

533.

ON THE BINOMIAL THEOREM, FACTORIALS, AND DERIVATIONS.

[From the *Oxford, Cambridge, and Dublin Messenger of Mathematics*, vol. v. (1869), pp. 102—114.]

THE following was part of my course of lectures in the year 1867.

The proof commonly called "Euler's" of the binomial theorem is as follows: the theorem is assumed to be true for positive integer indices; that is, it is assumed that for any positive integer m we have

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{1.2} x^2 + \&c.$$

This being so, since $(1+x)^m \cdot (1+x)^n = (1+x)^{m+n}$, the equation

$$\left\{ 1 + mx + \frac{m(m-1)}{1.2} x^2 + \&c. \right\} \left\{ 1 + nx + \frac{n(n-1)}{1.2} x^2 + \&c. \right\} \\ = 1 + \frac{m+n}{1} x + \frac{(m+n)(m+n-1)}{1.2} x^2 + \&c.$$

is true for any positive integer values whatever of the indices m, n ; the equation is therefore true identically; and it is consequently true for all values whatever of the indices m, n . But any function ϕm of m , satisfying the functional equation $\phi m \cdot \phi n = \phi(m+n)$, is an m^{th} power, $= C^m$ suppose; that is, we have

$$C^m = 1 + mx + \frac{m(m-1)}{1.2} x^2 + \&c.,$$

where C is a constant, viz. it is independent of m ; but the value of C will of course depend upon x ; and if, in order to determine it, we write $m=1$, the equation gives $C=1+x$; that is, we have

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{1.2} x^2 + \&c.,$$

which is the binomial theorem in its general form.

It is to be observed that there is not in the demonstration any employment of the so called "principle of equivalent forms," we do not from the truth of an equation for positive integer values of m, n , infer the truth of it for any values whatever of m, n ; there is the intermediate step, that being true for integers, it is true identically; and this identical truth of the equation depends on the circumstance that comparing on the two sides of the equation the coefficients of the successive powers $x^0, x^1, x^2, \&c.$, these coefficients are in every case finite, rational, and integral functions of m, n .

For instance, comparing the coefficients of x^2 , we have

$$(m+n)(m+n-1) = m(m-1) + 2mn + n(n-1),$$

and any such equation, being true for all positive integer values of m, n , will be true identically; developing the two sides, the equation is in fact

$$m^2 + 2mn + n^2 - m - n = m^2 - m + 2mn + n^2 - n.$$

The reasoning is thus perfectly good; but I remark that it is quite as easy to prove the general equation of which the last mentioned equation is an example, as it is to prove the binomial theorem for positive integer indices; and consequently that we can *without the aid of the binomial theorem for positive integer indices* prove the fundamental equation

$$\begin{aligned} \left\{ 1 + mx + \frac{m(m-1)}{1 \cdot 2} x^2 + \&c. \right\} \left\{ 1 + nx + \frac{n(n-1)}{1 \cdot 2} x^2 + \&c. \right\} \\ = 1 + \frac{m+n}{1} x + \frac{(m+n)(m+n-1)}{1 \cdot 2} x^2 + \&c. \end{aligned}$$

To show this I introduce the factorial notation and write

$$m(m-1) \dots (m-r+1) = [m]^r;$$

this being so, the equation obtained by comparing the coefficients of x^r is readily found to be

$$[m+n]^r = [m]^r + \frac{r}{1} [m]^{r-1} [n]^1 + \frac{r(r-1)}{1 \cdot 2} [m]^{r-2} [n]^2 + \&c.,$$

and I say that this, *the factorial binomial theorem for a positive integer index r* , is proved as easily and in the same manner as the binomial theorem for a like value of the index; or say as the equation

$$(m+n)^r = m^r + \frac{r}{1} m^{r-1} n^1 + \frac{r(r-1)}{1 \cdot 2} m^{r-2} n^2 + \&c.$$

To show how this is, I form the values of $[m+n]^1, [m+n]^2, [m+n]^3, \&c.$ successively, by what may be termed the process of *varied multiplication*: we have

$$[m+n]^1 = m+n;$$

But the method of varied multiplication may be applied to the demonstration of a much more general theorem; viz. we may use it to develop a product such as

$$(h - a)(h - b)(h - c)(h - d)(h - e),$$

according to a series of products

$$(h - \alpha)(h - \beta)(h - \gamma)(h - \delta)(h - \epsilon),$$

$$(h - \alpha)(h - \beta)(h - \gamma)(h - \delta),$$

$$(h - \alpha)(h - \beta)(h - \gamma),$$

$$(h - \alpha)(h - \beta),$$

$$(h - \alpha),$$

$$1.$$

For this purpose, starting with

$$h - a = h - \alpha + \alpha - a,$$

we multiply by $h - b$, written first under the form $h - \beta + \beta - b$, and then under the form $h - \alpha + (\alpha - b)$; we have thus

$$\begin{aligned} (h - a)(h - b) &= (h - \alpha)(h - \beta) \\ &+ (h - \alpha) \quad \{(\alpha - a) + (\beta - b)\} \\ &+ 1 \quad (\alpha - a)(\alpha - b), \end{aligned}$$

and so on. It is easy to see that we may for instance write

$$\begin{aligned} &(h - a)(h - b)(h - c)(h - d)(h - e) \\ &= (h - \alpha)(h - \beta)(h - \gamma)(h - \delta)(h - \epsilon) \\ &+ (h - \alpha)(h - \beta)(h - \gamma)(h - \delta) \quad \left(\begin{matrix} \alpha, \beta, \gamma, \delta, \epsilon \\ a, b, c, d, e \end{matrix} \right)_1 \\ &+ (h - \alpha)(h - \beta)(h - \gamma) \quad \left(\begin{matrix} \alpha, \beta, \gamma, \delta \\ a, b, c, d, e \end{matrix} \right)_2 \\ &+ (h - \alpha)(h - \beta) \quad \left(\begin{matrix} \alpha, \beta, \gamma \\ a, b, c, d, e \end{matrix} \right)_3 \\ &+ h - \alpha \quad \left(\begin{matrix} \alpha, \beta \\ a, b, c, d, e \end{matrix} \right)_4 \\ &+ 1 \quad \left(\begin{matrix} \alpha \\ a, b, c, d, e \end{matrix} \right)_5, \end{aligned}$$

where the symbols $()_{\theta}$ denote sums of products of the differences of an upper letter and a lower letter, θ factors in each product; and the several sums being formed as follows :

For $(\alpha, \beta, \gamma, \delta, \epsilon)_1$ write

α	a
β	b
γ	c
δ	d
ϵ	e ;

the expression is

$$(\alpha - a) + (\beta - b) + (\gamma - c) + (\delta - d) + (\epsilon - e).$$

For $(\alpha, \beta, \gamma, \delta)_2$ write

α, α	a, b
α, β	a, c
α, γ	a, d
β, β	b, c
α, δ	a, e
β, γ	b, d
β, δ	b, e
γ, γ	c, d
γ, δ	c, e
δ, δ	d, e ;

viz. the expression is

$$(\alpha - a)(\alpha - b) + (\alpha - a)(\beta - c) + \dots + (\delta - d)(\delta - e).$$

For $(\alpha, \beta, \gamma)_3$ write

α, α, α	a, b, c
α, α, β	a, b, d
α, α, γ	a, b, e
α, β, β	a, c, d
α, β, γ	a, c, e
β, β, β	b, c, d
α, γ, γ	a, d, e
β, β, γ	b, c, e
β, γ, γ	b, d, e
γ, γ, γ	c, d, e ;

viz. the expression is

$$(\alpha - a)(\alpha - b)(\alpha - c) + \dots + (\gamma - c)(\gamma - d)(\gamma - e).$$

For $\binom{\alpha, \beta}{a, b, c, d, e}_4$ write $\begin{array}{l|l} \alpha, \alpha, \alpha, \alpha & a, b, c, d \\ \alpha, \alpha, \alpha, \beta & a, b, c, e \\ \alpha, \alpha, \beta, \beta & a, b, d, e \\ \alpha, \beta, \beta, \beta & a, c, d, e \\ \beta, \beta, \beta, \beta & b, c, d, e; \end{array}$

viz. the expression is

$$(\alpha - a)(\alpha - b)(\alpha - c)(\alpha - d) \dots + (\beta - b)(\beta - c)(\beta - d)(\beta - e).$$

Finally for

$\binom{\alpha}{a, b, c, d, e}_5$ write $\alpha, \alpha, \alpha, \alpha, \alpha \mid a, b, c, d, e;$

viz. the expression is

$$(\alpha - a)(\alpha - b)(\alpha - c)(\alpha - d)(\alpha - e),$$

which explains the law of the formation of the several coefficients. It is to be observed that in forming the development of any symbol, for instance $\binom{\alpha, \beta, \gamma}{a, b, c, d, e}_3$, the first column contains the homogeneous products, 3 together, of α, β, γ ; the second column the combinations (that is, combinations without repetitions) 3 together of a, b, c, d, e : the top line is $\alpha, \alpha, \alpha \mid a, b, c$ and to form the subsequent lines we must for any advance α into β , &c. of a greek letter make the like advance a into b , b into c , or c into d , of the corresponding latin letter.

Two particular cases of the theorem may be noticed: if the latin letters all vanish, we have, for example,

$$\begin{aligned} h^5 &= (h - \alpha)(h - \beta)(h - \gamma)(h - \delta)(h - \epsilon). \\ &+ (h - \alpha)(h - \beta)(h - \gamma)(h - \delta) \quad \cdot H_1(\alpha, \beta, \gamma, \delta, \epsilon) \\ &+ (h - \alpha)(h - \beta)(h - \gamma) \quad \cdot H_2(\alpha, \beta, \gamma, \delta) \\ &+ (h - \alpha)(h - \beta) \quad \cdot H_3(\alpha, \beta, \gamma) \\ &+ (h - \alpha) \quad \cdot H_4(\alpha, \beta) \\ &+ 1 \quad \cdot H_5(\alpha), \end{aligned}$$

where the symbols H denote the sum of the homogeneous products of the annexed letters, taken together according to the suffix number: the last coefficient $H_5(\alpha)$ is of course $= \alpha^5$. And if the greek letters all vanish, then we have in like manner

$$\begin{aligned} (h - a)(h - b)(h - c)(h - d)(h - e) &= h^5 \\ &- h^4 C_1(a, b, c, d, e) \\ &+ h^3 C_2(a, b, c, d, e) \\ &- h^2 C_3(a, b, c, d, e) \\ &+ h C_4(a, b, c, d, e) \\ &- C_5(a, b, c, d, e), \end{aligned}$$

where the symbols C denote the combinations of the annexed letters taken together according to the suffix number; the last coefficient $C_5(a, b, c, d, e)$ is of course $= abcde$. This is the ordinary theorem giving the expression of an equation in terms of its roots.

Combining the two theorems, if in the first theorem we express the products $(h - \alpha)(h - \beta)(h - \gamma)(h - \delta)(h - \epsilon)$, &c. in powers of h by means of the second theorem; or if in the second theorem we express the powers h^5, h^4 , &c. in terms of the products $(h - a)(h - b)(h - c)(h - d)(h - e)$, &c. by means of the first theorem; then in either case we obtain certain identical relations connecting the C, H of (α, β, \dots) or of (a, b, \dots) .

I have mentioned the factorial notation

$$[m]^r = m(m-1) \dots (m-r+1),$$

where r is a positive integer; a consequence of this is

$$[m]^{r+s} = [m]^r [m-r]^s,$$

where r and s are positive integers; or as this may also be written

$$[m+r]^{r+s} = [m+r]^r [m]^s.$$

Assuming this to subsist for $s=0$, or a negative integer; first for $s=0$, we have $[m+r]^r = [m+r]^r [m]^0$; that is, $[m]^0$ is $= 1$; and then for $s = -r$, we have $1 = [m+r]^r [m]^{-r}$; that is,

$$[m]^{-r} = \frac{1}{[m+r]^r},$$

and in particular, $r=1, 2$, &c., we have

$$[m]^{-1} = \frac{1}{m+1},$$

$$[m]^{-2} = \frac{1}{(m+1)(m+2)},$$

&c.,

which explains the extension of the factorial notation to negative integer values of the index.

But the equation

$$[m]^{r+s} = [m]^r [m-r]^s$$

does not in any determinate manner lead to an extension of the factorial notation to fractional or other values of the index. In fact, assuming $[m]^r = \frac{\Pi m}{\Pi(m-r)}$, where Π is an arbitrary functional symbol, the equation in question becomes

$$\frac{\Pi m}{\Pi(m-r-s)} = \frac{\Pi m}{\Pi(m-r)} \frac{\Pi(m-r)}{\Pi(m-r-s)},$$

viz. the original equation is identically satisfied, without any condition whatever being imposed upon the function Π , and on this account we have not, in the notation of the factorial with an integer index, any sufficient basis for a theory of general differentiation.

A product

$$m(m - \alpha) \dots \{m - (r - 1)\alpha\}$$

can of course be expressed in the factorial notation, viz. it is

$$= \alpha^r \left[\frac{m}{\alpha} \right]^r,$$

and on this account it is not in general necessary to employ a notation such as $[m, \alpha]^r$ to denote such a factorial wherein the difference of the successive factors instead of being $= -1$ is $= -\alpha$; in particular cases where factorials of the kind in question are used, it may be convenient to employ such a notation. In particular it is sometimes convenient to use the notation $[m, -1]^r$ or better $\{m\}^r$ to denote the product

$$m(m + 1) \dots (m + r - 1),$$

where the successive factors instead of being diminished, are increased by unity. It may be noticed, that reversing in this last product the order of the factors, we find

$$\{m\}^r = [m + r - 1]^r;$$

a somewhat similar formula, but employing only the ordinary factorial notation, is obtained from the equation

$$[m]^r = m(m - 1) \dots (m - r + 1),$$

by first changing the sign of m and then reversing the order of the factors; viz. we have

$$[-m]^r = (-)^r [m + r - 1]^r.$$

Reverting to the process used for the development of the expressions $()_\theta$, where there are two columns, the one of greek, the other of latin letters; it is to be remarked that although the order in which the successive lines are evolved is not material for the purpose of the theorem, yet that a certain definite order of evolution has been made use of; thus in regard to $\begin{pmatrix} \alpha, \beta, \gamma, \delta \\ a, b, c, d, e \end{pmatrix}_2$, the column of greek letters, giving the homogeneous products of the second order in $(\alpha, \beta, \gamma, \delta)$, was

$$\begin{array}{c} \alpha \quad \alpha \\ \hline \alpha \quad \beta \\ \hline \alpha \quad \gamma \\ \hline \beta \quad \beta \\ \hline \alpha \quad \delta \\ \hline \beta \quad \gamma \\ \hline \beta \quad \delta \\ \hline \gamma \quad \gamma \\ \hline \gamma \quad \delta \\ \hline \delta \quad \delta \end{array}$$

this is evolved from the top term (α, α) by a process given implicitly in Arbogast's Calculus of Derivations, and which may be termed the rule of the last and the last but one. Let the direction "operate on any letter," be understood to mean that the letter in question is to be changed into that which immediately follows it, but in such wise that when the letter occurs more than once, e.g. as in α, α the operation affects only the letter in the right-hand place. Then operate on the α, α in regard to α , we obtain α, β ; operate on this in regard to β , we obtain α, γ ; and in regard to α , we obtain β, β . Again we operate on α, γ in regard to γ , and obtain α, δ ; we do not operate on it in regard to α for the reason that α is not the letter immediately preceding γ . Operate on β, β in regard to β , we obtain β, γ . The next step, if the series extended to ϵ would be to operate on α, δ in regard to δ , giving α, ϵ ; do not operate on it in regard to α , for the reason that α is not the letter immediately preceding δ . But in the example, since the series does not extend to ϵ , there is no operation on α, δ . Passing then to the next term β, γ , we operate in regard to γ , obtaining β, δ , and since β is the letter immediately preceding γ , we also operate in regard to β , obtaining γ, γ . Similarly if ϵ were admissible, β, δ would give β, ϵ , but it in fact gives nothing; γ, γ gives γ, δ ; thus if ϵ were admissible would give γ, ϵ and δ, δ , but it in fact gives only δ, δ , and, ϵ being inadmissible, the process is here concluded. The rule is, operate on the last letter, and when the last but one letter is that which, in alphabetical order, immediately precedes the last letter (but in this case only) operate on the last but one letter.

Taking another example, but with numbers instead of letters, and supposing the highest admissible number to be 5, then from 111 we derive as follows:

111	112	113	114	115	125	135	145	155	255	355	455	555
		122	123	124	134	144	235	245	345	445		
			222	133	224	225	244	335	444			
				223	233	234	334	344				
									333,			

the original single column being here for greater convenience broken up into distinct columns; but the order of the terms, when the columns are taken one after the other in order, each being read from the top to the bottom, being the same as before; it will be noticed that the successive divisions are the partitions into 3 parts (no part exceeding 5) of the numbers 3, 4, ..., 15 respectively; the partitions being in each case obtained without repetition, and those of the same number being given, say in their numerical order (corresponding with the alphabetical order when letters are employed). It is necessary to show that the partitions will be obtained without repetitions; and that all the partitions will be obtained; for this purpose consider, for example, the partitions of 9; any one of these is either a partition 135 where the last number 5 is not a repeated number; and in this case there is a partition of 8, viz. 134, from which operating on the last we obtain 135; but there is no other partition of 8 which would give 135, the only such partition would be 125, but here, as 2 is not the number which immediately precedes 5, there is no operation on the last but one, and we do not from it obtain 135. Or else a partition of 9 is of the form 144

where the last letter is repeated; there exists in this case no partition of 8, such that operating on the last we obtain from it 144, but there does exist a partition of 8, viz. 134, such that operating on the last but one we obtain from it 144. That is, for any partition whatever of 9 there exists one (and only one) partition of 8, such that operating on the last or the last but one, we obtain from it the partition of 9; that is, taking the entire system of the partitions of 8, and operating on the last and the last but one, we obtain, and that without repetitions, the entire series of the partitions of 9; and so in general.

Translating the example into letters, but using for greater convenience a^2 , &c. instead of a, a , &c. the process will be precisely the same; taking the letters to be a, b, c, d, e , we have

$$\begin{array}{cccccccccccc}
 a^3 & a^2b & a^2c & a^2d & a^2e & abe & ace & ade & ae^2 & be^2 & ce^2 & de^2 & e^3 \\
 & ab^2 & abc & abd & acd & ad^2 & bce & bde & cde & d^2e & & & \\
 & & b^3 & ac^2 & b^2d & b^2e & bd^2 & c^2e & d^3 & & & & \\
 & & & b^2c & bc^2 & bcd & c^2d & cd^2 & & & & & \\
 & & & & & & & & & & & & c^3.
 \end{array}$$

Attributing *weights* to the several letters, viz. to a, b, c, d, e the weights 1, 2, 3, 4, 5 respectively, the several columns show the terms of the weights 3, 4, ... 15 respectively.

I have said that the foregoing rule is given implicitly in Arbogast's Calculus of Derivations; this calculus includes in fact a process for the expansion of a function

$$\phi(a + bx + cx^2 + dx^3 + \&c.)$$

in powers of x ; the expansion in question may be obtained by means of Taylor's theorem, viz.

$$\begin{aligned}
 &\phi(a + bx + cx^2 + dx^3 + \&c.) \\
 &= \phi a \\
 &+ \frac{\phi'a}{1} (bx + cx^2 + dx^3 + \&c.) \\
 &+ \frac{\phi''a}{1.2} (bx + cx^2 + dx^3 + \&c.)^2 \\
 &+ \frac{\phi'''a}{1.2.3} (bx + cx^2 + dx^3 + \&c.)^3,
 \end{aligned}$$

viz. expanding the several powers of the polynomial increment, and arranging in powers of x , this is

$$\begin{aligned}
 &= \phi a \\
 &+ x (\phi'a . b) \\
 &+ x^2 \left(\phi'a . c + \phi''a . \frac{b^2}{2} \right) \\
 &+ x^3 \left(\phi'a . d + \phi''a . bc + \phi'''a . \frac{b^3}{6} \right) \\
 &+ x^4 \left\{ \phi'a . e + \phi''a . (bd + \frac{1}{2}c^2) + \phi'''a . \frac{b^2c}{2} + \phi''''a . \frac{b^4}{24} \right\} \\
 &+ \&c.
 \end{aligned}$$

but the object of the rule is to obtain this last-mentioned result directly, and understanding "operate in regard to a letter" to mean differentiate with regard to this letter and integrate with respect to the next succeeding letter, then the coefficients of the successive powers of x are all obtained from the first coefficient ϕa , by operating thereon according to the rule of the last and the last but one; thus ϕa , operating on a gives $\phi'a.b$; this operating in regard to b gives $\phi'a.c$, and in regard to a gives $\phi''a.\frac{b^2}{2}$; the term $\phi'a.c$ is to be operated upon in regard to c only, and it gives $\phi'a.d$; the other term $\phi''a.\frac{b^2}{2}$ operated on in regard to b gives $\phi''a.bc$, and in regard to a it gives $\phi'''a.\frac{b^3}{6}$; and so on. But attending only to the literal parts, the terms, for instance $b^2, bc, bd, \&c.$, which present themselves in the formula, are the homogeneous terms derived from b^2 , by the rule, as originally stated, with a view to the derivation of such terms.