

PROOF OF THE HITHERTO UNDEMONSTRATED FUNDAMENTAL THEOREM OF INVARIANTS.

[*Philosophical Magazine*, v. (1878), pp. 178—188.]

I AM about to demonstrate a theorem which has been waiting proof for the last quarter of a century and upwards. It is the more necessary that this should be done, because the theorem has been supposed to lead to false conclusions, and its correctness has consequently been impugned*. But, of the two suppositions that might be made to account for the observed discrepancy between the supposed consequences of the theorem and ascertained facts—one that the theorem is false and the reasoning applied to it correct, the other that the theorem is true but that an error was committed in drawing certain deductions from it (to which one might add a third, of the theorem and the reasoning upon it being both erroneous)—the wrong alternative was chosen.

* Thus in Professor Faà de Bruno's valuable *Théorie des Formes Binaires*, Turin, 1876, at the foot of page 150 occurs the following passage:—"Cela suppose essentiellement que les équations de condition soient toutes indépendantes entr'elles, ce qui n'est pas toujours le cas, ainsi qu'il résulte des recherches du Prof. Gordan sur les nombres des covariants des formes quintique et sextique."

The reader is cautioned against supposing that the consequence alleged above does result from Gordan's researches, which are indubitably correct. This supposed consequence must have arisen from a misapprehension on the part of M. de Bruno of the nature of Professor Cayley's rectification of the error of reasoning contained in his second memoir on Quantics, which had led to results discordant with Gordan's. Thus error breeds error, unless and until the pernicious brood is stamped out for good and all under the iron heel of rigid demonstration. In the early part of this year Mr Halsted, a Fellow of Johns Hopkins University, called my attention to this passage in M. de Bruno's book; and all I could say in reply was that "the extrinsic evidence in support of the independence of the equations which had been impugned rendered it to my mind as certain as any fact in nature could be, but that to reduce it to an exact demonstration transcended, I thought, the powers of the human understanding."

At the moment of completing a memoir, to appear in Borchardt's Journal, demonstrating my quarter-of-a-century-old theorem for enabling Invariants to procreate their species, as well by an act of self-fertilization as by conjugation of arbitrarily paired forms, the un hoped and unsought-for prize fell into my lap, and I accomplished with scarcely an effort a task which I had believed lay outside the range of human power.

An error was committed in reasoning out certain supposed consequences of the theorem; but the theorem itself is perfectly true, as I shall show by an argument so irrefragable that it must be considered for ever hereafter safe from all doubt or cavil. It lies at the basis of the investigations begun by Professor Cayley in his *Second Memoir on Quantics*, which it has fallen to my lot, with no small labour and contention of mind, to lead to a happy issue, and thereby to advance the standards of the Science of Algebraical Forms to the most advanced point that has hitherto been reached. The stone that was rejected by the builders has become the chief corner-stone of the building.

I shall for greater clearness begin with the case of a single binary quantic $(a, b, c, \dots, l \mathcal{X} x, y)^i$. Any rational integral function of the elements a, b, c, \dots, l which remains unchanged in value when for them are substituted the elements of the new quantic obtained by putting $x + hy$ instead of x in the original one, I call a Differentiant in x to the given quantic.

By a differentiant of a given weight w and order j , I mean one in every term of which the combination of the elements is of the j th order and the sum of their weights w , the weights of the successive elements (a, b, c, \dots, l) themselves being reckoned as $0, 1, 2, \dots, i$ respectively.

The proposition to be proved is, that the number of arbitrary constants in the most general expression for such differentiant is the difference between the number of ways in which w can be made up with j of the integers $0, 1, 2, 3, \dots, i$ (repetitions allowable), less the number of ways in which $w - 1$ can be made up with the same integers. We may denote these two numbers by $(w : i, j)$, $\{(w - 1) : i, j\}$ respectively, and their difference by $\Delta(w : i, j)$. Then, if we call the number of arbitrary constants in the differentiant of weight w and order j belonging to a binary quantic of the i th order $D(w : i, j)$, the proposition to be established is that $D(w : i, j) = \Delta(w : i, j)$.

Let us use Ω to denote the operator

$$a \frac{d}{db} + 2b \frac{d}{dc} + \dots + ik \frac{d}{dl},$$

and O to denote the operator

$$ib \frac{d}{da} + (i - 1) c \frac{d}{db} + \dots + l \frac{d}{dk}.$$

Then it is well known that the necessary and sufficient condition for D being a differentiant in x is that the identity $\Omega D = 0$ be satisfied.

Let us study the relations of Ω and O in respect to D .

In the first place, let U be any rational integral function of the elements of order j and weight w ; then I say that

$$\Omega . O . U - O . \Omega . U = (ij - 2w) U.$$

For if we use * to signify the act of pure differential operation, it is obvious that

$$\Omega . O . U = (\Omega \times O) U + (\Omega * O) U,$$

$$O . \Omega . U = (\Omega \times O) U + (O * \Omega) U;$$

Therefore $\Omega . O . U - O . \Omega . U = \{(\Omega * O) - (O * \Omega)\} U$

$$\begin{aligned} &= ia \frac{d}{da} + 2(i-1)b \frac{d}{db} + 3(i-2)c \frac{d}{dc} + \dots + ik \frac{d}{dk} \\ &\quad - ib \frac{d}{db} - 2(i-1)c \frac{d}{dc} - \dots - 2(i-1)k \frac{d}{dk} - il \frac{d}{dl} \\ &= ia \frac{d}{da} + (i-2)b \frac{d}{db} + (i-4)c \frac{d}{dc} - \dots - (i-2)k \frac{d}{dk} - 2l \frac{d}{dl}. \end{aligned}$$

If now $\rho a^p . b^q . c^r \dots l^t$, where ρ is a number, be any term in U , we have

$$\left. \begin{aligned} p + q + r + \dots + t &= j \\ q + 2r + \dots + it &= w \end{aligned} \right\} \text{by hypothesis}^2;$$

therefore
that is

$$\Omega . O . U - O . \Omega . U,$$

$$\begin{aligned} &= i \left(a \frac{d}{da} + b \frac{d}{db} + c \frac{d}{dc} \dots + l \frac{d}{dl} \right) U \\ &- 2 \left(b \frac{d}{db} + 2c \frac{d}{dc} \dots + il \frac{d}{dl} \right) U \\ &= \Sigma \rho (ij - 2w) (a^p . b^q . c^r \dots l^t) \\ &= (ij - 2w) U, \text{ as was to be proved.} \end{aligned}$$

If now for U we write D a differentiant in x , we have $\Omega D = 0$, and therefore

$$\Omega . O . D = \delta D,$$

where $\delta = ij - 2w$.

Again,

$$\Omega . O (O . D) - O . \Omega (O . D) = \{ij - 2(w + 1)\} O . D;$$

for $O . D$ is of the weight $w + 1$;

therefore

$$\begin{aligned} \Omega^2 . O^2 . D &= \Omega . O \delta D + (\delta - 2) \Omega . O . D \\ &= (2\delta - 2) \Omega . O . D \\ &= \delta (2\delta - 2) D. \end{aligned}$$

Similarly it will be seen that

$$\Omega^3 . O^3 . D = \delta (2\delta - 2) (3\delta - 6) D,$$

and in general

$$\begin{aligned} \Omega^q . O^q . D &= \delta (2\delta - 2) (3\delta - 6) \dots \{q\delta - (q^2 - q)\} D \\ &= (1 . 2 . 3 \dots q) \{\delta . (\delta - 1) (\delta - 2) \dots (\delta - q + 1)\} D, \end{aligned}$$

the successive numbers $\delta, 2\delta - 2, 3\delta - 6$, &c. being the successive sums of the arithmetical series $\delta, \delta - 2, \delta - 4, \delta - 6$, &c.

To find the most general differentiant in question, we must take every combination of the elements whose weight is w and order j , of which the number is obviously $(w : i, j)$, and prefix an indeterminate constant to each such combination; then operating upon this form with Ω , we shall reduce its weight by unity, and shall obtain as many combinations of this reduced weight (the order j remaining unchanged) as there are units in $\{(w-1) : i, j\}$. Each of these combinations will have for its coefficient a linear function of the assumed indeterminate coefficients; and in order to satisfy the identity $\Omega D = 0$, each such linear function must be made equal to zero. There are therefore $(w : i, j)$ quantities connected by $\{(w-1) : i, j\}$ homogeneous equations. *Supposing the equations to be independent*, the number of the indeterminate coefficients left arbitrary is obviously the difference between these quantities, namely, $\Delta(w : i, j)$. The difficulty consists in proving this independence—a difficulty so great that I think any one attempting to establish the theorem, as it were by direct assault, in this fashion, would find that he had another Plevna on his hands. But a position that cannot be taken by storm or by sap may be *turned* or starved into surrender; and this is how we shall take our Plevna. Be the equations of condition linearly independent or not, it is obvious that we must have $D(w : i, j)$ equal to or greater than $\Delta(w : i, j)$. I shall show by aid of a construction drawn from the resources of the Imaginative Reason, and founded on the reciprocal properties that have just been exhibited by the famous O and Ω , that this latter supposition, of the first member of the equation being greater than the second, is inadmissible and must be rejected. Observe that $(0 : i, j)$, the number of ways of making up 0 with j combinations of 0, 1, 2, ... i , is 1; also that $D(0 : i, j)$, the number of arbitrary constants in the most general differentiant in x to the quantic $(a, b, c, \dots \chi x, y)^i$ of order j and weight 0, is also 1; for such differentiant is obviously λa^n .

Thus we have for all values of w ,

$$D(w : i, j) = \text{or } > (w : i, j) - \{(w-1) : i, j\},$$

and also

$$D(0 : i, j) = (0 : i, j);$$

therefore

$$\begin{aligned} D(w : i, j) + D\{(w-1) : i, j\} + D\{(w-2) : i, j\} + \dots + D(0 : i, j) \\ = \text{or } > (w : i, j). \end{aligned}$$

If in the above condition, for any assumed value of w , $>$ is the sign to be employed, then the equation $D(w : i, j) = \Delta(w : i, j)$ cannot be satisfied for all values of w . If, on the other hand, $>$ is not the sign to be employed, then this equation, for *every value of w* , commencing with the assumed one down to 0, must be satisfied. The greatest value of w for given values of i, j , it is well known, is $\frac{ij}{2}$ for ij even, and $\frac{ij-1}{2}$ for ij odd. Let us give to w this

maximum value in the above "greater or equal" relation; for brevity, denote the differentiants whose types are $[w, i, j]$, $[(w-1), i, j]$... by $[w]$, $[w-1]$, $[w-2]$, &c. respectively, i and j being regarded as constants. It will be convenient to substitute for the number of arbitrary constants in any of these differentiants the same number of linearly independent specific values of them; so that we shall have $D(w:i, j)$ of linearly independent $[w]$'s, $D\{(w-1):i, j\}$ of linearly independent $[w-1]$'s, and so on. Now, instead of $D\{(w-q):i, j\}$ differentiants $[w-q]$, let us substitute the same number of the derived forms $O^q[w-q]$. I shall prove that the quantities (*all of the same weight w*) thus obtained are linearly independent of one another.

For suppose that those belonging to any one set $O^q.[w-q]$ are not independent, but are connected by a linear equation. Then, operating upon this equation with Ω^q , we shall obtain a linear equation between the quantities $[w-q]$, for each quantity $\Omega^q.O^q.[w-q]$ is a numerical multiple of $[w-q]$; which is contrary to the hypothesis. Again, let there be a linear equation between the quantities contained in any number of sets of the form $O^q.[w-q]$ for which m is the greatest value of q . Then, operating upon this with Ω^m , it is clear that all the quantities in the sets for which $q < m$ will introduce quantities of the form $\Omega^{m-q}[w-q]$ where $m-q > 0$, and which consequently vanish. There will be left, therefore, only quantities of the form $[w-q]$, between which a linear equation would exist, contrary to hypothesis, as in the preceding case. Therefore all the quantities in all the sets are linearly independent. But these are all of the weight w , that is,

$$\left[\frac{ij}{2} \text{ or } \frac{ij-1}{2} \right],$$

and are therefore linear functions of the number of ways in which the integers $0, 1, 2, 3, \dots i$ can be combined i and j together so as to give the weight w . Therefore being linearly independent, as just proved, their number cannot exceed this last-named number, that is, cannot exceed $(w:i, j)$. That is to say,

$$D(w:i, j) + D\{(w-1):i, j\} + \dots + D(O:i, j)$$

cannot exceed $(w:i, j)$. Therefore every one of the equations

$$D(w:i, j) = \Delta(w:i, j)$$

must be satisfied from the maximum value of w down to the value 0 , which proves the great hitherto undemonstrated fundamental theorem for a single quantic.

For any number of quantics the demonstration is precisely similar at all points: there will be as many systems of i, j as there are quantics. $(w:i, j; i', j'; \&c.)$ will denote the number of ways of making up w with j

of the integers $0, 1, 2, \dots i$, with j' of the integers $0, 1, 2, \dots i'$, and so on. The theorem to be demonstrated will be

$$D(w : i, j : i', j' : \dots) = \Delta(w : i, j : i', j' : \dots).$$

$$\Omega \text{ will become } \Sigma \left(a \frac{d}{db} + 2b \frac{d}{dc} + \dots \right),$$

$$O \quad \text{,,} \quad \text{,,} \quad \Sigma \left(ib \frac{d}{da} + (i-1)c \frac{d}{db} + \dots \right).$$

It will still be true that Ω^q, O^q, D —where D is a differentiant in x (that is, a function of the elements in all the given quantics which withstand change when these are transformed by writing $x + hy$ for x)—is a numerical multiple of D ; and D will be subject to the identity $\Omega D = 0$. We shall still have

$$D(w : i, j : i', j' : \dots) = \text{or} > \Delta(w : i, j : i', j' : \dots),$$

and

$$D(0 : i, j : i', j' : \dots) = (0 : i, j : i', j' : \dots),$$

and shall be able in precisely the same way as before to demonstrate the impossibility of $\sum_{k=w}^{k=0} D(w-k : i, j : i', j' : \dots)$ being greater than $(w : i, j : i', j' : \dots)$, and so shall be able to infer by the same logical scheme

$$\Delta(w : i, j : i', j' : \dots) = D(w : i, j : i', j' : \dots).$$

This is my extension of Professor Cayley's theorem, which leads direct to the Generating Fractions given in my recent papers in the *Comptes Rendus*.

In a series of articles which I hope to publish in the *American Journal of Pure and Applied Mathematics*, I propose to give a systematic development of the Calculus of Invariants, taking a differentiant as the primordial germ or unit. I have spoken of a differentiant in x , and of course might have done so equally of a differentiant in y . If we call the former D_x , it is capable of being shown, from the very natures of the forms O and Ω , that if the quantity $ij - 2w$, which may be called the *degree* of D_x , be called δ , then $O^\delta D_x$ becomes a differentiant in y . These may be termed simple differentiants; but the principle of continuity forbids that we should omit to comprise in the same scheme the intermediate forms $O^p D_x$ or $\Omega^q D_y$, through which simple differentiants in x and y pass into each other. These may be termed mixed differentiants; $O^p D_x$ may be termed a differentiant p removed (as we speak of *cousins* once, twice, &c. removed) from x , which will be the same thing as $O^q D_y$ (a differentiant q removed from y) if $p+q$ is equal to the degree, namely, $ij - 2w$. Now all these differentiants, whether simple or mixed, possess a wonderful property, which may be deduced by means of Salmon's Theorem, given in the *Philosophical Magazine* for August 1877. They are all, in an enlarged sense of the term, Invariants—in this sense to wit, that if the elements are made to undergo a substitution consequent upon or, as we may say, induced by a general linear substitution impressed on the variables, which for greater simplicity of enunciation may be

supposed to have unity for the determinant of its matrix, then every differentiant, whether single or double (the latter being equivalent to an invariant), and whether simple or mixed, will remain a Constant Function of the Coefficients of the impressed substitution. To wit, if the differentiant be p removes from x and q removes from y (so that its degree is $p + q$), and if the impressed substitution be $lx + \lambda y$ for x , and $mx + \mu y$ for y , where $l\mu - \lambda m = 1$, then will the differentiant be a constant bipartite quantic in the two sets of coefficients l, m and λ, μ , of the degree q in the former and p in the latter—a theorem which amounts almost to a revolution in the whole sphere of thought about Invariants.

I have borrowed the term “Imaginative Reason” from a recent paper of Mr Pater on Giorgione, in which, as in many of those of Mr Symonds (I will instance one on Milton in particular), I find a continued echo of my own ideas, and in the latter many of the very formulæ contained in my *Laws of Verse*, where versification in sport has been made æsthetic in earnest. Surely the claim of Mathematics (its “*Andersstreben*”) to take a place among the liberal arts must be now admitted as fully made good. Whether we look to the advances made in modern geometry, in modern integral calculus, or in modern algebra, in each of these a free handling of the material employed is now possible, and an almost unlimited scope left to the regulated play of the fancy. It seems to me that the whole of æsthetic (so far as at present revealed) may be regarded as a scheme having four centres, which may be treated as the four apices of a tetrahedron, namely Epic, Music, Plastic, and Mathematic. There will be found to be a *common* plane to every three of these, *outside* of which lies the fourth; and through every two may be drawn a common axis *opposite* to the axis passing through the remaining two. So far is certain and demonstrable. I think it also possible that there is a centre of gravity to each set of three, and that the lines joining each such centre with the outside apex will intersect in a common point the centre of gravity of the whole body of æsthetic; but what that centre is or must be I have not had time to think out.

Postscript.—In the first fervour of a new conception, I fear that in the manuscript which is now on its way to England I may have expressed myself with some want of clearness or precision on the subject of pure and mixed differentiants. I will therefore add a few more explanatory and vaticinatory words on this subject, through the medium of which I catch a glimpse of the possibility of obtaining a simple proof of Gordan’s theorem, just as through the medium of pure differentiants taken *per se* I caught a glimpse (almost immediately afterwards to be converted into a certainty) of the proof of Cayley’s theorem given in this memoir. I conceive that what the *ensemble* of pure differentiants have done for the one, the larger *ensemble* of all sorts of

And here for the present I end. The subject is, as it was, a vast one; and this conception of mixed differentials opens out still vaster horizons. Every thing grown on American soil, or born under the influence of its skies, as its lakes, its rivers, its trees, and its political system, seems to have a tendency to rise to colossal proportions.

I will merely add one remark which has occurred to me relating to Sturm's theorem and the process of Algebraical common measure in general. If $f(x, y)$ be a rational integral function of x, y , and $f'(x, y)$ its derivative in respect to x , and we perform the process of common measure between them regarded as functions of x , we know that the irreducible part of the successive remainders taken in ascending order, say U_0, U_1, U_2, \dots , will have for their leading coefficients (say $D_0, D_1, D_2 \dots$) the discriminants of f and of its successive derivatives in respect to x respectively.

Here D_0 is an invariant of the given form;

D_1 (a differential in x) will be the leading coefficient of the covariant

$$D_1 x^2 + O \cdot D_1 x y + \frac{O^2}{1 \cdot 2} D_1 y^2;$$

D_2 (another differential in x) will be the leading coefficient of the covariant

$$D_2 x^4 + O \cdot D_2 x^3 y + \frac{O^2}{1 \cdot 2} D_2 x^2 y^2 + \frac{O^3}{1 \cdot 2 \cdot 3} D_2 x y^3 + \frac{O^4}{1 \cdot 2 \cdot 3 \cdot 4} D_2 y^4,$$

and so on until we come back to the first Sturmian remainder of $(x, y)^i$, the irreducible part of which (or we may call it the Sturmian Auxiliary Proper) is the Hessian differentiated down from being of the degree $2i - 4$ to the degree $i - 2$, that is, to half of what it was at first; and so in like manner every Sturmian Auxiliary Proper is, so to say, a Covariant differentiated down to half its original dimensions.

The above invariant and the following covariants may be called V_0, V_1, V_2, \dots respectively. The interesting point in question is that (to numerical factors *près*)

$$U_0 = V_0, \quad U_1 = \frac{d}{dx} V_1, \quad U_2 = \left(\frac{d}{dx}\right)^2 V_2, \quad U_3 = \left(\frac{d}{dx}\right)^3 V_3,$$

and so on.

So more generally for any two functions $f(x, y), \phi(x, y)$, the irreducible part of the remainders obtained by common-measuring them with respect to x will all be derivatives in regard to x of covariants of the two given quantities. If we take for our quantities

$$(a, b, c, \dots h, k, l \chi(x, y)^i : (a', b', c', \dots h', k', l' \chi(x, y)^i),$$

the covariants in question will all be educts of (that is, functions having for their leading coefficients) the successive resultants of the forms

$$[(a, \dots h, k, l), (a', \dots h', k', l')],$$

of the forms $[(a, \dots h, k), (a', \dots h', k')],$

of the forms $[(a, \dots h), (a', \dots h')],$

and so on, the discriminants of which may be called *partial* resultants of the given forms; in a word, the simplified residues arising in the process of common-measuring in respect to one of their variables two given binary quantics are differential derivatives, in respect to that variable, of the educts of their partial resultants (of course with the understanding that the last simplified residue is the complete resultant itself).

This seems to point to the existence of some generalized statement of Sturm's theorem in which the same Educts as above referred to shall appear, but where, instead of their derivatives in respect to one of the variables being made use of, perfectly general Emanants of them shall be employed as the Criterion functions. For I need hardly add that all Educts (although not written so as to show it in what precedes) are in fact symmetrical in respect to the two sides of the quantic to which they belong.

On various *à priori* grounds I suspect the generalized theorem to be as follows. If $X_{2\mu}$ is the covariant (of degree 2μ) whose μ th derivative in respect to x is a Sturmian Auxiliary Proper to $F(x, y)$, we may substitute throughout for all the values of μ , instead of each such derivative, the more general one $(f \frac{d}{dx} - g \frac{d}{dy})^\mu X_{2\mu}$, where f and g are any assumed positive constants, of course with the understanding that the second criterion also is to be $(f \frac{d}{dx} - g \frac{d}{dy})f$ in lieu of $\frac{dF}{dx}$. And the method of Sturm will still be applicable for finding the positions of the real roots of $\frac{x}{y}$ in $f(x, y) = 0$ when we use these more general derivatives as the criteria instead of Sturm's own. When $g = 0$ the theorem is that of Sturm; when $f = 0$ it is an immediate deduction from this theorem applied to finding the positions of the root values of $\frac{y}{x}$, when it is borne in mind that the motions of $\frac{x}{y}$ and of $\frac{y}{x}$, as regards ascent and descent (excluding the moment for which either of these ratios is indefinitely near to zero) are inverse to each other. It is this that accounts for the negative sign which precedes g .

It is difficult to conceive by what theorem other than the assumed one the chasm between those extreme cases can be bridged over; and all analogy and all belief in continuity veto the supposition that no such bridge exists. "Divide *et impera*" is as true in algebra as in statecraft; but no less true and even more fertile is the maxim "auge *et impera*." The more to do or to prove, the easier the doing or the proof.