On the Reduction of a Linear Substitution to its Canonical Form.

By W. BURNSIDE. Received January 9th, 1899. Read January 12th, 1899. Received, in revised form, March 1st, 1899.

Introduction.

A canonical form, to which any linear substitution of non-vanishing determinant may be reduced, was first given by M. Jordan in his *Traité des Substitutions &c.* (pp. 114–126). M. Jordan's investigation applies directly to linear substitutions of the form

$$x'_s \equiv a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n, \quad (\text{mod. } p),$$
$$(s = 1, 2, \ldots, n),$$

where p is a prime; but the canonical form at which he arrives is known to hold equally when the congruences are replaced by equations.

The canonical form may be written as follows :—

$$\begin{array}{llll}
y'_1 = \lambda y_1, & y'_r = \lambda y_r + y_{r-1}, & (r = 2, 3, \ldots, a_1), \\
y'_{a_1+1} = \lambda y_{a_1+1}, & y'_{a_1+r} = \lambda y_{a_1+r} + y_{a_1+r-1}, & (r = 2, 3, \ldots, a_2), \\
y'_{a_1+a_2+1} = \lambda y_{a_1+a_2+1}, & y'_{a_1+a_2+r} = \lambda y_{a_1+a_2+r} + y_{a_1+a_2+r-1}, & (r = 2, 3, \ldots, a_3), \\
\ldots & \ldots & \ldots \\
z'_1 = \mu z_1, & z'_r = \mu z_r + z_{r-1}, & (r = 2, 3, \ldots, b_1), \\
z'_{b_1+1} = \mu z_{b_1+1}, & z'_{b_1+r} = \mu z_{b_1+r} + z_{b_1+r-1}, & (r = 2, 3, \ldots, b_2), \\
\ldots & \ldots & \ldots \\
w'_1 = \nu w_1, & w'_r = \nu w_r + w_{r-1}, & (r = 2, 3, \ldots, c_1). \\
\ldots & \ldots & \ldots
\end{array}$$

The numbers $a_1, a_2, \ldots, b_1, \ldots, c_1, \ldots$ are such that

$$a_1 + a_2 + a_3 + \ldots + b_1 + b_2 + \ldots + c_1 + \ldots + \ldots = n,$$

where n is the number of independent variables affected by the substitution. It will be assumed in what follows that the a's, b's, ..., are written so that

$$\begin{array}{l}
a_1 \geqslant a_2 \geqslant a_3 \geqslant \ldots\ldots, \\
b_1 \geqslant b_2 \geqslant \ldots\ldots\ldots\ldots, \\
c_1 \geqslant \ldots\ldots\ldots\ldots\ldots\ldots. \\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots
\end{array}$$

The quantities λ, μ, ν, ... are the *distinct* roots of the characteristic equation

$$\begin{vmatrix} a_{11}-\theta, & a_{12}, & a_{13}, & \dots, & a_{1n} \\ a_{21}, & a_{22}-\theta, & \dots\dots\dots, & & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1}, & a_{n2}, & \dots\dots\dots, & & a_{nn}-\theta \end{vmatrix} = 0.$$

The y's, z's, ... are independent linear functions of the original variables with coefficients depending on $\lambda, \mu, \dots$, respectively, and on the coefficients of the given substitution.

The necessary and sufficient conditions that two linear substitutions A and B in the same number of variables should be capable of being transformed the one into the other, *i.e.*, that it should be possible to find a third linear substitution C such that

$$C^{-1}AC = B,$$

were first given by Weierstrass in his memoir "Zur Theorie der bilinearen und quadratischen Formen" (*Berliner Monatsberichte*, 1868, and *Mathematische Werke*, Vol. II., pp. 19–44). The conditions are that, for each value of r from 0 to $n-1$, the greatest common measure of all the r^{th} minors of the characteristic determinant of A should be equal to the greatest common measure of all the r^{th} minors of that of B.

Two substitutions which can be transformed the one into the other clearly can be reduced to the same canonical form; and the invariant elements of either, namely, the greatest common measure of all the r^{th} minors of the characteristic determinant ($r = 0, 1, \dots, n-1$) must be expressible in terms of the various constants which enter into the canonical form. This identification has been carried out by Netto, "Zur Theorie der linearen Substitutionen" (*Acta Mathematica*, Vol. XVII., pp. 267–271). He shows in fact that, the characteristic determinant of the canonical form (p. 180) being

$$(\lambda-\theta)^{a_1+a_2+a_3+\dots}(\mu-\theta)^{b_1+\ _2+b_3+\dots}(\nu-\theta)^{c_1+c_2+c_3+\dots}\ \dots,$$

the greatest common measure of its first minors is

$$(\lambda-\theta)^{a_2+a_3+\dots}(\mu-\theta)^{b_2+b_3+\dots}(\nu-\theta)^{c_2+c_3+\dots}\ \dots,$$

that of its second minors

$$(\lambda-\theta)^{a_3+\dots}(\mu-\theta)^{b_3+\dots}(\nu-\theta)^{c_3+\dots}\ \dots,$$

and so on.

It is clear that, for the actual reduction of a substitution to its canonical form, the roots $\lambda, \mu, \nu, \ldots$ of the characteristic equation must be known; in other words, the operations necessary must be carried out in a region of rationality which contains $\lambda, \mu, \nu, \ldots$ as well as the coefficients of the substitution. In the general case, *i e.*, when all the roots of the characteristic equation are different, the operation presents no difficulty. If this case and that in which the substitution is of finite order are put on one side, no general method has hitherto been given, so far as I know, for effecting the reduction.

A method always effective for this purpose might be regarded from two points of view. Formally, it is equivalent to an independent proof of the accuracy of the canonical form; and, thence, on the lines of Herr Netto's determination just quoted, *a posteriori* to a verification of Weierstrass's conditions of equivalence for two substitutions. From this point of view it is quite immaterial whether the characteristic equation can be solved arithmetically or not. On the other hand, if the roots of the characteristic equation can be actually obtained by the four rules of arithmetic and the extraction of roots, such a method is an actual and feasible process of calculation by which the canonical form can be arrived at. It is the object of the present paper to give such a method. The leading idea is to consider the effect of the substitution on a linear function of the n variables with n arbitrary coefficients. This gives rise to a new substitution on a number of variables which cannot be greater than n, but may be less. The reduction of this new substitution to its canonical form can be carried out by a process which is the same in all cases; and on the reduction of this substitution that of the original substitution is made to depend.

It may be pointed out, as arising from the final result of the reduction, that the characteristic determinant of the new substitution spoken of is the result of dividing the characteristic determinant of the original substitution by the greatest common measure of its first minors; *i.e.*, it is the first "invariant factor" (*Elementar-theiler*) of the characteristic determinant of the original substitution. This is the same as saying that the characteristic equation of the new substitution is the equation of lowest degree which the matrix of the original substitution satisfies.*

---

* This Introduction has been recast at the suggestion of one of the referees; no other alteration has been made.

1. Let $$x'_s = \Sigma a_{st} x_t, \quad (s, t = 1, 2, \ldots, n),$$

be a linear substitution S, on n variables, whose determinant is not zero.

Let $$y_1 = \sum_1^n A_r x_r$$

be a homogeneous linear function of the variables with arbitrary coefficients; and let

$$y_1, y_2, y_3, \ldots$$

be a series of homogeneous linear functions of the variables, each of which is derived from the preceding one by carrying out the substitution S on the variables. Since there are only n variables, not more than n y's can be linearly independent; but it may happen that, though the coefficients in y_1 are arbitrary, the number of linearly independent y's is less than n.

Suppose that y_{m+1} is the first of the y's that can be expressed linearly in terms of those that precede it, and that

$$y_{m+1} + a_m y_1 + a_{m-1} y_2 + \ldots + a_1 y_m = 0.$$

The constants $a_1, a_2, \ldots, a_m$ are functions of the coefficients of the substitution S; and every one of the y's can be expressed linearly in terms of the first m.

So far as it affects the y's, the substitution S takes the form

$$\left.\begin{aligned} y'_1 &= y_2, \\ y'_2 &= y_3, \\ \ldots & \quad \ldots \\ y'_{m-1} &= y_m, \\ y'_m &= -a_m y_1 - a_{m-1} y_2 - \ldots - a_1 y_m \end{aligned}\right\}. \qquad \text{(i)}$$

Let $$z = \sum_1^m c_r y_r$$

be a linear function of the y's, which is changed into a multiple, λz, of itself by the substitution S. Then

$$\lambda \sum_1^m c_r y_r = \sum_1^{m-1} c_r y_{r+1} - c_m (a_m y_1 + a_{m-1} y_2 + \ldots + a_1 y_m);$$

so that the c's and λ are determined by

$$\begin{aligned}
0 &= c_m(\lambda+a_1)-c_{m-1},\\
0 &= c_m a_2 \quad + c_{m-1}\lambda - c_{m-2},\\
0 &= c_m a_3 \quad\quad + c_{m-2}\lambda - c_{m-3},\\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots\\
0 &= c_m a_{m-1} \quad\quad\quad + c_2\lambda - c_1,\\
0 &= c_m a_m \quad\quad\quad\quad + c_1\lambda.
\end{aligned}$$

The equation for λ, obtained by eliminating the c's, is

$$\lambda^m + a_1\lambda^{m-1} + a_2\lambda^{m-2} + \ldots + a_{m-1}\lambda + a_m = 0; \tag{ii}$$

and the ratios of the c's are given by

$$\frac{c_r}{c_m} = \lambda^{m-r} + a_1\lambda^{m-r-1} + \ldots + a_{m-r-1}\lambda + a_{m-r},$$
$$(r = 1, 2, \ldots, m-1).$$

These ratios are definite functions of the roots of (ii.), so that corresponding to any root λ of the equation there is a single linear function of the y's which is changed into λ times itself by S. If unity be written for c_m, this function is given by

$$z = \sum_1^m (\lambda^r + a_1\lambda^{r-1} + \ldots + a_{r-1}\lambda + a_r)\, y_{m-r},$$
$$= \sum_1^m \lambda^{m-r}(y_r + a_1 y_{r-1} + \ldots + a_{r-1}y_1).$$

Consider next the linear function defined by

$$z^{(s)} = \frac{1}{s!}\frac{\partial^s z}{\partial\lambda^s}.$$

Its value is given by

$$z^{(s)} = \sum_{r=1}^{r=m-s} \binom{m-r}{s} \lambda^{m-r-s}(y_r + a_1 y_{r-1} + \ldots + a_{r-1}y_1);$$

and

$$\begin{aligned}
\lambda z^{(s)} + z^{(s-1)} = &\sum_{r=1}^{r=m-s} \binom{m-r}{s} \lambda^{m-r-s+1}(y_r + a_1 y_{r-1} + \ldots + a_{r-1}y_1)\\
&+ \sum_{r=1}^{r=m-s+1} \binom{m-r}{s-1} \lambda^{m-r-s+1}(y_r + a_1 y_{r-1} + \ldots + a_{r-1}y_1)\\
= &\sum_{r=1}^{r=m-s+1} \binom{m-r+1}{s} \lambda^{m-r-s+1}(y_r + a_1 y_{r-1} + \ldots + a_{r-1}y_1).
\end{aligned}$$

But, denoting as before the operation of the substitution S by an accent,

$$\begin{aligned} z'^{(s)} &= \sum_{r=1}^{r=m-s} \binom{m-r}{s} \lambda^{m-r-s} (y'_r + a_1 y'_{r-1} + \dots + a_{r-1} y'_1) \\ &= \sum_{r=1}^{r=m-s} \binom{m-r}{s} \lambda^{m-r-s} (y_{r+} + a_1 y_r + \dots + a_{r-1} y_2) \\ &= \sum_{r=2}^{r=m-s+1} \binom{m-r+1}{s} \lambda^{m-r-s+1} (y_r + a_1 y_{r-1} + \dots + a_{r-2} y_1) \\ &= \lambda z^{(s)} + z^{(s-1)} - y_1 \sum_{r=1}^{r=m-s+1} \binom{m-r+1}{s} \lambda^{m-r-s+1} a_{r-1}. \end{aligned}$$

Hence, if λ is a $(s+1)$-ple root of the equation (ii), so that

$$\sum_{r=1}^{r=m-t+1} \binom{m-r+1}{t} \lambda^{m-r-t+1} a_{r-1} = 0, \quad (t = 0, 1, \dots, s),$$

then
$$\begin{aligned} z' &= \lambda z, \\ z'^{(1)} &= \lambda z^{(1)} + z, \\ z'^{(2)} &= \lambda z^{(2)} + z^{(1)}, \\ &\dots \quad \dots \quad \dots \quad \dots \\ z'^{(s)} &= \lambda z^{(s)} + z^{(s-1)}. \end{aligned}$$

So, if μ is a $(t+1)$-ple root of (ii), distinct from λ, and if

$$w = \sum_1^m \mu^{m-r} (y_r + a_1 y_{r-1} + \dots + a_{r-1} y_1),$$

then
$$\begin{aligned} w' &= \mu w, \\ w'^{(1)} &= \mu w^{(1)} + w, \\ &\dots \quad \dots \quad \dots \quad \dots \\ w'^{(t)} &= \mu w^{(t)} + w^{(t-1)}. \end{aligned}$$

That the m linear functions, z's, w's, ..., thus introduced are linearly independent may be shown as follows. Suppose, if possible, that a linear relation

$$\alpha z^{(s)} + \beta z^{(s-1)} + \dots + \gamma w^{(t)} + \delta w^{(t-1)} + \dots + \dots = 0$$

connects them. Then

$$\alpha z'^{(s)} + \beta z'^{(s-1)} + \dots + \gamma w'^{(t)} + \delta w'^{(t-1)} + \dots = 0,$$

or $\alpha (\lambda z^{(s)} + z^{(s-1)}) + \beta (\lambda z^{(s-1)} + z^{(s-2)}) + \dots + \gamma (\mu w^{(t)} + w^{(t-1)})$
$$+ \delta (\mu w^{(t-1)} + w^{(t-2)}) + \dots = 0.$$

Hence $\qquad \alpha z^{(s-1)} + \beta z^{(s-2)} + \dots + \gamma (\lambda - \mu) w^{(t)} + \dots = 0.$

Since $\lambda-\mu$ is different from zero, all the z's may thus be got rid of without the w's, &c., at the same time disappearing. Hence a linear relation such as that assumed above involves a linear relation among the functions of a single set, say the w's; and this would involve w itself being identically zero; which is not the case, since the coefficient of y_m in w is unity.

Finally then, the substitution S, so far as it affects the y's, *i.e.* the substitution (i), can be expressed in the form

$$\left.\begin{array}{lll} z' = \lambda z, & w' = \mu w, & u' = \nu u, \;\ldots \\ z'^{(1)} = \lambda z^{(1)} + z, & w'^{(1)} = \mu w^{(1)} + w, & \ldots \;\ldots \;\ldots \\ z'^{(2)} = \lambda z^{(2)} + z^{(1)}, & \ldots \;\ldots \;\ldots \;\ldots & \\ \ldots \;\ldots \;\ldots \;\ldots & w'^{(t)} = \mu w^{(t)} + w^{(t-1)}, & \\ z'^{(s)} = \lambda z^{(s)} + z^{(s-1)}, & & \end{array}\right\}, \qquad \text{(i)}'$$

where λ, μ, ν, ... are the distinct roots of

$$\lambda^m + a_1\lambda^{m-1} + a_2\lambda^{m-2} + \ldots + a_m = 0; \qquad \text{(ii)}'$$

$s+1$, $t+1$, ... are their multiplicities; and

$$z^{(p)} = \frac{1}{p!}\frac{\partial_p}{\partial\lambda^p}\sum_1^m \lambda^{m-r}(y_r + a_1 y_{r-1} + \ldots + a_{r-1}y_1),$$

$$w^{(q)} = \frac{1}{q!}\frac{\partial^q}{\partial\mu^q}\sum_1^m \mu^{m-r}(y_r + a_1 y_{r-1} + \ldots + a_{r-1}y_1),$$

... ...

The equations (i)′ are formally true whatever special values may be given to the A's. If the special values are such that z does not vanish identically, it is clear that no one of the functions $z^{(1)}, z^{(2)}, \ldots, z^{(s)}$ can vanish; and the same is true for the sets of functions in the other columns. It may, however, happen that, for some sets of values of the A's, $z, z^{(1)}, \ldots, z^{(k)}$ vanish identically, but that $z^{(k+1)}$ does not. No one of the functions $z^{(k+2)}, \ldots, z^{(s)}$ can then vanish; and for such special values the equations of the first column reduce to

$$\begin{array}{l} z'^{(k+1)} = \lambda z^{(k+1)}, \ldots \\ z'^{(k+2)} = \lambda z^{(k+2)} + z^{(k+1)}, \\ \ldots \;\ldots \;\ldots \;\ldots \;\ldots \\ z'^{(s)} = \lambda z^{(s)} + z^{(s-1)}. \end{array}$$

It is clear that in no case can the replacing of the arbitrary A's by special values lead to a linear relation between non-vanishing functions occurring in different columns.

2. Suppose, now, that the z of the preceding paragraph has been actually calculated as a function of the original variables, or x's, and the original arbitrary coefficients, or A's. We may then write

$$z = B_1 z_1 + B_2 z_2 + \ldots + B_p z_p,$$

where the z's are independent linear functions of the x's, and the B's are independent linear functions of the A's; with coefficients, in each case, which are functions of the coefficients a_{st} of the substitution S. In fact, if either the z's or the B's were not linearly independent, we could represent z as the sum of a smaller number of similar terms; and this process could be continued till the conditions in question are satisfied.

A similar form may be obtained for $z^{(1)}$, so that we may write

$$z^{(1)} = C_1 \zeta_1 + C_2 \zeta_2 + \ldots + C_q \zeta_q.$$

Let the A's be replaced by n linear functions of themselves, of which $B_1, B_2, \ldots, B_p$ are chosen for the first p. Then

$$C_r = \gamma_{r1} B_1 + \gamma_{r2} B_2 + \ldots + \gamma_{rp} B_p + D_r,$$

where D_r is independent of $B_1, B_2, \ldots, B_p$. Hence, if

$$z_a^{(1)} = \gamma_{ra} \zeta_1 + \gamma_{ra} \zeta_2 + \ldots + \gamma_{qa} \zeta_q,$$

$$z^{(1)} = B_1 z_1^{(1)} + B_2 z_2^{(1)} + \ldots + B_p z_p^{(1)} + D_1 \zeta_1 + D_2 \zeta_2 + \ldots + D_q \zeta_q.$$

The function $D_1 \zeta_1 + \ldots$ may again be expressed in the form

$$B_{p+1} z_{p+1}^{(1)} + B_{p+2} z_{p+2}^{(1)} + \ldots + B_{p+p_1} z_{p+p_1}^{(1)},$$

where $B_{p+1}, B_{p+2}, \ldots, B_{p+p_1}$ are linearly independent of each other, and of the preceding B's; and $z_{p+1}^{(1)}, z_{p+2}^{(1)}, \ldots, z_{p+p_1}^{(1)}$ are linearly independent of each other.

If the A's are chosen so that

$$B_1 = B_2 = \ldots = B_p = 0,$$

then the relation $$z'^{(1)} = \lambda z^{(1)} + z$$

gives $$B_{p+1} z_{p+1}'^{(1)} + \ldots + B_{p+p_1} z_{p+p_1}'^{(1)} = \lambda (B_{p+1} z_{p+1}^{(1)} + \ldots + B_{p+p_1} z_{p+p_1}^{(1)}).$$

It follows from this that $z^{(1)}_{p+1}, z^{(1)}_{p+2}, \ldots, z^{(1)}_{p+p_1}$ are linearly independent of $z^{(1)}_1, z^{(1)}_2, \ldots, z^{(1)}_p$. For an identical relation such as

$$Pz^{(1)}_1 + Qz^{(1)}_2 + \ldots + Rz^{(1)}_{p+1} + Sz^{(1)}_{p+2} + \ldots = 0$$

involves $P(\lambda z^{(1)}_1 + z_1) + Q(\lambda z^{(1)}_2 + z_2) + \ldots + R\lambda z^{(1)}_{p+1} + S\lambda z^{(1)}_{p+2} + \ldots = 0,$

or $$Pz_1 + Qz_2 + \ldots = 0,$$

which can only hold if P, Q, ... are zero. This, moreover, shows that $z^{(1)}_1, z^{(1)}_2, \ldots, z^{(1)}_p$ are linearly independent of each other. It must, however, be noticed that $z^{(1)}_{p+1}, z^{(1)}_{p+2}, \ldots, z^{(1)}_{p+p_1}$ are not in general linearly independent of $z_1, z_2, \ldots, z_p$.

Continuing this process, we may now express each $z^{(r)}$, $(r=1, 2, \ldots, s)$, in the form

$$z^{(r)} = B_1 z^{(r)}_1 + \ldots + B_{p+1} z^{(r)}_{p+1} + \ldots + B_{p+p_1+1} z^{(r)}_{p+p_1+1} + \ldots$$
$$\ldots + B_{p+p_1+\ldots+p_r} z^{(r)}_{p+p_1+\ldots+p_r};$$

where, from the mode of expression, the B's are necessarily independent, while it may be shown as above that the $p+p_1+\ldots+p_r$ functions involved in $z^{(r)}$ are linearly independent.

In the general equations that correspond to the root λ of the equation (ii), let all the B's be made zero except $B_1, B_2, \ldots, B_p$. The system of $s+1$ equations, with p arbitrary coefficients,

$$\left.\begin{aligned} B_1 z_1' + B_2 z_2' + \ldots + B_p z_p' &= \lambda (B_1 z_1 + B_2 z_2 + \ldots + B_p z_p), \\ B_1 z_1'^{(r)} + B_2 z_2'^{(r)} + \ldots + B_p z_p'^{(r)} & \\ = \lambda (B_1 z^{(r)}_1 + \ldots + B_p z^{(r)}_p) &+ B_1 z^{(r-1)}_1 + \ldots + B_p z^{(r-1)}_p, \end{aligned}\right\} \quad (r = 1, 2, \ldots, s), \quad \text{(iii)}$$

is thus obtained.

That the $p(s+1)$ functions contained in these equations, and therefore also the equations themselves, are linearly independent may be verified as before.

Next, make all the B's zero except $B_{p+1}, B_{p+2}, \ldots, B_{p+p_1}$. The system of s equations with p_1 arbitrary coefficients

$$\left.\begin{aligned} B_{p+1} z'^{(1)}_{p+1} + \ldots + B_{p+p_1} z'^{(1)}_{p+p_1} &= \lambda (B_{p+1} z^{(1)}_{p+1} + \ldots + B_{p+p_1} z^{(1)}_{p+p_1}), \\ B_{p+1} z'^{(r)}_{p+1} + \ldots + B_{p+p_1} z'^{(r)}_{p+p_1} & \\ = \lambda (B_{p+1} z^{(r)}_{p+1} + \ldots + B_{p+p_1} z^{(r)}_{p+p_1}) &+ B_{p+1} z^{(r-1)}_{p+1} + \ldots + B_{p+p_1} z^{(r-1)}_{p+p_1}, \end{aligned}\right\} \quad (r = 2, 3, \ldots, s), \quad \text{(iv)}$$

is then obtained. These equations and the functions contained in them may, as before, be shown to be independent among themselves. They are, however, not necessarily independent of the functions contained in equations (iii).

The totality of the functions contained in equations (iv) arise from $\zeta^{(s)}$, or

$$B_{p+1}z^{(s)}_{p+1}+B_{p+2}z^{(s)}_{p+2}+\dots+B_{p+p_1}z^{(s)}_{p+p_1},$$

and the results of operating repeatedly on this function with S. Hence, if

$$\zeta^{(s)}=\zeta^{(s)}_1+C_1\zeta^{(s)}_{p+1}+C_2\zeta^{(s)}_{p+2}+\dots+C_{\varpi_1}\zeta^{(s)}_{p+\varpi_1},$$

where $C_1, C_2, \dots, C_{\varpi_1}$ are independent linear functions of the B's; $\zeta^{(s)}_{p+1}, \zeta^{(s)}_{p+2}, \dots, \zeta^{(s)}_{p+\varpi_1}$ independent linear functions of the variables; and $\zeta^{(s)}-\zeta^{(s)}_1$ denotes what $\zeta^{(s)}$ reduces to when all the functions in equations (iii) are made zero, the terms arising from $\zeta^{(s)}_1$ in equations (iv) destroy each other in virtue of equations (iii). Hence equations (iv) may be replaced by the system of s equations

$$\left.\begin{aligned} C_1\zeta'^{(1)}_{p+1}+\dots+C_{\varpi_1}\zeta'^{(1)}_{p+\varpi_1}&=\lambda\,(C_1\zeta^{(1)}_{p+1}+\dots+C_{\varpi_1}\zeta^{(1)}_{p+\varpi_1}),\\ C_1\zeta'^{(r)}_{p+1}+\dots+C_{\varpi_1}\zeta'^{(r)}_{p+\varpi_1}&\\ =\lambda\,(C_1\zeta^{(r)}_{p+1}+\dots+C_{\varpi_1}\zeta^{(r)}_{p+\varpi_1})&+C_1\zeta^{(r-1)}_{p+1}+\dots+C_{\varpi_1}\zeta^{(r-1)}_{p+\varpi_1},\end{aligned}\right\}$$

$$(r=2, 3, \dots, s), \quad \text{(iv)}'$$

containing ϖ_1 arbitrary coefficients; in which the $\varpi_1 s$ variables are independent of each other, and of those in equations (iii).

Similarly, the system of $s-1$ equations with p_2 arbitrary coefficients resulting from making all the B's zero, except B_{p+p_1+1}, $B_{p+p_1+2}, \dots, B_{p+p_1+p_2}$, may be replaced by a system of $s-1$ equations with ϖ_2 arbitrary coefficients in which the variables are linearly independent of each other and of those in equations (iii) and (iv)$'$.

This process may be continued till finally the general system of $s+1$ equations which correspond to the root λ of the equation (ii), viz.,

$$z'=\lambda z,\quad z'^{(1)}=\lambda z^{(1)}+z,\quad \dots,\quad z'^{(s)}=\lambda z^{(s)}+z^{(s-1)},$$

is replaced by

a system of		$s+1$	equations	of similar form	with	p	arbitraries,		
"	"	s	"	"	"	ϖ_1	"		
"	"	$s-1$	"	"	"	ϖ_2	"		
...	...	...	...	...	...	...	...	...	...
"	"	1	"	"	"	ϖ_s	"		

such that the

$$p(s+1)+\varpi_1 s+\varpi_2(s-1)+\ldots+\varpi_s = n_\lambda$$

variables entering in these equations are linearly independent. Moreover, whatever values the original arbitrary coefficients $A_1, A_2, \ldots, A_n$ may have, any one of the variables $z^{(r)}$ $(r = 0, 1, \ldots, s)$ entering in the general system can be expressed linearly in terms of these n_λ linearly independent variables.

From the general form obtained above for z it is clear that the integer p cannot be less than unity; but, in particular cases, any one or more of the integers $\varpi_1, \varpi_2, \ldots, \varpi_s$ may be zero.

The general system of $t+1$ equations corresponding to the root μ of equation (ii) may be similarly treated; and they will give rise to a similar set of systems of equations containing n_μ linearly independent variables. That these are linearly independent of the n_λ variables that arise from the root λ has already been shown (§ 1). Let

$$n' = n_\lambda + n_\mu + \ldots,$$

where the sum is extended to all the roots of equation (ii). Since the n' variables thus obtained are linearly independent, n' cannot be greater than n. On the other hand, since y_1, an absolutely arbitrary function of the original variables, can be expressed in terms of the z's, w's, ..., while the z's can be expressed in terms of the n_λ variables arising from the root λ, the w's in terms of the n_μ variables arising from the root μ, and so on, it follows that n' cannot be less than n. Hence

$$n' = n,$$

and the set of systems of equations, giving the effect of S upon the

$$n_\lambda + n_\mu + \ldots$$

new variables which have been constructed, is linearly equivalent to the original equations

$$x'_s = \Sigma a_{st} x_t, \quad (s,\ t = 1,\ 2,\ \ldots,\ n).$$

The process by which the sets of quantities

$$\left.\begin{matrix} \lambda, & s+1, & p, & \varpi_1, & \varpi_2, & \ldots, & \varpi_s; \\ \mu, & t+1, & & \ldots & \ldots & & \ldots; \\ \ldots & \ldots & & \ldots & \ldots & & \ldots \end{matrix}\right\} \qquad \text{(v)}$$

have been derived from the equations of the substitution S is a definite and unique one. Any two substitutions A and B for which these sets of quantities are the same are conjugate in the general linear group. For, if $u, v, w, \ldots$ are the new variables which arise from A when reduced as above, and if $U, V, W, \ldots$ are the corresponding new variables which arise from B, then the substitution C, or

$$U = u, \quad V = v, \quad W = w, \quad \ldots,$$

is evidently such that

$$C^{-1}BC = A.$$

On the other hand, if C is any linear substitution, the sets of quantities in question are obviously the same for B and $C^{-1}BC$. Hence the necessary and sufficient conditions that two linear substitutions should be conjugate in the general linear group are that the sets of quantities (v) should be the same for both.

3. Since, as stated above, the method of reduction here explained is put forward as a practicable one, whenever the characteristic equation of the substitution can be actually solved, it seems proper to give a numerical example in which the reduction is really carried out.

Let S be

$$\begin{aligned} x'_1 &= -2x_1 - x_2 - x_3 + 3x_4 + 2x_5, \\ x'_2 &= -4x_1 + x_2 - x_3 + 3x_4 + 2x_5, \\ x'_3 &= x_1 + x_2 - 3x_4 - 2x_5, \\ x'_4 &= -4x_1 - 2x_2 - x_3 + 5x_4 + x_5, \\ x'_5 &= 4x_1 + x_2 + x_3 - 3x_4. \end{aligned}$$

The linear functions $y_1, y_2, y_3, \ldots$, as obtained by direct calculation, are given by the following table:—

	y_1	y_2	y_3	y_4	y_5
x_1	A_1	$-2A_1-4A_2+A_3$ $-4A_4+4A_5$	$3A_1-A_2-2A_3$ $-A_4+A_5$	$-4A_1-12A_2+3A_3$ $-12A_4+12A_5$	$5A_1-11A_2-4A_3$ $-11A_4+11A_5$
x_2	A_2	$-A_1+A_2+A_3$ $-2A_4+A_5$	$-4A_1 \quad +4A_3$ $-8A_4+4A_5$	$-12A_1-4A_2+12A_3$ $-24A_4+12A_5$	$-32A_1-16A_2+32A_3$ $-64A_4+32A_5$
x_3	A_3	$-A_1-A_2$ $-A_4+A_5$	$2A_1+2A_2-A_3$ $+2A_4-2A_5$	$-3A_1-3A_2+2A_3$ $-3A_4+3A_5$	$4A_1+4A_2-3A_3$ $+4A_4-4A_5$
x_4	A_4	$3A_1+3A_2-3A_3$ $+5A_4-3A_5$	$3A_1+3A_2-3A_3$ $+7A_4-3A_5$	$9A_1+9A_2-9A_3$ $+17A_4-9A_5$	$15A_1+15A_2-15A_3$ $+31A_4-15A_5$
x_5	A_5	$2A_1+2A_2-2A_3$ $+A_4$	$-A_1-A_2+A_3$ $-5A_4+5A_5$	$-3A_1-3A_2+3A_3$ $-15A_4+11A_5$	$-17A_1-17A_2+17A_3$ $-49A_4+33A_5$

No linear relation connects the first two, three, or four y's, but between the first five the relation

$$y_5-2y_4-3y_3+4y_2+4y_1=0$$

holds. The equation for λ is therefore

$$\lambda^4-2\lambda^3-3\lambda^2+4\lambda+4\equiv(\lambda+1)^2(\lambda-2)^2=0.$$

The coefficients in z, w, $z^{(1)}$, $w^{(1)}$ are given by the table:—

	$\lambda=-1$	$\lambda=2$
$\lambda^3-2\lambda^2-3\lambda+4$	4	-2
$\lambda^2-2\lambda-3$	0	-3
$\lambda-2$	-3	0
$3\lambda^2-4\lambda-3$	4	1
$2\lambda-2$	-4	2

Hence $z=4y_1-3y_3+y_4$

$$=-9(A_1+A_2-A_3+A_4-A_5)(x_1+x_3)=-9B_1(x_1+x_3),$$

$$z^{(1)}=4y_1-4y_2+y_3$$

$$=B_1(15x_1+6x_3-9x_4-9x_5)+9A_3(x_1+x_3),$$

$$w=-2y_1-3y_2+y_4$$

$$=-9(A_1+A_2-A_3+2A_4-A_5)(x_2+x_5)=-9C_1(x_2+x_5),$$

$$w^{(1)}=y_1+2y_2+y_3$$

$$=-9(A_2+A_4-A_5)x_1-(6C_1-9A_2)x_2$$

$$+9C_1x_4+(3C_1-9A_1+9A_5)x_5$$

$$=C_1(-6x_2+9x_4+3x_5)-9(A_2+A_4-A_5)(x_1-x_2)$$

$$-9(A_4-A_5)(x_2+x_5).$$

Finally then, if

$$\xi_1 = -9(x_1+x_3),$$
$$\xi_2 = 15x_1+6x_3-9x_4-9x_5,$$
$$\xi_3 = -9(x_2+x_5),$$
$$\xi_4 = -6x_2+9x_4+3x_5,$$
$$\xi_5 = x_1-x_2,$$

the substitution S can be expressed in the form

$$\xi_1' = -\xi_1,$$
$$\xi_2' = \xi_1-\xi_2,$$
$$\xi_3' = 2\xi_3,$$
$$\xi_4' = \xi_3+2\xi_4,$$
$$\xi_5' = 2\xi_5.$$

The linear functions which enter in a canonical form of a linear substitution are never unique, unless the roots of the characteristic equation are all distinct. It is not then to be expected that a definite process for reducing a substitution to its canonical form will always give the reducing equations in as simple a form as possible. In fact, in the above example, the simpler relations

$$\xi_1 = x_1+x_3, \quad \xi_2 = x_3+x_4+x_5, \quad \xi_3 = x_2+x_5, \quad \xi_4 = -x_4-x_5, \quad \xi_5 = x_1-x_2$$

will equally well reduce the substitution.

4. If the substitution which it is proposed to reduce happens to be a substitution of finite period, the application of the method is particularly simple. The substitution (i)′ of § 1 can, in fact, be of finite period only when $s = t = \ldots = 0$. The roots of the equation (ii)′ are therefore in this case all distinct, and the calculation of

$$\sum_1^m (\lambda^r + a_1\lambda^{r-1} + \ldots + a_r)\, y_{m-r}$$

for each root of the equation at once gives the reducing substitution.