot & \cdot & \cdot \\ a_0^{(k+1)} & \cdots & a_k^{(k+1)} & f^{(k+1)} \end{vmatrix}_\xi}{\begin{vmatrix} a_0 & \cdots & a_k & a_{k+1} \\ a'_0 & \cdots & a'_k & a'_{k+1} \\ \cdot & \cdot & \cdot & \cdot \\ a_0^{(k+1)} & \cdots & a_k^{(k+1)} & a_{k+1}^{(k+1)} \end{vmatrix}_\xi}$$

ξ being a value between the largest and the smallest of the arguments $x_0 \cdots x_k x$, and A being defined by (4.1). The last expression in (5.13) follows immediately from (1.1). In the polynomial case the expression (5.13) reduces to

$$(5.14) \qquad (x-x_0)(x-x_1)\cdots(x-x_k)\,\frac{f^{(k+1)}(\xi)}{(k+1)!}$$

which is the well known form of the remainder of LAGRANGE's formula.

The function $a_{k+1}(x)$ enters only in the remainder (5.13), and not in the principal term of (5.12). Therefore, if we put $a_\varkappa(x) = x^\varkappa$ for $\varkappa = 0, 1 \cdots k$, but put $a_{k+1}(x) = g(x)$. where $g(x)$ is a function later to be disposed of, we still get the ordinary LAGRANGE interpolation formula, but the remainder now appears in the form

$$(5.15) \qquad R_k = (x-x_0)(x-x_1)\cdots(x-x_k)\cdot g(x_0 \cdots x_k x)\frac{f^{(k+1)}(\xi)}{g^{(k+1)}(\xi)}$$

where $g(x_0 \cdots x_k x)$ is the ordinary $(k+1)$-order divided difference of the function $g(x)$. The presence of this arbitrary function $g(x)$ in the remainder raises the following minimizing problem:

Given a function $f(x)$ and $(k+1)$ points $x_0 \cdots x_k$ in a certain intervalle (i). Determine the function $g(x)$ which is such that the maximum of

$$\theta(x) = \frac{g(x_0 \cdots x_k) \cdot f^{(k)}(x)}{g^{(k)}(x)}$$

in the intervalle (i), (regardless of the sign of $\theta(x)$) is the least possible.

The solution of this problem is $g(x)=f(x)$. In fact, no function $g(x)$ is such that the maximum of $|\theta(x)|$ in the interval (i) is less than $|f(x_0 \cdots x_k)|$. By (1. 1) there namely always exists a value ξ in (i) such that $\theta(\xi)=f(x_0 \cdots x_k)$. No function $g(x)$ can therefore be better than the function $g(x)$ which keeps $\theta(x)=\text{constant}=f(x_0 \cdots x_k)$, that is $g(x)=f(x)$.

The solution of this problem for the $(k+2)$ points $x_0 \cdots x_k x$ ought to furnish the best possible error limit for the LAGRANGE formula, that is, it ought to give the rigorous remainder formula $(x-x_0) \cdots (x-x_k) f(x_0 \cdots x_k x)$. That this formula is actually furnished by the solution of the minimizing problem is immediately verified by putting $g(x)=f(x)$ in (5. 15).

In practice no advantage is thus gained by the general solution of the minimizing problem considered. If the type of solution is limited by introducing convenient side relations, practical formulae may, however, possibly be obtained, furnishing a better limitation than (5. 14) and being at the same time maniable.

If the manipulation which was used in order to obtain (5. 9) is carried out on the denominator and the first $(k+1)$ rows of the numerator in the right hand side of (5. 12), this formulae takes on the NEWTON[1] form:

[1] The problem of a general expansion ordered in the same way as NEWTON's formula (as distinguished from LAGRANGE's), that is, a general expansion where the introduction af new terms does not alter the coefficients of those already computed, has been suggested to me by professor MEIDELL. At the time of my conversation with professor MEIDELL he had himself attacked this problem, however, in a different form, where the points $x_0 \cdots x_k$ do not enter symmetrically. He had also determined the first few terms of his expansion, while I had worked out the general formulae of the present Section up to (5. 12). The formula (5. 16) to (5. 19) were worked out subsequently with the purpose of furnishing an expansion more similar to the one which professor MEIDELL had in view.

(5. 16) $$f(x) = -\begin{vmatrix} 1 & \frac{a_1(x_0)}{a_0(x_0)} & \frac{a_2(x_0)}{a_0(x_0)} & \cdots & \frac{a_k(x_0)}{a_0(x_0)} & \frac{f(x_0)}{a_0(x_0)} \\ 0 & 1 & a_2(x_0 x_1) & \cdots & a_k(x_0 x_1) & f(x_0 x_1) \\ 0 & 0 & 1 & \cdots & a_k(x_0 x_1 x_2) & f(x_0 x_1 x_2) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & f(x_0 x_1 \cdots x_k) \\ a_0(x) & a_1(x) & a_2(x) & \cdots & a_k(x) & 0 \end{vmatrix} + R_k.$$

This formula might be written out in either of the following two ways

(5. 17) $$f(x) = \frac{a_0(x)}{1} \cdot \frac{f(x_0)}{a_0(x_0)} + \frac{\begin{vmatrix} a_0(x_0) & a_1(x_0) \\ a_0(x) & a_1(x) \end{vmatrix}}{a_0(x_0)} \cdot f(x_0 x_1) + $$
$$+ \frac{\begin{vmatrix} a_0(x_0) & a_1(x_0) & a_2(x_0) \\ a_0(x_1) & a_1(x_1) & a_2(x_1) \\ a_0(x) & a_1(x) & a_2(x) \end{vmatrix}}{\begin{vmatrix} a_0(x_0) & a_1(x_0) \\ a_0(x_1) & a_1(x_1) \end{vmatrix}} f(x_0 x_1 x_2) + \cdots + R_k$$

(5. 18) $$f(x) = a_0(x) \frac{f(x_0)}{a_0(x_0)} + a_0(x) \cdot (\alpha(x) - \alpha(x_0)) \cdot f(x_0 x_1)$$
$$+ a_0(x) \cdot (\alpha(x) - \alpha(x_0)) \cdot (\alpha(x x_0) - \alpha(x_0 x_1)) \cdot f(x_0 x_1 x_2)$$
$$+ a_0(x) (\alpha(x) - \alpha(x_0)) \cdot (\alpha(x x_0) - \alpha(x_0 x_1)) \cdot$$
$$\cdot (\alpha(x x_0 x_1) - \alpha(x_0 x_1 x_2)) \cdot f(x_0 x_1 x_2 x_3)$$
$$+ \cdots + R_k$$

R_k being given by (5. 13).

The intuitiv significance of the successive steps of approximation which are expressed by either of these formulae, say (5. 17), can be illustrated as follows.

The first approximation in the case of the ordinary NEWTON's formula consists in the assumption that $f(x)$ is a constant, or as we shall prefer to express it: that $f(x)$ is proportional to a constant. On the contrary, the first approximation in the general case of formula (5. 17) consists in the assumption that $f(x)$ is proportional to some function $a_0(x)$. The factor of proportionality is determined in such a way that the expression obtained coincides with the given function in the particular point x_0. This gives the first term of (5. 17).

When we proceed to the second approximation we abandon the hypothesis that $\frac{f(x)}{a_0(x)}$ is constantly equal to $\frac{f(x_0)}{a_0(x_0)}$. We introduce instead the hypothesis that the deviation $\frac{f(x)}{a_0(x)} - \frac{f(x_0)}{a_0(x_0)}$ is proportional to the deviation $\frac{a_1(x)}{a_0(x)} - \frac{a_1(x_0)}{a_0(x_0)}$ where $a_1(x)$ is some new function introduced for the purpose of the second approximation. And the factor of proportionality is determined in such a way that the expression obtained conicides with the given function not only in the point x_0 but also in the new point x_1. This gives the second term of (5. 17).

Quite generally: The passage from one approximation to the next following, consists in putting the remainder equal to a linear form in the approximation functions already used, plus a term containing a new approximation function, and then determining the constants of the form in such a way that the remainder vanishes not only in all the points previously used but also in one new point.

If each approximation function is choosen so as to vanish in all the previously used points but not in the new point; i. e., if $a_\varkappa(x_\varkappa) \neq 0$ and $a_\varkappa(x_\lambda)=0$ for $\varkappa > \lambda$, then the expansion (5. 17) reduces to

(5. 19) $$f(x) =$$

$$= a_0(x) \cdot \frac{f(x_0)}{a_0(x_0)} + a_1(x) \cdot f(x_0 x_1) + a_2(x) \cdot f(x_0 x_1 x_2) + \cdots + R_k.$$

6. Special Case: Expansion in Terms of Power Moments.

I shall now consider the case where the functions used for approximating $f(x)$ or $g(x)$ in (2. 1) are polynomials. For convenience I shall assume that the polynomials in question are ordered in terms of powers taken about the beginning of the interval of integration. This means that we have to put the functions $a_\varkappa(x)$ of the preceding Sections equal to

$$a_\varkappa(x) = (x-\alpha)^\varkappa.$$

This gives for the numerical coefficients (2. 8)

$$a_{ij} = \int_\alpha^\beta (x-\alpha)^{i+j}\, dx = \frac{(\beta-\alpha)^{i+j+1}}{i+j+1}.$$

The determinant (2. 10) is now equal to

(6. 1) $$A^{(k+1)} = (\beta-\alpha)^{(k+1)^2}\, C^{(k+1)}$$

where $C^{(k+1)}$ is the determinant

(6. 2) $$C^{(k+1)} = \begin{vmatrix} \frac{1}{1} & \frac{1}{2} & \cdots & \frac{1}{k+1} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{k+2} \\ \frac{1}{k+1} & \frac{1}{k+2} & \cdots & \frac{1}{2k+1} \end{vmatrix}$$

the value of which, by (2. 11) is given by

$$(k+1)!\; C^{(k+1)} = \int_0^1 dx_0 \int_0^1 dx_1 \cdots \int_0^1 dx_k \prod_{j>i} (x_j - x_i)^2$$

where j runs through $j=1, 2 \cdots k$ and i through $i=0, 1 \cdots (j-1)$. The last formula shows that $C^{(k+1)}$ is positive not zero. P_k

is now a bilinear form in the ordinary power moments of $f(x)$ and $g(x)$ respectively, namely

$$(6.3)\quad P_k = \frac{1}{\beta-\alpha}\sum_{i=0}^{k}\sum_{j=0}^{k}\frac{c_{ij}^{(k+1)}}{(\beta-\alpha)^{i+j}}\int_\alpha^\beta (x-\alpha)^i f(x)\,dx\cdot\int_\alpha^\beta (x-\alpha)^j g(x)\,dx$$

where $c_{ij}^{(k+1)}$ are the elements of the reciprocal of the $(k+1)$ rowed determinant $C^{(k+1)}$.

If we put $a_\varkappa(x) = \frac{1}{\sqrt{\beta-\alpha}}\,\psi_\varkappa\left(\frac{x-\alpha}{\beta-\alpha}\right)$ where the $\psi_\varkappa(x)$ form a set of polynomials of degree $\varkappa$, which are orthogonal over the standard interval $(0, 1)$, i. e. polynomials such that

$$\int_0^1 \psi_i(x)\,\psi_j(x)\,dx = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases},$$

P_k will be identical with the expression (6. 3) but will appear in the form

$$(6.4)\quad P_k = \frac{1}{\beta-\alpha}\sum_{i=0}^{k}\int_\alpha^\beta \psi_i\left(\frac{x-\alpha}{\beta-\alpha}\right) f(x)\,dx \int_\alpha^\beta \psi_j\left(\frac{x-\alpha}{\beta-\alpha}\right) g(x)\,dx.$$

I now proceed to develop some formulae concerning the numbers $c_{ij}^{(k)}$.

For $C^{(k)}$ we have the following explicite formula

$$(6.5)\quad C^{(k)} = \begin{vmatrix} \frac{1}{1} & \frac{1}{2} & \cdots & \frac{1}{k} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{k+1} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{1}{k} & \frac{1}{k+1} & \cdots & \frac{1}{2k-1} \end{vmatrix} =$$

$$= \frac{1}{1\cdot 3\cdots(2k-1)\cdot\left(\binom{2}{1}\binom{4}{2}\binom{6}{3}\cdots\binom{2k-2}{k-1}\right)^2}.$$

This expression for the determinant $C^{(k)}$ is obtained as follows: Subtract the second column of $C^{(k)}$ from the first, the third from the second ... and the k-th from the $(k-1)$-th. This operation involves $(k-1)$ column subtractions. In the determinant thus obtained subtract the second column from the first, the third from the second ... and the $(k-1)$-th from the $(k-2)$-th. This operation involves $(k-2)$ column subtractions, and the first $(k-2)$ columns will now have the factor 2, which is taken out. Repeat the process. The next step involves $(k-3)$ column subtractions, and the first $(k-3)$ columns will get the factor 3, which is taken out. The last step will involve only one column subtraction, namely the second column subtracted from the first.

This being done, take the factor $1/k!$ out of the first row, the factor $1/(k+1)!$ out of the second row ... and the factor $1/(2k-1)!$ out of the last row. The result is

$$(6.6)\quad C^{(k)} = \frac{1^{k-1}2^{k-2}\cdots(k-1)^1}{k!\,(k+1)!\cdots(2k-1)!}\begin{vmatrix} 0! & 1! & 2! & \cdots & (k-1)! \\ 1! & 2! & 3! & \cdots & k! \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ (k-1)! & k! & (k+1)! & \cdots & (2k-1)! \end{vmatrix}.$$

In the determinant of (6. 6) take the factor $0!$ out of the first row and the first column, the factor $1!$ out of the second row and the second column ... and the factor $(k-1)!$ out of the last row and the last column. This gives the following expression for the determinant in (6. 6)

$$(0!\,1!\,2!\cdots(k-1)!)^2\begin{vmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \cdots & \binom{k-1}{0} \\ \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \cdots & \binom{k}{1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \binom{k-1}{k-1} & \binom{k}{k-1} & \binom{k+1}{k-1} & \cdots & \binom{2k-2}{k-1} \end{vmatrix}.$$

But the determinant in the last expression is equal to unity.[1]

[1] See for instance P. B. FISCHER: Determinanten, Leipzig 1908, p. 83.

We therefore have

$$C^{(k)}=\frac{(1!\,2!\cdots(k-1)!)^3}{k!\,(k+1)!\cdots(2k-1)!}$$

hence

$$C^{(k)}=(2k+1)\binom{2k}{k}^2 C^{(k+1)}. \tag{6. 7}$$

From this and $C^{(1)}=1$ we infer (6. 5).

The first of the $C^{(k)}$ have the values

	$C^{(k)}$
$k=1$	1
2	$\frac{1}{12}$
3	$\frac{1}{2\,160}$
4	$\frac{1}{6\,048\,000}$
5	$\frac{1}{266\,716\,800\,000}$

The formula (6. 5) is the special case $h=0$ of the following formula

$$C^{(kh)}=\begin{vmatrix} \frac{1}{h+1} & \frac{1}{h+2} & \cdots & \frac{1}{h+k} \\ \frac{1}{h+2} & \frac{1}{h+3} & \cdots & \frac{1}{h+k+1} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{1}{h+k} & \frac{1}{h+k+1} & \cdots & \frac{1}{h+2k-1} \end{vmatrix}= \tag{6. 8}$$

$$=\frac{1}{(h+1)(h+3)\cdots(h+2k-1)\left(\binom{h+2}{1}\binom{h+4}{2}\cdots\binom{h+2k-2}{k-1}\right)^2}.$$

This formula may be derived by the same process as that which has been used to establish (6. 5). For brevity we denote a k-rowed determinant whose elements are z_{ij} $(i,j=0,1\cdots(k-1))$ by $\|z_{ij}\|$. The process of column subtractions gives

$$C^{(kh)}=\frac{(1!\,2!\cdots(k-1)!)^2\,(h+1)!\,(h+2)!\cdots(h+k-1)!}{(h+k)!\,(h+k+1)!\cdots(h+2k-1)!}\cdot\left\|\binom{h+i+j}{i}\right\|.$$

By h times repeating the process of subtracting from each row the row which preceds, we see that the determinant $\left\|\binom{h+i+j}{i}\right\|$ is equal to $\left\|\binom{i+j}{i}\right\|$, which is the previously considered determinant whose value is 1. We therefore have

$$C^{(kh)}=(2k+h+1)\binom{2k+h}{k}^2 C^{(k+1,\,h)}.$$

This together with $C^{(1,\,h)}=1/(h+1)$ gives (6. 8).

The formula for $C^{(kh)}$ permits to give a simple explicite expression for the numbers $c^{(k)}_{00}$. We namely have

$$c^{(k)}_{00}=C^{(k-1,\,2)}/C^{(k)}=\left(\frac{\binom{2}{1}\binom{4}{2}\cdots\binom{2k-2}{k-1}}{\binom{2}{0}\binom{4}{1}\cdots\binom{2k-2}{k-2}}\right)^2.$$

Hence

$$c^{(k)}_{00}=k^2. \tag{6. 9}$$

Any minor in the reciprocal of a non singular matrix is equal to the algebraic complement of the similarly situated minor in the matrix itself, divided by the determinant value of the matrix. Since (6. 8) furnishes an explicite expression for any minor in $C^{(k)}$ which is formed by adjacent rows and adjacent columns, we see that we can give the explicite expression for any minor in the matrix of the numbers $c^{(k)}_{ij}$, which is such that its complement is formed by adjacent rows and adjacent columns. This gives in particular the explicite expression for the nnmbers $c^{(k)}_{0,\,k-1}$ and $c^{(k)}_{k-1,\,0}$. In fact we have

$$c_{00}^{(k)} c_{k-1,\,k-1}^{(k)} - c_{0,\,k-1}^{(k)} c_{k-1,\,0}^{(k)} = C^{(k-2,\,2)}/C^{(k)}.$$

Now $c_{ij}^{(k)} = c_{ji}^{(k)}$, because the matrix $(c_{ij}^{(k)})$, being the reciprocal of a symmetric matrix, must be symmetric. Furthermore

$$c_{k-1,\,k-1}^{(k)} = C^{(k-1)}/C^{(k)}$$

from which we infer by (6. 7)

$$c_{k-1,\,k-1}^{(k)} = (2k-1)\binom{2k-2}{k-1}^2. \tag{6. 10}$$

We therefore have

$$c_{0,\,k-1}^{(k)2} = (2k-1)^2\binom{2k-2}{k-1}^2$$

$(-)^{k-1} c_{0,\,k-1}^{(k)}$ must be positive because the $(k-1)$ rowed determinant $(-)^{k-1} C_{0,\,k-1}^{(k)}$ is the GRAM-ian determinant (for $\beta-\alpha=1$) of the product moments of $(k-1)$ linearly independent functions namely $(x-\alpha)^{i+1/2}$ $(i=0,\,1\cdots(k-2))$. By taking account of the sign, we therefore have

$$c_{0,\,k-1}^{(k)} = c_{k-1,\,0}^{(k)} = (-)^{k-1}(2k-1)\binom{2k-2}{k-1}. \tag{6. 11}$$

An independent explicite check on the computation for each value of k is furnished by the following formulae.

Developing the determinant $C^{(k)}$ after the row $\frac{1}{i+1}$, $\frac{1}{i+2}\cdots\frac{1}{i+k}$ we get

$$\sum_{j=0}^{k-1}\frac{1}{i+j+1}c_{ij}^{(k)} = 1. \qquad (i=0,\,1\cdots(k-1))$$

Evidently

$$\sum_{j=0}^{k-1}\frac{1}{\varkappa+j+1}c_{ij}^{(k)} = 0$$

if $\varkappa$ is any of the numbers $0, 1, \cdots (k-1)$ with exception of i.

Developing the determinant $C^{(k+1)}$ after the last row, and further expanding each of the terms thus obtained after the last column, we get by (6. 7)

$$\sum_{i=0}^{k-1}\sum_{j=0}^{k-1}\frac{1}{(k+i+1)(k+j+1)}c_{ij}^{(k)} = \frac{\binom{2k}{k}^2-1}{(2k+1)\binom{2k}{k}^2}.$$

Since the determinant value of the reciprocal of a non singular matrix is equal to the reciprocal of the determinant value of the matrix, we also have

$$\begin{vmatrix} c_{00}^{(k)} & \cdots & c_{0,\,k-1}^{(k)} \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ c_{k-1,\,0}^{(k)} & \cdots & c_{k-1,\,k-1}^{(k)} \end{vmatrix} = 1/C^{(k)}.$$

The first of the numbers $c_{ij}^{(k)}$ are

$c_{ij}^{(1)}$	$j=0$
$i=0$	1

$c_{ij}^{(2)}$	$j=0$	1
$i=0$	4	-6
1	-6	12

$c_{ij}^{(3)}$	$j=0$	1	2
$i=0$	9	-36	30
1	-36	192	-180
2	30	-180	180

Introducing the differences

$$d_{ij}^{(k)} = c_{ij}^{(k+1)} - c_{ij}^{(k)} \tag{6. 12}$$

$(i, j=0,\,1\cdots k)$ where by convention $c_{ij}^{(k)}=0$ for $i=k$ or $j=k$, we can write the expansion of the integral (2. 1) in the form of a series

$$\int_\alpha^\beta f(x)\,g(x)\,dx = T_0 + T_1 + \cdots + T_k + R_k$$

where

(6. 13) $T_k =$

$$\frac{1}{\beta-\alpha}\sum_{i=0}^{k}\sum_{j=0}^{k} d_{ij}^{(k)}\cdot(\beta-\alpha)^{-(i+j)}\int_\alpha^\beta (x-\alpha)^i f(x)\,dx\cdot\int_\alpha^\beta (x-\alpha)^j g(x)\,dx$$

and

(6. 14) $$P_k = T_0 + T_1 + \cdots + T_k.$$

The expression (6. 3) for P_k can of course also be written in terms of the moments taken about any other point than α. If the moments about $\alpha+\varrho\lambda$ (ϱ an arbitrary constant, $\lambda=\beta-\alpha$) are denoted

$$M_\varkappa = \int_\alpha^\beta (x-(\alpha+\varrho\lambda))^\varkappa f(x)\,dx \quad\text{and}\quad N_\varkappa = \int_\alpha^\beta (x-(\alpha+\varrho\lambda))^\varkappa g(x)\,dx$$

the expression for P_k will be

(6. 15) $$P_k = \frac{1}{(\beta-\alpha)}\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} b_{ij}^{(k+1)}\frac{M_i N_j}{(\beta-\alpha)^{i+j}}$$

where the numerical coefficients $b_{ij}^{(k)}$ are

$$b_{ij}^{(k)} = \frac{1}{\varrho^{i+j}}\sum_{r=i}^{k-1}\sum_{s=j}^{k-1} c_{rs}^{(k)}\binom{r}{i}\binom{s}{j}\varrho^{r+s}$$

or in determinant form

$$b_{ij}^{(k)} = \frac{-1}{\varrho^{i+j}C^{(k)}}\begin{vmatrix} 0 & \binom{0}{j}\varrho^0 & \binom{1}{j}\varrho^1 & \cdots & \binom{k-1}{j}\varrho^{k-1} \\ \binom{0}{i}\varrho^0 & \frac{1}{1} & \frac{1}{2} & \cdots & \frac{1}{k} \\ \binom{1}{i}\varrho^1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{k+1} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \binom{k-1}{i}\varrho^{k-1} & \frac{1}{k} & \frac{1}{k+1} & \cdots & \frac{1}{2k-1} \end{vmatrix}$$

Putting $\varrho=\frac{1}{2}$ we get an expansion in terms of the moments taken about the middle of the interval (α, β).

The first of the coefficients $b_{ij}^{(k)}$ in this case are

$b_{ij}^{(1)}$	$j=0$
$i=0$	1

$b_{ij}^{(2)}$	$j=0$	1
$i=0$	1	0
1	0	12

These are the first of the coefficients computed by professor Steffensen.[1]

Putting $a_\varkappa(x)=(x-\alpha)^\varkappa$ in (4. 5) we get by (6. 1) and (6. 5)

(6. 16) $$R_k = (\beta-\alpha)^{2k+3}\frac{f(\xi_0\cdots\xi_{k+1})\,g(\xi_0\cdots\xi_{k+1})}{(2k+3)\binom{2k+2}{k+1}^2}.$$

$f(\xi_0\cdots\xi_{k+1})$ and $g(\xi_0\cdots\xi_{k+1})$ being the ordinary $(k+1)$-th order divided differences of the functions $f(x)$ and $g(x)$ for a set of values $\xi_0\cdots\xi_{k+1}$ situated in the interval (α,β). By (4. 6) the remainder (6. 16) can also be written in the form

(6. 17) $$R_k = \frac{(\beta-\alpha)^{2k+3}}{(2k+3)}\,\frac{f^{(k+1)}(\xi)\,g^{(k+1)}(\eta)}{((k+2)(k+3)\cdots(2k+2))^2}.$$

[1] loc. cit., p. 58.

The following table gives an idea of the rapid decrease of the numerical factors in the remainder terms (6. 16) and (6. 17)

	$(2k+3)\binom{2k+2}{k+1}^2$	$(2k+3)((k+2)(k+3)\cdots(2k+2))^2$
$k=0$	12	12
1	180	720
2	2 800	100 800
3	44 100	25 401 600
4	798 544	10 059 033 600
5	11 099 088	5 753 767 219 200

7. The Damping Effect. Evaluation of the Remainder after the First Term.

The damping problem discussed by professor MEIDELL[1] may be outlined as follows.

Consider a function of the time $W(t)$, the *resultant*, which is generated through a cumulative process, defined by the equation

$$W(t) = \int_0^t n(\tau)\, w(t-\tau)\, d\tau = \int_0^t n(t-\xi)\, w(\xi)\, d\xi. \tag{7. 1}$$

The function of the time $n(t)$ may be looked upon as the *originator* (»Zugang») of the process through which $W(t)$ is generated. And the function of time elapsed $w(x)$ may be looked upon as the *distributor* which brings the effect over from n to W. Numerous instances, to which this point of view is applicable, might be drawn from the statistical and actuarial field.

[1] BIRGER MEIDELL: Über periodische und angenäherte Beharrungszustände. This Journal 1926, and: On damping effects and approach to equilibrium in certain general phenomena. Journal of The Washington Academy of Science. October 1928.

It is a well known fact that in many practical cases the course of the resultant $W(t)$ will be comparatively smooth even though the originator $n(t)$ has pronounced fluctuations. This is what is known as the damping effect of the process considered.

The idea therefore naturally presents itself to attempt an estimate of the resultant by computing it as if the originator had been constant, say equal to the value of the originator at the beginning of the interval considered, or better equal to the average over that interval. By this procedure one avoids the laborious, direct computation of $W(t)$ by (7. 1).

If the originator is rigorously constant, we have

$$W(t) = n\,\Omega_t \tag{7. 2}$$

where n is the constant value of the originator and $\Omega_x = \int_0^x w(\xi)\, d\xi$ the cumulated distributor.

The procedure suggested by the observed damping effect would therefore be to put n in (7. 2) equal to some plausible (mean or extremity) value of $n(\tau)$ in the interval previous to t, and adopt (7. 2) as an approximation to $W(t)$, even though $n(\tau)$ is not constant.

The main object of professor MEIDELL's papers was to show under what conditions such a procedure was legitimate. In particular he has arrived at important criteria for the *mean* deviation of $W(t)$. The object of the present Section is to establish criteria for the deviation of W *in a given point*, by utilizing some formulae from Section 2.

Let $N_t = \int_0^t n(\tau)\, d\tau$ be the cumulated originator an $m_t = \frac{N_t}{t}$ the mean value of the originator over the interval $(0, t)$. Similarly let $w_x = \frac{\Omega_x}{x}$ be the mean value of the distributor over the interval $(0, x)$.

By (2. 2) and (2. 6) we have the error limit

$$|W(t)-w_t N_t| \leqq p_t q_t \tag{7.3}$$

p_t and q_t being the deviation factors defined by

$$p_t^2 = \int_0^t (n(\tau)-n_t)^2\, dt \qquad q_t^2 = \int_0^t (w(\xi)-w_t)^2\, d\xi.$$

In (7. 2) no use is made of eventually known value of the resultant. If $t_0\, t_2 \cdots t_k$ are points of time where the resultant is known, and t a point of time where the resultant is sought, we get by taking the $(k+1)$ order divided difference of the expression $(W(t)-w_t N_t)$

$$W(t) = w_t N_t + \tag{7.4}$$

$$+\sum_{\varkappa=0}^{k} \frac{(t-t_0)(t-t_1)\cdots[t-t_\varkappa]\cdots(t-t_k)}{(t_\varkappa-t_0)(t_\varkappa-t_1)\cdots[t_\varkappa-t_\varkappa]\cdots(t_\varkappa-t_k)}\,(W(t_\varkappa)-w_{t_\varkappa} N_{t_\varkappa}) + R_k$$

where

$$R_k = (t-t_0)(t-t_1)\cdots(t-t_k)\, R(t_0\, t_1 \cdots t_k\, t)$$

and [ ] designates exclusion of the factor in question.

$R(t_0\, t_1 \cdots t_k\, t)$ is the $(k+1)$ order divided difference of

$$R(t) = \int_0^t (n(\tau)-n_t)(w(t-\tau)-w_t)\, d\tau. \tag{7.5}$$

The approximarion procedure (7. 3) can be improved also in another way. In most practical cases the originator $n(t)$ is an empirical function of the time, which it is not possible to characterize a priori. On the contrary, the distributor $w(x)$ is most frequently known a priori. The properties of $w(x)$ may therefore be characterized once for all, say be computing a certain number of parameters such as means and deviation factors over fixed subintervals.

The resultant can always, rigorously be written in the form $W(t)=n_t^*\,\Omega_t$. We only have to define

$$n_t^* = \int_0^t w(\xi)\, n(t-\xi)\, d\xi \Big/ \int_0^t w(\xi)\, d\xi.$$

By this definition n_t^* appears as the *weighted* average of $n(\tau)$ over the interval previous to t, the weights being the values of the distributor; (7. 3) can be looked upon as the approximation obtained by replacing the weighted average n_t^* be the ordinary arithmetic average n_t. This means attributing just as much weights to the values of the originator which occur in the remote past (until $\tau=0$) as to those which occur in the moments of time immediately before t. This procedure contains an obvious bias if $t \geqq a$, where a is a positive constant such that the distributor $w(x)$ is very small, or even rigorously equal to zero, for $x \geqq a$.

Instead of attributing equal weights to all the values of $n(\tau)$ in the total interval previous to t it would now have been better to neglect the values of $n(\tau)$ previous to $(t-a)$, and to attribute equal weights to the values of $n(\tau)$ in the rest of the interval. This procedure is equivalent with assuming $w(x)=0$ for $x \geqq a$, so that the expression for $W(t)$ will be for $t \geqq a$

$$W(t) = \int_{t-a}^{t} n(\tau)\, w(t-\tau)\, d\tau = \int_0^a n(t-\xi)\, w(\xi)\, d\xi$$

and then approximating this integral by (2. 2).

More generally, suppose that the nature of $w(x)$ has been characterized once for all by dividing the interval $(0, \infty)$ in $(k+1)$ subintervals (x_i, x_{i+1}) $(i = 0, 1 \cdots k)$, with $x_0 = 0$ and $x_{k+1} = \infty$, and computing the means w_i and the deviation factors q_i defined by

$$w_i = \int_{x_i}^{x_{i+1}} w(\xi)\, d\xi \Big/ (x_{i+1} - x_i)$$

$$q_i^2 = \int_{x_i}^{x_{i+1}} (w(\xi) - w_i)^2\, d\xi.$$

The principle for choosing the division points (their number being given) should be to have the deviation factors q_i approximately of the same magnitude (the last deviation factor i. e. q_k being, however, equal to zero it $w(x)=0$ for $x > x_k$).

Let j be the subscript defined by $x_j < t \leqq x_{j+1}$. If $W(t)$ is approximated by the expression

$$\overline{W}(t) = \frac{\Omega_t - \Omega_{x_j}}{t - x_j} N_{t-x_j} + \sum_{i=0}^{j-1} w_i (N_{t-x_i} - N_{t-x_{i+1}}) \tag{7.6}$$

we have the error limit

$$| W(t) - \overline{W}(t) | \leqq \sum_{i=0}^{j} \pi_i q_i \tag{7.7}$$

where π_i ist the maximum absolute value of the deviation of the originator from its mean over the interval $(t - x_{i+1},\, t - x_i)$, the interval corresponding to x_j being however $(0,\, t - x_j)$. The limit (7. 7) is readily proved by decomposing the integral $W(t) = \int_0^t n(t-\xi)\, w(\xi)\, d\xi$ in $\int_{x_0}^{x_1} + \cdots + \int_{x_j}^{t}$ and approximating each of these integrals by (2. 2). This gives the error limit $\sum_{i=0}^{j-1} \pi_i q_i + \pi_j q_j^*$ where q_j^* is the deviation factor of $w(\xi)$ over the interval $(x_j\, t)$. This deviation factor is, however, never greater than q_j, because a deviation factor never decreases by

an extention of the interval over which it is taken. In fact, let $f(\eta)$ be a function of η, and consider the square of its deviation factor over the interval (α, y). Taking the derivative with respect to y, we get

$$\frac{d}{dy} \int_\alpha^y \left(f(\eta) - \int_\alpha^y f(\xi)\, d\xi/(y-\alpha) \right)^2 d\eta = \left(f(y) - \int_\alpha^y f(\xi)\, d\xi/(y-\alpha) \right)^2$$

which is non negative.