

ELEMENTARY THEORY OF INVARIANTS

by

PROF. HERMANN WEYL

1935-36

1935 1936  
FOR ADVANCED STUDY

Notes by Prof. Weyl and  
Dr. Leonard M. Blumenthal

INSTITUTE  
FOR ADVANCED STUDY  
LIBRARY

The Institute For Advanced Study

TABLE OF CONTENTS

FOREWORD

<u>Literature and History</u> -----	Page 1
-------------------------------------	-----------

INTRODUCTION

Outline of the Main Ideas and Propositions

1. The special and the general problem of invariant theory -----	5
2. The symmetric group and the symmetric functions -----	7
3. Vector invariants of the orthogonal group -----	9
4. The second main problem -----	13
5. Invariants depending on contravariant and covariant vectors -----	20
6. Invariants of forms and the general problem. The linear and the algebraic stage -----	22
7. Absolute and relative invariants. Covariant quantities -----	26

CHAPTER I. VECTOR INVARIANTS

A. Covariant Argument Vectors

1. Formal preliminaries about forms -----	31
2. First example: the symmetric group -----	34
3. Capelli's identity -----	37
4. Reduction of the first main problem by means of Capelli's identities -----	42
5. Second example: the unimodular group -----	46
6. Third example: the orthogonal group -----	47
7. Fourth example: the complex group -----	51
8. The extension theorem -----	56

B. Covariant and Contravariant Vectors

1. Trivial examples -----	64
2. The general method -----	65
3. A formal identity -----	71
4. The unimodular group -----	75
5. The group of step transformations -----	75
6. Euclidean space: enumeration of basic invariants -----	79
7. Proof of completeness. Additional remarks -----	88

C. The Second Main Theorem

1. Statement of the problem for the unimodular group -----	97
2. Capelli's formal congruence -----	100
3. Proof of the second main theorem for the unimodular group -----	102
4. The orthogonal and complex groups -----	104

CHAPTER II. GENERAL THEORY OF INVARIANTS

A. Formal Methods in Classical Invariant Theory

	Page
1. The classical theory of invariants -----	110
2. The symbolic method -----	113
3. The binary quadratic form -----	115
4. Irrational methods -----	116
5. Some simple general consequences of the symbolic method and the Capelli identity -----	125
6. The adjunction theorem -----	128

B. Theory of Binary Vector Invariants and the Spin Theory of Valence Bonds

7. The Clebsch-Gordan expansion theorem -----	130
8. A basis for binary vector invariants. Rumer's Theorem -----	135
9. The Burnside-Frobenius-Schur Theorem, in elementary disguise ----	138
10. Sketch of the quantum-mechanical approach. Angular momentum and spin -----	141
11. Symmetry-spin-degeneracy -----	149
12. Spin states and binary invariants. Computation of binding energies -----	153

# THEORY OF INVARIANTS

## Literature and History

As the theory of invariants has dragged on a feeble existence for a number of decades and only recently came to life again, spurred by modern developments in physics and group theory, the literature of the subject is not in satisfactory shape. I venture to recommend the following books:

J. H. Grace and A. Young, The algebra of invariants, Cambridge, 1903  
Glenn, The theory of invariants, Boston, 1915  
L. E. Dickson, Algebraic invariants, New York, 1913

dealing more or less with the classical theory of invariants for forms under the influence of the projective group of all homogeneous linear transformations.

A freer attitude as to the underlying group of transformations is taken in:

E. Study, Einleitung in die Theorie der Invarianten linearer Transformationen auf Grund der Vektorenrechnung, Braunschweig, 1923, and  
R. Weitzenböck, Invariantentheorie, Groningen, 1923.

They are more in keeping with the program of this course.

As to the history of our subject, I may refer to the first volume of Felix Klein's *Entwicklung der Mathematik in neunzehnten Jahrhundert*.

-----

I should like in a brief survey to enumerate some of the leading names. The theory of invariants originated in England about the middle of the nineteenth century as a genuine analytic instrument for describing configurations and their inner geometric relations in projective geometry. The functions and algebraic relations expressing them in terms of projective coordinates are to be invariant under all homogeneous linear transformations. Cayley first passed from the consideration of determinants to more general invariants. This procedure accounts for the title of his paper, *Memoire sur les Hypordeterminants* (Crelle 30 (1846)),

which one may look upon as the birth certificate of invariant theory. In his later famous Memoires on Quantics (1854-1860) he succeeds, among other things, in obtaining a complete set of invariants for cubic and biquadratic forms. His work was taken up in England by Sylvester and Salmon. Sylvester is one of the earliest European scholars who came over to this continent; he taught at Johns Hopkins University for some years, and there founded the first American mathematical journal: The American Journal of Mathematics. The pages of its first volumes are filled with papers on invariant theory from Sylvester's prolific pen. In Germany, Aronhold, Clebsch and Gordan became adherents and promoters of the new discipline. In Italy, Brioschi, Cremona, Beltrami, and Capelli were attracted to the subject. This early period has a formal character throughout: the development of formal processes and the actual computation of invariants stand to the fore. Almost all papers refer to one group, the continuous group of all homogeneous linear transformations.

Another impulse, in a somewhat different direction, came from number theory, more particularly from the arithmetic theory of binary quadratic forms. Here one had been led to consider not a continuous but a discrete group, the group of unimodular linear substitutions with integral coefficients. Gauss, in his Disquisitiones arithmeticae, studied equivalence of quadratic forms with respect to this group. Besides, and after Gauss, we have Jacobi in Germany and Hermite in France, as outstanding men in this line of investigation.

The formal period of classic invariant theory is followed by a more critical and conceptual one which solves the general problems of finiteness less by explicit computations than by developing suitable general notions and their general properties along such abstract lines as have lately come into fashion all over the whole field of algebra. Here there is only one man to mention, - Hilbert. His papers (Mathematische Annalen 36 (1890), 42 (1892)) mark a turning

point in the history of invariant theory. He solves the main problems and thus almost kills the whole subject.

But its life lingers on, however flickering, during the next decades. A. Hurwitz makes a new and important contribution by introducing integral processes extending over the group manifold (1897); in England A. Young, working more or less alone in this field, obtains far-reaching results about the representations of the symmetric group and uses them for invariant-theoretic purposes (1900 and later). In recent times, however, the tree of invariant theory shows new life, and has begun to blossom again, chiefly as a consequence of the interest in invariant-theoretic questions awakened by the revolutionary developments in mathematical physics (relativity theory and quantum mechanics) but also due to the connection of invariant theory with the extension of the theory of representations to continuous groups and algebras. This has become quite a central subject in our present day mathematics: as to its bearing upon invariant theory, the most important name to be mentioned is that of I. Schur (several memoirs in the Berliner Berichte of 1925).

The rise of projective geometry made such an overwhelming impression on the geometers of the first half of the nineteenth century that they tried to fit all geometric considerations into the projective scheme. The narrowing down of the projective group to the affine group or to the group of Euclidean motions of metric geometry was accordingly effected by adjoining some so-called "absolute" entities: the plane at infinite, the absolute involution. The same attitude is expressed when one treats metric geometry in vector space by allowing arbitrary affine coordinate systems and their transformations, and adding the fundamental metric form  $x_1^2 + x_2^2 + \dots + x_m^2$  as something absolute instead of sticking to the metrical equivalent Cartesian coordinate system only

and the corresponding group of orthogonal transformations. This way of procedure, as it easily admits of extension into infinitesimal geometry, has remained in use with great success, particularly for the purpose of general relativity theory. In group theory it amounts to considering each group of linear transformations as a subgroup of and in relation to the total linear group.

The dictatorial regime of the projective idea in geometry was first successfully broken by the German astronomer and geometer Möbius, but the classical and all-embracing document of the democratic platform in geometry establishing the group of transformations as the ruling principle in any kind of geometry and yielding equal rights of independent consideration to each and any such group, is given by F. Klein's Erlangen program (1872). The adjustment of invariant theory to this standpoint has been slow; it could not be made without recognizing that the study of the groups themselves and their representations necessarily has to precede the study of their invariants. We come across a similar phenomenon in all branches of mathematics; for instance, functions, let us say algebraic functions, had long been studied before the field of variability for their arguments, the manifolds, the Riemannian surfaces, the soil on which they grow, were subjected to an independent investigation. We shall follow the modern standpoint, which treats all groups on an equal footing and does not, at the outset, concede royal prerogatives to the full linear group.

INTRODUCTION: OUTLINE OF THE MAIN IDEAS AND PROPOSITIONS

1. The special and the general problem of invariant theory

An invariant is a function whose value does not change when an arbitrary transformation  $S$  of a given group  $\gamma$  of transformations is performed on its arguments. In the beginning we restrict ourselves to groups of homogeneous linear transformations, so that the arguments of the functions we are going to study may be described as vectors in a vector space (of a finite number of dimensions). Under such circumstances it seems natural first of all to consider forms, i.e., functions which are homogeneous polynomials with respect to the components of the several independent arguments (vectors)  $x, y, \dots$  on which the function depends. These arguments  $x, y, \dots$  may vary independently within a vector space of  $m, n, \dots$  dimensions, respectively, and the given form  $f(x, y, \dots)$  will be of certain degree  $\mu, \nu, \dots$  with respect to each of the argument vectors  $x, y, \dots$ . Indeed, this property of being a form of certain assigned degrees, is invariant under arbitrary linear transformation of the arguments (i.e. when the components of  $x$  undergo a linear transformation  $S$  among themselves, the components of  $y$  a linear transformation  $T$ , etc.).

But even with these restrictions one can distinguish a more elementary problem of invariants from the general one. The elementary problem refers to a given group  $\gamma$  of linear transformations  $S$  in an  $n$ -dimensional vector space  $\mathcal{R}$ ;  $S$  changes the arbitrary vector  $x$  into  $x' = Sx$ . The functions  $f$  to be considered depend on a certain number of independent vectors  $x, y, \dots$  in the same space  $\mathcal{R}$ .  $f$  is carried into a new function  $f' = Sf$  by means of the transformation  $S$ , the transform  $f'$  being defined by the equation

$$(1.1) \quad f'(Sx, Sy, \dots) = f(x, y, \dots).$$

One will observe that here all the argument vectors undergo the same transforma-

tion  $S$ , that is to say, they are transformed cogrediently. The above definition (1.1) rather than the one which might have suggested itself as more natural at first glance, namely,

$$Sf(x, y, \dots) = f(Sx, Sy, \dots),$$

is in agreement with the composition rule

$$T(Sf) = (TS)f$$

if the compound transformation  $TS$  operating on an arbitrary vector  $x$  is defined by

$$x' = Sx, \quad x'' = Tx', \quad x'' = (TS)x.$$

$f$  is invariant under  $\gamma$ , if  $Sf = f$  for all transformations  $S$  of  $\gamma$ .

The general problem, on the other hand, deals with a given abstract group  $\gamma$ , and a certain number of representations

$$s \rightarrow S, \quad s \rightarrow T, \dots,$$

of  $\gamma$  by linear transformations  $S, T, \dots$  in an  $m, n, \dots$ -dimensional vector space, respectively.  $s$  is an arbitrary element of our group  $\gamma$ ;  $S, T, \dots$  denote the linear transformations corresponding to  $s$  in the given representations

$\gamma, \mathcal{Z}, \dots$ ; if a more detailed notation is desirable, the element  $s$  may be added as an index:  $S_s, T_s, \dots$ . Let  $x, y, \dots$  be arbitrary vectors varying in the corresponding vector spaces of  $m, n, \dots$  dimensions. A function

$f(x, y, \dots)$  gives rise to a transform  $f' = sf$  as defined by the equation

$$f'(S_s x, T_s y, \dots) = f(x, y, \dots).$$

When  $sf = f$  for all elements  $s$  of the given group  $\gamma$ ,  $f$  is called an invariant for the representations  $\gamma, \mathcal{Z}, \dots$  of  $\gamma$ .

The elementary problem is of course a particular case of this more general conception.

To leave no gap, I repeat the definition of the terms "group" and "representation". A group  $\gamma$  is a system of elements such that any two elements

$a, b$  give rise to a compound element  $ab = a \cdot b$ , this composition fulfilling the associative law

$$(ab)c = a(bc).$$

The group contains a unit element  $1$  such that  $1a = a1 = a$ , for every element  $a$ ; and each  $a$  has an inverse  $a^{-1}$  within the group:  $aa^{-1} = a^{-1}a = 1$ .

A group  $\gamma$  is represented by linear transformations  $S$  in an  $n$ -dimensional vector space if there is associated an  $S = S_a$  with every element  $a$  of  $\gamma$ :  $a \rightarrow S_a$ , in such a way that composition of elements induces composition of the corresponding transformations:

$$ab \rightarrow S_a S_b.$$

We add that the unit element  $1$  is to be represented by the identical transformation  $I$  which carries every vector  $x$  into itself:  $Ix = x$ . A representation in  $n$ -dimensional vector space is called "of degree  $n$ ".

## 2. The symmetric group and the symmetric functions

It will be convenient before going on to illustrate our general notions and problems by some familiar examples.

In the theory of algebraic equations one is led to consider symmetric functions  $f(x_1, x_2, \dots, x_n)$  of  $n$  arguments  $x_1, \dots, x_n$ , i.e. functions invariant under the group  $\gamma$  of all  $n!$  possible permutations of the  $n$  arguments.

These permutations obviously are linear transformations of the  $n$ -dimensional vector  $x = (x_1, x_2, \dots, x_n)$ . The elementary symmetric functions

$\phi_1, \phi_2, \dots, \phi_m$  are the coefficients of the polynomial  $\Phi(t)$  of the indeterminate  $t$ :

$$(2.1) \quad \Phi(t) = (t-x_1)(t-x_2)\dots(t-x_n) = t^m - \phi_1(x) t^{m-1} + \phi_2(x) t^{m-2} - \dots \pm \phi_m(x)$$

whose roots are  $x_1, x_2, \dots, x_n$ .

$$\varphi_1(x) = \sum_i x_i,$$

$$\varphi_2(x) = \sum_{i < k} x_i x_k,$$

$$\varphi_3(x) = \sum_{i < k < p} x_i x_k x_p$$

...

$$\varphi_n(x) = x_1 x_2 \cdots x_n.$$

The main fact concerning symmetric functions is their being expressible in terms of the elementary symmetric functions  $\varphi_i(x)$ ; or more explicitly:  $f$

$f(x_1, x_2, \dots, x_n)$  being any symmetric function of  $n$  arguments  $x_1, \dots, x_n$  there exists a function  $F(\xi_1, \xi_2, \dots, \xi_n)$  of  $n$  arguments  $\xi_1, \xi_2, \dots, \xi_n$  such that

$$f(x) = F(\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)).$$

We say, the functions  $\varphi_i(x)$ , ( $i = 1, 2, \dots, n$ ) form a functional basis for the symmetric functions. This is almost trivial if we take the notion of function in its widest scope; for then it simply states the fact that the values of the elementary symmetric functions  $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$  determine the values of the arguments  $x_1, x_2, \dots, x_n$  uniquely but for their order. Indeed, the equation (2.1),  $\Phi(t) = 0$  determines uniquely the set of its roots. But if  $f(x_1, x_2, \dots, x_n)$  is a polynomial in  $x_1, x_2, \dots, x_n$ , the question arises whether  $f$  is expressible in terms of the functions  $\varphi_i(x)$ , ( $i=1, 2, \dots, n$ ), in the same algebraic fashion; that is, whether  $F$  is also a polynomial. This is stated to be true by the so-called fundamental theorem of symmetric functions: the functions  $\varphi_i(x)$  ( $i = 1, 2, \dots, n$ ), constitute an integral rational basis for the symmetric forms. The restriction of the hypothesis, namely that the given symmetric function is integral-rational, is thus counterbalanced by a corresponding narrowing of the inference: the functional expression  $F$  of  $f$  by means of the basis  $\varphi$  is also of integral rational nature. Thus the "algebraic theorem" referring to forms  $f$  is not a particular case of the "functional theorem"

in which the functional dependence in  $f$  and  $F$  is understood in the widest possible sense; on the contrary, it is the algebraic theorem alone that needs an elaborate proof.

A similar situation prevails in many cases. "All invariants are expressible in terms of a finite number among them": this so-called first main theorem of invariant theory seems to be suggested by our present example. We cannot claim its validity for every group  $\gamma$  of linear transformations; rather, it will be our chief task to investigate for each particular group whether a finite number of its invariants form a basis or not; the answer, to be sure, will turn out affirmative in the most important cases. In those cases, and here is the point I wish to emphasize, one will find the purely functional part -- asserting that the values of all invariants are determined by the values of the basic invariants -- almost trivial; the essential difficulties lie in the algebraic part only.

### 3. Vector invariants of the orthogonal group

I choose the group which rules the classic Euclidean geometry, the group  $\gamma$  of orthogonal transformations, as a further instance to throw more light on this point. The defining property of an orthogonal transformation  $S$  is its leaving invariant the scalar product of two vectors  $x, y$ :

$$(x y) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n ,$$

when performed on both vectors cogrediently. Let us consider functions of two arbitrary vectors  $x, y$ , which are invariant under all (proper and improper) orthogonal transformations. The first fundamental theorem asserts that the scalar products which may be constructed for these two vectors, namely the three products

$$(3.1) \quad (x x) , (x y) = (y x) , (y y) ,$$

form a basis. The functional part of this statement is nothing else than the fundamental proposition about the congruence of triangles: "Two triangles ABC and A'B'C' are congruent when two sides and the included angle of one triangle coincide with the corresponding elements of the other", or "Two figures each consisting of a couple of vectors  $x, y$ , and  $x', y'$ , are congruent, i.e. are changeable into each other by an appropriate orthogonal transformation if and only if

$$(x, x) = (x', x'), \quad (x, y) = (x', y'), \quad (y, y) = (y', y').$$

Deeper lying but still true is the algebraic proposition that every orthogonally invariant form  $f(x, y)$  is expressible as a polynomial or an "aggregate" (as one often says) of the three scalar products (3.1). The proof, as it will be given in requires formal tools entirely different from those on which the proof of the congruence theorem rests. If we follow the latter procedure to obtain the expression of an arbitrary invariant form  $f(x, y)$  in terms of the three scalar products we get will contain irrationalities, or at least denominators will be present whose disappearance cannot be ascertained in this way.

Indeed, how does one go about demonstrating the congruence theorem in analytic  $n$ -dimensional geometry? The process is as follows: orthogonal transformation amounts to replacing one Cartesian coordinate system by another. Let the two vectors  $x, y$  be numerically fixed. One may choose a new Cartesian coordinate system  $e^1, e^2, \dots, e^n$  such that  $x$  lies in the direction of the first fundamental vector  $e^1$  and  $y$  lies in the plane  $(e^1, e^2)$ . Then

$$\begin{aligned} x &= \alpha e^1, \\ y &= \beta e^1 + \gamma e^2. \end{aligned}$$

The integral rational invariant  $f(x, y)$  then equals  $f(x', y')$  where

$$\begin{aligned}x' &= (\alpha, 0, 0, \dots, 0), \\y' &= (\beta, \gamma, 0, \dots, 0).\end{aligned}$$

Thus,  $f(x', y')$  is a polynomial in the three quantities  $\alpha, \beta, \gamma$ . We have

$$\begin{aligned}\alpha^2 &= (x'x') = (xx), \\ \alpha \cdot \beta &= (x'y') = (xy), \\ \beta^2 + \gamma^2 &= (y'y') = (yy); \end{aligned}$$

hence

$$\alpha = \sqrt{(xx)}, \quad \beta = \frac{(xy)}{\sqrt{(xx)}}, \quad \gamma = \sqrt{\frac{(xx)(yy) - (xy)^2}{(xx)}}.$$

In this way, square roots and the denominator  $(xx)$  creep in.

It is fairly easy, however, to get rid of the square roots. We found  $f(x, y)$  equal to a certain polynomial  $F$  of the quantities  $\alpha, \beta, \gamma$ . Invariance of  $f$  for the particular orthogonal transformations which consist in changing the direction of the first or the second fundamental axis shows that  $F$  remains unaltered under the two substitutions

$$(1) \quad \gamma \rightarrow -\gamma; \quad (2) \quad \alpha \rightarrow -\alpha, \quad \beta \rightarrow -\beta.$$

The polynomial  $F$  is a linear combination of monomials

$$M = \alpha^a \beta^b \gamma^c;$$

because of the invariance just mentioned the exponent  $c$  must be even in all terms of  $F$  and the two exponents  $a$  and  $b$  of equal parity, i.e. either both even, or both odd. According to the two cases  $M$  is a monomial of the squares  $\alpha^2, \beta^2, \gamma^2$  or  $\alpha\beta$  times such a monomial. Hence  $F$  can be written as a polynomial in

$$\alpha^2, \beta^2, \gamma^2 \quad \text{and} \quad \alpha\beta \quad \text{or in}$$

$$(xx), (xy), (yy), \frac{(xy)^2}{(xx)}.$$

$F$  is thus rationally expressible by the scalar products with a power of  $(xx)$  as denominator.

In a similar manner one may find a rational expression of  $f(x, y)$  in terms of the scalar products containing a power of  $(yy)$  as its denominator. No direct way is visible by which, from these two expressions, one with a power of  $(xx)$ , the other with a power of  $(yy)$  as denominator, one may derive an expression without denominator. See, however, the next section.

The problem of orthogonal invariants may be generalized to an arbitrary number, let us say  $h$  argument vectors  $x, y, \dots, z$ . The result is analogous. The symmetric matrix of all the scalar products

$$(3.2) \quad \left| \begin{array}{cccc} (xx) & (xy) & \dots & (xz) \\ (yx) & (yy) & \dots & (yz) \\ \cdot & \cdot & \cdot & \cdot \\ (zx) & (zy) & \dots & (zz) \end{array} \right|$$

is a complete table of basic invariants.

This suggests the possibility of assigning to a given group  $\gamma$  of linear transformations a finite number of typical basic invariants independent of the number of argument vectors to be considered. Such a table would consist of certain invariants depending on some "typical" argument vectors  $u, v, \dots$ ; and it would yield a basis for the invariants of an arbitrary number of argument vectors  $x, y, z, \dots$  after one substituted those argument vectors  $x, y, z, \dots$  in all possible combinations (repetitions not excluded) for the typical ones  $u, v, \dots$ . In this sense the orthogonal group possesses this scalar product  $(u, v)$  as its only typical basic invariant. For one gets a basis of invariants of  $h$  independent vectors  $x, y, \dots, z$ , whatever this number  $h$  may be, in forming all the scalar products.

#### 4. The second main problem; relations between basic invariants.

When one is called upon to express certain functions like the invariants in terms of given quantities like the basic invariants, it is essential to the question whether these quantities are dependent or not. Thus the basic invariants for the symmetric group in § 2, the elementary symmetric functions

$\varphi_i(x)$ , are independent in the strict functional sense that they can assume simultaneously arbitrarily assigned values  $a_i : \varphi_i(x) = a_i$ . Indeed the components  $x_i$  of the vector  $x$  are to be taken as the roots of the equation

$$t^m - a_1 t^{m-1} + a_2 t^{m-2} \dots \pm a_m = 0,$$

with the given coefficients  $a_i$ ; the main theorem about algebraic equations guarantees their existence. The algebraic independence, the fact that there exists no rational relation among the functions  $\varphi_i(x)$  is an immediate consequence of this strict functional independence. It is desirable though to have a purely algebraic proof for this purely algebraic proposition: that a polynomial

$F(\varphi_1, \varphi_2, \dots, \varphi_m)$  of  $n$  independent variables  $\varphi_1, \varphi_2, \dots, \varphi_m$  vanishes identically provided it vanishes identically in  $x_1, x_2, \dots, x_m$  after the  $\varphi_i$ 's are replaced by the elementary symmetric functions  $\varphi_i(x)$ . Such a proof becomes indispensable when one operates in an arbitrary number field (as the abstract algebraists are wont to do) rather than in the domain of ordinary real or complex numbers.

For the study of relations one always has to distinguish two standpoints, according as the  $\varphi_i$  are taken as independent variables or as some functions, here the elementary symmetric functions  $\varphi_i(x)$ , of other variables

$(x) = (x_1, x_2, \dots, x_m)$ , (first and second standpoint). Identical vanishing for the first and second standpoint shall be denoted by  $\equiv \circ$  and  $= \circ$ , respectively.

We give a demonstration by means of induction from  $n-1$  to  $n$ . On cutting off

the one variable  $x_1 = z$  and denoting the elementary symmetric functions of the remaining variables  $x_2, \dots, x_m$  by  $\varphi'_1(x), \dots, \varphi'_{m-1}(x)$  one has the recursive formulae:

$$(4.1) \quad \begin{aligned} \varphi_1(x) &= z + \varphi'_1(x), \\ \varphi_2(x) &= z \varphi'_1(x) + \varphi'_2(x), \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \varphi_m(x) &= z \varphi'_{m-1}(x). \end{aligned}$$

Let  $F(\varphi_1, \varphi_2, \dots, \varphi_m) \equiv 0$ , be a polynomial which vanishes after the "substitution"  $\varphi_i = \varphi_i(x)$  -- in contradiction to our proposition. We then find

$$F(z + \varphi'_1(x), z \varphi'_1(x) + \varphi'_2(x), \dots, z \varphi'_{m-1}(x)) = 0$$

identically in  $z; x_2, \dots, x_m$ . On putting  $z = 0$  one gets

$$(4.2) \quad F(\varphi'_1(x), \dots, \varphi'_{m-1}(x), 0) = 0$$

Supposing as we did  $\varphi'_1(x), \dots, \varphi'_{m-1}(x)$  to be algebraically independent, we infer from (4.2):

$$F(\varphi_1, \dots, \varphi_{m-1}, 0) \equiv 0,$$

or the polynomial  $F(\varphi_1, \dots, \varphi_m)$  of the independent variables  $\varphi_1, \dots, \varphi_m$  must be of the form:  $\varphi_m \cdot G(\varphi_1, \dots, \varphi_m)$ . The total degree of  $G$  is less by one than that of  $F$ . The identity

$$F(\varphi_1(x), \dots, \varphi_m(x)) = 0$$

at once leads to the "lower" equation

$$G(\varphi_1(x), \dots, \varphi_m(x)) = 0$$

since the other factor  $\varphi_m(x)$  in  $F(\varphi(x)) = \varphi_m(x) \cdot G(\varphi(x))$

does not vanish identically. Iteration of this process would result in identities of lower and lower degree -- but this cannot go on indefinitely unless

$F(\varphi) \equiv 0$ . This proof holds whatever the underlying number-field may be.

One cannot expect that this same situation, which we encountered here also in the case of the symmetric group, will prevail in general; there may exist

algebraic dependences among the basic invariants though their set is not redundant. This happens for instance if  $\mathcal{G}$  is the alternating rather than the symmetric group containing only the even permutations of the arguments  $x_1, x_2, \dots, x_m$ . A set of basic invariants for the alternating group consists of the elementary symmetric functions  $\varphi_1(x), \dots, \varphi_m(x)$  together with the "difference product"

$$\Delta(x) = \prod_{i < k} (x_i - x_k) .$$

The square  $\Delta^2$ , the "discriminant", is a symmetric function and therefore is expressible by means of the  $\varphi_i(x)$ . This relation is the "only" one to which our set of invariants is bound.

We prove now that the invariants  $\varphi_1(x), \dots, \varphi_m(x); \Delta(x)$  form a finite basis for the alternating group.

A form  $f(x_1, \dots, x_m)$  that is invariant with respect to the even permutations is changed into the same second form  $f'$  by all odd permutations. The sum  $f + f' = F$  is symmetric whereas  $f - f' = g$  is alternating, i.e. changes its sign under the influence of transposition of two variables and therefore vanishes if the values of two variables coincide, e.g. for  $x_2 = x_1$ . The polynomial  $g$  therefore must contain the factor  $x_2 - x_1$ ; whence

$$g(x_1, x_2, x_3, \dots) = (x_2 - x_1) \cdot g^*(x_1, x_2, x_3, \dots) ,$$

as one readily sees on considering  $g$  as the polynomial of  $x_2$  only and substituting  $x_2 = z + x_1$ . Furthermore  $g$  vanishes identically for  $x_3 = x_1$ ; and therefore

$$(x_2 - x_1) \cdot g^*(x_1, x_2, x_1, \dots) = 0$$

As the first factor is not identically zero, the second cannot escape this fate, and  $g^*(x_1, x_2, x_3, \dots)$  vanishes for  $x_3 = x_1$ , and hence contains the factor  $x_3 - x_1$ . Proceeding in the same manner one concludes that  $g$  is divisible

by the whole difference product:

$$g = \Delta \cdot G,$$

where  $G$  is a polynomial again and obviously a symmetric one. After expressing the symmetric forms  $F$  and  $G$  in terms of the elementary symmetric functions one gets  $f$  expressed by the same  $\varphi_i(x)$  and  $\Delta$ :

$$f = \frac{1}{2} \{ F + G \cdot \Delta \}$$

(It is not surprising that  $\Delta$  appears in the first power only, since  $\Delta^2$  can be expressed as a polynomial  $D$  of the  $\varphi_i(x)$ .)

To prove the second part of our statement, we observe that in two senses, the "functional" and the "algebraic" sense, it can be maintained that the quadratic equation

$$(4.3) \quad \Delta^2 - D(\varphi_1, \dots, \varphi_m) = 0$$

mentioned above is the only one holding between our basic invariants

$\varphi_1(x), \dots, \varphi_m(x); \Delta(x)$ . For the functional aspect, we remark that the invariants may take on arbitrary values  $\varphi_1, \dots, \varphi_m; \Delta$  only if those values satisfy our equation (4.3). Indeed, the "coefficients"  $\varphi_1, \dots, \varphi_m$  determine the roots  $x_1, \dots, x_m$  but for their order, and according to this order  $\Delta(x)$  will take on both signs  $\pm \Delta$  as allowed by equation (4.3). Secondly, every polynomial  $H(\varphi_1, \dots, \varphi_m; \Delta)$  of the independent variables  $\varphi_1, \dots, \varphi_m; \Delta$  when vanishing identically in  $x_1, \dots, x_m$  after the substitution

$\varphi_i = \varphi_i(x), \Delta = \Delta(x)$  is a multiple of the left side of (4.3):

$$H(\varphi_1, \dots, \varphi_m; \Delta) \equiv L(\varphi_1, \dots, \varphi_m; \Delta) \cdot \{ \Delta^2 - D(\varphi_1, \dots, \varphi_m) \},$$

$L$  being a polynomial again. To prove this, consider  $H$  as a polynomial in  $\Delta$  and divide by  $\Delta^2 - D(\varphi_1, \dots, \varphi_m)$ . The remainder is linear:

$A(\varphi_1, \dots, \varphi_m) + \Delta \cdot B(\varphi_1, \dots, \varphi_m)$ . The "substitution" yields both equations  $A \pm B\Delta = 0$ , according to the arrangement of the variables

$x_1, \dots, x_m$ . Therefore  $A$  and  $B$  vanish individually.

The underlying number field may be any field of characteristic  $\neq 2$ .

Not less instructive, regarding our present consideration of interrelations connecting the basic invariants, than this example of the finite group, is the continuous group of all orthogonal transformations. We considered above the invariants of two vectors  $x, y$ . Their basic invariants  $(xx), (xy), (yy)$  -- at least if the number of dimensions is  $\geq 2$  -- are capable of all numerical values satisfying the inequality

$$(4.4) \quad (xy)^2 \leq (xx) \cdot (yy);$$

for the lengths of two sides of the triangle and the angle included may be assigned arbitrarily. The inequality (4.4) is surely to be counted as a relation from the general functional standpoint; from the algebraic standpoint, however,  $(xx), (xy), (yy)$  are independent since they are not bound by any algebraic equation. Here again one first has to look upon  $xx, xy, yy$  as independent variables -- which for convenience are denoted in this somewhat unusual fashion involving the empty symbols  $x, y$  -- whereas in the second instance they are replaced by the scalar products  $(xx), (xy), (yy)$  of two arbitrary vectors  $x, y$ .

What is the behavior in this respect of an arbitrary number  $h$  of independent vectors  $x, y, \dots, z$  and their table of scalar products (3.2)? The scalar products are algebraically independent as long as  $h$  is less than or equal to the dimensionality  $n$ , but not so if  $h > n$ . The scalar products of  $m+1$  vectors  $x, y, \dots, z$  satisfy, for instance, the equation

$$\begin{vmatrix} (xx) & (xy) & \dots & (xz) \\ (yx) & (yy) & \dots & (yz) \\ \cdot & \cdot & \cdot & \cdot \\ (zx) & (zy) & \dots & (zz) \end{vmatrix} = 0$$

In case  $h \leq n$  the problem of determining the  $h$  vectors  $x, y, \dots, z$  such that the matrix (3.2) of their scalar products coincides with a given symmetric matrix  $\|a_{ik}\|$  of  $h$  rows and columns has a solution if and only if the quadratic form with the coefficients  $a_{ik}$  is positive definite. Our statement is merely a different formulation of the well-known fact that such a form may be linearly transformed into the squared sum of the  $h$  independent variables. One sees that the table (3.2) of the basic invariants is bound only by inequalities when algebraic equations, however, appear as soon as the number  $h$  of vectors surpasses the dimensionality  $n$ .

We proved in the case  $h = 2$  that every invariant form  $f(x, y)$  depending on the two vectors  $x, y$ , is expressible either as

$F((xx), (xy), (yy)) / (xx)^\alpha$  or as  $G((xx), (xy), (yy)) / (yy)^\beta$ , where  $F$  and  $G$  are polynomials. We may assume that  $F$ , considered as a function of the independent variables  $xx, xy, yy$  does not contain a factor  $xx$  nor  $G$  a factor  $yy$ . The equation

$$(yy)^\beta \cdot F(xx, xy, yy) = (xx)^\alpha \cdot G(xx, xy, yy)$$

holds after substitution; since (for  $n \geq 2$ ) the scalar products  $(xx), (xy), (yy)$  are algebraically independent, it must also hold before substitution, i.e. with  $xx, xy, yy$  as independent variables. Neither of the two factors of the left side however, is divisible by  $xx$ . Hence  $xx$  cannot be a divisor of the left side, and hence the exponent  $\alpha \geq 0$  must be equal to zero. This completes the proof that  $f(x, y)$  is expressible as a polynomial  $F((xx), (xy), (yy))$  in the three scalar products.

In spite of our thus succeeding in proving the first fundamental theorem for orthogonal invariants of two vectors  $x, y$ , we should get into serious trouble if we attempted to deal in the same manner with  $h \geq 3$  independent vec-

tors. The procedure becomes entirely hopeless if  $h$  surpasses the dimensionality  $n$  and the scalar products are no longer algebraically independent. To overcome these difficulties a new formal apparatus is needed as we shall see in ample detail in Chapter I.

The first main problem in the theory of vector invariants of a given group  $\gamma$  of linear transformations is the determination of a set of basic invariants and the first main theorem (which we cannot assert, however, for all groups  $\gamma$ ) states the finiteness of such a basis. The second main problem consists in determining "all" algebraic relations holding between the basic invariants

$$\varphi_1(x, y, \dots), \varphi_2(x, y, \dots), \dots, \varphi_n(x, y, \dots),$$

or rather, to find a number of such relations of which all others are algebraic consequences. The finiteness of this number is averred by the second main theorem. It holds for any group whatsoever; for it is a special case of Hilbert's general theorem about the finite basis of an arbitrary polynomial ideal. It was

first developed by Hilbert exactly in this context of the theory of invariants; the reader will find the general proof in Chapter II. Indeed, all such polynomials  $\mathcal{R}(\varphi_1, \varphi_2, \dots, \varphi_n)$  ("relations") of  $r$  independent variables

$\varphi_1, \varphi_2, \dots, \varphi_n$  that vanish after the "substitution",  $\varphi_i = \varphi_i(x, y, \dots)$  identically in  $x, y, \dots$  form an ideal  $\mathcal{J}$ . In other words, if  $\mathcal{R}$  is a relation, so is  $L \cdot \mathcal{R}$  where  $L$  is an arbitrary polynomial of the variables  $\varphi_i$ , and with two relations  $\mathcal{R}_1, \mathcal{R}_2$  their sum  $\mathcal{R}_1 + \mathcal{R}_2$  is also a relation. The ideal

$\mathcal{J}$  has a finite basis according to Hilbert's theorem. In other words, it contains a finite number of polynomials  $\mathcal{R}_1, \dots, \mathcal{R}_s$  such that every polynomial  $\mathcal{R}$  of  $\mathcal{J}$  is expressible in the form

$$\mathcal{R} = L_1 \cdot \mathcal{R}_1 + \dots + L_s \cdot \mathcal{R}_s, \quad (L_j, \text{ a polynomial}).$$

All relations  $R = 0$  holding between the  $r$  basic invariants  $\Phi_i(x, y, \dots)$  are thus consequences of the  $S$  relations

$$R_1 = 0, \dots, R_s = 0.$$

This general solution of our problem does not free us from the duty of ascertaining an actual basis  $R_j = 0$  for the relations in each particular case that may come under our consideration.

### 5. Invariants depending on contravariant as well as covariant vectors

The elementary problem of invariants is capable of a certain extension which it is convenient to consider along with the elementary problem proper before attacking the general one. With a given  $n$ -dimensional vector space we may associate a dual space; the "product"

$$(x\xi) = x_1\xi_1 + \dots + x_n\xi_n$$

of a vector  $x = (x_1, \dots, x_n)$  of the original space  $R$  and a vector

$\xi = (\xi_1, \dots, \xi_n)$  of the dual space  $P$  has an invariant significance, independent of the choice of the coordinate system. If the components of the

"Latin vector"  $x$  in  $R$  undergo a linear transformation with matrix  $\|a_{ik}\| = A$ , under the influence of a change of the coordinate system, then the components of a "Greek vector"  $\xi = (\xi_1, \dots, \xi_n)$  in  $P$  are subject to the "contragredient transformation" with matrix  $\|a_{ik}\| = \check{A}$ :

$$x'_i = \sum_R a_{iR} x_R, \quad \xi'_i = \sum_R a_{iR} \xi_R, \\ \sum_i x'_i \xi'_i = \sum_i x_i \xi_i.$$

One may look upon the dual space  $P$  as having for its elements the linear forms depending on an arbitrary vector  $x$  in  $R$ .

The extension we have in mind consists in studying functions

$f(x, y, \dots; \xi, \eta, \dots)$ , and in particular those invariant forms depending on several covariant or Latin vectors  $x, y, \dots$  and several Greek or contra-

variant vectors  $\xi, \eta, \dots$ . A group  $\gamma$  of linear transformations  $\mathcal{S}$  is given. We subject each Latin vector  $x, y, \dots$  to the same arbitrary transformation  $S$  of  $\gamma$  and at the same time each Greek vector  $\xi, \eta, \dots$  to the contragredient transformation  $\check{S}$ . The transform  $\mathcal{S}f = f'$  is defined by

$$f'(Sx, Sy, \dots; \check{S}\xi, \check{S}\eta, \dots) = f(x, y, \dots; \xi, \eta, \dots).$$

is invariant when  $Sf = f$  for all transformations  $S$  of  $\gamma$ .

The orthogonal group is most unsuitable to illustrate this generalized elementary problem; for by its very nature and definition an orthogonal transformation coincides with its contragredient. Let us therefore take as a more elucidating example, the unimodular group, i.e. the group  $\gamma$  of all linear transformations  $S$  of determinant 1. This group lies at the basis of affine and projective geometry. To the projective geometer the covariant and contravariant vectors and their components are known as point-coordinates and plane-coordinates respectively; only their proportions

$$x_1 : x_2 : \dots : x_m, \quad \xi_1 : \xi_2 : \dots : \xi_m$$

matter in the projective interpretation.

So long as we consider functions of covariant vectors only, there is but one single typical basic invariant for the unimodular group, namely the "determinant" or "bracket factor",

$$[x^1, x^2, \dots, x^m],$$

which depends on  $n$  arbitrary vectors

$$x^i = (x_1^i, x_2^i, \dots, x_m^i)$$

and is the determinant of their components  $x_k^i$ . If, however, an arbitrary number of contravariant vectors are allowed to enter as arguments, in addition to the covariant ones, the complete table of basic invariants will consist of three types: the Latin and the Greek bracket-factors

$$[x^1, x^2, \dots, x^m], \quad [\xi^1, \xi^2, \dots, \xi^m]$$

depending on  $n$  covariant and  $n$  contravariant vectors  $x^i$  and  $\xi^i$  respectively, and the product

$$(x \xi)$$

of a covariant and a contravariant vector  $x$  and  $\xi$ .

### 6. Invariants of forms and the general problem. The linear and the algebraic stage.

The unimodular group  $\gamma$  may also serve to give us an idea of the implications of the general problem. Let us consider an arbitrary form,

$$f(x) = \sum \alpha_{i k \dots l} x_i x_k \dots x_l$$

of degree  $h$  depending on a variable vector  $x = (x_1, x_2, \dots, x_m)$ ; it is natural to suppose that the coefficients  $\alpha_{i k \dots l}$  will depend in a symmetric manner on their  $h$  indices  $i, k, \dots, l$  ranging from 1 to  $n$ . Under the influence of an arbitrary unimodular transformation  $s$ :

$$x_i = \sum_k s_k^i x'_k$$

$f$  goes over into a new form

$$f' = s f : f(x) = f'(x') = \sum \alpha'_{i k \dots l} x'_i x'_k \dots x'_l.$$

An invariant

$$J = J(f) = J(\alpha_{i k \dots l})$$

of  $f$  is a homogeneous polynomial of a certain degree  $\nu$  depending on the

$$\frac{(m+h-1)!}{h!(m-1)!} \text{ arbitrary coefficients } \alpha_{i k \dots l} \text{ such that } J(sf) = J(f)$$

for every unimodular transformation  $s$ .

In this form, fitting into the frame of projective geometry, was the question of invariants first raised by Cayley and his immediate followers in the middle of the nineteenth century: to determine all invariants  $J(f, g, \dots)$  that depend on one or several arbitrary forms  $f, g, \dots$  of given degrees. In the work of Cayley, however, the arguments  $x_1, x_2, \dots, x_m$  of the forms  $f, g, \dots$

were interpreted as being projective point coordinates, and in agreement with this homogeneous interpretation, the group under consideration was the group of all non-singular linear transformations rather than the unimodular group, and the definition of invariance for  $J$  was modified accordingly (see next section). These early investigators focussed their attention not on the forms  $f, g, \dots$

themselves and the invariant  $J(f, g, \dots)$ , but on the equations

$f = 0, g = 0$  and  $J(f, g; \dots) = 0$ . The latter expresses a relation between the arbitrary algebraic "surfaces"  $f = 0, g = 0, \dots$  in  $(n-1)$ -dimensional projective space, that has an invariant meaning in the sense of projective geometry. And for this purpose, the discussion of the possible projective relations between algebraic surfaces (curves in the projective plane for  $n = 3$ ), were the invariants invented.

The coefficients  $\alpha'_{ik\dots l}$  of the transform  $f'$  arise from the coefficients  $\alpha_{ik\dots l}$  of  $f$  by means of the linear transformation

$$(6.1) \quad \mathcal{S} : \alpha'_{ik\dots l} = \sum_{u,v,\dots,w} s_i^u s_k^v \dots s_l^w \alpha_{u,v,\dots,w}$$

The geometric object on which our invariant  $J(f)$  depends may hence as well be described as a contravariant tensor of rank  $h$ ; its components  $\alpha_{ik\dots l}$  are arbitrary but for the postulate of symmetry:  $\alpha_{ik\dots l}$  shall not change if the indices  $i, k, \dots, l$  are subject to any permutation. (The arguments  $x_1, \dots, x_n$  of the form are here considered as components of a covariant vector.)

There is no reason, however, why we should restrict our tensor of rank  $h$ , that appears as the argument of the invariants to be studied, by this rather than by some other symmetry conditions. We may, instead, get interested in invariants  $J(\alpha_{ik\dots l})$  depending on an arbitrary antisymmetric contravariant tensor  $\alpha_{ik\dots l}$  of rank  $h$ . It is by no means an easy matter to ascer-

tain the different ways in which a tensor may be bound by symmetry conditions; the question is closely tied up with representations of the symmetric group. But the transition from the coordinate transformation  $s$  to the corresponding transformation  $S$ , (6.1), in the domain of all tensors  $\alpha_{i k \dots l}$  of certain prescribed symmetry properties:  $s \rightarrow S$ , is always a representation of the group  $\gamma$ . Hence the invariants  $J(f, g, \dots)$  depending on one or several tensors  $f, g, \dots$  of prescribed rank and symmetry fall under the general notion of an invariant as formulated in § 2.

As a matter of fact, the latter notion is not broader than the former in the case of the unimodular group  $\gamma$ . For it has been shown that the only representations (to be exact, the only irreducible representations) of the unimodular group, are those whose substratum is the manifold of all tensors of prescribed rank and symmetry. But, be this as it may, the general definition of § 2, at any rate, obviously gives the notion its natural breadth, the specialization to invariants of tensors or forms  $f, g, \dots$ , to which the historical development first led, is to be stripped off as being too formal and accidental. Where there are other representations (fortunately there are none) besides those of the tensors of given rank and symmetry, the invariants belonging to them would be as worth studying as the classical invariants. The general formulation of § 2 fully evinces its superiority in the case of other groups by enabling us to carry the classical notion over from the unimodular group to any abstract group

Let us then consider this notion in its widest scope. There are associated with a given abstract group  $\gamma$  a number of representations:

$$(6.2) \quad s \rightarrow S, \quad s \rightarrow T, \quad \dots,$$

of degree  $m, n, \dots$  respectively. The arguments  $x, y, \dots$  of the forms

$J(x, y, \dots)$  that we are concerned with are vectors undergoing those linear

transformations  $S, T, \dots$  "under the influence of  $s$ ", and  $J$  is thus changed into  $sJ$ :

$$sJ(Sx, Sy, \dots) = J(x, y, \dots).$$

$J$  is an invariant if  $sJ = J$  for every  $s$  in  $\gamma$ . Let us first collect and envisage as a whole all invariant forms  $J$  of definite given degrees  $\mu, \nu, \dots$  with respect to  $x, y, \dots$ , respectively. These forms may be thought of as elements of the linear manifold or vector space given by the totality of forms of given degrees  $\mu, \nu, \dots$  in  $x, y, \dots$ , respectively. Such a form is a linear combination of the monomials

$$Z(\mu_1, \dots, \mu_m; \nu_1, \dots, \nu_n; \dots) = (x_1^{\mu_1} \dots x_m^{\mu_m})(y_1^{\nu_1} \dots y_n^{\nu_n}) \dots,$$

the non-negative integral exponents  $\mu_i, \nu_k, \dots$  being limited by the conditions

$$\mu_1 + \dots + \mu_m = \mu, \quad \nu_1 + \dots + \nu_n = \nu, \quad \dots$$

With the  $x_i, y_k, \dots$  undergoing their respective transformations  $S, T, \dots$  these monomials experience a certain linear transformation  $U$ ; the representation  $\mathcal{U}: s \rightarrow U$  is composed in some way of the given representations  $s \rightarrow S, s \rightarrow T, \dots$ . The algebraic problem of constructing all invariants to the given set of representations is thus reduced to a much simpler question when one restricts oneself to forms of given degrees  $\mu, \nu, \dots$  in the arguments  $x, y, \dots$ ; namely to the linear problem of determining a complete set of linearly independent linear forms  $J$  of the variables  $Z(\mu_1, \dots, \mu_m; \nu_1, \dots, \nu_n; \dots)$  which are invariant under a given group  $\mathcal{U}$  of linear transformations. In particular, one would like to ascertain the number  $M$  of basic linear invariants  $J_1, J_2, \dots, J_M$  by whose linear combinations all invariant linear forms  $J$  of the kind described arise.  $M$  cannot exceed the dimensionality  $N$  of that vector space  $Z$  to which the transformations  $U$  refer, i.e. the number  $N$  of variables

$Z(\mu_1, \dots, \mu_m; \nu_1, \dots, \nu_m)$ . Looking upon  $Z$  as an affine space of  $N$  dimensions, we may feel free to replace the coordinates  $Z(\mu_i, \nu_k; \dots)$  by a set of  $N$  independent linear forms of the  $z$ ; and then one could choose one or even  $M$  of these new coordinates among the basic linear invariants  $J_1, J_2, \dots, J_M$ . If we use this new coordinate system for expressing the linear transformations  $U$ , they appear in reduced form: they contain ( $M$  times) the identical transformation  $J' = J$  of a single variable  $J$ . The question of linear invariants is thus shown to be identical with the question as to how often the representation  $\mathcal{U}$ , if fully reduced, contains the identical representation.

It seems natural to make the study of linear invariants preliminary to the higher investigation of algebraic nature where one wants to survey the invariants  $J(x, y, \dots)$  of all possible degrees in  $x, y, \dots$  associated with the given representations of  $\chi$  and not only those of given fixed degrees. These invariants do not form a linear manifold (or "vector space") but an algebraic structure called a "ring", as they allow multiplication as well as addition. The words "dependence", "basis", then change their meaning accordingly from linear to algebraic, i.e. integral rational expressions. It is in no way obvious that there exists at all a finite basis in this algebraic sense; but we shall succeed in finding extensive classes of groups  $\chi$  for which the "first main theorem" contending the finiteness of an algebraic basis holds for any representations (6.2) of  $\chi$ .

### 7. Absolute and relative invariants. Covariant quantities

The form  $J$  itself is devoid of significance in projective geometry

where only the ratios of coordinates matter; what does count is the equation  $J=0$ , as it expresses a relation among the geometrical objects on which  $J$  depends. This relation will be of invariant character if the transformed relation  $sJ=0$  is identical with  $J=0$  for each element  $s$  of the group; or, what amounts to the same thing, if  $sJ$  differs from  $J$  only by a constant factor  $\lambda = \lambda(s)$ , depending on  $s$ . We therefore call the form  $J(x, y, \dots)$  a relative invariant of multiplier  $\lambda(s)$  if

$$sJ(x, y, \dots) = \lambda(s) \cdot J(x, y, \dots).$$

The invariants in the strict sense are then distinguished as absolute invariants.

The multiplier  $\lambda(s)$  has the value 1 for the unit element  $s = \underline{1}$  of the group. Furthermore, when  $s$  and  $t$  are two elements of the group, one has

$$(7.1) \quad \lambda(st) = \lambda(s) \lambda(t).$$

These two facts can be combined into the statement that the correspondence

$s \rightarrow \lambda(s)$  is a representation of degree 1 of  $\gamma$ . Every relative invariant belongs to a certain such representation.

In the case in which the projective geometer is interested, where the group  $\gamma$  consists of all non-singular linear transformations  $s$  of  $n$  (homogeneous) variables:

$$x_i = \sum_{k} s_{ik}^i x'_k,$$

we have as the only possible representations  $\lambda(s)$  of degree 1, the powers of the determinant

$$|s| = \det. (s_{ik}^i).$$

Proof: The arguments of  $J$  are now tensors whose components are subject to linear transformations with coefficients depending on the  $s_{ik}^i$  in an integral rational fashion. Hence  $\lambda(s)$  is a homogeneous polynomial of the transformation coefficients. The adjoint transformation has the minors of the matrix  $\|s_{ik}^i\|$  as its coefficients; hence  $\lambda(\bar{s}) = \bar{\lambda}(s)$  is also a poly-

nomial in  $S_k^i$ . The compound of the two transformations  $S, \bar{S}$  which one gets by first performing  $s$  and then  $\bar{s}$  is the multiplication of all  $m$  variables  $x_i$  with the same factor  $|s|$ . The arguments of  $J$ , and hence  $J$  itself, are multiplied under the influence of this multiplication by a factor  $|s|^q$ ; that is, a power of  $|s|$  with a non-negative integral exponent  $q$ . The relation

$$\lambda(s) \cdot \bar{\lambda}(s) = |s|^q,$$

implied by (7.1) shows that  $\lambda(s)$  itself must needs be a power of  $|s|$ :

$\lambda(s) = |s|^g$ , because the determinant  $|s|$  is an irreducible polynomial of its arguments  $S_k^i$ . The integer  $g$  is called the weight of the relative invariant.

Later on we shall give another proof for the same fact, making no use of the assumption that  $\lambda(s)$  depends rationally on  $s$ .

Our theorem has the consequence that all relative invariants of the full linear group  $\gamma$  turn into absolute invariants when one restricts oneself to the unimodular transformations within  $\gamma$ . By this simple trick the study of relative invariants for the full linear group can be replaced by that of absolute invariants for the unimodular group -- the standpoint we adopted in our exposition in sec. 6.

If one once starts considering invariant relations, one cannot help realizing at the outset that such a relation need not be defined by one equation,  $J=0$ , alone. On dealing with a set of equations

$$(7.2) \quad J_1 = 0, J_2 = 0, \dots, J_n = 0,$$

it is natural to assume that all the left sides  $J_p = J_p(x, y, \dots)$  are forms of the same degrees  $\mu, \nu, \dots$  with respect to the argument vectors. The forms of these degrees constitute a linear manifold -- a vector space one might call it -- of  $N$  dimensions; we operated in this vector space when we

talked of the "linear problem" in sec. 6. The forms  $J_1, \dots, J_h$  span a linear space within the total space, or a modul, as the algebraists say, consisting of all linear combinations

$$J(x, y, \dots) = \lambda_1 J_1(x, y, \dots) + \dots + \lambda_h J_h(x, y, \dots)$$

with arbitrary constant coefficients  $\lambda_p$ . It is natural to assume, furthermore, that the basis does not contain any redundant terms, i.e. that the  $J_p$  are linearly independent. Any relation  $J = 0$ , with a  $J$  not included in our modul, would add a further restriction to the equations (7.2). Consequently, if these equations establish an invariant relation among the arguments  $x, y, \dots$  every transform  $S J_p$  of any one of them must lie in the modul:

$$(7.3) \quad S J_p(x, y, \dots) = \sum_{q=1}^h J_q \cdot a_{qp}(s)$$

The correspondence  $S \rightarrow \|a_{qp}(s)\|$  is a representation  $\mathcal{O}$  of degree  $h$  of our group  $\gamma$ .

A quantity with  $h$  components  $J_1, \dots, J_h$ , transforming under the influence of  $s$  according to the equations (7.3), where  $S \rightarrow \|a_{qp}(s)\| = A_s$  is a given representation  $\mathcal{O}$ , shall be named a covariant quantity of kind  $\mathcal{O}$ . If  $\mathcal{F}$  designates the identical representation of degree 1, associating the number 1 with each element  $s$  of the group, the invariants are to be described as covariant quantities of kind  $\mathcal{F}$ . This new generalization from (absolute and relative) invariants to covariant quantities of any given kind is urged upon us when we take the general idea of invariant relations seriously. The result is this: to the representations (6.2) of  $\gamma$  which regulate the transformations of the argument vectors  $x, y, \dots$  is added a further representation

$$\mathcal{O}: s \rightarrow A_s,$$

of degree  $h$ ,  $h$  linearly independent forms  $J_1(x, y, \dots), \dots, J_h(x, y, \dots)$

of prescribed degrees  $\mu, \nu, \dots$  in  $x, y, \dots$  constitute a covariant quantity of kind  $\mathcal{O}$  if they transform among each other according to  $A_S$  under the influence of an arbitrary element  $s$  of  $\gamma$ , as in (7.3). The simultaneous equations

$$J_1(x, y, \dots) = 0, \dots, J_n(x, y, \dots),$$

then constitute an invariant relation of kind  $\mathcal{O}$  among the vectors  $x, y, \dots$ .

$J_1(x, y, \dots), \dots, J_n(x, y, \dots)$  form a partial set of coordinates in the vector space  $Z$  we mentioned in sec. 6 as the carrier of the compound representation  $\mathcal{U}$ ; in using these coordinates, which may be supplemented to a full coordinate system, one "reduces" the representation  $\mathcal{U}$  such that  $\mathcal{O}$  is put in evidence as a part of  $\mathcal{U}$ ; and the construction of covariant quantities is thus identical with the reduction of our representation  $\mathcal{U}$ . One sees how the theory of invariants becomes by degrees submerged into the theory of representations.

The aims of this introduction were :

- 1) to explain the general notion of invariants as we see it today,
- 2) to assign to the classical, more specialized, concept of invariant as it emerged from algebraic geometry, its place within a modern frame;
- 3) to enable the reader, from the beginning, to understand the general disposition of the subject as it will be presented in this course.

Noticeable, above all, is the partition into an elementary part (vector invariants or invariants for a given group of linear transformations) and a higher part dealing with the invariants for a given abstract group relative to certain of its representations.

CHAPTER I. VECTOR INVARIANTS

A. Covariant Argument Vectors Only

1. Formal preliminaries about forms

When defining a form or homogeneous polynomial of degree  $h$  of the "vector"  $u = (u_1, \dots, u_n)$  by the formal expression

$$(1.1) \quad A(u) = \sum a_{i_1 \dots i_h} u_{i_1} u_{i_2} \dots u_{i_h}$$

one is led to introduce the "corresponding"  $h$ -fold multilinear form  $A(x, y, \dots, z)$  of  $h$  independent vectors  $x, y, \dots, z$  by the equation

$$(1.2) \quad A(x, y, \dots, z) = \sum a_{i_1 \dots i_h} x_{i_1} y_{i_2} \dots z_{i_h}.$$

The given form  $A(u)$  arises from it by identifying all  $h$  vectors  $x, y, \dots, z$  with  $u$ :

$$(1.3) \quad x = y = \dots = z = u.$$

This is an essential reduction because linear form  $L(x)$  is easily characterized by simple functional properties independently of the coordinate system in our affine vector space:

$$L(x + x') = L(x) + L(x'); \quad L(\alpha x) = \alpha \cdot L(x), (\alpha, \text{ a number}).$$

Hence the definition: a form  $A(u)$  of degree  $h$  is one arising from a form  $A(x, y, \dots, z)$  linear in its  $h$  independent argument vectors  $x, y, \dots, z$  by the "identification" (1.3).

If the multilinear form  $A(x, y, \dots, z)$  leads to the form  $A(u)$ , then any form obtained from  $A(x, y, \dots, z)$  by any permutation  $p$  of its  $h$  arguments also leads to  $A(u)$ . The sum extending over all permutations  $p$ :

$$\frac{1}{h!} \sum_p A(x, y, \dots, z) = A^*(x, y, \dots, z)$$

is therefore a symmetric multilinear form corresponding to the given  $A(u)$ :

$$p A^*(x, y, \dots, z) = A^*(x, y, \dots, z) \text{ for all permutations } p.$$

The coefficients  $a_{ik\dots l}^*$  of

$$A^*(x, y, \dots, z) = \sum a_{ik\dots l}^* x_i y_k \dots z_l$$

then depend symmetrically on their  $h$  indices  $i, k, \dots, l$ . The process of polarization will show that this symmetric multilinear form is uniquely determined by  $A(u)$ ; we call it "the corresponding multilinear form" and the transition from

$$A(u) = \sum a_{ik\dots l} u_i u_k \dots u_l \text{ to } A(x, y, \dots, z) = \sum a_{ik\dots l} x_i y_k \dots z_l$$

with symmetric coefficients  $a_{ik\dots l}$  is "complete polarization".

The individual step of polarization consists in performing the differential operation  $\mathcal{D}_{xu}$ :

$$\mathcal{D}_{xu} A(u) = \sum_i x_i \frac{\partial A}{\partial u_i}.$$

The operator  $\mathcal{D}$  has the following effect upon  $A$ : (1) replacing the first factor  $u$  in all the terms of  $A(u)$  by  $x$ ; (2) adding to the expression thus gained

$$\sum a_{ik\dots l} x_i u_k \dots u_l,$$

further expressions ( $h-1$  in number) in which the second, ..., the  $h^{\text{th}}$  factor respectively experiences the same fate. But when the coefficients are written symmetrically, these further contributions evidently equal the first one:

$$(1.4) \quad \mathcal{D}_{xu} A(u) = h \cdot \sum a_{ik\dots l} x_i u_k \dots u_l.$$

If one re-substitutes  $u$  for  $x$  in (1.4), one falls back on the original form except for the factor  $h$  (Euler's relation):

$$\left\{ \mathcal{D}_{xu} A(u) \right\}_{x=u} = \sum u_i \frac{\partial A}{\partial u_i} = h \cdot A(u).$$

Owing to the formula (1.4), the process

$$(1.5) \quad \frac{1}{h!} \mathcal{D}_{xu} \mathcal{D}_{yu} \dots \mathcal{D}_{zu}$$

carries  $A(u)$  into the corresponding symmetric multilinear form  $A(x, y, \dots, z)$

(1.5) is the sought-for process of complete polarization.

The polar process  $\mathcal{D} = \mathcal{D}_{x u}$  has the following formal properties of differentiation:

$$\mathcal{D}(f(u) + g(u)) = \mathcal{D}f(u) + \mathcal{D}g(u),$$

$$\mathcal{D}(c \cdot f(u)) = c \cdot \mathcal{D}f(u), \quad (c, \text{ a number}),$$

$$\mathcal{D}(f(u) \cdot g(u)) = g(u) \cdot \mathcal{D}f(u) + f(u) \cdot \mathcal{D}g(u).$$

The polarized form  $\mathcal{D}_{x u} A(u)$  is the coefficient of the linear term in Taylor's expansion of  $A(u + \lambda x)$  by the parameter  $\lambda$ :

$$A(u + \lambda x) \equiv A(u) + \lambda \sum_i x_i \frac{\partial A}{\partial u_i} \pmod{\lambda^2}.$$

From this fact we derive at once the following statement on which the importance of polarization for invariant theory rests:

An invariant  $A(u)$  is changed into an invariant by the polarization

$\mathcal{D}_{x u}$ . Here,  $x$  is to be looked upon as a vector transforming cogrediently with  $u$ .

There exists a second method by which to show that the coefficients  $a_{i k \dots l}$  of a symmetric multilinear form are uniquely determined by the values of the corresponding form  $A(u)$ . As a polynomial,  $A(u)$  may thus be written:

$$A(u) = \sum b(h_1, \dots, h_m) u_1^{h_1} \dots u_m^{h_m};$$

the non-negative exponents  $h_1, \dots, h_m$  are bound by the equation

$$h_1 + \dots + h_m = h.$$

Since the values of a polynomial (of a single variable) determine its coefficients uniquely, it is certain that there is no ambiguity about these coefficients  $b(h_1, \dots, h_m)$ . But their relation to the symmetric coefficients  $a_{i k \dots l}$  is described as follows:

$$b(h_1, \dots, h_m) = \frac{h!}{h_1! \dots h_m!} a_{i k \dots l},$$

if  $h_1$  of the indices  $i, k, \dots, l$  are equal to 1,  $h_2$  of them are equal to 2,  $\dots$ ,  $h_m$  equal to  $n$ .

Trivial as these well-known preliminaries concerning forms of degree  $h$  may sound, it should not pass unnoticed that their validity is bound to one assumption about the underlying number field: it must not have a prime characteristic dividing one of the first  $h$  numbers  $1, 2, \dots, h$ .

## 2. First example: the symmetric group

We mentioned in § 2 of the Introduction the fundamental algebraic theorem, that the elementary symmetric functions

$$(2.1) \quad \begin{aligned} \varphi_1(x) &= \sum_i x_i, \\ \varphi_2(x) &= \sum_{i < k} x_i x_k, \\ &\dots \\ \varphi_n(x) &= x_1 x_2 \cdots x_n \end{aligned}$$

form a complete set of basic invariants in the realm of forms  $f(x)$  depending on a single vector  $x = (x_1, \dots, x_n)$  and invariant under the symmetric group of all permutations of the variables  $x_1, x_2, \dots, x_n$ . We are now going to solve the same problem for forms  $f(x^{(1)}, x^{(2)}, \dots, x^{(m)})$  depending on an arbitrary number  $x^{(1)}, x^{(2)}, \dots, x^{(m)}$  of independent vectors. The conjecture which offers itself at once is that (2.1) by full polarization will yield a complete table of typical basic invariants. If one adds the factor  $i!$  to  $\varphi_i$ , this table reads as follows:

$$(2.2) \quad \begin{aligned} \varphi_1(u) &= \sum u_i \\ \varphi_2(u, v) &= \sum_{i \neq k} u_i v_k \\ \varphi_3(u, v, w) &= \sum_{\substack{i, k, l \\ \text{different}}} u_i v_k w_l \\ &\dots \\ \varphi_n(u, v, \dots) &= \sum_{\substack{i, k, \dots \\ \text{all permutations} \\ \text{of } 1, 2, \dots, n}} u_i v_k \dots \end{aligned}$$

Our statement means, as may be recalled here, that one obtains a complete algebraic basis for the invariants  $f(x^{(1)}, x^{(2)}, \dots, x^{(n)})$  depending on  $m$  argument vectors  $x^{(1)}, \dots, x^{(m)}$  by substituting these arguments for the "typical" arguments  $u, v, w, \dots$  in all possible combinations (repetitions included) into the forms (2.2):

$$\varphi_i(x^{(\alpha_1)}, x^{(\alpha_2)}, \dots, x^{(\alpha_i)}) \quad [\alpha_1, \alpha_2, \dots, \alpha_i = 1, 2, \dots, m; i = 1, 2, \dots, m].$$

In order to avoid the crowding of indices, the  $m$  argument vectors shall now be designated by  $x, y, \dots, z$ . The proof will be given by complete induction with respect to dimensionality  $n$ . For this purpose each vector

$x = (x_1, x_2, \dots, x_m)$  is to be considered as combining a number  $x_1 = \xi$  with an  $(n-1)$ -dimensional vector  $x' = (x_2, \dots, x_m)$ . The elementary symmetric functions of the  $(n-1)$ -dimensional vectors (unpolarized or polarized) are designated by  $\varphi'_1, \varphi'_2, \dots, \varphi'_{m-1}$ . Every form  $f(x, y, \dots, z)$  is an aggregate (linear combination) of terms

$$\xi^\alpha \eta^\beta \dots \zeta^\gamma \cdot f_{\alpha\beta\dots\gamma}(x', y', \dots, z').$$

If  $f$  is symmetric, so is  $f_{\alpha\beta\dots\gamma}$ , and according to our assumption of the truth of our theorem for  $n-1$  dimensions,  $f_{\alpha\beta\dots\gamma}$  is expressible (in an integral rational way) by the polarized elementary symmetric forms  $\varphi'_1, \dots, \varphi'_{m-1}$ . The given  $f(x, y, \dots, z)$  appears thus as an aggregate of terms\*

$$\xi^\alpha \eta^\beta \dots \zeta^\gamma \cdot (\varphi'_1)^{p_1} (\varphi'_2)^{p_2} \dots (\varphi'_{m-1})^{p_{m-1}}$$

---

\*  $(\varphi'_1)^{p_1}$  here indicates a product of  $p_1$  factors  $\varphi'_1$  each of which may depend on different arguments.

---

We now make use of the equations (2.2) in their polarized form:

$$\begin{aligned} \Phi_i(x^{(1)}, x^{(2)}, \dots, x^{(i)}) &= \Phi_i'(x^{(1)}, x^{(2)}, \dots, x^{(i)}) \\ &\quad + \sum_{\alpha=1}^i x_1^{(\alpha)} \Phi_{i-1}'(x^{(1)}, \dots, x^{(\alpha-1)}, x^{(\alpha+1)}, \dots, x^{(i)}), \\ [i &= 1, \dots, m-1; \Phi_0' = 1], \end{aligned}$$

and thus express  $\Phi_i'$  by  $\Phi_i, \Phi_{i-1}'$  and the Greek variables

$x_1 = \xi, y_1 = \eta, \dots, z_1 = \zeta$ . In this way we eliminate the quantities

$$\Phi_{m-1}', \Phi_{m-2}', \dots, \Phi_1',$$

one after the other in replacing them by the Greek variables and

$\Phi_{m-1}, \Phi_{m-2}, \dots, \Phi_1$ . We end up expressing  $f(x, y, \dots, z)$  as an aggregate of terms

$$(2.3) \quad \xi^\alpha \eta^\beta \cdot \zeta^\gamma \cdot \Phi_1^{p_1} \cdot \Phi_{m-1}^{p_{m-1}} = x_1^\alpha y_1^\beta \cdot z_1^\gamma \cdot \Phi_1^{p_1} \cdot \Phi_{m-1}^{p_{m-1}}.$$

The second part of this term following the large dot is symmetric even when the first component  $x_1$  of each vector  $x$  is interchanged with the components

$x_2, \dots, x_m$  of  $x'$ . As the whole function  $f$  is symmetric in all  $n$  components,

the term (2.3) may be replaced by

$$\Phi_1^{p_1} \cdot \Phi_{m-1}^{p_{m-1}} \cdot \frac{1}{m} \sum_{i=1}^m x_i^\alpha y_i^\beta \cdot z_i^\gamma.$$

The sum  $\sum$  appearing here arises by consecutive polarizations from the power sum

$$\sigma_\nu(x) = \sum_{i=1}^m x_i^\nu, \quad (\nu = \alpha + \beta + \dots + \gamma),$$

and Newton's well-known formulae show how to express these in terms of the elementary symmetric functions  $\Phi_1, \Phi_2, \dots, \Phi_m$ .

To leave no gap we add their simplest derivation. The polynomial

$$\psi(\lambda) = \prod_{i=1}^m (1 - \lambda x_i) = 1 - \varphi_1(x) \lambda + \varphi_2(x) \lambda^2 - \dots \pm \varphi_m(x) \lambda^m$$

has as its logarithmic derivative

$$-\frac{\psi'(\lambda)}{\psi(\lambda)} = \sum_{i=1}^m \frac{x_i}{1 - \lambda x_i} = \sigma_1(x) + \sigma_2(x) \lambda + \sigma_3(x) \lambda^2 + \dots$$

The Taylor expansion on the right side is to be understood in the formal sense such that no questions of convergence arise:

$$-\psi'(\lambda) \equiv \psi(\lambda) \cdot \sum_{v=1}^N \sigma_v(x) \lambda^{v-1} \pmod{\lambda^N}$$

Hence the recursive formulae

$$\begin{aligned} \varphi_1(x) &= \sigma_1(x) , \\ -2\varphi_2(x) &= -\varphi_1(x)\sigma_1(x) + \sigma_2(x) , \\ 3\varphi_3(x) &= \varphi_2(x)\sigma_1(x) - \varphi_1(x)\sigma_2(x) + \sigma_3(x) , \\ &\dots \end{aligned}$$

[  $\varphi_v(x)$  is to be put = 0 for  $v > m$  ].

The underlying number field must not be of prime characteristic dividing one of the numbers 1, 2, ..., n.

### 3. Capelli's identity

Though we have just succeeded in proving the first main theorem for the symmetric group in a direct way by completely polarizing, as it were, one of the known proofs of the fundamental theorem on symmetric functions, we shall not succeed in a similar fashion in other cases. Here we need a certain powerful instrument: Capelli's identity. It is concerned with the result of successive polarizations.

Let  $x, y, z, \dots$  be a row of independent vectors in an  $n$ -dimensional vector space,  $x', y', z', \dots$  the same vectors in (the same or) a different ar-

rangement.  $\Delta_{x'x}$  as well as  $\mathcal{D}_{x'x}$  may symbolize the polar process. By performing several polarizations like  $\mathcal{D}_{x'x}$ ,  $\mathcal{D}_{y'y}$ ,  $\mathcal{D}_{z'z}$  in succession on a form  $f(x, y, z, \dots)$  one will get

$$\mathcal{D}_{z'z} \mathcal{D}_{y'y} \mathcal{D}_{x'x} f = \sum_{i,k,l} x'_i y'_k z'_l \frac{\partial^3 f}{\partial x_i \partial y_k \partial z_l},$$

provided  $x'$  does not coincide either with  $y$  or  $z$ , nor  $y'$  with  $z$ . The auxiliary symbols  $\Delta_{x'x}$ ,  $\Delta_{y'y}$ , . . . may be used instead of  $\mathcal{D}_{x'x}$ ,  $\mathcal{D}_{y'y}$ , . . . to indicate by their composition this result regardless of whether the coincidences just mentioned occur or not:

$$\Delta_{z'z} \Delta_{y'y} \Delta_{x'x} f = \sum_{i,k,l} x'_i y'_k z'_l \frac{\partial^3 f}{\partial x_i \partial y_k \partial z_l}.$$

We propose to compute how the composite operator  $\mathcal{D}_{z'z} \mathcal{D}_{y'y} \mathcal{D}_{x'x}$  differs from this "pseudo-composition". The symbol  $\delta_{x'x}$  will be used to indicate, whether  $x, x'$  are two different or identical vectors of our row:  $\delta_{x'x} = 1$  for  $x' = x$ ,  $\delta_{x'x} = 0$  for  $x' \neq x$ .

We obtain by definition

$$\begin{aligned} \mathcal{D}_{z'z} \Delta_{y'y} \Delta_{x'x} f &= \sum_l z'_l \frac{\partial}{\partial z_l} \left( \sum_{i,k} x'_i y'_k \frac{\partial^2 f}{\partial x_i \partial y_k} \right) \\ (3.1) \quad &= \sum_{i,k,l} x'_i y'_k z'_l \frac{\partial^3 f}{\partial x_i \partial y_k \partial z_l} + \delta_{x'z} \sum_{i,k} z'_i y'_k \frac{\partial^2 f}{\partial x_i \partial y_k} \\ &\quad + \delta_{y'z} \sum_{i,k} x'_i z'_k \frac{\partial^2 f}{\partial x_i \partial y_k} \\ &= \Delta_{z'z} \Delta_{y'y} \Delta_{x'x} f + \delta_{x'z} \Delta_{y'y} \Delta_{z'x} f + \delta_{y'z} \Delta_{z'y} \Delta_{x'x} f. \end{aligned}$$

Our chief interest is in the alternating sum

$$\sum_{(x', y', z')} \pm \mathcal{D}_{z'z} \Delta_{y'y} \Delta_{x'x} = \begin{vmatrix} \mathcal{D}_{z'z} & \Delta_{z'y} & \Delta_{z'x} \\ \mathcal{D}_{y'z} & \Delta_{y'y} & \Delta_{y'x} \\ \mathcal{D}_{x'z} & \Delta_{x'y} & \Delta_{x'x} \end{vmatrix}$$

extending to the 3! permutations of  $x'$ ,  $y'$ ,  $z'$ . The individual terms of the determinant of operators are to be written in such a way that the factors follow each other from left to right as in the determinant itself: first the factor taken from the first column, second the factor from the second column, etc.; the same rule is to be observed throughout the following. In computing this determinant we may exchange an  $x'$  and  $z'$  in the second term on the right side of our equation (3.1) as well as  $y'$  and  $z'$  in the third term, provided we change sign at the same time:

$$\sum_{(x', y', z')} \pm \mathcal{D}_{z'z} \Delta_{y'y} \Delta_{x'x} = \sum \pm \Delta_{z'z} \Delta_{y'y} \Delta_{x'x} - 2 \sum \pm \delta_{z'z} \Delta_{y'y} \Delta_{x'x}$$

or

$$\sum \pm (\mathcal{D}_{z'z} + 2 \delta_{z'z}) \Delta_{y'y} \Delta_{x'x} = \sum \pm \Delta_{z'z} \Delta_{y'y} \Delta_{x'x}.$$

The result reads, when written in determinant form:

$$(D_3) \begin{vmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \mathcal{D}_{z'z} + 2 & \Delta_{z'y} & \Delta_{z'x} \\ \mathcal{D}_{y'z} & \Delta_{y'y} & \Delta_{y'x} \\ \mathcal{D}_{x'z} & \Delta_{x'y} & \Delta_{x'x} \end{vmatrix} = \begin{vmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \Delta_{z'z} & \Delta_{z'y} & \Delta_{z'x} \\ \Delta_{y'z} & \Delta_{y'y} & \Delta_{y'x} \\ \Delta_{x'z} & \Delta_{x'y} & \Delta_{x'x} \end{vmatrix}$$

The meaning of this equation being that the corresponding minors (of order 3) on both sides coincide.

This equation for three consecutive polarizations should have been preceded by those holding for one or two such operators:



with the inner sum extending alternately to all permutations  $x', y', \dots, z'$  of  $x, y, \dots, z$ . This sum is zero unless all  $m$  indices  $i, k, \dots, \lambda$  are different. As this is impossible when  $m > n$ , the total result will in this case be zero. When  $m = n$ , the inner sum is  $\pm [x y \dots z]$  or 0, according as the sequence  $i, k, \dots, \lambda$  is an even or odd permutation of  $1, 2, \dots, n$ , or contains equal indices. The "bracket"  $[x y \dots z]$  designates the determinant

$$\begin{vmatrix} x_1 & x_2 & \dots & x_m \\ y_1 & y_2 & \dots & y_m \\ \cdot & \cdot & \cdot & \cdot \\ z_1 & z_2 & \dots & z_m \end{vmatrix}$$

as before. Hence one obtains for  $m = n$ :

$$[x y \dots z] \cdot \Omega f,$$

where

$$\Omega f = \sum_{(i,k,\dots,\lambda)} \pm \frac{\partial^n f}{\partial x_i \partial y_k \dots \partial z_\lambda} = \begin{vmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \dots & \frac{\partial}{\partial x_m} \\ \frac{\partial}{\partial y_1} & \frac{\partial}{\partial y_2} & \dots & \frac{\partial}{\partial y_m} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial}{\partial z_1} & \frac{\partial}{\partial z_2} & \dots & \frac{\partial}{\partial z_m} \end{vmatrix} f$$

is derived from  $f$  by the so-called Cayley- $\Omega$ -process. The sum extends alternately over all permutations  $i, k, \dots, \lambda$  of  $1, 2, \dots, n$ .

We thus arrive at Capelli's identity. In its final form the notation  $x, y, \dots, z$  may be replaced by the more pedantic  $x^1, x^2, \dots, x^m$  and the symbols  $\mathcal{D}_{\beta \alpha}$  by  $\mathcal{D}_{\beta \alpha}$ ; it then reads

$$(3.2) \quad \begin{vmatrix} D_{mm+(m-1)} & \dots & D_{m2} & D_{m1} \\ \dots & \dots & \dots & \dots \\ D_{2m} & \dots & D_{22+1} & D_{21} \\ D_{1m} & \dots & D_{12} & D_{11} \end{vmatrix} f = \begin{cases} 0, & \text{if } m > n, \\ [x^1 x^2 \dots x^n] \cdot \Omega f, & m = n. \end{cases}$$

We shall refer to the two cases  $m > n$ ,  $m = n$  of this formula as Capelli's general and special identity, respectively.

4. Reduction of the first main problem by means of Capelli's identities

The way in which Capelli's identity is used for the investigation of invariants depending on  $m$  argument vectors  $x^1, x^2, \dots, x^m$  can thus be described for an arbitrary group  $\gamma$  of linear transformations. Polarization  $D_{\beta\alpha}$  carries an (absolute or relative) invariant  $f(x^1, x^2, \dots, x^m)$  into an invariant  $D_{\beta\alpha} f$  (of the same multiplier). The special Capelli identity shows that in the case  $m = n$ ,  $\Omega f$  is a relative invariant whose multiplier equals the multiplier of  $f$  divided by the transformation determinant; in particular the operator  $\Omega$  applied to an absolute invariant  $f$  will yield an absolute invariant provided  $\gamma$  consists of unimodular transformations.

A form  $f(x^1, x^2, \dots, x^m)$  is of a certain degree  $r_1, r_2, \dots, r_m$  in each of its arguments  $x^1, x^2, \dots, x^m$ . The sum  $r = r_1 + \dots + r_m$  is its total degree. We arrange forms  $f(x^1, x^2, \dots, x^m)$  first according to their total degree, i.e.  $f$  is of lesser rank than  $f^*$  in the hierarchy of forms if the total degree of  $f$  is lower than that of  $f^*$ . For instance  $\Omega f$  is lower than  $f$  (in the case  $m = n$ ) since the  $\Omega$ -process diminishes the total degree by  $n$ . Within the set of all forms of given total degree which constitutes a linear totality of a finite number of dimensions, lexicographic arrangement according to the individual degrees  $r_1, \dots, r_m$  is followed; i.e.  $f$  is lower than  $f^*$  if the first of the degrees  $r_1, r_2, \dots, r_m$  in which  $f$  and  $f^*$

differ, has a lower value for  $f$ . Forms coinciding in all their  $m$  degrees  $r$  are considered of equal rank; we abstain from imposing an order of precedence upon them.

The main term in the operator determinant of Capelli's identity,

$$(\mathcal{D}_{m_1} + m - 1) \cdot \cdot \cdot (\mathcal{D}_{22} + 1) f,$$

equals

$$r_1 (r_2 + 1) \cdot \cdot \cdot (r_m + m - 1) f = \rho f.$$

The numerical factor  $\rho$  multiplying  $f$  is  $\neq 0$  when  $r_1 > 0$ , i.e. when  $f$  actually contains the first variable  $x^1$ . Writing  $\mathcal{D}_{\beta\alpha}^* = \mathcal{D}_{\beta\alpha}$ , ( $\beta \neq \alpha$ ),  $\mathcal{D}_{\alpha\alpha}^* = \mathcal{D}_{\alpha\alpha} + \alpha - 1$ , we observe that for a term different from that given by the principal diagonal, we may drop the factors  $\mathcal{D}_{\alpha\alpha}^*$  with two equal indices. Their effect is merely to multiply  $f$  by certain constants which we collect in the factor  $\rho^*$ . Then such a term is of the form

$$\rho^* \cdot \mathcal{D}_{\beta_n \alpha_n} \cdot \cdot \cdot \mathcal{D}_{\beta_2 \alpha_2} \mathcal{D}_{\beta_1 \alpha_1} f,$$

where  $\alpha_n > \alpha_{n-1} > \cdot \cdot \cdot \alpha_2 > \alpha_1$ ,  $\beta_i \neq \alpha_i$ , and  $\beta_1, \beta_2, \cdot \cdot \cdot, \beta_n$  is a permutation of  $\alpha_1, \alpha_2, \cdot \cdot \cdot, \alpha_n$ . Hence, in particular,  $\beta_1 > \alpha_1$ ;  $n$  is at least equal to 2. The main term being the only one in which no factor  $\mathcal{D}_{\beta\alpha}^*$  with different indices  $\alpha, \beta$  appears, on putting

$$P = \mathcal{D}_{\beta_n \alpha_n} \cdot \cdot \cdot \mathcal{D}_{\beta_2 \alpha_2} \quad ; \quad f^* = -\rho^* \cdot \mathcal{D}_{\beta_1 \alpha_1} f, \quad (\beta_1 > \alpha_1),$$

the left side of Capelli's identity takes the form

$$\rho \cdot f - \sum P f^*.$$

The degree of  $\mathcal{D}_{\beta_1 \alpha_1} f$  with respect to the  $\alpha_1$ -th argument is diminished by one, while the degree in  $x^{(\beta_1)}$ , ( $\beta_1 > \alpha_1$ ), is increased by one. The first effect is, however, decisive, as it places  $\mathcal{D}_{\beta_1 \alpha_1} f$  in a lower rank than  $f$ .

We keep in mind that (1)  $P$  is a succession of polar operations, and (2)  $f^*$  is of lower rank than  $f$  and is an invariant when  $f$  is an invariant. The fact that  $f^*$  is derived from  $f$  by a single polar operation  $D_{\beta_1 \alpha_1}$ , ( $\beta_1 > \alpha_1$ ), need not burden our memory.

In case  $m > n$  we put all terms except the main term on the right side of our equation. The result is an equation

$$(4.1) \quad \rho f = \sum P f^* \quad , \quad (m > n),$$

whose right side consists of terms arising by successive polarizations  $P$  from certain forms  $f^*$  of lower rank than  $f$ . Invariance of  $f$  implies invariance of  $f^*$ . The same holds true in the limiting case  $m = n$ ; one has merely to add to the right side of (4.1) the supplementary term  $[x^1 x^2 \dots x^m] \cdot \Omega f$ :

$$\rho f = \sum P f^* + [x^1 x^2 \dots x^m] \cdot \Omega f \quad , \quad (m = n).$$

The numerical factor  $\rho$  is  $> 0$  when  $f$  actually contains  $x^1$ .

Let us now suppose we are given a finite set of invariants

$$(4.2) \quad \varphi_1, \varphi_2, \dots, \varphi_\ell,$$

for which we want to prove "the first main theorem", that they form a complete set of basic invariants for our group  $\gamma$ . They depend of course on our  $m$  argument vectors  $x^1, x^2, \dots, x^m$ . We assume that the table (4.2) is closed under polarization:

Assumption I: Every  $D_{\beta\alpha} \varphi_j$  is itself one of the  $\varphi_j$ 's or at least is expressible by the set of the  $\varphi$ 's. ("Expressible" now always means expressible in an integral rational manner.)

I contend: this assumption once granted, the first main theorem holds good for  $m > n$  arguments provided it holds for  $m = n$  arguments,

Indeed, owing to Capelli's general identity (3.2),  $f$  is expressible by the set (4.2) if the invariants  $f^*$  occurring on the right side are so express-

ible; for in view of Assumption I, and the properties of polar operators given in § 1 of this chapter, such an expression of  $f^* = F^*(\varphi_1, \varphi_2, \dots, \varphi_\rho)$  leads to a similar expression of  $Pf^*$ , as one is able to turn the polar processes  $D$ , of which  $P$  is composed, upon the arguments  $\varphi_1, \dots, \varphi_\rho$  of  $F^*$ . Our reasoning presupposes that  $f$  actually contains  $x^1$ , for then  $\rho$  does not vanish. The  $f^*$  are of lower rank than  $f$ . In using complete induction with respect to the rank, we shall be stopped only when the invariant  $f$  under consideration becomes of degree  $\rho_1 = 0$  with respect to  $x^1$ . But this means that the first main theorem is true for  $m > n$  arguments  $x^1, x^2, \dots, x^m$  when its truth is granted for  $m-1$  arguments  $x^2, \dots, x^m$ . The induction with respect to the number of arguments will not come to an end before  $m$  has decreased to  $n$ .

One can go one step further and cut the number of arguments down to  $n-1$  by means of Capelli's special identity, if one adds the

Assumption II: The determinants  $[x^1 x^2 \dots x^m]$  occur among the  $\varphi$ 's or are expressible in terms of them.

It is perhaps more satisfactory to pronounce our result as dealing with a complete set of typical basic invariants. A given finite table of typical invariants which are linear in their typical arguments will yield a complete set for any number  $m$  of arguments if this can be shown to be true for  $n$  arguments; even  $n-1$  arguments will suffice provided the determinant  $[u^1 u^2 \dots u^m]$  appears in the table or is at least expressible by the invariants of the table.

In almost all cases the proof of the first main theorem consists of two parts: a formal part in which the reduction to  $n$  or  $n-1$  arguments takes place by dint of Capelli's identities, and another more conceptual part in which the proof for this restricted number of argument vectors is accomplished by considerations similar to those we used in proving the theorem of congruence in the

Introduction. The formal part could be carried through, as we have seen here in a general way for every possible group, whereas the other part cannot be brought down to such a general mechanized procedure and remains specific for each particular group. How this combination works out shall here be shown for some of the most important continuous groups.

### 5. Second example: Unimodular group $\gamma$ .

Every vector invariant of the full unimodular group is expressible by bracket factors; or the complete table of typical basic invariants consists of one term only: the determinant

$$[u^1 u^2 \dots u^m],$$

(of  $n$  typical argument vectors  $u^1, u^2, \dots, u^m$ ).

According to the general result obtained in the last section, we must prove this theorem merely for invariants  $f(x^1, \dots, x^{m-1})$  depending on  $n-1$  vectors  $x^1, x^2, \dots, x^{m-1}$ . We do this by showing that such an invariant must needs be a constant -- depending on no vector.

Let the vectors  $x^1, x^2, \dots, x^{m-1}$  be given numerically such that the determinant of their first  $n-1$  components does not vanish;

$$(5.1) \quad \Delta = \begin{vmatrix} x_1^1 & \dots & x_{m-1}^1 \\ x_1^2 & \dots & x_{m-1}^2 \\ \dots & \dots & \dots \\ x_1^{m-1} & \dots & x_{m-1}^{m-1} \end{vmatrix} \neq 0.$$

On introducing by means of a unimodular transformation

$$x^1, x^2, \dots, x^{m-1}, \frac{1}{\Delta} e^m,$$

as a new coordinate system instead of the original fundamental vectors

$e^1, e^2, \dots, e^m$ , our  $n-1$  vectors  $x^1, \dots, x^{m-1}$  are changed into the first  $n-1$  unit vectors:

$$\begin{aligned}(1, 0, \dots, 0, 0) &= \dot{x}^1, \\ (0, 1, \dots, 0, 0) &= \dot{x}^2, \\ \vdots & \\ (0, 0, \dots, 1, 0) &= \dot{x}^{m-1}.\end{aligned}$$

As  $f$  is an (absolute) invariant, we have

$$(5.2) \quad f(x^1, x^2, \dots, x^{m-1}) = c,$$

where  $c$  denotes the value of  $f$  for the particular arguments  $\dot{x}^1, \dot{x}^2, \dots, \dot{x}^{m-1}$ .

For general algebraic reasons an algebraic equation like (5.2) is a formal identity if it holds for all values of the arguments satisfying an algebraic inequality like (5.1). (This I call the argument of the "algebraic irrelevance of inequalities".)

### 6. Third example: Orthogonal group

In investigating the group  $\gamma$  of all proper and improper orthogonal transformations (rotations) it is convenient to take into consideration besides the absolute or "even" invariants, also the special kind of relative invariant called an odd invariant whose multiplier  $\mu(s)$  is +1 for proper, and -1 for improper rotations. An orthogonal transformation  $s$ , as one knows, is proper or improper according as its determinant is +1 or -1. The determinant  $[x^1, \dots, x^m]$  of  $n$  vectors is an odd invariant. Capelli's special identity shows that  $\Omega f$  is an odd or even invariant according as the invariant  $f$  is even or odd.

The even as well as the odd invariants are absolute invariants for the proper orthogonal group  $\bar{\gamma}$ , i.e. the group of all proper rotations. Vice versa: an absolute invariant  $f$  of  $\bar{\gamma}$  is carried by all proper rotations into itself, by all improper rotations into one and the same new form  $f'$ , and  $f$  is therefore the sum of an even and an odd invariant with respect to the total



in  $n-1$  dimensions, or more precisely, since they depend on exactly  $n-1$  vectors, to

$\Gamma_{m-1}^{m-1}$ . In view of this situation, it seems best first to pass from

$$(6.2) \quad \Gamma_{m-1}^{m-1} \rightarrow \Gamma_m^{m-1} \rightarrow \Gamma_m^m,$$

and then to generalize  $\Gamma_m^m$  to  $\Gamma_n^m$ . The two steps into which the transition

$\Gamma_{m-1}^{m-1} \rightarrow \Gamma_m^m$  breaks up, according to (6.2), are performed by the "non-formal"

argument and Capelli's special identity respectively, whereas the transition

$\Gamma_m^m \rightarrow \Gamma_n^m$ , ( $m > n$ ) rests on Capelli's general identity. As it is obvious how

to carry out the second part, we turn to the inductive proof of  $\Gamma_m^m$  according to

the scheme (6.2). Let us first restate:

$\Gamma_m^m$ . An even invariant depending on  $n$  vectors  $x^1, \dots, x^m$  in  $n$ -dimensional space is expressible in terms of their  $m^2$  scalar products: every odd invariant arises from an even one by multiplication with the bracket factor  $[x^1 \dots x^m]$ .

Proof: Step  $\Gamma_{m-1}^{m-1} \rightarrow \Gamma_m^{m-1}$ . Let

$$f(x, \dots, y) = f \left( \begin{matrix} x_1, \dots, x_{m-1}, x_m \\ y_1, \dots, y_{m-1}, y_m \end{matrix} \right),$$

be an even or odd invariant depending on  $n-1$  vectors  $x, \dots, y$  in  $n$  dimensions.

The function

$$f_c \left( \begin{matrix} x_1, \dots, x_{m-1} \\ y_1, \dots, y_{m-1} \end{matrix} \right) = f \left( \begin{matrix} x_1, \dots, x_{m-1}, 0 \\ y_1, \dots, y_{m-1}, 0 \end{matrix} \right),$$

is an orthogonal invariant of the same parity in  $n-1$  dimensions. If  $x, \dots, y$  are

numerically given and linearly independent, one can introduce a new coordinate sys-

tem  $e^1, \dots, e^{m-1}, e^m$ , by a proper orthogonal transformation such that

$x, \dots, y$  lie in the  $(n-1)$ -dimensional subspace  $\{e^1, \dots, e^{m-1}\}$ :

$$\begin{aligned} x &= \xi_1 e^1 + \dots + \xi_{m-1} e^{m-1}, \\ y &= \eta_1 e^1 + \dots + \eta_{m-1} e^{m-1}. \end{aligned}$$

Hence

$$(6.3) \quad f(x, \dots, y) = f_0 \left( \begin{matrix} \xi_1, & \dots, & \xi_{m-1} \\ \eta_1, & \dots, & \eta_{m-1} \end{matrix} \right) = f_0(\xi, \dots, \eta).$$

When  $f$  is odd, one gets the same equation with a minus sign in front of the right side upon replacing  $e^m$  by  $-e^m$ . Simultaneous validity of both equations implies

$$(6.4) \quad f(x, \dots, y) = 0.$$

When  $f$  is even, the even  $(n-1)$ -dimensional invariant  $f_0$ , (6.3), is expressible as a polynomial  $F$  in terms of the scalar products of its arguments

$(T_{m-1}^{m-1})$ . In this case our equation (6.3) furnishes

$$f(x, \dots, y) = F \left\{ \begin{matrix} (\xi\xi), & \dots, & (\xi\eta) \\ \vdots & & \vdots \\ (\eta\xi), & \dots, & (\eta\eta) \end{matrix} \right\}$$

and since  $(\xi\eta) = (x y)$ , we finally obtain:

$$(6.5) \quad f(x, \dots, y) = F \left\{ \begin{matrix} (xx), & \dots, & (xy) \\ \vdots & & \vdots \\ (yx), & \dots, & (yy) \end{matrix} \right\}$$

We proved the algebraic equations (6.4), (6.5) for numerical vectors  $x, \dots, y$  whose first  $n-1$  components have a non-vanishing determinant:

$$\begin{vmatrix} x_1, & \dots, & x_{m-1} \\ \vdots & & \vdots \\ y_1, & \dots, & y_{m-1} \end{vmatrix} \neq 0.$$

But such a restriction by an algebraic inequality is irrelevant, and consequently (6.4) and (6.5) are identities, i.e.,  $\Gamma_m^{m-1}$ : there does not exist any odd invariant form of  $n-1$  vectors in  $n$  dimensions while every even invariant of  $n-1$  vectors is expressible by their scalar products.

The other step  $\Gamma_m^{m-1} \rightarrow \Gamma_m^m$  is taken care of by Capelli's special identity applied to invariants  $f(x^1, \dots, x^m)$  depending on  $n$  vectors. Its right side

$$(6.6) \quad [x^1 \dots x^m] \cdot \Omega f,$$

contains the factor  $\Omega f$  of lower rank than  $f$ . If  $f$  is even, this  $\Omega f$  is odd and after it has been brought into the form:  $[x^1 \dots x^m]$  times (an aggregate of scalar products) one must resort to the equation

$$[x^1 \dots x^m]^2 = \det_{\alpha, \beta=1, \dots, m} (x^\alpha x^\beta),$$

in order to express the even invariant (6.6) in terms of scalar products only. It should be noticed that merely this special case of the equation (6.1) enters into our proof.

The reader is asked to compare the proof thus finished with the preliminary clumsy attempts at achieving the same end in our Introduction; the promise as given there is now redeemed.

#### 7. Fourth example: The complex group

A counterpart of the orthogonal group is what I am accustomed to call the complex group -- because of its relation to "complexes" of straight lines. Whereas an orthogonal transformation is characterized by the property of leaving unaltered the non-singular symmetric bilinear form

$$(7.1) \quad (xy) = x_1 y_1 + x_2 y_2 + \dots + x_m y_m,$$

a complex transformation  $s$  has the same property with respect to the non-degen-

erate anti-symmetric bilinear form,

$$(7.2) \quad \{xy\} = (x_1 y'_1 - x'_1 y_1) + \dots + (x_m y'_m - x'_m y_m),$$

where  $(x_1, x'_1, \dots, x_m, x'_m)$  are the components of the arbitrary vector  $x$  in our  $2n$ -dimensional space subject to the transformation  $s$ ; the vector  $y$  is, of course, to be transformed cogrediently. The set of all complex transformations is the complex group.  $\{xy\}$  may be called the skew product; it plays the same role for the complex group as the scalar product for the orthogonal group. The normal form (7.2) of the skew product is immaterial: a non-degenerate anti-symmetric bilinear form  $\{xy\}$  can exist only in a space of even dimension  $2n$ , and is carried into the normal form (7.2) by an appropriate linear transformation -- just as an arbitrary non-degenerate symmetric bilinear form can be normalized into (7.1). Note, however, that while the latter fact is true only on the assumption that the square root of every number can be extracted within the underlying number field, the corresponding proposition about the complex group is a purely rational matter and is, therefore, not bound to any restriction as to the number field.

There are a number of instances in which the complex group behaves "rationally" where the orthogonal group requires extraction of square roots. We have to avail ourselves of two additional such occurrences:

(1) Every complex transformation is of determinant 1; hence the distinction between proper and improper transformations that caused some complications for the orthogonal group is here lacking.

(2) The bracket factor, the determinant  $[x^1, \dots, x^m]$  of  $2n$  vectors  $x^1, x^2, \dots, x^{2m}$  is integral-rationally expressible in terms of the skew products  $\{x^\alpha x^\beta\}$  -- whereas in the case of the orthogonal group, but the square of the determinant could be expressed by the scalar products.

To prove these two statements we form the alternating sum

$$(7.3) \quad \frac{1}{n! 2^n} \sum \pm \{x^1 x^2\} \{x^3 x^4\} \dots \{x^{2n-1} x^{2n}\},$$

extending to all  $(2n)!$  permutations of the  $2n$  independent vectors  $x^1, x^2, \dots, x^{2n}$ . The factor  $1/n! 2^n$  was added because a group of that number among these  $(2n)!$  permutations leaves unchanged the leading term of our sum, namely the group generated first by the transpositions of two arguments like  $x^1$  and  $x^2$  or  $x^3$  and  $x^4$  united in the same parenthesis  $\{ \}$  and secondly by the  $n!$  permutations of the  $n$  parentheses.  $\{xy\}$  may here at first designate an arbitrary skew symmetric form:

$$(7.4) \quad \{xy\} = \sum_{i,k} g_{ik} x_i y_k.$$

(7.3) then becomes

$$\frac{1}{n! 2^n} \sum_{i_1, \dots, i_{2n}} \left\{ g(i_1, i_2) \dots g(i_{2n-1}, i_{2n}) \sum \pm x_{i_1}^1 x_{i_2}^2 \dots x_{i_{2n}}^{2n} \right\},$$

with the inner sum running again over all permutations of  $x^1, \dots, x^{2n}$ .

This inner sum vanishes unless  $i_1, \dots, i_{2n}$  is a permutation  $p$  of  $1, \dots, 2n$ , and it equals  $\pm [x^1 x^2 \dots x^{2n}]$  according as  $p$  is an even or odd permutation.

We are thus led to introduce the "Pfaffian"

$$\text{Pf} \{g_{ik}\} = \frac{1}{n! 2^n} \sum \pm g(i_1, i_2) \dots g(i_{2n-1}, i_{2n}),$$

in which the sum runs alternately over all permutations  $i_1, i_2, \dots, i_{2n-1}, i_{2n}$  of  $1, \dots, 2n$ . The Pfaffian plays the same role for the anti-symmetric form

(7.4) as the determinant plays for symmetric forms. Our result is the formula

$$(7.5) \quad \frac{1}{n! 2^n} \sum \pm \{x^1 x^2\} \dots \{x^{2n-1} x^{2n}\} = \text{Pf} \{g_{ik}\} \cdot [x^1 x^2 \dots x^{2n}].$$

When (7.4) goes over into

$$\sum \bar{g}_{ik} \bar{x}_i \bar{y}_k$$

by a linear transformation

$$x_i = \sum s_j^i \bar{x}_j,$$

(to be cogrediently performed on  $x$  and  $y$ ) (7.5) leads to the further relation

$$(7.6) \quad \text{Pf}\{\bar{g}(ik)\} = \text{Pf}\{g(ik)\} \cdot \det.(s_j^i).$$

(It corresponds to the relation

$$|\bar{g}(ik)| = |g(ik)| \cdot (\det s_j^i)^2$$

for symmetric forms  $g(ik)$  -- which, however, involves the square of the transformation determinant!)

When the normal form: (7.1) or

$$(7.7) \quad \sum_{i,k} e(ik) x_i y_k, \quad (i,k = 1, 1', \dots, m, m'),$$

is adopted for  $xy$ , the Pfaffian  $\text{Pf}\{e(ik)\}$  becomes = 1. Hence (7.6) shows that a substitution  $S_j^i$ , leaving this form  $\{xy\}$  unaltered, must be of determinant 1. That was the first point. And the second point: on clinging to the same normal form for the skew product, the equation (7.5) simplifies to

$$[x^1 \dots x^{2m}] = \frac{1}{m! 2^m} \sum \pm \{x^1 x^2\} \dots \{x^{2m-1} x^{2m}\},$$

and thus gives the desired expression of the bracket factor in terms of the skew products.--

The first main theorem concerning invariants of the complex group asserts that a full table of basic invariants consists of one type only: the skew product  $\{xy\}$ . The bracket factor need not be added because of its being expressible by the skew products as shown above. The proof of the main theorem may be carried through exactly along the same lines as for the orthogonal group -- with the simplification arising from this redundancy of the bracket factor which eliminates at the same time the distinction of proper and improper transformations. There is, however, one further observation to be made lest the induction with respect to the dimensionality  $2n$  would be stopped: Then consider-

ing an invariant  $f$  depending on  $2n-1$  vectors  $x, y, \dots$ , one introduces a new "normal" coordinate system by means of an appropriate complex transformation such that the first components  $x_1, y_1, \dots$  of  $x, y, \dots$  vanish relatively to the new coordinate system. Here  $x, y, \dots$  are supposed to be numerically given and linearly independent. After thus being led to introduce the function

$$f_0 \left( \begin{matrix} x'_1, x_2, x'_2, \dots \\ y'_1, y_2, y'_2, \dots \end{matrix} \right) = f \left( \begin{matrix} 0, x'_1, x_2, x'_2, \dots \\ 0, y'_1, y_2, y'_2, \dots \end{matrix} \right),$$

we should get into trouble if the arguments  $x'_1, y'_1, \dots$  did not disappear from  $f_0$  along with  $x_1, y_1, \dots$ ; the induction from  $2(n-1)$  dimensions to  $2n$  would not work. Fortunately this obstacle can be overcome by the following simple reasoning. We shall not touch the components  $x_2, x'_2, \dots$ , and on the only ones we put in evidence,  $x_1, x'_1$ , may be performed an arbitrary unimodular transformation

$$\begin{aligned} x_1 &\rightarrow \alpha x_1 + \beta x'_1, \\ x'_1 &\rightarrow \gamma x_1 + \delta x'_1, \end{aligned} \quad (\alpha\delta - \beta\gamma = 1).$$

As this is a complex transformation and  $f$  an invariant, we have

$$(7.8) \quad f(0, x'_1; 0, y'_1; \dots) = f(\beta x'_1, \delta x'_1; \beta y'_1, \delta y'_1; \dots).$$

(7.8) holds for arbitrary numbers  $\beta$  and  $\delta$  if only  $\delta \neq 0$ . For then we may choose  $\alpha = 1/\delta$ ,  $\gamma = 0$ . The argument of the algebraic irrelevance of inequalities (like  $\delta \neq 0$ ) establishes (7.8) as an identity in the variables  $\beta$  and  $\delta$ . Consequently (7.8) remains true for the values  $\beta = 0, \delta = 0$ . The equation thus arising

$$f(0, x'_1; 0, y'_1; \dots) = f(0, 0; 0, 0; \dots)$$

proves that  $f_0(x'_1, y'_1, \dots)$  is in fact independent of  $x'_1, y'_1, \dots$ .

### 8. The extension theorem

In  $n$ -dimensional Euclidean space a plane is represented by a linear equation

$$x_1 u_1 + \dots + x_m u_m + \xi = 0 ,$$

in the coordinates  $u_1, \dots, u_m$  of a generic point; the coefficients  $x_1, \dots, x_m, \xi$  figure as the (homogeneous) coordinates of the plane.

Transition to another Euclidean coordinate system is performed in terms of these plane coordinates by means of a transformation of the following type:

$$\begin{aligned} x'_i &= \sum O_{ik} x_k , \quad (i, k = 1, 2, \dots, m) , \\ \xi' &= c_1 x_1 + \dots + c_m x_m + \xi , \end{aligned}$$

where  $\|O_{ik}\|$  is an orthogonal matrix while the constants  $c_1, \dots, c_m$  are subject to no restriction. The orthogonal matrix  $\|O_{ik}\|$  appears bordered with a rim of width 1 in the following fashion

$$\left| \begin{array}{cccc|c} O_{11} & \dots & \dots & O_{1m} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ O_{m1} & \dots & \dots & O_{mm} & 0 \\ \hline c_1 & \dots & \dots & c_m & 1 \end{array} \right|$$

The small width 1 of the rim is due to the fact that the  $n$ -dimensional Euclidean space, when viewed from the projective standpoint, has an  $(n-1)$ -dimensional plane and not a linear manifold of less dimensions as its infinite locus. Our "extension" of the orthogonal group may thus suitably be subsumed under the following general bordering procedure: Given a group  $\gamma$  of unimodular linear transformations  $s$ :

$$x'_i = \sum S^i_k x_k , \quad (i, k = 1, 2, \dots, m) ,$$

in  $n$ -dimensional vector space  $\mathcal{R}_m$ . We consider transformations  $\bar{s}$  of  $n+v$  variables  $x_1, \dots, x_m, \xi_1, \dots, \xi_v$  of the following structure



The transformation that leaves all the components  $x_1, \dots, x_m, \xi_1, \dots, \xi_{v-1}$  unaltered but replaces  $\xi_v$  by

$$\xi'_v = c_1 x_1 + \dots + c_m x_m + \sigma_1 \xi_1 + \dots + \sigma_{v-1} \xi_{v-1} + \xi_v,$$

belongs to the extended group  $\bar{\gamma}_v$ . The constants  $c_1, \dots, c_m, \sigma_1, \dots, \sigma_{v-1}$  may be determined in such a manner that the last component  $\xi'_v$  ( $\eta'_v, \dots$ ) of each of the numerically given vectors  $x(y, \dots)$  vanishes. Because of the invariance of  $f$  one then obtains

$$f \left( \begin{array}{c} x_1, \dots, x_m, \xi_1, \dots, \xi_{v-1}, \xi_v \\ y_1, \dots, y_m, \eta_1, \dots, \eta_{v-1}, \eta_v \end{array} \right) = f \left( \begin{array}{c} x_1, \dots, x_m, \xi_1, \dots, \xi_{v-1}, 0 \\ y_1, \dots, y_m, \eta_1, \dots, \eta_{v-1}, 0 \end{array} \right)$$

The restricting algebraic inequality (8.1) may here be disregarded according to the principle of the algebraic irrelevance of such inequalities. Our result, that  $f$  does not actually depend on the  $v^{\text{th}}$  Greek component  $\xi_v$  ( $\eta_v, \dots$ ) applies to the other Greek components as well.

Examples: 1) Affine space. If one takes the ~~name~~ "affine transformation" in  $n$ -dimensional vector space in its historically primordial sense as it was used for instance by Euler (a linear transformation which preserves the volumes of  $n$ -dimensional parallelepipeda), then the affine transformations of point space, when described in terms of plane coordinates, form a group  $\bar{\gamma}_1$  derived from the unimodular group  $\gamma$  in  $n$  dimensions by the bordering process with a rim of width 1. Consequently all invariants of an arbitrary number of planes  $x = (x_1, \dots, x_m, \xi)$  are expressible by the two types: the  $n$ -rowed bracket factor

$$[x \dots y]_n = \begin{vmatrix} x_1 & \dots & x_m \\ \dots & \dots & \dots \\ y_1 & \dots & y_n \end{vmatrix}$$

depending on the "Latin" components of  $n$  planes  $x, \dots, y$  and the  $(n+1)$ -rowed bracket factor

$$[x \dots y z]_{n+1} = \begin{vmatrix} x_1, \dots, x_m, \xi \\ \cdot & \cdot & \cdot & \cdot \\ y_1, \dots, y_m, \eta \\ z_1, \dots, z_m, \zeta \end{vmatrix}$$

depending on all components of  $n+1$  planes  $x, \dots, y, z$ .

2) Affine space of rank  $\nu$ . An  $(m + \nu - 1)$ -dimensional projective space with the plane coordinates  $x_1, \dots, x_m, \xi_1, \dots, \xi_\nu$  is changed into an affine space of rank  $\nu$  by designating as the infinite locus a certain linear  $(n-1)$ -dimensional manifold; namely, that manifold in which all the planes

$$(0, \dots, 0, \xi_1, \dots, \xi_\nu)$$

intersect. Its group of automorphisms is the extension of order  $\nu$  of the unimodular group in  $n$  variables. Ordinary affine space is the space of rank 1 according to this nomenclature. The application of our extension theorem is obvious.

3) On starting with the unimodular group and repeating the bordering process several times, one gets what may suitably be called the group of step transformations. Let us suppose, for definiteness, that our staircase consists of three steps of width  $p, q-p, m-q$ :

$$0 < p < q < m.$$

The components of an arbitrary vector are split up into three sections:

$$(8.2) \quad x_1, \dots, x_p, | x_{p+1}, \dots, x_q, | x_{q+1}, \dots, x_m,$$

and a "step transformation" is one transforming the first  $p$  components  $x_1, \dots, x_p$  among themselves, then the first  $q$  components  $x_1, \dots, x_q$  (and finally all the  $n$  components). The matrices of such transformations are of the following form

0	*	0	0
p		*	0
q			*
n			

where the rectangles on the right of the main diagonal marked by 0 are void, i.e. filled with zeros, whereas the squares \* along the main diagonal bear unimodular matrices. Denoting by  $X_1, X_2, X_3$  the three parts into which the column of the  $n$  components  $x_1, \dots, x_m$  breaks up according to (8.2), these transformations are of this form

$$\begin{aligned} X'_1 &= A_{11} X_1, \\ X'_2 &= A_{21} X_1 + A_{22} X_2, \\ X'_3 &= A_{31} X_1 + A_{32} X_2 + A_{33} X_3, \\ \{ |A_{11}| = |A_{22}| = |A_{33}| = 1 \}. \end{aligned}$$

The typical basic invariants forming a complete set are the bracket factors of order  $p, q, n$ , respectively, namely the determinant of the first  $p$  components of  $p$  arbitrary vectors, that of the first  $q$  components of  $q$  vectors, and finally the determinant of all  $n$  components of  $n$  vectors.

4) A special case of the step transformations is the translation group in plane coordinates  $(x_1, \dots, x_m, \xi)$ ,

$$\begin{aligned} x'_i &= x_i, \quad (i=1, \dots, m) \\ \xi' &= c_1 x_1 + \dots + c_m x_m + \xi \end{aligned}$$

$(c_1, \dots, c_m, \text{arbitrary})$ . A full set of typical basic invariants is furnished by the  $n$  variables  $x_1, x_2, \dots, x_m$ , together with the determinant

$$\begin{vmatrix} x_1, \dots, x_m, \xi \\ \cdot \cdot \cdot \cdot \cdot \\ z_1, \dots, z_m, \xi \end{vmatrix} = [x \cdot \cdot z]$$

of  $n+1$  vectors  $x, \dots, z$ .

5) The "semi-invariants" are another special case; they refer to that step group for which the partition (8.3) separates all the individual components  $x_1, x_2, \dots, x_m$  from each other. In other words,  $\gamma$  now consists of all "recursive" transformations

$$\begin{aligned}x'_1 &= a_{11} x_1, \\x'_2 &= a_{21} x_1 + a_{22} x_2, \\&\dots \\x'_m &= a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mm} x_m,\end{aligned}$$

whose coefficients  $a_{11}, a_{22}, \dots, a_{mm}$  along the main diagonal equal to 1. Invariants with respect to this group are called semi-invariants. A complete table is made up by the "brackets":

$$x_1, [xy]_2, \dots, [xy \dots z]_m,$$

$[x \dots y]_i$  being the determinant of the first  $i$  components of  $i$  vectors  $x, \dots, y$ :

$$\begin{vmatrix} x_1 & \dots & x_i \\ \vdots & \ddots & \vdots \\ y_1 & \dots & y_i \end{vmatrix}$$

6) The case of  $n$ -dimensional Euclidean space from which we started is easily generalized to the notion of  $(n + \nu - 1)$ -dimensional Euclidean space of rank  $\nu$ : here the infinite consists of a linear manifold of  $\nu$  dimensions less than the total space. Its group of automorphisms, when written in homogeneous plane coordinates  $(x_1, \dots, x_n, \xi_1, \dots, \xi_\nu)$  is derived from the  $n$ -dimensional orthogonal group by bordering it with a rim of width  $\nu$ . A full table of typical basic invariants accordingly consists first of the scalar product

$$(x|y) = x_1 y_1 + \dots + x_n y_n,$$

second, of the "Latin" bracket factor depending on  $n$  vectors  $x, \dots, y$ :

$$[x \dots y]_n = \begin{vmatrix} x_1 & \dots & x_n \\ \vdots & & \vdots \\ y_1 & \dots & y_n \end{vmatrix},$$

and third, of the "total" bracket factor depending on  $m+v$  arbitrary vectors

$x, \dots, z$ :

$$[x \dots z]_{m+v} = \begin{vmatrix} x_1 & \dots & x_m & \xi_1 & \dots & \xi_v \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ z_1 & \dots & z_m & \xi_1 & \dots & \xi_v \end{vmatrix}.$$

When we remove the restrictions

$$(8.3) \quad \det.(O_{ik}) = 1, \quad |\sigma| = \begin{vmatrix} \sigma_{11} & \dots & \sigma_{1v} \\ \vdots & & \vdots \\ \sigma_{v1} & \dots & \sigma_{vv} \end{vmatrix} = 1,$$

imposed upon our bordered transformations s:

$$\begin{array}{c|c} O_{11}, \dots, O_{1m} & 0, \dots, 0 \\ \vdots & \vdots \\ O_{m1}, \dots, O_{mm} & 0 \dots 0 \\ \hline C_{11}, \dots, C_{1m} & \sigma_{11}, \dots, \sigma_{1v} \\ \vdots & \vdots \\ C_{v1}, \dots, C_{vm} & \sigma_{v1} \dots \sigma_{vv} \end{array}$$

and thus allow the rotational part to be improper  $\{\det.(O_{in}) = -1\}$ , as well as proper, and permit any value  $\neq 0$  for the determinant of the  $\sigma$ 's, then the basic invariants which were absolute before, become relative ones, even or odd in the head components, of a multiplier  $|\sigma|^g$  and hence of a certain integral weight  $g$  in the tail components. Our table remains a complete table of basic relative invariants because other types of "relativity" are impossible.

The group taken in this slightly widened sense may be described as the group of all transformations in  $(m + \nu)$ -dimensional vector space leaving invariant a given quadratic form

$$(8.4) \quad \sum g_{ik} x_i x_k .$$

The form depends on all  $m + \nu$  components  $x_i$  of the arbitrary vector  $x$ , and  $\nu$  may be its index of degeneracy. Indeed, such a form can be linearly transformed into the square sum

$$(8.5) \quad x_1^2 + \dots + x_m^2 ,$$

of a number of the variables, the full set of which we now denote by  $x_1, \dots, x_m, \xi_1, \dots, \xi_\nu$ , and the linear transformations of all these  $m + \nu$  variables, which leave (8.5) unaltered, constitute our group derived from the orthogonal group in the  $n$  variables  $x_1, \dots, x_m$  by extension of rank  $\nu$ . The study of Euclidean geometry of any rank in terms of plane coordinates is thus the study of the group of transformations that leaves a given quadratic form unchanged; the differences in "rank" arise from the extent to which such a form may degenerate.

7) In similar fashion one may consider a  $(2m + \nu)$ -dimensional vector space, whose group of automorphisms is the complex group in  $2n$  dimensions, as bordered with a rim of width  $\nu$ . The "skew product" and the "total bracket factor" are the two types of invariants sufficient to furnish a complete basis of all invariants.

One may justly rejoice at the simplicity of these considerations, but then one should not forget that they solve only the question of invariants for figures consisting of planes; and it becomes quite a different matter when we are to consider point as well as plane coordinates, as we shall now do in Part B.

INSTITUTE  
FOR ADVANCED STUDY  
PRINCETON

B. Covariant as well as Contravariant Vectors

1. Trivial examples

In this part we stick to the convention that covariant vectors are designated by Latin, contravariant ones by Greek letters.

1) Orthogonal group. In the case of the first orthogonal group, every contravariant vector  $\xi$  is a covariant vector  $x$ : their transformation laws coincide, since an orthogonal transformation is identical with its conjugredient. The complete list of basic typical invariants of an arbitrary number of covariant vectors consists, as we know, of the scalar product only  $(xy)$ , if  $\gamma$  includes the improper as well as the proper transformations, and only absolute invariants are taken into account. Consequently if both kinds of vectors, covariant and contravariant ones, are admitted, there appear three types in the complete table: first, the scalar product  $(xy)$  of two covariant vectors  $x, y$ , second the product  $(x\xi)$  of a covariant vector  $x$  with a contravariant  $\xi$ , third the scalar product  $(\xi\eta)$  of two contravariant vectors  $\xi, \eta$ .

2) Hardly less trivial is the case of the complex group. We start with the skew product  $\{xy\}$  of two covariant vectors  $x, y$ :

$$(1.1) \quad \{xy\} = (x_1 y_1' - x_1' y_1) + \dots + (x_n y_n' - x_n' y_n)$$

The relations

$$(1.2) \quad x_1 = \xi_1', \quad x_1' = -\xi_1; \quad \dots; \quad x_n = \xi_n', \quad x_n' = -\xi_n$$

link up a covariant vector  $x$  with a contravariant  $\xi$ , because these equations can be summed up in the one relation

$$\{xy\} = \{\xi y\}$$

holding identically in  $y$ , and this relation is invariant under all complex transformations. This transition (1.2) from  $\xi$  to  $x$  may be indicated by  $x = \xi'$ .

When considering both kinds of vectors we are therefore free to replace each contravariant vector  $\xi$  by its corresponding covariant  $\xi'$ . Hence

$$\{\xi'x\}, \{\xi'\eta'\}$$

are to be taken into account besides the original  $\{xy\}$ ; the complete table one thus obtains consists of the three types

$$\{xy\}, (x\xi), \{\xi\eta\}.$$

Remark. When

$$\{xy\} = \sum g_{ik} x_i x_k$$

is the skew product of two Latin vectors, the skew product of two Greek vectors  $\xi, \eta$  is the form  $\sum \gamma_{ik} \xi_i \xi_k$ , whose coefficients  $\gamma_{ik}$  form the inverse matrix to  $\|g_{ik}\|$ ; for the normal form (1.1) however, both matrices are  $\|e_{ik}\|$ .]

## 2. The general method

For any group  $\gamma$  whatsoever the product  $(x\xi)$  of an arbitrary Latin with an arbitrary Greek vector will surely appear as a standard figure among the basic invariants. But this trivial remark does not lead us very far; and only in the simplest cases shall we get off as cheaply as with the orthogonal or complex group. One is however in possession of a general method allowing free passage from contravariant to covariant vectors. It was developed for the purpose of the problem at hand, above all by Weitzenböck, under the name of "complex calculus". The basic fact is that  $n-1$  covariant vectors  $x^{(1)}, \dots, x^{(n-1)}$  in  $n$ -dimensional vector space give rise to a contravariant vector  $\xi$  whose components  $\xi_i$  are the  $n-1$  rowed minors of the matrix

$$\left\| \begin{array}{cccc} x_1^{(1)} & \cdot & \cdot & \cdot & x_n^{(1)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_1^{(n-1)} & \cdot & \cdot & \cdot & x_n^{(n-1)} \end{array} \right\|$$

More exactly,  $\xi_i$  equals the minor which is left after dropping the  $i^{\text{th}}$  column with the sign  $(-1)^{n-1}$  attached to it. This shall be indicated by the formula

$$\xi = [\alpha^{(1)}, \dots, \alpha^{(m-1)}].$$

Our process -- corresponding to the fundamental constructions in  $(n-1)$ -dimensional projective space of determining a plane by  $n-1$  incident points or a point by  $n-1$  incident planes -- is invariant under all unimodular transformations, and may thus be applied to any group consisting of such transformations. For the definition of  $\xi$  may be condensed into one identity:

$$\begin{vmatrix} \alpha_1^{(1)} & \dots & \alpha_m^{(1)} \\ \vdots & \ddots & \vdots \\ \alpha_1^{(m-1)} & \dots & \alpha_m^{(m-1)} \\ z_1 & \dots & z_m \end{vmatrix} = \xi_1 z_1 + \dots + \xi_m z_m,$$

involving the indeterminate covariant vector  $z$ , or

$$[\alpha^{(1)}, \dots, \alpha^{(m-1)} z] = (\xi z).$$

In this form the invariance is obvious. Another suitable form of expressing the same relation is

$$(2.1) \quad \begin{vmatrix} \alpha_{i_1}^{(1)} & \dots & \alpha_{i_{m-1}}^{(1)} \\ \vdots & \ddots & \vdots \\ \alpha_{i_1}^{(m-1)} & \dots & \alpha_{i_{m-1}}^{(m-1)} \end{vmatrix} = \delta(i_1, \dots, i_{m-1}, i_m) \cdot \xi_{i_m}$$

where  $i_1, i_2, \dots, i_{m-1}, i_m$  is any permutation of  $1, \dots, n$  and  $\delta(i_1, \dots, i_{m-1}, i_m)$  the linear character = +1 or -1, according as  $i_1, \dots, i_{m-1}, i_m$  is an even or odd permutation.

We now proceed to investigate invariants depending on an arbitrary number of covariant vectors and a given number of contravariant ones  $\xi, \eta, \dots$ , in the following manner. Let us suppose that the invariant  $f$  under considera-

tion is a homogeneous polynomial of degree  $h$  with respect to the components of  $\xi$ . We first perform the polar process:

$$\frac{1}{h} \sum_{i=1}^m \xi_i^* \frac{\partial f}{\partial \xi_i},$$

introducing a new auxiliary contravariant vector  $\xi^*$ . The polarized invariant, linear in  $\xi^*$ , is changed back into  $f$  by identifying  $\xi^*$  with  $\xi$ . We now replace  $\xi^*$  by  $n-1$  auxiliary covariant vectors  $\alpha^{(1)}, \dots, \alpha^{(n-1)}$ , according to

$$\xi^* = [\alpha^{(1)}, \dots, \alpha^{(n-1)}].$$

On contracting the two steps into one, the transition we accomplished is then from  $f(\xi)$  to

$$(2.2) \quad \bar{f}(\alpha^{(1)}, \dots, \alpha^{(n-1)}; \xi) = \begin{vmatrix} \alpha_1^{(1)}, \dots, \alpha_n^{(1)} \\ \vdots \\ \alpha_1^{(n-1)}, \dots, \alpha_n^{(n-1)} \\ \frac{\partial f}{\partial \xi_1}, \dots, \frac{\partial f}{\partial \xi_n} \end{vmatrix}$$

The degree of  $\bar{f}$  in  $\xi$  is less by unity than the degree of  $f(\xi)$  in  $\xi$ ; this advantage is, however, to a certain degree compensated for by the introduction of  $n-1$  new Latin vectors  $\alpha^{(1)}, \dots, \alpha^{(n-1)}$ . This is why we have to insist that no limitations should be put upon the number of Latin arguments (whereas the number of Greek arguments may be fixed in our investigation once for all). We call the new vectors  $\alpha^{(1)}, \dots, \alpha^{(n-1)}$  the symbolic vectors, and the whole process of changing  $f$  into  $\bar{f}$  we name "declination". After one repeats the process  $h$  times (each time with new symbolic covariant vectors) the vector  $\xi$  has been eliminated entirely. Then  $\eta$  may be thrown out in the same fashion. But when we finally have got rid of all the Greek arguments a lot of new "symbolic" Latin arguments will have made their appearance.

Let us now assume we are given a finite set  $T$  of invariants depending on the given Greek vectors  $\xi, \eta, \dots$ , and some typical Latin vectors. We wish to show that  $T$  provides a complete basis for all invariants of these Greek vectors  $\xi, \eta, \dots$ , and any number of Latin vectors  $x, y, \dots$ ,-- in the sense as indicated by the term "typical", namely that the typical Latin arguments in the table are to be replaced by the actual arguments  $x, y, \dots$  in all possible combinations. Let us further assume at the outset that our table  $T$  of "basic invariants" contains the complete table of typical basic invariants for covariant vectors only (which we suppose to be known to us). We get at our proof by means of induction with respect to the degree  $h$  of the invariant  $f$  in one of the Greek arguments, say  $\xi$ . We then are allowed to suppose that the declined  $\bar{f}$ , of order  $h-1$ , is expressible in the desired fashion and we must try to wring from it a similar expression for  $f$ .

$\bar{f}$  is a linear combination of "terms"; an individual term is a product of basic invariants that contains the symbolic vectors  $\alpha^{(1)}, \dots, \alpha^{(n-1)}$ , linearly. The symbolic arguments are spread in some way over the factors of the term; no symbolic argument appears twice and one assumes, as it is convenient to do, that the typical basic invariants are fully polarized with respect to their symbolic Latin arguments and thus contain them in linear form. Whereas  $\bar{f}$  is skew symmetric in the symbolic arguments, its individual terms may not be. Hence, in reversing the process of declination we first take care of insuring skew symmetry through alternation with respect to the symbolic arguments and afterwards we replace the determinants (2.1) by the components of the Greek argument  $\xi$ . Thus the process of restitution to be performed on a polynomial  $g(\alpha^{(1)}, \dots, \alpha^{(n-1)})$ , depending linearly on the  $n-1$  symbolic Latin vectors  $\alpha^{(1)}, \dots, \alpha^{(n-1)}$ , is carried out in two steps:

1) Alternation  $g \rightarrow g^*$ :

$$(2.3) \quad g^*(x^{(1)}, \dots, x^{(n-1)}) = \frac{1}{(n-1)!} \sum_{\pm} g(x^{(1)}, \dots, x^{(n-1)});$$

the sum extending alternately to all permutations of the  $n-1$  arguments

$x^{(1)}, \dots, x^{(n-1)}$ . The skew symmetric multilinear  $g^*$  can be written in the form of a determinant

$$(2.4) \quad g^*(x^{(1)}, \dots, x^{(n-1)}) = \begin{vmatrix} x_1^{(1)} & \dots & x_m^{(1)} \\ \dots & \dots & \dots \\ x_1^{(n-1)} & \dots & x_m^{(n-1)} \\ A_1 & \dots & A_m \end{vmatrix}$$

where the  $A$ 's do not contain the symbolic arguments.

2) In the second step (2.3) it replaced by

$$(2.5) \quad \tilde{g} = A_1 \xi_1 + \dots + A_m \xi_m.$$

The contravariant vector  $\xi$  is called the restituent. The sign  $\rightarrow\rightarrow$  may serve as the symbol of restitution:  $g \rightarrow\rightarrow \tilde{g}$ . It is by no means excluded that  $g$ , and hence  $g^*$  depend on some other arguments, even on the restituent  $\xi$  besides the symbolic vectors  $x^{(1)}, \dots, x^{(n-1)}$ . The process of restitution changes  $\bar{f}$ , (2.2), back into  $f$ , when  $\bar{f}$  arose from  $f$  by declination.

In our inductive proof of the first main theorem we suppose that we have succeeded in expressing  $\bar{f}$  in terms of the basic invariants. Our goal is to carry out restitution in each term of this expression. Let us pick out a single term, a product of basic invariants with the symbolic arguments  $x^{(1)}, \dots, x^{(n-1)}$  spread over its several factors and then try to show that such a term by the process of restitution goes over into some aggregate of the basic invariants again.

The essentials of this section can be summarized by two definitions and a general proposition.

I. Definitions: a) Process of declination, applicable to any polynomial  $f(\xi)$  : eq. (2.2); b) The inverse process of restitution  $g \rightarrow \tilde{g}$  applicable to any  $g$  multilinear in the symbolic arguments  $x^{(1)}, \dots, x^{(m-1)}$  eqs. (2.3), (2.4), (2.5) combined.

II. The general proposition: We suppose that a finite set  $T$  of "basic invariants" depending on a given number of Greek arguments  $\xi, \eta, \dots$  and some typical Latin arguments is given. The table  $T$  is complete when two things can be shown:

1) It contains the complete table of typical basic invariants that depend on covariant vectors alone.

2) Any product of basic invariants containing the symbolic arguments  $x^{(1)}, \dots, x^{(m-1)}$  in linear form is changed by restitution into a linear combination of such products. (This second assumption may be referred to as: closure of  $T$  under restitution.)

To this general summary, I add a supplementary remark and the two simplest examples of restitution; they serve as an illustration, but we need them for later applications as well.

Remark: When applying restitution to a product of the kind just described, one is at liberty to discard factors that do not contain a symbolic argument, in particular factors depending on Greek arguments only.

Example 1. The bracket factor  $[x^{(1)} \dots x^{(m-1)} x]$  when it contains all  $n-1$  symbolic vectors  $x^{(1)}, \dots, x^{(m-1)}$ , is carried by restitution into the product  $(\xi x)$  of the contravariant restituent  $\xi$  with the covariant vector  $x$ :

$$[x^{(1)} \dots x^{(m-1)} x] \rightarrow (\xi x).$$

Example 2.

$$(x^{(1)} \xi^{(1)}) \dots (x^{(m-1)} \xi^{(m-1)}) \rightarrow \frac{1}{(m-1)!} [\xi^{(1)} \dots \xi^{(m-1)} \xi].$$

Proof: The left-hand product is changed by alternation into

$$\frac{1}{(m-1)!} \sum_{i_1, \dots, i_{m-1}} \left\{ \xi_{i_1}^{(1)} \dots \xi_{i_{m-1}}^{(m-1)} S x_{i_1}^{(1)} \dots x_{i_{m-1}}^{(m-1)} \right\}.$$

The inner sum indicated by S extends alternately to all permutations of  $x^{(1)}, \dots, x^{(m-1)}$ , whereas in the outer sum all indices  $i_1, \dots, i_{m-1}$  range independently from 1 to n. The inner sum is different from zero only if the numbers  $i_1, \dots, i_{m-1}$  are pairwise distinct; by the process of restitution it then gets changed into

$$\delta(i_1 \dots i_{m-1} i_m) \cdot \xi_{i_m}$$

$i_m$  being the one number of the row 1, 2, ..., n missing in the sequence  $i_1, \dots, i_{m-1}$ . Hence the result is

$$\frac{1}{(m-1)!} \sum \delta(i_1 \dots i_{m-1} i_m) \xi_{i_1}^{(1)} \dots \xi_{i_{m-1}}^{(m-1)} \xi_{i_m},$$

with the sum running over all permutations  $i_1, \dots, i_{m-1}, i_m$  of 1, 2, ..., n, or  $1/(n-1)!$  times the determinant  $[\xi^{(1)} \dots \xi^{(m-1)} \xi]$ .

### 3. A formal identity

As a necessary instrument for carrying out restitution on a product of brackets like  $[x y \dots z]$  and inner products like  $(x \xi)$  we need a certain formal identity concerning determinants. With  $n+1$  Latin vectors  $x^{(1)}, \dots, x^{(i)}$ ,  $y, z, u, \dots$  and  $n$  further numbers  $\pi_1, \dots, \pi_n$  which may conveniently be looked upon as the components of a Greek vector  $\pi$ , we form the equation

$$\begin{vmatrix} (x^{(1)} \pi), & x_1^{(1)}, & \dots, & x_n^{(1)} \\ \dots & \dots & \dots & \dots \\ (x^{(i)} \pi), & x_1^{(i)}, & \dots, & x_n^{(i)} \\ \dots & \dots & \dots & \dots \\ (y \pi), & y_1, & \dots, & y_n \\ (z \pi), & z_1, & \dots, & z_n \end{vmatrix}$$

This determinant is zero because the first column is a linear combination of the  $n$  subsequent columns by means of the numbers  $\pi_1, \pi_2, \dots, \pi_n$ . On expanding the determinant on the left side according to the elements of the first column, and leaving the terms above the dotted line on the left, while those recurring below are thrown on the right side of our equation, we obtain (perhaps after a change of sign):

$$(3.7) \quad [\alpha^{(1)} \cdot x^{(i-1)} y z \dots] (x^{(i)} \pi) - \dots = [\alpha^{(1)} \cdot x^{(i)} u \dots] (y \pi) \\ - [\alpha^{(1)} \cdot x^{(i)} y u \dots] (z \pi) + \dots$$

On the right-hand side stand the terms whose bracket factor includes all  $i$  symbolic vectors  $\alpha^{(1)}, \dots, \alpha^{(i)}$ . The left side is made up of the alternating sum of all terms arising from the symbol

$$[\alpha^{(1)} \cdot \dots \alpha^{(i-1)} y z \dots]$$

by throwing out one of the  $x$ 's and multiplying the remaining bracket factor by the corresponding  $(x \pi)$ . If  $S$  again denotes alternation:  $1/i!$  times the alternating sum over all permutations of the "symbolic" vectors  $\alpha^{(1)}, \dots, \alpha^{(i)}$ , it may be written as

$$i \cdot S \left\{ [\alpha^{(1)} \cdot \dots \alpha^{(i-1)} y z \dots] (x^{(i)} \pi) \right\}.$$

One now sees how our identity will serve the purpose of restitution: a product like

$$[\alpha^{(1)} \cdot \dots \alpha^{(i-1)} y z \dots] (x^{(i)} \pi),$$

in which the symbolic argument  $x^{(i)}$  stands outside the bracket  $[ ]$  is transformed, after alternation, into a sum of terms whose bracket factors have absorbed the stray argument  $x^{(i)}$ . In this way all symbolic arguments may be drawn into the bracket factor  $[ ]$ , one after the other,-- providing, of course, that such a factor is present. In the applications  $(\alpha^{(i)} \pi)$  need not be the product of  $x^{(i)}$  with a Greek vector  $\pi$ ; rather it may be anything linearly de-

pending on  $x^{(i)}$ , e.g. another bracket factor containing  $x^{(i)}$  as one of its  $n$  arguments.

The terms of the alternating sum of the left side of (3.7) are derived from the "leading term"

$$[x^{(1)} \cdot x^{(i-1)} y z \dots] (x^{(i)}) ,$$

if one replaces the sequence  $x^{(1)}, \dots, x^{(i-1)}, x^{(i)}$  by that arising from it on making  $x^{(i)}$  hop over the preceding arguments  $x^{(1)}, \dots, x^{(i-1)}$ , one after the other (from right to left). We shall say that we draw  $x^{(i)}$  (alternatingly) through the sequence  $x^{(1)}, \dots, x^{(i-1)}$ . In the same way, the argument  $y$  in the leading term on the right side:

$$[x^{(1)} \cdot \dots x^{(i)} z u \dots] (y \pi)$$

is drawn through the sequence of arguments  $z, u, \dots$  (from left to right).

It is convenient for later purposes to establish explicitly the result of a repeated absorption of symbolic arguments into one and the same bracket factor. We are to deal with an expression:

$$(3.2) \quad [x^{(1)} \cdot x^{(i)} y^{(i+1)} y^{(n)}] \cdot F(x^{(i+1)}, \dots, x^{(k)}) ,$$

depending on  $k$  "symbolic" vectors  $x$ :

$$x^{(1)}, \dots, x^{(i)}, x^{(i+1)}, \dots, x^{(k)} ,$$

and  $n-i$  further vectors  $y$ :

$$y^{(i+1)}, \dots, y^{(n)} ,$$

( $0 \leq i \leq k \leq n$ ).  $F$  is a multilinear form and skew-symmetric in its  $k-i$  arguments. From (3.2) we obtain the sum

$$\sum_x \pm [x^{(1)} \cdot \dots x^{(i)} y^{(i+1)} \dots y^{(n)}] \cdot F(x^{(i+1)}, \dots, x^{(k)})$$

by first drawing  $x^{(i+1)}$  through the sequence of arguments  $x^{(1)}, \dots, x^{(i)}$ , then drawing  $x^{(i+2)}$  through the sequence  $x^{(1)}, \dots, x^{(i)}, x^{(i+1)}$ , and so on.

Each step is to be performed on the whole expression that resulted from the preceding step. The whole sum equals the analogous sum

$$\sum_y \pm [x^{(1)} \cdot x^{(k)} y^{(k+1)} \cdot y^{(m)}] \cdot F(y^{(i+1)}, \dots, y^{(k)}),$$

in which each argument  $y^{(i+1)}, \dots, y^{(k)}$  of the leading term is drawn through the subsequent arguments in the whole sequence  $y^{(i+1)}, \dots, y^{(k)}, y^{(k+1)}, \dots, y^{(m)}$ :

$$(3.3) \quad \sum_x \pm [x^{(1)} \cdot x^{(i)} y^{(i+1)} \cdot y^{(m)}] \cdot F(x^{(i+1)}, \dots, x^{(k)}) = \sum_y \pm [x^{(1)} \cdot x^{(k)} y^{(k+1)} \cdot y^{(m)}] \cdot F(y^{(i+1)}, \dots, y^{(k)})$$

The sum  $\sum_x$  on the left side contains the arguments

$$x^{(1)}, \dots, x^{(i)} \mid x^{(i+1)}, \dots, x^{(k)},$$

in all arrangements preserving natural order with the first part  $x^{(1)}, \dots, x^{(i)}$ ; but the  $x$ 's of the second part can be permuted and are allowed to be scattered between those of the first part in any possible way. It is formally simpler and permissible to sum over all permutations  $x^{(1)}, \dots, x^{(k)}$ . The effect of this wholesale summation will be that each term of  $\sum_x$  appears  $i!$  times,  $i!$  being the numbers of ways in which  $x^{(1)}, \dots, x^{(i)}$  may be arranged. Hence

$$\sum_x = \frac{k!}{i!} S_x$$

where  $S_x$  means alternation with respect to all  $x$ 's. The right side of (3.3) can be written in the same manner as

$$\frac{(m-i)!}{(m-k)!} S_y,$$

and thus our identity takes on the equivalent form

$$(3.4) \quad \binom{m}{i} S_x = \binom{m}{k} S_y$$

where  $\binom{n}{i}$  is the binomial factor  $\frac{n!}{i!(n-i)!}$ .  $k$  is the number of vectors  $x$  and  $n-i$  the number of the  $y$ 's.

#### 4. Unimodular group

Proposition: The three types

$$[x y \dots z], (\alpha \xi), [\xi \eta \dots \xi]$$

form a complete base for the unimodular group.

Proof. 1) The table contains the typical basic invariant

$$[x y \dots z]$$

which suffices in the case of Latin vectors only.

2) Consider a product of some factors of types 1 and 2 involving the symbolic arguments  $x^{(1)}, \dots, x^{(n-1)}$  in a linear way. If the product contains at least one factor of type 1, all these symbolic arguments can be gathered into this factor by repeated application of last section's formal identity. Restitution then changes the factor  $[x^{(1)} \dots x^{(n-1)} x]$  of type 1 into the factor  $(\xi x)$  of type 2. However, if the product contains only factors of type 2, it is necessarily of the form

$$(x^{(1)} \xi^{(1)}) \dots (x^{(n-1)} \xi^{(n-1)}),$$

and is then carried by restitution into a factor

$$[\xi^{(1)} \dots \xi^{(n-1)} \xi]$$

of type 3.

Roughly speaking, restitution effects a shift from type 1 to type 2, and from type 2 to type 3; here it stops short since type 3 involves no more Latin arguments.

#### 5. Group of step transformations

For definiteness we again consider three steps of width  $p, q-p, m-q$  (cf. 8, example 3). A step transformation transforms the following components of an arbitrary vector  $x$  only among each other:

$$(x_1, \dots, x_p),$$

$$(x_1, \dots, x_q), \quad (0 < p < q < m),$$

$$(x_1, \dots, x_m).$$

The contragredient transformation affects in the same manner the following sections of the components  $\xi$  of a contravariant vector:

$$\begin{aligned} & (\xi_{q+1}, \dots, \xi_m), \\ & (\xi_{p+1}, \dots, \xi_m), \\ & (\xi_1, \dots, \xi_m). \end{aligned}$$

They are of length  $p' = m - q$ ,  $q' = m - p$ ,  $m$ , respectively. Hence we shall expect the following typical basic invariants

$$(5.1) \quad \begin{cases} [\alpha^{(1)} \dots \alpha^{(p)}]_p, [\alpha^{(1)} \dots \alpha^{(q)}]_q, [\alpha^{(1)} \dots \alpha^{(m)}]_m; \\ (\alpha \xi); \\ [\xi^{(1)} \dots \xi^{(p')}]_{p'}, [\xi^{(1)} \dots \xi^{(q')}]_{q'}, [\xi^{(1)} \dots \xi^{(m)}]_m. \end{cases}$$

The three quantities of the first row were defined before: they are the determinants of the first  $p$  or  $q$  or  $n$  components of  $p$  or  $q$  or  $n$  vectors  $x$  respectively. Similarly we have in the third row the determinants of the last  $p'$  or  $q'$  or  $m$  components of  $p'$  or  $q'$  or  $n$  contravariant vectors  $\xi$ . As we can think of no other fundamental invariant offhand, we are bold enough to guess that our table is complete. This is what we are now going to prove.

It certainly is complete when none but covariant vectors are taken into account: the types in the first row suffice for this purpose. Secondly, we must show that a product  $P$  of factors of the types in the first and second row linearly containing the symbolic vectors  $\alpha^{(1)}, \dots, \alpha^{(m-1)}$  is carried into a combination of factors of all three types by restitution. (1) If  $P$  contains no other factors than those of type  $(\alpha \xi)$  it is of form ( ) and hence changes into a factor of type  $[\xi^{(1)} \dots \xi^{(m)}]$ . Let us therefore assume that  $P$  involves at least one of the types in row 1. (2) If it contains a factor of type  $[\alpha \gamma \dots \zeta]_m$  we know that it is possible to throw all the symbolic arguments into this bracket factor:  $[\alpha^{(1)} \dots \alpha^{(m-1)} \alpha]$ , which then is changed by restitution into  $(\alpha \xi)$ . (3) If a bracket factor  $[ ]_m$  of or-

der  $n$  is not present, we might have one  $[ \ ]_q$  of order  $q$ . The product  $P$  splits into two partial products, the first  $P'$  containing all factors of types row 1, the second being a product of factors of type  $(\alpha \xi)$ . Only the first  $q$  components of all the factors engaged occur in  $P'$ ; we thus may here operate in the  $q$ -dimensional space of the first  $q$  components; and by repeatedly applying our formal identity to determinants in this space we are able to gather all the symbolic arguments  $x^{(1)}, \dots, x^{(i)}$  present in  $P'$  into the one factor  $[ \dots ]_q$  whose existence we assumed. After discarding factors free from the symbolic arguments  $x$  we then are faced by an expression of this kind

$$(5.2) \quad [ x^{(1)} \dots x^{(i)} y^{(i+1)} \dots y^{(q)} ]_q \cdot (x^{(i+1)} \xi^{(i+1)}) \dots (x^{(n-1)} \xi^{(n-1)}).$$

We introduce the unit vector  $e^{(i)}$  whose  $i^{\text{th}}$  component = 1 and all other components vanish. The unit vectors enable us to express a bracket factor of order  $q$ :  $[ z^{(1)} \dots z^{(q)} ]_q$  as an  $n$ -dimensional determinant:

$$[ z^{(1)} \dots z^{(q)} e^{(q+1)} \dots e^{(n)} ]$$

Let us do this with the first factor in (5.2) and then make use of formula

(3.3). We learn from it that after alternation with respect to the symbolic arguments (5.2) goes over into

$$\sum \pm [ x^{(1)} \dots x^{(n-1)} z^{(n)} ] \cdot (z^{(i+1)} \xi^{(i+1)}) \dots (z^{(n-1)} \xi^{(n-1)})$$

(a numerical factor has been omitted) where the alternating sum runs over all permutations  $z^{(i+1)}, \dots, z^{(n)}$  of

$$y^{(i+1)}, \dots, y^{(q)}, e^{(q+1)}, \dots, e^{(n)}.$$

Restitution replaces the first factor by  $(z^{(n)} \xi^{(n)})$  if the restituent is called  $\xi^{(n)}$ . The result is the determinant

$$\left| \left( \xi^{(i+1)} z \right), \dots, \left( \xi^{(n)} z \right) \right|$$

whose different lines are derived from the one written out by substituting

$y^{(i+1)}, \dots, y^{(q)}, e^{(q+1)}, \dots, e^{(n)}$  for  $z$ , one after the other. This determinant is

$$\begin{vmatrix} (\xi^{(i+1)} y^{(i+1)}), & \dots & (\xi^{(m)} y^{(i+1)}) \\ (\xi^{(i+1)} y^{(q)}), & \dots & (\xi^{(m)} y^{(q)}) \\ \xi_{q+1}^{(i+1)} & \dots & \xi_{q+1}^{(m)} \\ \vdots & \vdots & \vdots \\ \xi_n^{(i+1)} & \dots & \xi_n^{(m)} \end{vmatrix}$$

When expanded in terms of the last  $n-q$  rows, only determinants of the type

$$[\xi', \xi'', \dots]_{p'}$$

depending on the last  $p' = n - q$  components occur besides the factors of type  $(\xi y)$ . The leading term for instance reads

$$\begin{vmatrix} (\xi^{(i+1)} y^{(i+1)}), & \dots & (\xi^{(q)} y^{(i+1)}) \\ \vdots & \vdots & \vdots \\ (\xi^{(i+1)} y^{(q)}), & \dots & (\xi^{(q)} y^{(q)}) \end{vmatrix} \cdot \begin{vmatrix} \xi_{q+1}^{(q+1)} & \dots & \xi_{q+1}^{(m)} \\ \vdots & \vdots & \vdots \\ \xi_n^{(q+1)} & \dots & \xi_n^{(m)} \end{vmatrix}$$

It is at this point where one needs the "iterated" formula (3.3) explicitly.

(4) If no bracket factor either of order  $n$  or  $q$  is present, but one of order  $p$ , the argumentation is essentially the same. Restitution then gives rise to a Greek bracket factor of order  $m - p = q'$ . This completes the proof of our conjecture:

Theorem. Table (5.1) provides a full set of typical basic invariants for the group of step transformations with three steps of widths  $p, q-p, n-q$ .

It is obvious how to generalize the theorem to an arbitrary number  $\sigma$  of steps; the number of basic invariants will be  $2\sigma + 1$ . In its general form it answers among others the question of invariants depending on points and planes in affine space, and the question of semi-invariants for any number of co- and contra-variant vectors.

### 6. Euclidean space: enumeration of basic invariants

The group  $\chi$  of linear transformations in  $n$ -dimensional vector space with which we now propose to deal, will be the group of proper orthogonal transformations in  $\mu$  dimensions bordered with a rim of width  $\nu$ ;  $\mu + \nu = n$ . The components of a covariant vector  $x$  shall be denoted by

$$x_1, \dots, x_\mu; x_{\mu+1}, \dots, x_n.$$

The two groups of components may be referred to as the head and tail of  $x$ .

We must deviate in this manner from the notations used in § 8 because Greek letters are now reserved for contravariant vectors. Interpreting

$x_1 : x_2 : \dots : x_n$  as homogeneous plane coordinates in an  $(n-1)$ -dimensional projective space, our group is that of Euclidean space of rank  $\nu$  (in particular of the Euclidean space proper if  $\nu = 1$ ). The components

$\xi_1, \dots, \xi_\mu; \xi_{\mu+1}, \dots, \xi_n$  are the homogeneous point coordinates in this interpretation, and the infinite locus is described by the simultaneous equations

$$\xi_{\mu+1} = 0, \dots, \xi_n = 0.$$

A determinant like

$$\begin{vmatrix} x_{i_1} & \dots & x_{i_p} \\ \dots & \dots & \dots \\ y_{i_1} & \dots & y_{i_p} \end{vmatrix}$$

involving  $p$  vectors  $x, \dots, y$ , may be denoted by

$$[x \dots y]_{i_1, \dots, i_p}.$$

It is zero unless all indices  $i_1, \dots, i_p$  are different.

We start with the table of basic invariants depending on Latin vectors only:

a) The full determinant of  $n$  vectors  $x, y, \dots, z,$

$$D: [x \ y \ \dots \ z] ;$$

b) the determinant of the  $\mu$  head components of  $\mu$  vectors  $x, \dots, y,$

$$D^*: [x \ \dots \ y]_{1,2, \dots, \mu} ;$$

c) the scalar product

$$\sum_{i=1}^{\mu} x_i y_i = (x | y)$$

which also depends on the head components of  $x$  and  $y$  only, and for reasons that will soon become clear, shall be denoted by  $(x | y)$ .

We now try to make up a similar table for Greek vectors. It will of course, by analogy with a) and b), contain:

d) the full determinant of  $n$  contravariant vectors  $\xi, \eta, \dots, \zeta,$

$$\Delta: [\xi \ \eta \ \dots \ \zeta] ;$$

e) the determinant of the  $\nu$  tail components of  $\nu$  contravariant vectors  $\xi, \dots, \eta,$

$$\Delta^*: [\xi \ \dots \ \eta]_{\mu+1, \dots, n}$$

To obtain the analogue of c) we first consider Euclidean space proper:  $\nu = 1, \mu = n - 1$ ; then  $\Delta^*$  is the component  $\xi_m$ . The simplest conceivable rational point invariant in Euclidean space is the square of the distance of two points  $\xi, \eta$ :

$$\sum_i \left( \frac{\eta_i}{\eta_m} - \frac{\xi_i}{\xi_m} \right)^2.$$

The sum over  $i$  ranges from 1 to  $n-1$  or from 1 to  $n$  as one prefers. On putting the common denominator  $\xi_m^2 \eta_m^2$  in evidence, the numerator turns out as

$$\sum_i (\eta_i \xi_m - \xi_i \eta_m)^2 = \sum_i [\xi \eta]_{i,m}^2 = (\xi \eta || \xi \eta);$$

the denominator, and hence the numerator as well, are invariants. The polarized expression

$$\delta_1: (\xi \eta || \xi' \eta') = \sum_i [\xi \eta]_{i,m} [\xi' \eta']_{i,m}$$

is the numerator of the scalar product of two line elements  $\vec{\xi}_m \eta_m$ ,  $\vec{\xi}'_m \eta'_m$  in our  $(n-1)$ -dimensional Euclidean space, its denominator being  $\xi_m \eta_m \xi'_m \eta'_m$ .

The square of the size of a two-dimensional element (parallelogram) as spanned by two line elements  $\vec{\xi}_m \eta_m$ ,  $\vec{\xi}'_m \eta'_m$  is likewise given by a fraction whose denominator equals  $\xi_m^2 \eta_m^2 \xi'^2_m \eta'^2_m$ , and whose numerator is the sum

$$\sum [\xi \eta \xi']_{i,k,m}^2$$

extending either over all pairs  $i, k$  of head indices  $i, k = 1, 2, \dots, m-1$  with  $i < k$ , or  $i$  and  $k$  run independently from 1 to  $n$ ; in the latter case the factor  $1/2!$  is to be added. The duly polarized invariant is

$$\delta_2: (\xi \eta \xi' \parallel \xi' \eta' \xi') = \sum_{(i < k)} [\xi \eta \xi']_{i,k,m} \cdot [\xi' \eta' \xi']_{i,k,m}$$

The square of the area of a parallelogram can be expressed by means of the scalar products of the two spanning line elements  $\vec{\xi}_m \eta_m$ ,  $\vec{\xi}'_m \eta'_m$  as:

$$\begin{vmatrix} \frac{(\xi \eta \parallel \xi \eta)}{\xi_m^2 \eta_m^2} & \frac{(\xi \eta \parallel \xi \xi')}{\xi_m^2 \eta_m \xi'_m} \\ \frac{(\xi \xi' \parallel \xi \eta)}{\xi_m^2 \xi'_m \eta_m} & \frac{(\xi \xi' \parallel \xi \xi')}{\xi_m^2 \xi'^2_m} \end{vmatrix}$$

or

$$\frac{1}{\xi_m^4 \eta_m \xi'^2_m} \cdot \begin{vmatrix} (\xi \eta \parallel \xi \eta) & (\xi \eta \parallel \xi \xi') \\ (\xi \xi' \parallel \xi \eta) & (\xi \xi' \parallel \xi \xi') \end{vmatrix}$$

Hence the relation

$$\begin{vmatrix} (\xi \eta \parallel \xi \eta) & (\xi \eta \parallel \xi \xi') \\ (\xi \xi' \parallel \xi \eta) & (\xi \xi' \parallel \xi \xi') \end{vmatrix} = \xi_m^2 \cdot (\xi \eta \xi' \parallel \xi \eta \xi')$$

or in polarized form

$$\begin{vmatrix} (\xi \eta \parallel \xi' \eta') & (\xi \eta \parallel \xi' \xi') \\ (\xi \xi' \parallel \xi' \eta') & (\xi \xi' \parallel \xi' \xi') \end{vmatrix} = \xi_m \xi'_m \cdot (\xi \eta \xi' \parallel \xi' \eta' \xi')$$

It is therefore not true, as one might have expected, that the type  $\delta_2$  is integral-rationally expressible by invariants of type  $\delta_1$ . What we get is merely a relation between invariants of types  $\delta_1, \delta_2$  and  $\Delta^*$ :  $\delta_2$  should hence not be omitted from a full table of basic invariants.

From the 1- and 2-dimensional elements one may proceed to those of higher dimensionalities up to  $n-1$ . Generalization to Euclidean space of rank  $\nu$  is obvious. We thus get as an analogue to (c), on the Latin side, a whole string of invariants  $\delta_0, \delta_1, \dots, \delta_\mu$  on the Greek side:

$$f) \quad \delta_\rho: (\xi \eta \dots \xi \parallel \xi' \eta' \dots \xi')_\rho =$$

$$\sum_{\substack{[ \xi \eta \dots \xi ]_{i_1, \dots, i_\rho, \mu+1, \dots, \mu} \\ \{ \rho = 0, 1, \dots, \mu \}}} [ \xi' \eta' \dots \xi' ]_{i_1, \dots, i_\rho, \mu+1, \dots, \mu}$$

$\delta_\rho$  depends on two sets of  $\rho + \nu$  Greek vectors  $\xi, \eta, \dots, \xi; \xi', \eta', \dots, \xi'$ .

The range for the indices of summation is described by

$$i_1 < i_2 < \dots < i_\rho; \quad i_1, i_2, \dots, i_\rho = 1, 2, \dots, \mu;$$

or if one prefers, one may run all  $\rho$  into  $i$  independently from 1 to  $n$ , but

then one has to divide the whole sum by  $\rho!$ . The two extreme types  $\delta_0$  and

$\delta_\mu$  are products of two invariants of types  $\Delta^*$  and  $\Delta$  respectively:

$$(\xi \dots \eta \parallel \xi' \dots \eta')_0 = \Delta^*(\xi \dots \eta) \cdot \Delta^*(\xi' \dots \eta');$$

$$(\xi \dots \xi \parallel \xi' \dots \xi')_\mu = \Delta(\xi \dots \xi) \cdot \Delta(\xi' \dots \xi').$$

It is easy to generalize the relations; but we have no immediate need for them and so leave their establishment to the reader! Another expression, however, for the invariant  $\delta_\rho$  will prove useful.

The series of the  $\rho + \nu$  arguments

$$\xi, \dots, \eta; \xi', \dots, \xi'$$

consists of two sections of  $\rho$  and  $\nu$  vectors respectively; the division is

marked by the semicolon.  $i_1, \dots, i_\rho$  may run independently from 1 to  $n$ .

An expression for  $\rho! \delta_\rho$  can be derived from the term

$$\sum_{i_1, \dots, i_\rho} \xi_{i_1} \dots \eta_{i_\rho} \vartheta_{\mu+1} \dots \xi_m \cdot \xi'_{i_1} \dots \eta'_{i_\rho} \vartheta'_{\mu+1} \dots \xi'_m$$

$$= (\xi\xi') \dots (\eta\eta') \cdot \vartheta'_{\mu+1} \dots \xi'_m \cdot \vartheta_{\mu+1} \dots \xi_m$$

by first summing alternately over all permutations of  $\xi'_1, \dots, \xi'_m$  and then performing the same operation with respect to  $\xi_1, \dots, \xi_m$ . The first step leads to

$$\left| \begin{array}{cccc} (\xi\xi'), \dots, (\eta\xi'), \xi'_{\mu+1}, \dots, \xi'_m \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (\xi\xi'), \dots, (\eta\xi'), \xi'_{\mu+1}, \dots, \xi'_m \end{array} \right| \cdot \vartheta_{\mu+1} \dots \xi_m$$

and we thus find

$$(6.1) \quad \delta_\rho = \sum_{\pm} \left| \begin{array}{cccc} (\xi\xi'), \dots, (\eta\xi'), \xi'_{\mu+1}, \dots, \xi'_m \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (\xi\xi'), \dots, (\eta\xi'), \xi'_{\mu+1}, \dots, \xi'_m \end{array} \right| \cdot [\vartheta \dots \xi]_{\mu+1, \dots, m}$$

where the sum extends alternately to all possible ways in which the  $\nu$  arguments  $\vartheta, \dots, \xi$  may be mixed among the  $\rho$  first ones  $\xi, \dots, \eta$  (without changing the order in either of the two parts). This may finally be condensed into the neat form of a single  $(2\nu + \rho)$ -rowed determinant:

$$(6.2) \quad \delta_\rho = (-1)^\nu \left| \begin{array}{ccc|ccc} (\xi\xi'), \dots, (\xi\xi') & \xi'_{\mu+1}, \dots, \xi'_m \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (\xi\xi'), \dots, (\xi\xi') & \xi'_{\mu+1}, \dots, \xi'_m \\ \hline \xi_{\mu+1}, \dots, \xi_{\mu+1} & 0, \dots, 0 \\ \xi_m, \dots, \xi_m & 0, \dots, 0 \end{array} \right|$$

It is immaterial whether we here interpret the products like  $(\xi \xi')$  as meaning

$$(\xi \xi') = \sum_{i=1}^m \xi_i \xi_i' \quad \text{or as} \quad (\xi | \xi') = \sum_{i=1}^m \xi_i \xi_i'$$

For by subtracting from the first column of the determinant the last  $\nu$  ones multiplied by  $\xi_{\mu+1}, \dots, \xi_m$ , respectively, the terms

$$(\xi \xi'), \dots, (\xi \xi')$$

of the first column are changed into

$$(\xi | \xi'), \dots, (\xi | \xi')$$

After the Latin and the Greek invariants, we now turn to the mixed ones involving both covariant and contravariant vectors. First the product

$$(\alpha \xi)$$

When one writes the scalar product

$$(\alpha | \gamma) = \alpha_1 \gamma_1 + \dots + \alpha_\mu \gamma_\mu$$

as

$$(\xi \gamma) = \xi_1 \gamma_1 + \dots + \xi_\mu \gamma_\mu + \xi_{\mu+1} \gamma_{\mu+1} + \dots + \xi_m \gamma_m$$

one must put

$$(\xi_1, \dots, \xi_\mu; \xi_{\mu+1}, \dots, \xi_m) = (\alpha_1, \dots, \alpha_\mu; 0, \dots, 0).$$

Hence every covariant vector  $\alpha = (\alpha_1, \dots, \alpha_m)$  gives rise to a contravariant one

$$\xi = \alpha | = (\alpha_1, \dots, \alpha_\mu, 0, \dots, 0)$$

by cutting off its tail; the process | of striking off the last  $\nu$  components is invariant under all transformations of  $\gamma$ . We use both notations  $\alpha |$  or  $| \alpha$  without difference, so that  $(\alpha | \gamma)$  may be interpreted either as the product of the contravariant vector  $\alpha |$  with the covariant  $\gamma$  or as the product of the covariant vector  $\alpha$  with the contravariant  $| \gamma$ . One sees how this usage of the "off stroke" | justifies the symbol  $(\alpha | \gamma)$  adopted for the scalar product.

On replacing some of the  $n$  Greek arguments in  $\Delta = [\xi \eta \dots \zeta]$  by arguments of the type  $\alpha|$  we obtain new mixed invariants; and hence we supplement our table by the following new members of mixed character:

g)  $(\alpha \xi)$

h)  $\Delta_a: [\alpha|, \dots, y|, \xi, \dots, \eta]$ ,  $(a = 0, 1, \dots, \mu)$ .

$\xleftarrow{a}$   $\xrightarrow{m-a}$

$a$  denotes the number of Latin arguments  $x, \dots, y$  in this expression, so that  $m - a$  is the number of the Greek ones.

The limiting cases  $a = 0$  and  $a = \mu$  are not new:

$$\Delta_0 = \Delta; \quad \Delta_\mu = [\alpha, \dots, y]_{1, \dots, \mu} [\xi, \dots, \eta]_{\mu+1, \dots, m} = D^* \Delta^*$$

I contend the completeness of our table at this stage -- in spite of its having been collected in a rather haphazard way. It consists of the three Latin invariants

(a, b, c)  $D, D^*, (\alpha|y),$

of the  $\mu + 1$  Greek invariants

(d, e, f)  $\Delta, \Delta^*, \delta_1, \dots, \delta_{\mu-1},$

and the  $\mu$  mixed invariants

(g, h)  $(\alpha \xi), \Delta_1, \dots, \Delta_{\mu-1};$

the total number of different types thus amounting to  $2\mu + 4$ . The statement might at first sound far from convincing at the process of replacing Greek arguments by Latin ones carrying the "off stroke" can be applied to all the Greek invariants, i.e. to  $\Delta^*, \delta_1, \dots, \delta_{\mu-1}$  as well as to  $\Delta$ .

$\Delta^*$  can here be disposed of at once; for it involves none but the tail components and thus is turned into zero if one of its arguments is replaced by a  $\alpha|$  whose tail components all vanish. But what about  $\hat{c}_\rho$ ? We apply formula (6.1) or (6.2) in substituting an  $\alpha|$  for  $\xi$ . In the first column will

then appear the expressions

$$(x | \xi'), \dots, (x | \zeta')$$

On adding to it, however, the last  $\nu$  columns multiplied by  $x_{\mu+1}, \dots, x_n$ , respectively, they will be changed into the more reasonable products

$$(x \xi'), \dots, (x \zeta').$$

After expanding the determinant (6.2), or the determinant in (6.1) by the first column, one gets

$$(x | \eta, \dots, \xi || \xi', \eta', \dots, \zeta')_p = \sum^{\pm} (x \xi') (\eta, \dots, \xi || \eta', \dots, \zeta')_{p-1},$$

where in the leading term on the right side  $\xi'$  must be drawn through the sequence  $\eta', \dots, \zeta'$ . A similar formula holds when instead of  $\xi$  the first dashed argument  $\xi'$  is an  $x'$ . If some arguments among the  $\xi', \eta', \dots, \zeta'$  are of this type  $x'$

$$x', \dots, y', \vartheta', \dots, \zeta',$$

the factor  $(x \xi')$  in formula which ran over the values

$$(x \xi'), (x \eta'), \dots, (x \zeta'),$$

takes on the consecutive forms

$$(x | x'), \dots, (x | y'), (x | \vartheta'), \dots, (x | \zeta').$$

By repeated application of this formula we are thus able to foresee what will happen to a  $\delta_p$  of rank  $p$  when some, let us say  $\underline{a}$  of the  $\xi, \eta, \dots, \zeta$  and  $\underline{b}$  of the  $\xi', \eta', \dots, \zeta'$  are replaced by arguments of type  $x'$ :

$$(x^{(1)} |, \dots, x^{(\alpha)} |, \xi^{(1)}, \dots, \xi^{(\alpha)} || y^{(1)} |, \dots, y^{(b)} |, \eta^{(1)}, \dots, \eta^{(\beta)}),$$

$$(a + \alpha = b + \beta = p + \nu).$$

This invariant vanishes of necessity unless  $a$  and  $b$  are  $\leq p$ . These conditions being fulfilled, it is expressible in terms of the scalar products

$$(x^{(i)} | y^{(k)}), \quad (i = 1, \dots, a; k = 1, \dots, b),$$

the products

$$(x^{(i)} \eta^{(\kappa)}), \quad (\xi^{(i)} y^{(k)}), \quad (i = 1, \dots, \alpha; \kappa = 1, \dots, \beta),$$

and certain  $\delta_\sigma$  of rank  $\sigma$  lower than  $\rho$ , involving certain of the Greek arguments  $\xi^{(\nu)}$  and  $\eta^{(\kappa)}$ . One may derive this result without iteration as well from the following explicit expression of (6.2):

$$\begin{array}{c}
 \left( \begin{array}{ccc|ccc}
 (x^{(1)} | y^{(1)}), & \dots, & (x^{(1)} | y^{(b)}) & (x^{(1)} | \eta^{(1)}), & \dots, & (x^{(1)} | \eta^{(\beta)}) & 0, & \dots, & 0 \\
 (x^{(a)} | y^{(1)}), & \dots, & (x^{(a)} | y^{(b)}) & (x^{(a)} | \eta^{(1)}), & \dots, & (x^{(a)} | \eta^{(\beta)}) & 0, & \dots, & 0 \\
 \hline
 (\xi^{(1)} | y^{(1)}), & \dots, & (\xi^{(1)} | y^{(b)}) & (\xi^{(1)} | \eta^{(1)}), & \dots, & (\xi^{(1)} | \eta^{(\beta)}) & \xi_{\mu+1}^{(1)}, & \dots, & \xi_m^{(1)} \\
 (\xi^{(a)} | y^{(1)}), & \dots, & (\xi^{(a)} | y^{(b)}) & (\xi^{(a)} | \eta^{(1)}), & \dots, & (\xi^{(a)} | \eta^{(\beta)}) & \xi_{\mu+1}^{(a)}, & \dots, & \xi_m^{(a)} \\
 \hline
 0, & \dots, & 0, & \eta_{\mu+1}^{(1)}, & \dots, & \eta_{\mu+1}^{(\beta)} & 0, & \dots, & 0 \\
 \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
 0, & \dots, & 0, & \eta_m^{(1)}, & \dots, & \eta_m^{(\beta)} & 0, & \dots, & 0
 \end{array} \right)
 \end{array}$$

arising from (6.2) by using the same trick as before for turning the unreasonable  $(\alpha | \xi)$  into the reasonable  $(\alpha \xi)$ .

We thus realize that the  $\delta_\rho$  do not lead to new basic invariants by conversions of the kind  $\xi \rightarrow \alpha |$ .

The product of two  $\Delta' s$  :

$$[\xi \dots \xi] \cdot [\xi' \dots \xi']$$

is a  $\delta_\mu$ . On replacing some of the Greek vectors by an  $\alpha |$  it turns into a product of two  $\Delta'_a s$  (with equal or different  $a$ ). Consequently such a product

$$[x^{(1)} |, \dots, x^{(a)} |, \xi^{(a+1)}, \dots, \xi^{(a)}] \cdot [y^{(1)} |, \dots, y^{(b)} |, \eta^{(b+1)}, \dots, \eta^{(m)}]$$

is expressible in terms of invariants of types  $(\alpha | y)$ ,  $(\alpha \xi)$  and  $\delta_\rho$ .

7. Proof of completeness. Additional remarks

We now master the relations holding among our basic invariants well enough to be able to enter upon the proof of our table's completeness. Restitution not only will lead to this result, but it will provide at the same time a systematic method of constructing the whole list that was more or less casually gathered from the four winds in the preceding section. We start with the typical Latin invariants which are known to constitute a complete table:

$$a) D = [x \dots z]_m, \quad b) D^* = [x \dots y]_\mu, \quad c) (x|y),$$

and at once add the most trivial mixed type:

$$g) (x\xi).$$

(1) A product of  $n-1$  factors of types c) and g) each of which involves one of the symbolic vectors  $x^{(1)}, \dots, x^{(n-1)}$ :

$$(x^{(1)}|y^{(1)}) \cdot \dots \cdot (x^{(a)}|y^{(a)}) (x^{(a+1)}|\eta^{(a+1)}) \cdot \dots \cdot (x^{(n-1)}|\eta^{(n-1)})$$

leads by restitution with the restituent  $\xi^{(n)}$  to

$$h) \Delta_a = [y^{(1)}|, \dots, y^{(a)}|, \eta^{(a+1)}, \dots, \eta^{(n)}].$$

This becomes evident at once if one interprets  $(x|y)$  as the product of  $x$  and

$|y$ .  $\Delta_a$  includes the type

$$d) \Delta = [\xi \dots \xi]_m,$$

for  $a = 0$ . One sees how the  $\Delta_a$  of necessity make their appearance in the course of the restituting process.

(2) We know from investigation of the affine space that the following invariant

$$[x^{(1)} \dots x^{(\mu)}]_\mu \cdot (x^{(\mu+1)}|\xi') \cdot \dots \cdot (x^{(n-1)}|\xi^{(v-1)})$$

gives rise to the type

$$e) \Delta^* = [\xi' \dots \xi^{(v)}]_v$$

by restitution with the restituent  $\xi^{(v)}$ .

(3) The product of a factor  $\Delta_a$  with  $m-a-1$  factors of type  $(\alpha \xi)$  in which all the  $m-1$  Latin arguments are symbolic is carried into  $\delta_{\mu-a}$ , type f):

$$[\alpha^{(1)} | \dots \alpha^{(a)} | \eta^{(1)}, \dots, \eta^{(m-a)}] \cdot (\alpha^{(a+1)} \xi^{(1)}) \dots (\alpha^{(m-1)} \xi^{(m-1-a)}) \rightarrow$$

$$(\xi^{(1)} \dots \xi^{(m-a)} | \eta^{(1)} \dots \eta^{(m-a)})_{\mu-a}$$

( $\xi^{(m-a)}$  being the restituent). The proof, though it needs a bit of calculation, is absolutely straightforward.

To prove completeness one must show that restitution in an aggregate of our basic invariants leads to no other invariants. Let us examine all possibilities! We are to consider a product P of factors taken from our table and each containing at least one of the symbolic vectors  $\alpha^{(i)}$ . The factors therefore are of the purely Latin or the mixed type: a), b), c), g), h). We can disregard the possibility of factors like  $(\alpha^{(1)} | \alpha^{(2)})$  where two symbolic vectors  $\alpha^{(1)}$ ,  $\alpha^{(2)}$  are united to form a scalar product. A product P containing such a factor would be annihilated by restitution or rather by its first step, the transposition of  $\alpha^{(1)}$  and  $\alpha^{(2)}$ , since  $(\alpha^{(1)} | \alpha^{(2)})$  is symmetric in its two members  $\alpha^{(1)}$ ,  $\alpha^{(2)}$ .

(1) If P contains a factor D, all symbolic arguments may be drawn into this factor which then by restitution is changed into an  $(\alpha \xi)$ . We assume henceforth that no factor D occurs in P.

(2) If P contains a factor  $D^* = [\alpha \dots y]_{\mu}$  depending on the head components of  $x, \dots, y$  only, we break up P into two partial products, the first P' containing all factors of types b), c), h) which involve but the head components of Latin arguments. Hence the second part P'' is made up of factors of type  $(\alpha \xi)$  alone. The product P' can be handled by considering the  $\mu$ -dimensional space of the head components, and all the symbolic arguments  $\alpha^{(1)}, \dots, \alpha^{(i)}$  occurring in P' may thus be drawn into the first factor  $D^*$  of P'. (Their num-

ber must not exceed  $\mu$  or else P will be annihilated by the alternating process.) On canceling factors that no longer involve symbolic arguments, we are left with an expression

$$\left[ x^{(1)} \cdot x^{(2)} y^{(2+1)} \cdot y^{(\mu)} \right]_{\mu} \cdot (x^{(2+1)} \xi^{(2+1)}) \cdot \dots \cdot (x^{(m-1)} \xi^{(m-1)}).$$

As was shown in studying affine space, restitution turns it into an invariant built up of types  $(x \xi)$  and  $\Delta^*$

After we have got the better of factors a) and b), taking into consideration that a product of two  $\Delta_a$  is expressible in terms of types c), g), and the purely Greek  $\delta_p$ , there remain but two cases to be examined:

(3) A product p of types  $(x \xi)$  and  $(x|y)$ ; as we saw before, it goes into a  $\Delta_a$  by restitution,

$$(4) \Delta_a \cdot p, \text{ or in explicit form exhibiting the symbolic arguments } [x^{(1)}|, \dots, x^{(p)}|, y^{(1)}|, \dots, y^{(\sigma)}|, \eta^{(\sigma+1)}, \dots, \eta^{(m-p)}] \text{ times } (x^{(\sigma+1)}|z^{(1)}) \cdot \dots \cdot (x^{(\rho+\tau)}|z^{(\tau)}) \cdot (x^{(\rho+\tau+1)} \xi^{(\tau+1)}) \cdot \dots \cdot (x^{(m-1)} \xi^{(m-p-1)}).$$

We write for a moment

$$(7.1) \quad \begin{array}{l} \eta^{(1)}, \dots, \eta^{(\sigma)} \quad \text{instead of} \quad y^{(1)}|, \dots, y^{(\sigma)}|, \\ \xi^{(1)}, \dots, \xi^{(\tau)} \quad \text{instead of} \quad z^{(1)}|, \dots, z^{(\tau)}|. \end{array}$$

Then restitution turns our expression into

$$\delta_{\mu-p} = (\eta^{(1)}, \dots, \eta^{(m-p)} || \xi^{(1)}, \dots, \xi^{(m-p)})_{\mu-p}$$

as we mentioned above;  $\xi^{(m-p)}$  denotes the restituent. On replacing

$\eta^{(1)}, \dots, \eta^{(\sigma)}; \xi^{(1)}, \dots, \xi^{(\tau)}$  again by what they stand for according to (7.1), we do what was discussed at the end of the last section: some of the Greek arguments in our  $\delta_{\mu-p}$  are converted into arguments of type  $x|$ . As this was seen to lead to no new invariants, our proof of completeness is finished.

Our group consisted, as far as the head components of covariant vectors were concerned, of the proper orthogonal transformations. On adding the improper orthogonal transformations, some of our basic invariants will stay unaltered under their influence, some will take on the factor  $-1$  (even and odd invariants). The types c), e), f), g) are even, whereas the odd types are found under a), b), d) and h):

$$D = D(x \dots z) = [x \dots z],$$

$$D^* = D^*(x \dots y) = [x \dots y]_{\mu},$$

$$\Delta_a = [x_1, \dots, y_1, \xi, \dots, \eta]$$

$\leftarrow \quad a \quad \rightarrow \quad \leftarrow \quad n-a \quad \rightarrow$

(including  $\Delta_0 = \Delta$ ). A full table of basic typical invariants for the complete group can be made up by adding to the even invariants in our list the products of any two of the odd ones. Let us examine a little closer the latter sort of even invariants.

First, the product of two D's:

$$D(x \dots z) \cdot D(x' \dots z').$$

We have proved the formula

$$[\xi \dots \xi] [\xi' \dots \xi'] = \delta_{\mu}(\xi \dots \xi | \xi' \dots \xi') = (-1)^{\nu} \begin{vmatrix} (\xi \xi'), \dots, (\xi \xi'), \xi_{\mu+1}, \dots, \xi_m \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (\xi \xi'), \dots, (\xi \xi'), \xi_{\mu+1}, \dots, \xi_m \\ \xi'_{\mu+1}, \dots, \xi'_{\mu+1}, 0, \dots, 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \xi'_m, \dots, \xi'_m, 0, \dots, 0 \end{vmatrix}$$

On subtracting appropriate multiples of the last  $m - \mu = \nu$  rows from the first  $n$ , one can change all products like  $(\xi \xi')$  into  $(\xi | \xi')$ . We thus may put the product of two D's into the following determinant form:

$$\begin{vmatrix} (x|x'), \dots, (x|z'), x_{\mu+1}, \dots, x_m \\ (z|x'), \dots, (z|z'), z_{\mu+1}, \dots, z_m \\ \hline x'_{\mu+1}, \dots, z'_{\mu+1}, 0, \dots, 0 \\ \hline x'_m, \dots, z'_m, 0, \dots, 0 \end{vmatrix} = D D(x \dots z | x' \dots z').$$

According to the well-known formula

$$\begin{vmatrix} x_1, \dots, x_\mu \\ \vdots \\ y_1, \dots, y_\mu \end{vmatrix} \cdot \begin{vmatrix} x'_1, \dots, y'_1 \\ \vdots \\ x'_\mu, \dots, y'_\mu \end{vmatrix} = \begin{vmatrix} (x|x'), \dots, (x|y') \\ \vdots \\ (y|x'), \dots, (y|y') \end{vmatrix}$$

the product of two  $D^*$  is expressible by means of scalar products. However  $D$  times  $D^*$ :

$$\begin{vmatrix} x_1 & \dots & x_m \\ \vdots & & \vdots \\ z_1 & \dots & z_m \end{vmatrix} \cdot \begin{vmatrix} x'_1 & \dots & y'_1 & 0 & \dots & 0 \\ \vdots & & \vdots & & & \vdots \\ x'_\mu & \dots & y'_\mu & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{vmatrix}$$

yields a new result

$$D D^* \{ x \dots z | x' \dots y' \} = \begin{vmatrix} (x|x'), \dots, (x|y'), x_{\mu+1}, \dots, x_m \\ \vdots \\ (z|x'), \dots, (z|y'), z_{\mu+1}, \dots, z_m \end{vmatrix}$$

( $n$  vectors  $x, \dots, z$ ;  $\mu$  vectors  $x', \dots, y'$ ).

The product of two  $\Delta_a$ , as we know, is expressible by types  $(x\xi)$ ,  $(x|y)$  and  $\delta_p$ ; in particular, the product of two  $\Delta$ 's is  $\delta_\mu$ . Nor does the product  $D \cdot \Delta_a$  result in anything new:

$$\begin{vmatrix} x_1 & \dots & x_m \\ \vdots & & \vdots \\ z_1 & \dots & z_m \end{vmatrix} \cdot \begin{vmatrix} x'_1 & \dots & y'_1 & \xi_1 & \dots & \eta_1 \\ \vdots & & \vdots & & & \vdots \\ x'_\mu & \dots & y'_\mu & \xi_\mu & \dots & \eta_\mu \\ 0 & \dots & 0 & \xi_{\mu+1} & \dots & \eta_{\mu+1} \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & 0 & \xi_m & \dots & \eta_m \end{vmatrix} = \begin{vmatrix} (x|x'), \dots, (x|y'), (x\xi), \dots, (x\eta) \\ \vdots & & \vdots & & & \vdots \\ (z|x'), \dots, (z|y'), (z\xi), \dots, (z\eta) \end{vmatrix}$$

(a vectors  $x', \dots, y'$ ). As to the products  $D' \cdot \Delta_a$  we find

$$\begin{array}{c} \uparrow \\ \mu \\ \downarrow \end{array} \begin{vmatrix} x_1 & \dots & x_m \\ \vdots & & \vdots \\ z_1 & \dots & z_m \\ \hline 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{vmatrix} \cdot \begin{vmatrix} x'_1 & \dots & y'_1 & \xi_1 & \dots & \eta_1 \\ \vdots & & \vdots & & & \vdots \\ x'_\mu & \dots & y'_\mu & \xi_\mu & \dots & \eta_\mu \\ 0 & \dots & 0 & \xi_{\mu+1} & \dots & \eta_{\mu+1} \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & 0 & \xi_m & \dots & \eta_m \end{vmatrix} = \begin{vmatrix} (x|x'), \dots, (x|y'), (x\xi), \dots, (x\eta) \\ \vdots & & \vdots & & & \vdots \\ (z|x'), \dots, (z|y'), (z\xi), \dots, (z\eta) \\ 0 & \dots & 0 & \xi_{\mu+1} & \dots & \eta_{\mu+1} \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & 0 & \xi_m & \dots & \eta_m \end{vmatrix}$$

On expanding the latter determinant by the last  $\nu$  columns one sees that it is expressible in terms of types  $(x|x')$ ,  $(x\xi)$  and  $\Delta^* = [\xi' \dots \xi^{(\mu)}]_{\mu+1, \dots, m}$

We thus arrive at the following complete table of basic invariants for the full Euclidean group:

Latin invariants:  $\bar{a}$ ).  $(x|y)$ ,

$\bar{b}$ ).  $DD(x \cdots z | x' \cdots z')$ ,

$\bar{c}$ ).  $DD^* \left\{ \begin{matrix} x & \cdots & z \\ x' & \cdots & y' \end{matrix} \right\}$ ;

Greek invariants:

$\bar{d}$ ).  $\Delta^* = [\xi \cdots \eta]_{\mu+1, \dots, n}$ ,

$\bar{e}$ ).  $\delta_1, \dots, \delta_\mu$

$\{\delta_\nu(\xi \cdots \eta || \xi' \cdots \eta') = \Delta^*(\xi \cdots \eta) \cdot \Delta^*(\xi' \cdots \eta')\}$ ;

Mixed invariant:

$\bar{f}$ ).  $(x\xi)$ .

Another point which deserves mention is the weight of our invariants.

When we remove the restriction  $|\sigma| = 1$  imposed on the rim of the transformations s. (p. 57) of our group, most types become relative rather than absolute invariants: under the influence of the transformation s they multiply with a certain power  $|\sigma|^g$  of  $|\sigma|$ ; the exponent g is what we called the weight. The weights of our basic invariants of type

(a), (b), (c), (d), (e), (f), (g), (h), are  
1, 0, 0, -1, -1, -2, 0, -1, respectively.

It is worth while to set forth explicitly what the result of our investigation tells about mere point-invariants in a Euclidean space of dimensionality n-1 and rank  $\nu$ . The basic invariants are

1) when reflections are excluded:

$\Delta^*$ ,  $\Delta$ ;  $\delta_1, \dots, \delta_{\mu-1}$ ;

2) when reflections are included:

(7.2)  $\Delta^*$ ;  $\delta_1, \dots, \delta_{\mu-1}, \delta_\mu$ .

Let us dwell a little more in detail upon the meaning of this result in the latter case of the full group including reflections. This indeed is the first, and most

elementary, question a naive geometer will ask the invariant theorist: Can you give me a complete set of basic invariants for an arbitrary number of points

$\varphi, \psi$  in Euclidean geometry? And he would probably expect the simple answer to be: The square of the distance of two points  $\varphi, \psi$  -- or rather, the corresponding polarized expression, the scalar product of two line elements  $\overrightarrow{\varphi\psi}, \overrightarrow{\varphi'\psi'}$  is the only existing fundamental invariant. Is he right or wrong? Our more complicated table (7.2) seems to indicate that he is wrong.

Let it be understood that we now talk about Euclidean space proper ( $V = 1$ ) in  $n-1$  dimensions. In this case our table reads:

$$\xi_m; \delta_1, \dots, \delta_{m-1}.$$

Until now we have used homogeneous coordinates. However, it is natural in Euclidean space to represent a point  $\varphi$  by its non-homogeneous coordinates

$$(7.3) \quad \varphi_1 = \xi_1 / \xi_m, \dots, \varphi_{m-1} = \xi_{m-1} / \xi_m.$$

A homogeneous polynomial depending on the coordinates  $\xi, \eta, \dots$  of several points is changed into the corresponding expression in terms of the non-homogeneous coordinates either by putting the last coordinate  $\xi_m(\eta_m, \dots) = 1$  or through division by appropriate powers of  $\xi_m$  and  $\eta_m \dots$ . Vice versa, a given polynomial  $f^*(\varphi, \psi, \dots)$  of the non-homogeneous coordinates of several points  $\varphi, \psi, \dots$  becomes a quotient of a homogeneous form  $f(\xi, \eta, \dots)$  and a product  $\xi_m^a \eta_m^b \dots$  on introducing the homogeneous coordinates  $\xi(\eta, \dots)$  by (7.3).  $f$  is of degree  $a$  in  $\xi$ , or degree  $b$  in  $\eta$ , etc.  $a$  is the highest degree of terms occurring in the non-homogeneous polynomial  $f^*$  with respect to the set of variables  $\varphi_1, \dots, \varphi_{m-1}$ ; we call it the degree of  $f^*$  in  $\varphi$ . Similarly for  $b, \dots$

The invariants derived from (7.2) which depend on the non-homogeneous coordinates only:

$$\varphi_i = \xi_i / \xi_m, \quad \psi_i = \eta_i / \eta_m, \dots \\ (i = 1, \dots, m-1),$$

are

$$\delta_p^* (\varphi \dots \eta \parallel \varphi' \dots \eta') = \frac{\delta_p (f \cdot \eta \parallel f' \cdot \eta')}{\xi_m \dots \eta_m \xi'_m \cdot \eta'_m}, \quad (p = 1, \dots, \mu).$$

They replace the former  $\delta_p$ ; the first invariant  $\xi_m$ , of course, has disappeared. Any integral rational invariant depending on the coordinates  $\varphi, \eta, \dots$  of a number of points is expressible in an integral rational manner by invariants of type  $\delta_p^*$ . But as we saw before, and as is well known from the elements of analytic geometry, the higher  $\delta_p^*$ , ( $p = 2, \dots, \mu$ ), can be represented by means of  $\delta_1^*$  only, i.e. by the scalar product of two line elements  $\overrightarrow{\varphi\eta}$  and  $\overrightarrow{\varphi'\eta'}$ . Nor is this result unexpected; as a matter of fact we could have got it much cheaper by arguing thus: An invariant  $f$  in Euclidean space depending on several points  $\varphi, \eta, \zeta, \dots$  is an orthogonal invariant of the  $(n-1)$ -dimensional vectors  $\eta - \varphi, \zeta - \varphi, \dots$ ; to see this one only needs to make  $\varphi$  the origin of our coordinate system by a suitable translation. But then the main theorem about the orthogonal group tells us that  $f$  is an aggregate of the scalar products of those vectors.

So the naive geometer is right and our long list (7.2) appears to be useless. But let us not judge too hastily! They all become indispensable again as soon as we put the question in a slightly more sophisticated form. An invariant  $f(\varphi, \eta, \dots)$  of degree  $a, b, \dots$  in  $\varphi, \eta, \dots$  being given, we wish to express it in terms of the basic invariants such that no single term of that expression is of higher degree in any of the variable points  $\varphi, \eta, \dots$  than the invariant itself. This is a reasonable requirement; and when thus put, our problem is identical with the one we solved for the homogeneous coordinates: in this sharpened sense every invariant is expressible by  $\delta_1^*, \delta_2^*, \dots, \delta_\mu^*$  whereas the requirement is not met by the formulas that express the higher through  $\delta_1^*$ . For instance, in

$$\delta_2^*(\varrho_{xy}g || \varrho'_{xy}g') = \begin{vmatrix} \delta_1^*(\varrho_{xy} || \varrho'_{xy}) & \delta_1^*(\varrho_{xy} || \varrho'g') \\ \delta_1^*(\varrho g || \varrho'_{xy}) & \delta_1^*(\varrho g || \varrho'g') \end{vmatrix} \quad 97$$

the left side is of degree 1 in  $\varrho$  and  $\varrho'$  while the two terms on the right side are of degree 2 in both  $\varrho$  and  $\varrho'$ .

A treatment of the "extended" complex group similar to that given here for the extended orthogonal group, is to be found in Wanner's Zürich thesis:

Volle Systeme von Grundinvariantentypen, Zürich, 1926.

### C. The Second Main Theorem

#### 1. Statement of the proposition for the unimodular group

In  $n$ -dimensional vector space the typical basic invariants with respect to the group of all unimodular linear transformations are the Latin bracket factor  $[\alpha_1 \cdots \alpha_m]$  of  $n$  covariant vectors  $\alpha_i$ , the Greek bracket factor  $[\xi_1 \cdots \xi_m]$  of  $n$  contravariant vectors  $\xi$ , and the mixed factor  $(\alpha \xi) = (\xi \alpha)$ , the product of a covariant vector  $\alpha$  by a contravariant  $\xi$ . Among these basic invariants there exist relations of the following five types:

$$(I) \quad \sum_x^{\pm} [\alpha_1 \alpha_2 \cdots \alpha_m] (\alpha_0 \xi) = 0,$$

$$(II) \quad \sum_x^{\pm} [\alpha_1 \alpha_2 \cdots \alpha_m] [\alpha_0 \gamma_2 \cdots \gamma_m] = 0,$$

$$(III) \quad \sum_{\xi}^{\pm} [\xi_1 \xi_2 \cdots \xi_m] (\xi_0 \alpha) = 0,$$

$$(IV) \quad \sum_{\xi}^{\pm} [\xi_1 \xi_2 \cdots \xi_m] [\xi_0 \eta_2 \cdots \eta_m] = 0,$$

$$(V) \quad [\alpha_1 \alpha_2 \cdots \alpha_m] [\xi_1 \xi_2 \cdots \xi_m] - \begin{vmatrix} (\alpha_1 \xi_1) & \cdots & (\alpha_1 \xi_m) \\ \cdot & \cdot & \cdot \\ (\alpha_m \xi_1) & \cdots & (\alpha_m \xi_m) \end{vmatrix} = 0.$$

The Latin letters denote covariant vectors, the Greek ones contravariant vectors.

The sum  $\sum_x^{\pm}$  in (I) and (II) is to be understood as drawing  $\alpha_0$  alternately through the series  $\alpha_c \alpha_1 \cdots \alpha_m$ ; in similar fashion  $\xi_0$  in (III) and (IV)

is to be drawn through the series  $\xi_0, \xi_1, \dots, \xi_m$ . The second main theorem states that all relations holding between the basic invariants are an algebraic consequence of relations of these five types. For the sake of a precise formula-

tion of the second main theorem, one will first have to consider quantities like

$$(1.1) \quad [x_1 \dots x_m], [\xi_1 \dots \xi_m], (x\xi) = (\xi x)$$

as independent variables ("formal standpoint"); the Latin and Greek "symbols"  $x$  and  $\xi$  are here devoid of any independent significance. Nevertheless, it is to

be understood that a bracket factor containing two identical symbols is zero, and

that a bracket factor like  $[x_1 x_2 \dots x_m]$  changes into  $\pm [x_1 x_2 \dots x_m]$

by a permutation of the  $x_i$  -- with the positive sign for even, the negative for

odd permutations. Let  $F$  be an integral rational function of such variables com-

posed of certain Latin symbols  $x_1, x_2, \dots$  and certain Greek ones  $\xi_1, \xi_2, \dots$ .

All the functions  $J$  obtained by substituting into the expressions on the left

side of (I) to (V) those symbols in all possible combinations for the Latin and

Greek letters used there -- of course Latin symbols should be substituted for Latin

letters only, Greek symbols for Greek letters -- form the basis of an ideal

$$\mathcal{J} = \{ J_t \}. \text{ One returns to the old standpoint by replacing each of the}$$

Latin and Greek symbols  $x_1, x_2, \dots; \xi_1, \xi_2, \dots$  by a variable covariant

or contravariant vector respectively, and then interpreting the symbols (1.1) in

their old meaning, as determinants and inner product; this procedure is what I

call the substitution. The second main theorem contends: If  $F$  goes over into 0

by substitution, then it belongs to the ideal  $\mathcal{J}$ .

Besides (I) to (V), I make use of the following typical expression, which vanishes by substitution:

$$(VI) \quad \begin{vmatrix} (x_0 \xi_0), (x_0 \xi_1), \dots, (x_0 \xi_m) \\ (x_1 \xi_0), (x_1 \xi_1), \dots, (x_1 \xi_m) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ (x_m \xi_0), (x_m \xi_1), \dots, (x_m \xi_m) \end{vmatrix} \cdot$$

It belongs to the ideal  $\mathcal{J}$ ; for on expanding by the first column and replacing determinants like  $\det. (\alpha_i \xi_k)$ ,  $[i, k = 1, \dots, n]$ , which appear as factors, by the product,

$$[\alpha_1 \dots \alpha_n] [\xi_1 \dots \xi_n]$$

modulo (V), one obtains (I) (where one has to take  $\xi = \xi_0$ ) multiplied by  $[\xi_1 \dots \xi_n]$ .

Without loss of generality we may suppose  $F$  to be homogeneous in each of the Latin and Greek symbols. For  $F$  may be decomposed in such homogeneous parts according to the degrees in those symbols; and if  $F$  changes into zero by substitution, the same holds true for each individual part. (When computing the degree of a monomial term of  $F$ , one has of course to consider each variable ( ) of degree 1 in the symbols that occur, of degree 0 in those that do not occur.) A single term of  $F$  may contain  $\lambda$  Latin,  $\mu$  Greek bracket factors; then the total degree as to the Latin symbols minus the total degree in the Greek symbols, amounts to  $n(\lambda - \mu)$  for this term. Consequently under the assumption of a homogeneous  $F$ , the difference  $\lambda - \mu$  has the same value for all terms of  $F$ . The product of a Latin and a Greek bracket factor,  $[\alpha_1 \dots \alpha_n]$  and  $[\xi_1 \dots \xi_n]$  may be replaced, mod. (V), by an aggregate of variables of type  $(\alpha \xi)$ : namely, the determinant of the  $(\alpha_i \xi_k)$ . Hence we are allowed to suppose that  $F$  contains either Latin or Greek bracket factors only, and each term of  $F$  the same number of them. Since our table of fundamental relations is symmetric with respect to the part played by the Latin and Greek symbols, we confine ourselves to the case where at most Latin but no Greek bracket factors occur in  $F$ . After these preparations, the following sharper formulation of the second main theorem will hold:

T<sub>0</sub>. A homogeneous  $F$  when becoming zero by substitution and containing only variables of type  $(\alpha \xi)$  is  $\equiv 0$  mod. (expressions shaped after the type (VI) alone).

$T_x$ . If, however, the homogeneous F involves Latin, but no Greek, bracket factors besides variables of type  $(\alpha \xi)$ , then it is  $\equiv 0 \text{ mod.}$  (expressions of type (I) and (II)).

## 2. Capelli's formal congruence

Upon the homogeneous F we are going to apply that Capelli identity which involves  $n+1$  Latin arguments in  $n$ -dimensional space. However, we now stand on the "new standpoint" where the Latin and Greek symbols are not vectors but simply ingredients of the notations

$$(2.1) \quad [x_1 \cdots x_m], [\xi_1 \cdots \xi_m], (\alpha \xi)$$

of variables. Therefore we first ought to define the polar process according to this interpretation ("formal polarization"); Capelli's relation will then hold as a congruence mod.  $\mathcal{J}$  rather than as an equation.

$x$  and  $y$  being two of the Latin symbols, the polar process  $\mathcal{D} = \mathcal{D}_{yx}$  is as before bound to satisfy the formal laws:

$$(2.2) \quad \begin{aligned} \mathcal{D}(f+g) &= \mathcal{D}f + \mathcal{D}g, \\ \mathcal{D}(cf) &= c \cdot \mathcal{D}f, & (c, \text{ being a number}), \\ \mathcal{D}(f \cdot g) &= g \mathcal{D}f + f \cdot \mathcal{D}g. \end{aligned}$$

We derive from them how  $\mathcal{D}_{yx}$  affects any polynomial  $f$  as soon as we know how it affects the arguments (2.1) of  $f$ . This shall occur according to the rules:

1) A variable is changed into zero by  $\mathcal{D}_{yx}$  if its symbol does not contain the letter  $x$ ,

$$2) \quad \mathcal{D}_{yx} [x x_2 \cdots x_m] = [y x_2 \cdots x_m], \quad \mathcal{D}_{yx} (\alpha \xi) = (y \xi).$$

To prove Capelli's congruence we proceed exactly as before (pp. 37-42) distinguishing, however, from the beginning the two cases that our homogeneous  $F$  contains, either (a) variables of type  $(\alpha\xi)$  only, or (b) bracket factors besides. One introduces symbols  $x'_0, x'_1, \dots, x'_m$  not occurring in  $F$ , and forms the sum

$$(2.3) \quad \sum_{x'} \pm \mathcal{D}_{x'_m x'_m} \cdots \mathcal{D}_{x'_1 x'_1} \mathcal{D}_{x'_0 x'_0} F$$

extending alternately to the  $(n+1)!$  permutations of  $x'_0, x'_1, \dots, x'_m$ . In the case (a) this expression obviously is made up of terms

$$Q \cdot \sum_{x'} \pm (x'_0 \xi_0)(x'_1 \xi_1) \cdots (x'_m \xi_m)$$

where  $Q$  is a monomial of the variables occurring in  $F$ . Hence (2.3) is congruent 0 modulo determinants of type (VI). After forming (2.3) one substitutes

$x_0, x_1, \dots, x_m$  for  $x'_0, x'_1, \dots, x'_m$ . That this has the same effect as in the earlier less formal interpretation one sees by the same procedure employed there, when one now makes use of the following two facts, instead of the differentiation formula  $(\alpha)$ ,  $x, x'$  being two different symbols, we have

$$\mathcal{D}_{y x} \{f(x', x)\}_{x'=x} = \{\mathcal{D}_{y x'} f(x', x) + \mathcal{D}_{y x} f(x', x)\}_{x'=x};$$

( $\beta$ ), if  $f(x')$  be linear in  $x'$ , then  $\mathcal{D}_{y x'} f(x') = f(y)$ . According to the rules (2.2) it is sufficient to give the proof for the case when  $f$  is one of the variables. The only possibility of both symbols  $x, x'$  actually occurring together is then a bracket factor

$$[x' x x_3 \cdots x_m]$$

The left side is 0 as an  $[x x x_3 \cdots x_m]$  with two equal  $x$  means 0 by definition, whereas the right side equals

$$[y x x_3 \cdots x_m] + [x y x_3 \cdots x_m].$$

In this manner one finds that (2.3) goes into

$$(2.4) \quad \sum_{x'} \pm (\mathcal{D}_{x'_m x'_m} + m \delta_{x'_m x'_m}) \cdots (\mathcal{D}_{x'_2 x'_2} + 2 \delta_{x'_2 x'_2}) (\mathcal{D}_{x'_1 x'_1} + 1 \cdot \delta_{x'_1 x'_1}) \mathcal{D}_{x'_0 x'_0} F$$

after the original symbols  $x_0, x_1, \dots, x_m$  were substituted for the new ones  $x'_0, x'_1, \dots, x'_m$ . The alternating sum in (2.4) has now to be interpreted such that  $x'_0, x'_1, \dots, x'_m$  is replaced by all the permutations of  $x_0, x_1, \dots, x_m$  one after the other.  $\sum_{x'x}$  means 1 or 0 according as the symbols  $x'$  and  $x$  coincide or not. The result is that (2.4) is congruent 0 modulo type (VI).

In the case (b) we proceed as follows. We have to apply (2.4) on a monomial  $F$  which is a product of variables of type  $(x \xi)$  and of Latin bracket factors; at least one bracket factor is present. On extending the sum first only to the permutations of  $x'_1, \dots, x'_m$  we are capable of successively drawing all symbols  $x'_1, \dots, x'_m$  into one bracket factor. This is done by means of the identity (3.1) of § 3, p. 71, where one has to choose as the linear form  $(x^{(i)} \pi)$  either a variable of type  $(x^{(i)} \xi)$  or of type  $[x^{(i)} y_1 \dots y_m]$ . One should be aware that that equation in these two cases is either the identity (I) or (II); hence our transformation is not an identical transformation but a transformation modulis expressions of type (I) and (II). After dropping factors which do not contain the symbols  $x'_0, x'_1, \dots, x'_m$  the sum (2.4) is now either an expression of type (I) or (II). We thus arrive at the result that (2.4) is congruent modulis types (I) and (II) provided the homogeneous  $F$  involves Latin bracket factors.

### 3. Proof of the second main theorem for the unimodular group

If  $F$  is of degrees  $n_0, n_1, \dots, n_m$  with respect to the symbols

$x_0, x_1, \dots, x_m$ , we may write Capelli's congruence just proved in this form:

$$(3.1) \quad \rho F \equiv \sum \mathcal{P} F^*$$

Here

$$\rho = n_0(n_1+1) \dots (n_m+m)$$

and hence  $\rho \neq 0$  if  $F$  actually involves the symbol  $x_0$ . The polynomial  $F^*$  is of lesser rank than  $F$  and derived from  $F$  by polarization;  $\mathcal{P}$  is a succession of polar processes.

We defined "formal polarization" such that it does not matter whether polarization on a given  $F$  is carried out ("formally") before, or ("not formally") after substitution. Hence if  $F$  vanishes by substitution, the same will hold for any  $G$  derived from  $F$  by polarization, in particular for the forms  $F^*$  occurring in (3.1). The expressions (VI) as well as (I) and (II) change into expressions of the same structure by polarization. We have therefore, by means of the congruence (3.1), reduced the validity of the respective theorems  $\overline{T}_0$  and  $\overline{T}_\alpha$  for an  $F$  to their validity for the lower  $F^*$  as long as  $F$  actually contains the symbol  $\alpha_0$ . The inductive procedure thus started will end in entirely eliminating  $\alpha_0$  from  $F$ . The same procedure may be repeated as long as  $F$  still contains more than  $n$  Latin symbols by assigning to  $n+1$  of these symbols the role played by  $\alpha_0, \alpha_1, \dots, \alpha_m$  just now. In this manner one finally comes down to  $F$ 's involving not more than  $n$  Latin symbols  $\alpha_1, \dots, \alpha_m$ . In the case (a) such an  $F$  is a function of  $m \nu$  variables of form

$$(\alpha_i \xi_\kappa), \quad [i = 1, \dots, m; \kappa = 1, \dots, \nu];$$

the number  $\nu$  is not subject to any limitation. Theorem  $\overline{T}_0$  will be proved by showing: If in case (a) the number of Latin symbols amounts to  $n$ , our  $F$  cannot become zero through the substitution unless it was zero before the substitution.

This is readily done, as one can find at once  $n$  covariant vectors  $\alpha_1, \dots, \alpha_m$  and  $\nu$  contravariant vectors  $\xi_\kappa$  such that the inner products  $(\alpha_i \xi_\kappa)$  become equal to arbitrarily preassigned numbers  $z_{i\kappa}$ . For this purpose one only has to take

$$\begin{aligned} \alpha_1 &= (1, 0, 0, \dots, 0), \\ \alpha_2 &= (0, 1, 0, \dots, 0), \\ &\dots \\ \alpha_m &= (0, 0, 0, \dots, 1), \end{aligned} \quad \xi_\kappa = (z_{1\kappa}, z_{2\kappa}, \dots, z_{m\kappa}).$$

In case (b) the homogeneous  $F$ , if containing not more than  $n$  Latin symbols, is necessarily of the form

$$[x_1 x_2 \dots x_m]^\lambda \cdot G\{(x_i; \xi_\kappa)\}$$

where the second factor  $G$  again depends on variables of type  $(x_i; \xi_\kappa)$  alone.

Not more than one such term can occur because the exponent  $\lambda$  is fixed by the difference between the total degree of  $F$  in the Latin and Greek symbols. Hence, here again it is true that  $F$  can vanish after the substitution only if it would before the substitution; indeed the vanishing of  $F$  after the substitution implies that of  $G$ .

#### 4. The second main theorem for the orthogonal and complex group

If  $\gamma$  is the group of all proper and improper orthogonal transformations, then we have only one basic type of invariant, namely the scalar product  $(xy)$ . A typical relation among scalar products is the following involving  $n+1$  vectors  $x$  and  $n+1$  vectors  $y$ :

$$J = \begin{vmatrix} (x_0 y_0), & (x_0 y_1), & \dots, & (x_0 y_n) \\ (x_1 y_0), & (x_1 y_1), & \dots, & (x_1 y_n) \\ \cdot & \cdot & \cdot & \cdot \\ (x_n y_0), & (x_n y_1), & \dots, & (x_n y_n) \end{vmatrix}$$

(This relation follows in the simplest way from the fact that  $J$  is a skew-symmetric multilinear form of the  $n+1$  vectors  $x_0, x_1, \dots, x_n$  and such a form is always zero in  $n$ -dimensional space.)

Second main theorem. Every relation between scalar products is an algebraic consequence of relations of type  $J$ .

$m$  Latin "symbols"  $x_1, \dots, x_m$  being given, a relation is a polynomial in the  $\frac{m(m+1)}{2}$  variables  $(x_\alpha x_\beta)$  that becomes zero when one replaces the symbols  $x_1, \dots, x_m$  by arbitrary vectors, and the variable  $(x_\alpha x_\beta)$  by

the scalar product of the two vectors  $\alpha_\alpha, \alpha_\beta$  ("substitution"). We agree that even from the formal standpoint  $(\alpha_\alpha \alpha_\beta)$  and  $(\alpha_\beta \alpha_\alpha)$  shall be considered one and the same variable. When replacing in the expression  $J$  the "letters"  $x_0, x_1, \dots, x_m, y_0, y_1, \dots, y_n$  by any of the "symbols"  $x$ , one must allow a letter  $y$  to be replaced by the same symbol  $\alpha_\alpha$  as a letter  $x$ ; it is useless though to replace different letters  $x$  or  $y$  by the same symbol  $\alpha_\alpha$  since the whole expression  $J$  is skew-symmetric in  $x$  as well as in  $y$ . The expressions derived from  $J$  by the described substitution of the basis of an ideal, and the exact rendering of the second main theorem states that every relation  $R$  is congruent  $0 \pmod{\dots}$

The proof is given again by means of Capelli's congruence. The formal definition of the polar process  $\mathfrak{D} = \mathfrak{D}_{y,x}$  is here as follows:  $\mathfrak{D}(uv) = 0$  if neither  $u$  nor  $v$  equals  $x$ ;  $\mathfrak{D}(\alpha u) = (\gamma u)$ , if  $u$  is different from  $x$ ;  $\mathfrak{D}(\alpha x) = 2(\alpha y)$ . The rules  $(\alpha)$  and  $(\beta)$  remain valid; in particular, one may check the case  $\mathfrak{f}(\alpha'x) = (\alpha'x)$  for  $(\alpha)$ . In forming (2.3) we shall obtain some terms that are built quite similarly to the foregoing (p 101):

$$(41) \quad Q \cdot \sum_{\alpha'} \pm (\alpha'_0 y_0) (\alpha'_1 y_1) \dots (\alpha'_m y_m),$$

but just these terms are  $\equiv 0 \pmod{\text{the ideal}}$  defined above. However, this is not the only possibility now. Suppose that a term of  $F$ , for instance, contains the factor  $(\alpha_c \alpha_1)$  and that the first polarization  $\mathfrak{D}_{x'_0 x_0}$  is performed on this factor; it then goes into  $(\alpha'_0 \alpha_1)$ . The second polarization  $\mathfrak{D}_{x'_1 x_1}$  if performed again on this factor in its new form will change it into  $(\alpha'_0 \alpha'_1)$ . Hence we should be prepared for the possibility that instead of the alternating sum  $\sum_{\alpha'}$  in (4.1), another might occur whose leading term in addition to variables of the kind  $(\alpha'_0 y_0)$  involves variables of the kind  $(\alpha'_0 \alpha'_1)$  joining two of the new symbols  $x'$ . But  $\sum_{\alpha'}$  will then certainly be  $\equiv 0$  according to the convention  $(\alpha y) = (\gamma x)$ . Capelli's congruence therefore proves to be true

modulo the ideal here introduced, whose basis consists of expressions of type J, alone.

Using this congruence the number of Latin symbols may gradually be reduced to  $n$ . To finish the proof we must show: a polynomial  $F$  of the  $\frac{n(n+1)}{2}$  variables

$$(4.2) \quad (x_\alpha x_\beta), \quad [\alpha, \beta = 1, 2, \dots, n],$$

must needs be equal to zero before the substitution if changed into zero by the substitution. In the Introduction we alluded to a demonstration of this based on the fact that (in the realm of real numbers) vectors  $x_\alpha$  can be ascertained such that the scalar products (4.2) form an arbitrarily preassigned symmetric matrix if only the quadratic form with this matrix of coefficients is positive definite.

We shall here prefer to give instead a direct algebraic proof based on induction from  $n-1$  to  $n$ , and valid in any infinite number field. For this purpose we depend upon the following two trivial lemmata concerning the vanishing of polynomials:

1) A polynomial  $\Phi(\lambda)$  of the variable  $\lambda$  vanishes identically if  $\Phi(\lambda + a) = 0$ , (a being a number of the ring from which the coefficients of  $\Phi$  are taken);

2) A polynomial  $\Phi(\lambda_1, \dots, \lambda_n)$  of  $n$  variables  $\lambda_i$  is zero provided  $\Phi(\lambda'_1, \dots, \lambda'_n) = 0$  holds identically in  $\lambda$  when the  $\lambda'_i$  are derived from the variables  $\lambda_i$  by a non-singular linear transformation:

$$\lambda'_i = \sum_{k=1}^n a_{ik} \lambda_k, \quad |a_{ik}| \neq 0.$$

We assume  $x_1, \dots, x_{n-1}$  to be numerically given and linearly independent vectors in the subspace  $\mathbb{R}_{n-1}$  whose vectors have their last component = 0. Hence the determinant  $\Delta$  of the  $n-1$  vectors

$$(4.3) \quad x_i = (a_{i1}, \dots, a_{i,n-1}), \quad [i = 1, \dots, n-1]$$

in  $\mathbb{R}_{m-1}$  is  $\neq 0$ . We consider the vector

$$\alpha = (\lambda_1, \dots, \lambda_{m-1}, \lambda)$$

as variable. It is

$$(\alpha \alpha_i) = \sum a_{ik} \lambda_k, \quad [i = 1, \dots, m-1]$$

$$(\alpha \alpha) = \lambda^2 + (\lambda_1^2 + \dots + \lambda_{m-1}^2)$$

The given polynomial  $F$  depending on  $(\alpha \alpha)$  and the  $(\alpha \alpha_i) = \xi_i$  may first be looked upon as a polynomial in  $(\alpha \alpha)$  alone with coefficients lying in the ring of polynomials of  $\xi_1, \dots, \xi_{m-1}$ :

$$(4.4) \quad F = \sum_h (\alpha \alpha)^h \cdot \Phi_h(\xi_1, \dots, \xi_{m-1}).$$

One here has to think of the variables

$$(4.5) \quad (\alpha_\alpha \alpha_\beta), \quad [\alpha, \beta = 1, \dots, m-1]$$

as being replaced by the scalar products of the  $m-1$  vectors (4.3) in  $\mathbb{R}_{m-1}$ .

Putting for the moment

$$\lambda_1^2 + \dots + \lambda_{m-1}^2 = \alpha, \quad \text{and}$$

$$\Phi_h(\sum_k a_{ik} \lambda_k) = \Psi_h(\lambda_1, \dots, \lambda_{m-1}),$$

we have, according to the above,

$$\sum_h (\lambda^2 + \alpha)^h \cdot \Psi_h(\lambda_1, \dots, \lambda_{m-1}) = 0$$

Hence

$$(4.6) \quad H(\mu + \alpha) = 0$$

identically in the variable  $\mu$  where

$$H(\mu) = \sum_h \mu^h \cdot \Psi_h(\lambda_1, \dots, \lambda_{m-1}).$$

The vanishing of the polynomial  $H(\mu)$  now follows from (4.6) by the first lemma, and the vanishing of all its coefficients  $\Psi_h(\lambda_1, \dots, \lambda_{m-1})$  implies that of  $\Phi_h(\xi_1, \dots, \xi_{m-1})$  identically in  $\xi$  according to the second lemma.

The coefficients of the polynomial  $F$  of  $(x_1, x_m), \dots, (x_{n-1}, x_m), (x_n, x_m)$  are polynomials  $f$  in (4.5). Concerning such a coefficient  $f$ , we learned that it vanishes if one substitutes for  $(x_\alpha, x_\beta)$  the scalar products of  $n-1$  vectors  $x_\alpha$  in the space  $\mathcal{R}_{m-1}$  whose determinant  $\Delta \neq 0$ . The restriction by this algebraic inequality is irrelevant. On assuming the proposition under test to hold good in  $\mathcal{R}_{m-1}$  we are able to infer from the vanishing of  $f$  after this substitution its vanishing before the substitution, and that concludes the argument leading to the formal identity  $F = 0$ .

In the second place, let us consider the group  $\gamma$  of all proper orthogonal transformations. To the type  $(x, y)$  of invariants has then to be added as a further fundamental invariant, the bracket factor

$$[x_1 \dots x_m]$$

and to the relations of type  $J = J_1$ , the following two further types:

$$J_2 \equiv [x_1 \dots x_m][y_1 \dots y_m] - \begin{vmatrix} (x_1, y_1) & \dots & (x_1, y_m) \\ \dots & \dots & \dots \\ (x_m, y_1) & \dots & (x_m, y_m) \end{vmatrix} = 0$$

$$J_3 \equiv \sum \pm [x_1 \dots x_m](x_0, y) = 0.$$

In the last equation  $x_0$  is drawn through the series  $x_1, \dots, x_m$ . The second main theorem asserts that this énumération is exhaustive for the group  $\gamma$ .

Proof: A given relation  $R$  can first be reduced modulo type  $J_2$  to such a form that no two bracket factors ever appear as multiplied together; i.e.

$$R \equiv F + G = 0,$$

where  $F$  is a function of the

$$(x_\alpha, x_\beta), \quad [\alpha, \beta = 1, \dots, m],$$

and  $G$  a linear combination of terms of the shape

$$[x_{\alpha_1} \dots x_{\alpha_m}] \cdot F_x \{ (x_\alpha, x_\beta) \}.$$



The proof is essentially like that for the full orthogonal group. However, attention has to be paid to the possibility, now not excluded, that two of the "new symbols"  $x'$  may join in a single factor like  $\{x'_0 x'_1\}$ ; for this product is now skew-symmetric instead of symmetric, and hence not annihilated by alternation! This is the reason why in addition to the relation  $J_m$  that corresponds to the relation  $J$ , the types  $J_1, J_2, \dots, J_{m-1}$  appear. Again, Capelli's congruence reduces the theorem under examination to the fact that no relation holds between the skew-products of  $2m$  vectors  $x_i$ . This follows from the possibility of ascertaining  $2m$  vectors  $x_i$  such that the matrix of their skew-products  $\{x_i x_j\}$  coincides with an arbitrarily preassigned skew-symmetric matrix  $\|Z_{ij}\|$ ; the construction of the  $x_i$  can be accomplished in a purely rational way. But one may proceed also in an analogous fashion as before in the case of the orthogonal group.

These few examples for the second main theorem may suffice to illustrate the method whose clue is the recourse to a formal Capelli congruence where the proof of the first main theorem made use of the ordinary Capelli identity. It is remarkable though that the more complicated case  $m = n$  involving the Cayley- $\Omega$ -process does not show up in this connection.

## CHAPTER II. THE GENERAL PROBLEM OF INVARIANT THEORY

### 1. The classical theory of invariants

The classical theory of invariants is concerned with the invariants of algebraic forms under the group  $\chi$  of all non-singular, homogeneous, linear transformations. Here the invariants are relative (multiplied by a power of the transformation determinant) rather than absolute. The latter type is obtained by restricting the group of transformations considered to the unimodular group.

Consider, by way of illustration, a two-dimensional vector space  $(x_1, x_2)$

and let

$$f = a_0 x_1^h + h a_1 x_1^{h-1} x_2 + \binom{h}{2} a_2 x_1^{h-2} x_2^2 + \dots + a_h x_2^h$$

be an arbitrary binary form of degree  $h$ . It depends on its  $h+1$  coefficients  $a_i$ , ( $i = 0, 1, \dots, h$ ).

Suppose that under the influence of an arbitrary non-singular, homogeneous, linear transformation

$$\begin{aligned} x_1 &= \alpha x'_1 + \beta x'_2, \\ x_2 &= \gamma x'_1 + \delta x'_2, \end{aligned} \quad \Delta = \alpha\delta - \beta\gamma \neq 0,$$

$f$  is transformed into  $f'$ ,  $f \rightarrow f'$ :

$$f' = a'_0 x_1'^h + h a'_1 x_1'^{h-1} x_2' + \dots + a'_h x_2'^h.$$

If for every such transformation a homogeneous polynomial  $J(a_0, a_1, \dots, a_h)$  has the property that

$$J(a'_0, a'_1, \dots, a'_h) = \Delta^G \cdot J(a_0, a_1, \dots, a_h)$$

identically in  $a_0, a_1, \dots, a_h$  after the coefficients  $a'_i$  have been replaced by their values in terms of the coefficients  $a_i$ , then  $J(a_0, a_1, \dots, a_h)$  is an invariant of weight  $G$  of the form  $f$ .

The discriminant  $a_0 a_2 - a_1^2$  of a binary form of degree 2 is the simplest case of such an invariant. It is of weight two.

More generally, one may consider a vector space  $(x_1, x_2, \dots, x_m)$  of  $n$  dimensions and an  $n$ -ary form  $f$  of degree  $h$ . An invariant of  $f$  under the group  $\gamma$  is defined as above.

An extension to invariants (sometimes called simultaneous invariants) of several forms  $f, g, \dots$  arose early in the history of the subject. Thus, for example, the resultant  $a\beta - bA$  of the two linear forms  $ax_1 + bx_2$ ,  $Ax_1 + Bx_2$  is a simultaneous invariant of weight 1 of the two forms.

An extension of the concept of invariants in still another direction is obtained by permitting the polynomial  $J$  to contain the variables  $x_1, x_2, \dots, x_m$  of  $f$  as well as the coefficients  $a_0, a_1, \dots, a_n$ . Such invariants are called covariants of weight  $G$ .

Let  $J$  be an invariant of  $k$  forms  $f_1, f_2, \dots, f_k$  of degrees  $\mu_1, \mu_2, \dots, \mu_k$  respectively in the  $n$  variables  $x_1, x_2, \dots, x_m$ . Suppose a particular term  $j$  of  $J$  be of degree  $m_1$  in the coefficients of  $f_1, m_2$  in the coefficients of  $f_2$ , etc. Apply now the transformation

$x_1 = \alpha x'_1, x_2 = \alpha x'_2, \dots, x_m = \alpha x'_m$  whose determinant  $\Delta = \alpha^m$ . Then each form  $f_i$  is transformed into a form whose coefficients are multiples by  $\alpha^{\mu_i}$  of the coefficients of  $f_i$ . In the invariant  $J$  formed for the transformed coefficients, the term corresponding to  $j$  is the product of  $j$  by

$$(\alpha^{\mu_1})^{m_1} \cdot (\alpha^{\mu_2})^{m_2} \cdot \dots \cdot (\alpha^{\mu_k})^{m_k} = \alpha^{\sum_{i=1}^k \mu_i m_i}.$$

If now  $G$  is the weight of  $J$ , we have, evidently  $\sum_{i=1}^k \mu_i m_i = m \cdot G$ .

If  $k=1$ , we have that for an invariant  $J$  of weight  $G$  of a single  $n$ -ary form  $f$ , of degree  $\nu$ , each term is of constant degree  $N$ , where  $N\nu = m \cdot G$ .

## 2. Examples of covariants

(a) The Hessian of an  $n$ -ary form  $f(x_1, x_2, \dots, x_m)$  is defined by the determinant

$$\left| \frac{\partial^2 f}{\partial x_i \partial x_j} \right|, \quad (i, j = 1, 2, \dots, m).$$

If  $f(x_1, \dots, x_m) \rightarrow F(x'_1, \dots, x'_m)$  by virtue of a transformation of the group  $\gamma$ , it is readily found that

$$\left| \frac{\partial^2 F}{\partial x'_i \partial x'_j} \right|_{(i, j = 1, 2, \dots, m)} = \Delta^2 \cdot \left| \frac{\partial^2 f}{\partial x_i \partial x_j} \right|_{(i, j = 1, 2, \dots, m)}.$$

Hence the Hessian of a form  $f$  is a covariant of  $f$  of weight 2.

(b) Let  $f_1, f_2, f_3$  be three forms in  $x_1, x_2, x_3$ . The Jacobian of these three forms is defined by the determinant

$$\left| \frac{\partial f_i}{\partial x_j} \right|, \quad (i, j = 1, 2, 3).$$

It is easily seen that the Jacobian is a covariant of weight 1.

## 2. Symbolic method

The symbolic method of invariant theory, due to Clebsch and Aronhold, was foreshadowed by the hyperdeterminants of Cayley. The chief feature of the method is that it reduces the problem of determining all invariants of a form to that of determining vector invariants, and while the first fundamental theorem for invariants of forms cannot be derived by this reduction, it furnishes a good starting point for this goal.

Let 
$$A = \sum_{n_1 + \dots + n_m = r} \frac{r!}{n_1! \dots n_m!} a_{n_1 \dots n_m} x_1^{n_1} \dots x_m^{n_m} \quad \text{denote}$$

an  $n$ -ary form of degree  $r$ , and consider an invariant  $J(A)$  of the coefficients of this form. If  $A$  be specialized as  $(a_1 x_1 + \dots + a_m x_m)^r$ , the invariant  $J(A)$  is a vector invariant depending on one contravariant vector. This suggests writing symbolically

$$\sum_{n_1 + \dots + n_m = r} \frac{r!}{n_1! \dots n_m!} a_{n_1 \dots n_m} x_1^{n_1} \dots x_m^{n_m} = (a_1 x_1 + \dots + a_m x_m)^r = a_x^r,$$

where  $a_1, \dots, a_m$  are mere symbols, certain combinations of which stand for certain of the coefficients  $a_{n_1 \dots n_m}$ ; i.e.,  $a_{n_1 \dots n_m}$  is replaced by  $a_1^{n_1} \dots a_m^{n_m}$ .

One difficulty immediately arises: if  $J(A)$  is linear in the coefficients of  $A$ , it is evident that from the symbolic representation of  $J(A)$ , say

$$\sum c_{n_1 \dots n_m} a_1^{n_1} \dots a_m^{n_m} \quad \text{the non-symbolic}$$

$J(A) = \sum c_{n_1 \dots n_m} a_{n_1} \dots a_{n_m}$  is uniquely determined. On the other hand, this is obviously not the case when  $J$  is not linear. To remove this difficulty, we first replace  $J_h(A)$ , an invariant of degree  $h$  of  $A$ , by the completely polarized invariant  $J_h(A^{(1)}, \dots, A^{(h)})$ . (This is known as the Aronhold polarization process.) Replacing, now,  $A$  by the  $r$ -th power of a linear form, we obtain an invariant  $J_h(a^{(1)}, \dots, a^{(h)})$  of  $h$  contravariant vectors. Inversely, if  $J_h(a^{(1)}, \dots, a^{(h)})$  is invariant, then  $J_h(A^{(1)}, \dots, A^{(h)})$  is an invariant, and this leads to  $J_h(A)$ .

As examples, we write in symbolic form the invariants and covariants exhibited in §1:

(a) Let  $A = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = a_x^2$ , and consider the invariant  $J_2(A) = a_{11}a_{22} - a_{12}^2$ . Polarizing first; we obtain  $J_2(A^{(1)}, A^{(2)}) = a'_{11}a_{22} + a_{11}a'_{22} - 2a_{12}a'_{12}$ . Now, specializing,

$$J_2(a', a) = a_1'^2 a_2^2 + a_2'^2 a_1^2 - 2a_1 a_2 a_1' a_2' = (a_1 a_2' - a_2 a_1')^2 = [aa']^2.$$

It is clear that from  $(a_1 a_2' - a_2 a_1')^2$  one may derive the invariant  $2(a_{11}a_{22} - a_{12}^2)$ .

(b) The Jacobian

$$\left| \frac{\partial f_i}{\partial x_k} \right|, \quad (i, j = 1, 2, 3),$$

depends linearly on the three forms  $f_1, f_2, f_3$ . Hence specializing at once, we write

$$f_1 = a_x^g, \quad f_2 = b_x^h, \quad f_3 = c_x^k.$$

The first row of the determinant is, then,  $a_1 g a_x^{g-1}, a_2 g a_x^{g-1}, a_3 g a_x^{g-1}$  and we have

$$\left| \frac{\partial f_i}{\partial x_k} \right|_{(i,j=1,2,3)} = g \cdot h \cdot k \cdot a_x^{g-1} b_x^{h-1} c_x^{k-1} [abc].$$

(c) The Hessian

$$\left| \frac{\partial^2 f}{\partial x_i \partial x_j} \right|, (i, j = 1, 2, 3)$$

of a ternary form  $f$  is of degree 3 in the coefficients. Polarizing, we introduce

three functions

$$f' = a_x^2, f'' = b_x^2, f''' = c_x^2 \quad \text{where } a_x^2, b_x^2, c_x^2 \text{ are equivalent symbols.}$$

We have

$$\frac{1}{6} \sum_{(a,b,c)} \begin{vmatrix} f'_{11} & f'_{12} & f'_{13} \\ f''_{21} & f''_{22} & f''_{23} \\ f'''_{31} & f'''_{32} & f'''_{33} \end{vmatrix} = \frac{1}{6} \{n(n-1)\}^3 \cdot (a_x b_x c_x)^{n-2} \cdot [abc]^2,$$

where the summation is taken over all permutations of  $a, b, c$  and  $f'_{ij} = \frac{\partial^2 f'}{\partial x_i \partial x_j}$ .

### 3. The binary quadratic form

We have seen that the discriminant  $\delta = a_{11}a_{22} - a_{12}^2$  is an invariant of the binary quadratic form  $f = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$ .

We will prove that it is the only invariant; i.e. any invariant of  $f$  can be expressed as a polynomial in  $\delta$ .

Let  $J_h$  be an invariant of degree  $h$  of  $f$ . Its symbolic expression  $J(a, b, c, \dots)$  is an invariant depending on  $h$  contravariant vectors. It follows, then, that  $J(a, b, c, \dots)$  is a sum of products of bracket factors. In a single term each letter  $a, b, c, \dots$  will occur twice. Hence one may easily convince oneself that the term is a product of closed chains like  $K = [ab][bc][cd] \dots [fa]$ . If the length of a chain  $K$  is odd, then  $K = 0$ , for, evidently  $[fa] = -[af]$  and reversing the order of the factors yields  $K = -[af] \dots [dc][cb][ba]$ , if the length of the chain is odd. All the letters  $a, b, c, \dots$  being equivalent symbols and hence permutable, this is identical with  $-K$ , and hence  $K = -K$ ,  $K = 0$ .

Suppose the length of the chain is even. If the chain consists merely of  $[ab][ba]$  then, since  $[ab] = -[ba]$ , we have  $-[ab]^2$ , which is equal to  $-2\delta$ . If the length of the chain is even and exceeds two, by use of the identity

$$[ab][cd] + [ca][bd] + [bc][ad] = 0$$

we write

$$\begin{aligned} K &= -[ca][bd][bc][de] \cdots - [bc]^2[ad][de] \cdots, \\ &= -[ac][cb][bd][de] \cdots - [bc]^2[ad][de] \cdots, \end{aligned}$$

or,

$$K = -K - \delta \times (\text{chain } K_1 \text{ of 2 links less than the original chain}).$$

Then  $2K = -\delta K_1$  and \_\_\_\_\_

hence  $K$  is by reduction a power of the discriminant  $\delta$ . Therefore

$\mathcal{J}(a, b, c, \dots)$  is a polynomial in  $\delta$ . In an analogous fashion one may show that every concomitant (i.e. invariant or covariant) of the binary quadratic form  $f$  may be expressed as a polynomial in  $\delta$  and  $f$ .

#### 4. Irrational methods

We may write  $f = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = (x - \alpha)(x - \alpha')$ .

Equated to zero,  $f$  is represented by two points on the projective line, and the problem of invariants is to find all projectively invariant relations holding for the two points. Obviously their coincidence is the only such relation. This is expressed by  $(\alpha - \alpha')^2 = 0$ .

For the consideration of covariants, one wishes to ascertain the projectively invariant relations existing between two fixed points and a variable point on a projective line. Evidently the only such relation is that the variable point coincides with either of the two fixed points; i.e.,  $(x - \alpha)(x - \alpha') = 0$ . This suggests that the form itself is the only covariant.

We show now that every invariant of an  $n$ -ary quadratic form is a power of the discriminant.

Let

$$f = \sum_{i,k=1}^n a_{ik} x_i x_k$$

be a form for which the numerically given coefficients  $a_{ik}$  have a non-vanishing determinant,  $|a_{ik}| \neq 0$ . By means of a linear transformation

$$\xi_i = \sum_{k=1}^n \alpha_i^k x_k \quad \text{of determinant } \Delta, \quad f \rightarrow \xi_1^2 + \dots + \xi_m^2. \quad \text{Denoting by}$$

$J(a)$  an arbitrary invariant of weight  $r$ , depending on the coefficients  $a_{ik}$  of  $f$ , we have

$$(4.1) \quad J(a) = \Delta^r \cdot J(\delta_{ik}),$$

where  $\delta_{ik}$  is the Kronecker delta. Since  $D = |a_{ik}| = \Delta^2$ ,

$$(4.2) \quad J^2(a) = \text{const.} \times D^r.$$

This relation, established under the assumption that  $|a_{ik}| \neq 0$  is, by the principle of the irrelevance of algebraic inequalities, valid as a formal identity in the variables  $a_{ik}$ .

To complete the proof we must show that the symmetric determinant  $D$  is an irreducible function of its  $\frac{1}{2}n(n+1)$  arguments.

Consider the symmetric determinant

$$A_m = \begin{vmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mm} \end{vmatrix},$$

and make the induction hypothesis that the symmetric determinant  $A_{n-1}$  of order  $n-1$ , is irreducible (the induction being "anchored" by the determinant of a single element). Regarding  $A_n$  as a function of  $a_{nn}$ , we write

$$(4.3) \quad A_m = A_{m-1} a_{mm} + A_m(a_{mm}=c),$$

where the notation  $A_{m-1}(a_{mm}=c)$  denotes the determinant obtained from  $A_m$  by

putting  $a_{nn} = 0$ . If, now, we suppose  $A_n$  resolvable into two polynomial factors,

$$A_n = (Ba_{nn} + B') \cdot C,$$

when  $B, C$  do not contain  $a_{nn}$ . Upon identifying with (4.3) we have  $A_{n-1} = B \cdot C$ ,

$A_n(a_{nn}=0) = B'C$ . By the inductive hypothesis either  $C = 1$  or  $B = 1$ . The case

$C = 1$  does not represent a real decomposition of  $A_n$ . If  $B = 1$ , then  $C = A_{n-1}$

and we have  $A_n(a_{nn}=0) = B' \cdot A_{n-1}$ .

That this last relation, showing  $A_n(a_{nn}=0)$  to be divisible by  $A_{n-1}$  is not, in general, possible, is seen by the following specialization ( $a, b$  independent variables)

$$A_n(a_{nn}=0) = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & a & b \\ 0 & 0 & 0 & \dots & 0 & b & 0 \end{vmatrix} = -b^2,$$

$$A_{n-1} = a.$$

We observe that the method by which (4.1) is obtained is irrational, involving the adjunction to the given number field of certain square roots. The relation of invariance  $J(\alpha) = \Delta^2 \cdot J(\alpha')$ , first postulated in the rational field only, holds as a formal identity in the transformation coefficients in the extended number field. This justifies (4.1). In (4.2) the irrationalities have again been eliminated.

Another example of the use of irrational methods is given by the following treatment of the binary cubic, which, in non-homogeneous form, we write as  $x^3 - ax^2 + bx - c = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$ . Our operations are upon the roots  $\alpha_1, \alpha_2, \alpha_3$  rather than upon the coefficients. We

wish to determine all the invariants and covariants of the cubic.

Consider an invariant  $j(\alpha_1, \alpha_2, \alpha_3)$ , a symmetric polynomial in  $\alpha_1, \alpha_2, \alpha_3$ . Specializing, we suppose  $\alpha_1 = 0, \alpha_2 = 1$ ; then  $j(\alpha_1, \alpha_2, \alpha_3) = 0$  is a projectively invariant relation of the three points  $0, 1, \alpha_3$  of a line. But the only such relation is one that expresses the coincidence of  $\alpha_3$  with 0 or 1, and hence

$$j(\alpha_3) = c \cdot \alpha_3^f (\alpha_3 - 1)^g,$$

where we write  $j(\alpha_3) = j(0, 1, \alpha_3)$ .

We make, now, a transformation

$$x' = (\alpha_2 - \alpha_1)x + \alpha_1$$

which transforms the point 0 to the point  $\alpha_1$  and the point 1 to the point  $\alpha_2$ , with  $\alpha_1 - \alpha_2 \neq 0$ . Then,

$$(4.4) \quad j(\alpha_1, \alpha_2, \alpha_3) = c \cdot (\alpha_3 - \alpha_1)^f (\alpha_3 - \alpha_2)^g (\alpha_2 - \alpha_1)^h.$$

The condition  $\alpha_1 - \alpha_2 \neq 0$  is an algebraic inequality, and hence immaterial to the validity of (4.4). By symmetry, it is clear that  $f = g = h =$  an even number (since  $j$  is invariant). Now  $27\Delta^2 = D$  where

$$\Delta = (\alpha_1 - \alpha_2)(\alpha_2 - \alpha_3)(\alpha_3 - \alpha_1), \quad \text{and } D = \text{the discriminant of the}$$

cubic. Hence  $j = \text{const.} \times D^{\frac{n}{2}}$ .

In considering the covariants of the cubic we first regard absolute covariants that are rational functions of  $x$ . Evidently the only projective invariant of four points  $x, \alpha_1, \alpha_2, \alpha_3$  of a line is their cross ratio:

$$\lambda = \frac{(x - \alpha_1)(\alpha_3 - \alpha_2)}{(x - \alpha_2)(\alpha_3 - \alpha_1)} = \frac{\lambda_1}{\lambda_2}.$$

Though this is indeed invariant, the order of the points is essential, as the cross ratio may assume six values, depending upon the six permutations of the

points  $\alpha_1, \alpha_2, \alpha_3$ . These six values are  $\lambda, \frac{1}{\lambda}, 1 - \lambda, \frac{1}{1 - \lambda},$

$\frac{\lambda}{\lambda - 1}, \frac{\lambda - 1}{\lambda}$ . The six substitutions transforming  $\lambda$  into these values form a

group. The functions we seek are rational functions  $\mathcal{R}(\lambda)$  with numerical coefficients (independent of  $a, b, c$ ) invariant under these six substitutions.

We form a certain primitive such function of degree 6 (all others are functions of this one) by taking the function  $\lambda - \gamma$ , substituting the six possible values of  $\lambda$  and taking their product. We obtain a function that vanishes at  $\gamma$  and the six equivalent points. If  $\gamma$  is a fixed point, the number of equivalent points may be only 2 or only 3.

Thus for the fixed point  $\frac{1}{2}$  we have  $\frac{1}{2}, 2, -1$  as equivalent points, for the fixed point  $\infty$ , equivalent points are  $\infty, 0, 1$  while the fixed points of  $\lambda' = \frac{1}{1-\lambda}$ , namely the two conjugate complex cube roots  $\epsilon, \bar{\epsilon}$  of  $-1$ , are equivalent points.

Using the homogeneous  $\frac{\lambda_1}{\lambda_2}$  in place of  $\lambda$  we have for the three cases

$$\begin{aligned} \varphi &= \lambda_1(\lambda_1 - \lambda_2)\lambda_2 ; & (0, 1, \infty) , \\ \tau &= (2\lambda_1 - \lambda_2)(\lambda_1 + \lambda_2)(-\lambda_1 + 2\lambda_2) ; & (-1, 2, \frac{1}{2}) , \\ H &= \lambda_1^2 - \lambda_1\lambda_2 + \lambda_2^2 ; & (\epsilon, \bar{\epsilon}) . \end{aligned}$$

consider  $z(\lambda) = \varphi^2/H^3$ . Then the only  $\mathcal{R}(\lambda)$  invariant with respect to the six substitutions are the rational functions of  $z$ . This may be established by function-theoretic considerations (the reader is referred to Felix Klein's Icosaeder for a treatment of similar questions), as well as algebraically.

In place of  $z = \varphi^2/H^3$  we might also use  $\tau^2/H^3$ . Hence it is obvious that  $\varphi^2, \tau^2, H^3$  are connected by a linear relation with constant coefficients. We verify this and compute the coefficient explicitly by developing

$$\begin{aligned} H^3 &= \lambda_1^6 - 3\lambda_1^5\lambda_2 + 6\lambda_1^4\lambda_2^2 + \dots , \\ \tau^2 &= 4\lambda_1^6 - 12\lambda_1^5\lambda_2 - 3\lambda_1^4\lambda_2^2 + \dots , \\ \varphi^2 &= \lambda_1^4\lambda_2^2 + \dots , \end{aligned}$$

from which may be deduced the "syzygy"

$$4H^3 = \tau^2 + 27\varphi^2 .$$

We may readily write  $\mathcal{Q}, H, T$  in terms of the roots  $\alpha_1, \alpha_2, \alpha_3$  as well as the coefficients  $a, b, c$ . Thus,

$$\mathcal{Q} = \lambda_1 (\lambda_1 - \lambda_2) \lambda_2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)(\alpha_1 - \alpha_2)(\alpha_3 - \alpha_2)(\alpha_3 - \alpha_1),$$

or 
$$\mathcal{Q} = -f \cdot \Delta;$$

$$H = (a^2 - 3b)x_1^2 - (ab - 9c)x_1x_2 + (b^2 - 3ac)x_2^2$$

$$= -\frac{1}{4}(\text{Hessian of } f);$$

and  $T$  is the Jacobian of the form  $f$  and  $H$ .

$$T = \begin{vmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \\ \frac{\partial H}{\partial x_1} & \frac{\partial H}{\partial x_2} \end{vmatrix},$$

or

$$T = (2a^3 - 9ab + 27c)x_1^3 - (a^2b + 27ac - 12b^2)x_1^2x_2 + 3(ba^2c - ab^2 - 9bc)x_1x_2^2 - (9abc - 2b^3 - 27c^2)x_2^3$$

Denote by  $K(x_1, x_2; a_0, a_1, a_2, a_3)$  a covariant (relative) of the cubic  $f = a_0x_1^3 + a_1x_1^2x_2 + a_2x_1x_2^2 + a_3x_2^3$ , homogeneous in the coefficients  $a_i$ . We show that  $K$  may be expressed as a polynomial in  $D, f, T, H$ , with numerical coefficients -- indeed, as a polynomial of a certain type.

Introducing the roots  $\alpha_1, \alpha_2, \alpha_3$  we write

$$\lambda_1 = (x_1 - \alpha_1x_2)(\alpha_3 - \alpha_2),$$

$$\lambda_2 = (x_1 - \alpha_2x_2)(\alpha_3 - \alpha_1),$$

and  $K(x_1, x_2; a_0, a_1, a_2, a_3) = P(\lambda_1, \lambda_2)$ , a polynomial homogeneous in  $\lambda_1, \lambda_2$ . Now, roots of the equation  $P(\lambda_1, \lambda_2) = 0$  are independent of the coefficients  $a_i$ , and hence,

$$(4.5) \quad P(\lambda_1, \lambda_2) = C(a_0, \dots, a_3) \cdot \prod_{\rho} (\lambda_1 - \rho \lambda_2),$$

where  $\rho$  are numerical constants. (The factor  $\lambda_1 - \rho \lambda_2$  is to be inter-

preted as  $\lambda_2$  if  $\rho = \infty$ .)

In general, together with one root  $\rho$  there will be the six equivalent roots arising from the six substitutions on  $\lambda = \lambda_1 / \lambda_2$ . But for special cases these six numbers are, as we have seen, not distinct. Thus we may have (1),  $\rho = 0, 1, \infty$ ; (2),  $\rho = \frac{1}{2}, -1, 2$ ; (3)  $\rho = 1, \bar{i}$  (the two complex conjugate cube roots of -1).

The first case gives rise to a factor  $f$  on the right side of (4.5); the second case, to a factor  $T$ , and the third to a factor  $\cdot$ . In all other cases the six values of  $\rho$  are distinct. The corresponding value of

$$T^2 / H^3, \quad [(2\rho-1)(\rho+1)(-\rho+2)]^2 / (\rho^2 - \rho + 1)^3,$$

may be  $\kappa$ . ( $\kappa$  is  $\neq 4, 0, \infty$ , and <sup>is</sup> a numerical constant since  $\rho$  is numerical.)

The product of the six equivalent factors  $\prod (\lambda_1 - \rho \lambda_2)$  equals  $T^2 - \kappa H^3$ ,

but for a factor independent of the variables  $\lambda_1, \lambda_2$ . Hence we obtain

$$(4.6) \quad P(\lambda_1, \lambda_2) = C \cdot f^\beta T^\gamma H^\delta \cdot \prod_{\kappa} (T^2 - \kappa H^3),$$

where  $\prod_{\kappa} (T^2 - \kappa H^3)$  may be written as  $f_{\mu}(T^2, H^3)$ , a homogeneous polynomial in  $T^2, H^3$  with numerical coefficients. In (4.6),  $C$  is independent of  $\lambda_1, \lambda_2$  or  $x_1, x_2$ . It is clear that the exponent  $\gamma$  in (4.6) may be "reduced" to 0 or 1, while  $\delta$  may be restricted to the values 0, 1, 2.

Now since  $C$  is clearly an invariant (and rational) it follows, from an argument similar to that used on page 119, that  $C = c \cdot D^\alpha$  where  $c$  is a numerical constant, and  $\alpha = 0, \pm 1, \pm 2, \dots$ .  $D(a)$  is surely an irreducible function of  $a_i$ , according to the general principle that the prime factors of an invariant are invariant themselves. For let  $J_1(a), J_2(a)$  be prime factors of the invariant  $J(a)$  of weight  $g$ :  $J(a) = J_1(a) \cdot J_2(a)$ . After the transformation,  $J_1(a') \cdot J_2(a') = \Delta^g \cdot J_1(a) \cdot J_2(a)$ , where  $\Delta$  is the determinant of the transformation. It follows then from the irreducibility of

the factors on the right-hand side that  $J_1(a') = \Delta^9 J_1(a)$  (for a transformation with unit determinant shows  $J_1(a') = \Delta^9 J_1(a) \cdot J_2(a)$  to be impossible). In our case, however, the only invariants are the integral powers of  $D$ .

$\alpha$  cannot be  $< 0$ . If this were the case all coefficients of

$$f^\beta T^\gamma H^\delta \cdot \prod_{\kappa} (\pi^2 - \kappa H^3)$$

(cross indexed as a polynomial in  $x_1, x_2$ ) would be divisible by the prime factor  $D$ . By a fundamental theorem of Gauss this may happen only if one of the factors

$$f, T, H, \pi^2 - \kappa H^3$$

is divisible by  $D$ . (Here we are operating ~~within~~ within the ring of polynomials in  $a_0, a_1, a_2, a_3$ ). The first three possibilities are to be discarded since  $f, T, H$  are of lower degree in the  $a$ 's than  $D$  itself.  $\pi^2 - 4H^3$  is divisible. Hence if  $\pi^2 - \kappa H^3$  were divisible, so would be the difference  $(\kappa - 4)H^3$  and hence  $H$ , (Gauss's theorem again!) which already has been refuted. Hence we have arrived at the result

$$(4.7) \quad P(\lambda_1, \lambda_2) = c \cdot D^\alpha f^\beta T^\gamma H^\delta \cdot f_\mu(\pi^2, H^3),$$

where  $\alpha$  may have the values  $0, 1, 2, \dots$ ;  $\gamma$ , the values  $0, 1$ ; and  $\delta$ , the values  $0, 1, 2$ . If  $\alpha \geq 1$  and  $\beta \geq 2$ , then, by  $D f^2 = 4H^3 - \pi^2$ , we may absorb the part  $D f^2$  of the factor  $D^\alpha f^\beta$  into  $f_\mu(\pi^2, H^3)$ . Hence we may add the (somewhat artificial) normalizing condition:  $\alpha = 0$  or  $\beta = 0, 1$ .

It might appear somewhat surprising that  $P(\lambda_1, \lambda_2)$  can be exhibited in the form (4.7), instead of merely as "some" polynomial in  $f, D, T, H$ . That this form is, however, a consequence of the isobaric property of all the terms of a covariant will appear later.

We give the following table of degrees and weights:

	Degree in $x_1, x_2$	Degree in coefficients	Weight
D	0	4	6
<del>f</del>	3	1	0
H	2	2	2
T	3	3	3
$4H^2 + Df^2 = T^2$	6	6	6

An arbitrary polynomial in D, f, H, T is a linear combination of  $D^\alpha f^\beta H^\gamma T^\delta$ , with numerical coefficients. The degree of this term in  $x_1, x_2$  is, by the above table,  $3\beta + 2\gamma + 3\delta$ ; its degree in the coefficients  $a_i$  is  $4\alpha + \beta + 2\gamma + 3\delta$  and its weight equals  $6\alpha + 2\gamma + 3\delta$ . Since the invariant polynomial is isobaric it follows that

$$\begin{aligned} 3\beta + 2\gamma + 3\delta &= 3\beta' + 2\gamma' + 3\delta' \\ 4\alpha + \beta + 2\gamma + 3\delta &= 4\alpha' + \beta' + 2\gamma' + 3\delta' \\ 6\alpha + 2\gamma + 3\delta &= 6\alpha' + 2\gamma' + 3\delta' \end{aligned}$$

where  $\alpha', \beta', \gamma', \delta'$  are the corresponding exponents occurring in any other term of the invariant. (We checked this up in our table for the terms of the syzygy!) From these equalities it follows readily upon taking into account the syzygy that the polynomial can be brought into the form (4.6).

In closing this section, we remark that the method employed here for the binary cubic (based upon the projective relation of two sets of four points of a line with equal cross ratios) is not extensible to binary forms of higher order.

5. Some simple general consequences of the symbolic method and the Capelli identity

(a) Gram's Theorem.

Let  $f, g, \dots$  be a set of forms and let

$$(5.1) \quad F_1(a, b, \dots) = 0, \quad F_2(a, b, \dots) = 0, \quad \dots, \quad F_R(a, b, \dots) = 0,$$

be a system of relations existing between the coefficients (a) of  $f$ , (b) of  $g$ ,

$\dots$ . Projectively invariant properties of the forms  $f, g, \dots$  are given by

the system of equations (5.1) if we suppose the collection is unaltered under a

transformation  $x' \rightarrow x$  of the forms  $f, g, \dots$ . In other words if, when  $x' \rightarrow x$ ,

$F_i \rightarrow F'_i$ , then (5.1) is equivalent to

$$F'_1 = 0, \quad F'_2 = 0, \quad \dots, \quad F'_R = 0.$$

In case the system (5.1) consists merely of one form  $F_1$ , then the hypothesis amounts to assuming that  $F_1$  is an invariant.

Suppose, now, that  $J(a, b, \dots; x, y, \dots)$  is a covariant of the forms  $f, g, \dots$ . We allow  $J$  to depend on several vectors  $x, y, \dots$  which transform cogrediently. If it is required that  $J \equiv 0$  (identically in the vectors  $x, y, \dots$ ) then all the coefficients in  $J$  vanish, and the set of equations so obtained is a projectively invariant system of the kind (5.1). It might be conjectured, therefore, that any system (5.1) signifies that one or more covariants of the forms  $f, g, \dots$ , vanish identically. This conjecture was first proved by Gram (Math. Ann. 7 (1874)), who showed that any invariant equation system (5.1) can always be represented by the identical vanishing of a covariant.

To prove this theorem we apply the symbolic method. Suppose

$F_1(a, b, c, \dots)$  is of degree  $\mu$  in the coefficients (a) of  $f$ , degree  $\nu$  in the coefficients (b) of  $g, \dots$ . Polarizing, and replacing the forms  $f, g, \dots$  by

powers of linear forms  $(a'_x, \dots, a_x^{(\mu)})$  for  $f$ ;  $(b'_x, \dots, b_x^{(\nu)})$  for  $g; \dots$ )

$F_1(a, b, \dots)$  becomes a polynomial  $\Phi_1 = \Phi_1(a'_x, \dots, a_x^{(\mu)}; b'_x, \dots, b_x^{(\nu)}; \dots)$

dependent on a number of groups of equivalent contravariant vectors

$a', \dots, a^{(\mu)}$ ;  $b', \dots, b^{(\nu)}$ ; ... . Supposing the vectors in an  $n$ -dimensional vector space, with unit vectors  $e^{(1)}, e^{(2)}, \dots, e^{(n)}$ , we write the components  $a_i$  as  $(ae^{(i)})$ .

Apply now the linear transformation carrying  $e^{(1)}, e^{(2)}, \dots, e^{(n)}$  into  $n$  arbitrarily given but linearly independent vectors  $\xi, \eta, \dots, \zeta$ :

$a_1 \rightarrow (a\xi), a_2 \rightarrow (a\eta), \dots, a_n \rightarrow (a\zeta)$ ; ... : Then

$\Phi_1(a'_1, \dots, a'_n; \dots; a^{(\mu)}_1, \dots, a^{(\mu)}_n; b'_1, \dots, b^{(\nu)}_n)$  transforms into the absolute covariant

$$\Phi_1((a'\xi), \dots, (a'\zeta), \dots).$$

If, now, we reverse the whole process,  $\Phi_1$  changes into an absolute covariant

$$K_1(a, b, \dots; \xi, \eta, \dots, \zeta),$$

depending on  $n$  covariant vectors  $\xi, \eta, \dots, \zeta$ , in addition to the forms  $f, g, \dots$ , and  $K_1 = 0$  is the equation arising from  $F_1 = 0$  by the described linear transformation. The linear independence of  $\xi, \eta, \dots, \zeta$  may be omitted as being algebraically irrelevant.

If we allow relative covariants, we may apply Capelli's special identity to reduce the number of independent vector arguments of

$K_1(a, b, \dots; \xi, \eta, \dots, \zeta)$  from  $n$  to  $n-1$ . For, supposing  $K_1$  actually to contain  $\xi$ , we may write

$$\rho K_1 = \sum \mathcal{P} K_1^* + [\xi \eta \dots \zeta] \cdot \Omega K_1$$

where  $\sum \mathcal{P} K_1^*$  arises from  $K_1^*$  by successive polarizations  $\mathcal{P}$  of  $K_1^*$  ( $K_1^*$  being of lower rank (in  $\xi$ ) than  $K_1$ ). If, now,  $K_1 = 0$ , then  $K_1^* = 0$ ,  $\Omega K_1 = 0$

(each of lower rank than  $K_1$ ). We may continue the process until  $K_1$  no longer

contains  $\xi$ . We observe that  $\Omega K_1$  is a relative invariant.

Gram's Theorem is concerned with a set of invariant relations in the

coefficients of certain forms, and the theorem states that such a set means that a number of tensors algebraically dependent on the forms vanish. A more general theorem than Gram's may be stated as follows:

Every representation of the linear group, of algebraic character, breaks up into irreducible parts. The substratum of each irreducible representation is the set of all tensors of given rank and given symmetry properties.

(b) Pascal's Theorem.

As an example of the content of this theorem, suppose one wishes to study the simultaneous invariants of a set of binary cubics. Then Pascal's Theorem asserts that such an investigation can be restricted to studying the simultaneous invariants of four cubics; that is, any simultaneous invariant of more than four binary cubics is expressible by means of invariants depending on only four cubics.

More generally, consider a definite representation  $\Gamma$  of any group of linear transformations:  $\xi = \xi_1, \dots, \xi_\nu$  in the representation space. These variables undergo a definite transformation under the influence of the linear transformations. If, now,  $J(\xi, \xi', \dots)$  is an invariant of  $\xi, \xi', \dots$  then Pascal's Theorem states that  $J$  is expressible by means of invariants depending on  $\nu$  arguments only. Such an expression is obtained by the Aronhold process (polarization with respect to the arguments) and addition.

This is an immediate consequence of the Capelli general identity, for we may write

$$J_{\xi} = \sum P J^*$$

where  $J^*$  is of lower rank than  $J$ . We may repeat this as long as  $J$  contains more than  $\nu$  arguments.

## 6. The adjunction theorem

As a special case, consider the problem of invariants  $J(f, g, \dots)$  depending in an integral, rational manner on the coefficients of forms  $f, g, \dots$ , where the group of transformations is the orthogonal group. By means of the adjunction theorem, this problem, stated within the framework of euclidean space, can be treated in an affine or even a projective setting.

Suppose  $f = \sum a_{i,k} \xi^i \xi^k$ , (the Greek letters representing contravariant vectors) and consider a euclidean vector space (a euclidean space with one point distinguished as "origin"). Attached to this space we have the fundamental form

$$\Phi = (\xi^{(1)})^2 + (\xi^{(2)})^2 + \dots + (\xi^{(m)})^2.$$

In arbitrary affine form, instead of  $\Phi$  we have  $\sum g_{i,k} \xi^i \xi^k$ .

ADJUNCTION THEOREM. Every orthogonal invariant depending on several

forms is an affine invariant depending on the same forms and the fundamental

form  $g = \sum g_{i,k} \xi^i \xi^k$ .

That is, we can find an affine invariant  $J(f, f^*, \dots, g)$  such that specializing  $g_{ik} = \delta_{ik}$  (the Kronecker delta),  $J(f, f^*, \dots, g)$  reduces to the orthogonal invariant  $J(f, f^*, \dots)$ . We may thus treat a problem of a restricted group by imbedding in a wider group, provided we adjoin an "absolute" to the forms considered. This is, in essence, the epoch-making device of Klein in his Erlanger Program.

For the proof of this theorem we consider an orthogonal invariant

$J(f, f^*, \dots)$  of degree  $\mu$  in the coefficients of the form

$$f = \sum a_{i,k} \dots \xi^i \xi^k \dots \xi^l, \text{ of degree } m, \text{ of degree } \nu \text{ in the coefficients of } f^*, \dots$$

Applying the symbolic method, we introduce the equivalent linear forms  $f^{(1)}, \dots, f^{(\mu)}$  (polarizing with respect to  $f$ ), the

equivalent linear forms  $f^{*(1)}, \dots, f^{*(\nu)}$ , (polarizing with respect to  $f^*, \dots$  .  
 This process reduces  $J$  to an invariant depending on  $\mu$  vectors  $a$ ,  $\nu$  vectors  
 $a^*, \dots$  .

$$J(a^{(1)}, \dots, a^{(\mu)}; a_*^{(1)}, \dots, a_*^{(\nu)}; \dots)$$

But we have already seen (p. 48) that if the group is the group of proper orthogonal transformations, then  $J$  is expressible in terms of the bracket factor  $[a, b, \dots, c]$  and the scalar product  $(a, b)$ . Denoting by  $\gamma^{ik}$  the matrix inverse to  $g_{ik}$ ,  $(a, b) = \sum \gamma^{ik} a_i b_k$ . We obtain an affine invariant dependent on the absolute and the coefficients of the form .

The following more general problem arises: Let  $\gamma$  be a certain group of linear transformations, and suppose  $J(f, f^*, \dots)$  is an invariant of forms  $f, f^*, \dots$ , for this group. Can  $J(f, f^*, \dots)$  be expressed in terms of invariants of the full linear group by adjoining to the forms  $f, f^*, \dots$ , certain other forms?

By means of the usual procedure, we obtain  $J(f, f^*, \dots) \rightarrow J(a, a', \dots)$ , dependent on a number of covariant vectors. Consider the group of the conjugredient transformations operating on the  $a$ 's, and suppose  $j_1, j_2, \dots, j_N$  is a complete set of typical invariants of  $\gamma$ , depending covariant vectors.

If these invariants are subjected to an arbitrary linear transformation (not contained in the group  $\gamma$ ), then  $j_1, j_2, \dots, j_N \rightarrow j_1^*, j_2^*, \dots, j_N^*$ . For all such transformations the system  $j_1^*, j_2^*, \dots, j_N^*$  varies within a certain set of forms -- this set is invariant under arbitrary linear transformations.

From the assumption made concerning  $j_1, \dots, j_N$ , it follows that  $J(a, a', \dots)$  is expressible in terms of  $j_1, \dots, j_N \rightarrow j_1^*, j_2^*, \dots, j_N^*$  in which the arguments are replaced by some of our symbolic vector arguments  $a, a', \dots$ . Hence the given  $J$  arises from an affine invariant  $J(f, f^*, \dots, j_1^*, \dots, j_N^*)$ , by specializing  $j_a^*$ .

7. Theory of binary vector invariants and the spin theory of valence bonds. The Clebsch-Gordan expansion theorem.

The Clebsch-Gordan expansion theorem permits the development of a form  $f$ , depending on  $k$  vectors, in terms of powers of bracket products of the vectors, with coefficients that arise by polar processes from forms containing at most  $k-1$  vectors. We are concerned in this section with forms  $f(x, y, \dots)$  depending on binary vectors  $x = (x_1, x_2), y = (y_1, y_2), \dots$ .

We have

$$(7.1) \quad \mathcal{D}_{xy} \{ [xy] \cdot g \} = [xy] \cdot \mathcal{D}_{xy} g, \quad \text{where}$$

$$\mathcal{D}_{xy} = \sum_1^2 x_i \frac{\partial}{\partial y_i}, \quad \text{and, by Capelli's special identity}$$

$$(7.2) \quad \begin{vmatrix} \mathcal{D}_{yy} + 1 & \mathcal{D}_{yx} \\ \mathcal{D}_{xy} & \mathcal{D}_{xx} \end{vmatrix} f = [xy] \cdot \Omega f \quad \text{where} \quad \Omega = \frac{\partial^2}{\partial x_1 \partial y_2} - \frac{\partial^2}{\partial x_2 \partial y_1}$$

Supposing  $f$  to be of degree  $m > 0$  in the vector  $x$ , and of degree  $n$  in the vector  $y$ , we obtain from the above

$$(m+1)m f - \mathcal{D}_{xy} \mathcal{D}_{yx} f = [xy] \cdot \Omega f,$$

which we write in the form

$$f = \mathcal{D}_{xy} f_1 + [xy] g_1,$$

observing that  $f_1$  is of degree  $m-1$  in  $x$ ,  $n+1$  in  $y$ ; while  $g_1$  is of degree  $m-1$  in  $x$  and  $n-1$  in  $y$ . We may apply Capelli's special identity to this result, and continue the application until the vector  $x$  vanishes from the function to which the polar operator  $\mathcal{D}_{xy}$  is applied. We will obtain

$$(7.3) \quad f = \mathcal{D}_{xy}^m \varphi(y) + [xy] g,$$

with  $\varphi(y)$  of degree  $m+n$ .

For, if, now, we suppose this result valid when  $f$  is of degree  $m-1$  in  $x$ ,

$$f_1 = \mathcal{D}_{xy}^{m-1} \varphi(y) + [xy] \cdot g',$$

we obtain, by substituting  $f_1$  for  $f$  above, and making use of (7.1)

$$f = \mathcal{D}_{xy}^m \varphi(y) + [xy] g, \quad \text{with} \quad g = g_1 + \mathcal{D}_{xy} g'$$

which is in the form (7.3). This is the Clebsch-Gordan identity. Both of the summands in the expression are uniquely determined. To determine the first one, we have only to identify the variables  $x$  and  $y$ . Then  $f(x, y) = \frac{(m+n)!}{n!} \varphi(y)$ . Thus, except for a numerical constant,  $\varphi(y) = f(x, y)$ . The second summand is then determined as the difference of  $f(x, y)$  and the first summand.

We consider the group of all linear unimodular transformations in two variables. A homogeneous polynomial

$$f(x) = a_0 x_1^n + a_1 x_1^{n-1} x_2 + \dots + a_n x_2^n$$

of degree  $n$  is a linear combination of  $n+1$  monomials  $x_1^i x_2^k$ , ( $i+k = n$ ). Under the influence of a transformation  $s$ , these monomials undergo a linear transformation; these monomials constitute the substratum of a representation  $\Gamma_n$  in  $n+1$  dimensions.

The totality of forms  $f(x, y)$  of degrees  $n$  and  $m$  in  $x, y$ , respectively, form a linear manifold of  $(m+1)(n+1)$  dimensions: they constitute the substratum of a certain representation  $\Gamma_m \times \Gamma_n$ . [To describe this product, suppose the  $x_i$  undergo a linear transformation of matrix  $A$ ,  $y_k$  a linear transformation of matrix  $B$ . The products  $x_i y_k$  undergo a linear transformation  $A \times B$  where if  $A = \| a_{i,i'} \|$ , ( $i, i' = 1, \dots, n$ ),  $B = \| b_{k,k'} \|$ , ( $k, k' = 1, \dots, m$ ),  $A \times B = C = \| c_{ik,i'k'} \|$ ,  $c_{ik,i'k'} = a_{ii'} a_{kk'}$ .]

Turning to (1), consider  $\varphi(y)$  an arbitrary polynomial of one binary vector, of degree  $m+n$ ; all polynomials  $g$  of degrees  $m-1, n-1$  in  $x, y$ , respectively, form the substratum of the representation  $\Gamma_{m-1} \times \Gamma_{n-1}$ . We obtain, thus,

$$(7.4) \quad \Gamma_m \times \Gamma_n = \Gamma_{m+n} + (\Gamma_{m-1} \times \Gamma_{n-1}),$$

which gives the true significance of the Clebsch-Gordan expression (7.3) [The "+" occurring in (7.4) is in the sense of the decomposition of a vector space into two partial spaces.]

A different proof of (7.4) is suggested by the symbolic method. Representing the forms as powers of linear forms, we have  $\varphi(y) = \alpha_y^{m+n}$ ,

$$\mathcal{D}_{xy}^m \varphi(y) = \frac{(m+n)!}{m!} \alpha_x^m \alpha_y^n.$$

Consider an arbitrary linear

form  $\alpha_x = \alpha_1 x_1 + \alpha_2 x_2$ ; then

$$\alpha_x^m \alpha_y^n = \sum_{i+k=p} \frac{p!}{i! k!} \alpha_1^i \alpha_2^k \cdot \varphi_i(x, y), \quad (p=m+n),$$

the  $\varphi_i(x, y)$  being of degrees  $m, n$  in  $x, y$ , respectively, and reducing to  $x_1^i x_2^k$  when  $x$  and  $y$  are identified.

Identifying  $x$  and  $y$ ,

$$(\alpha_1 x_1 + \alpha_2 x_2)^p = \sum \frac{p!}{i! k!} \alpha_1^i \alpha_2^k x_1^i x_2^k.$$

consider all linear combinations of  $\varphi_i$  with constant coefficients. This manifold is invariant with respect to unimodular linear transformations, for by

making such a transformation  $x'_1 = ax_1 + bx_2$ ,  $x'_2 = cx_1 + dx_2$ , ( $ad - bc = 1$ ) (and transforming the vector  $y$  cogrediently),

$$\alpha_x^m \alpha_y^n = \alpha_{x'}^m \alpha_{y'}^n; \quad \sum \frac{p!}{i! k!} \alpha_1^i \alpha_2^k \varphi_i(x, y) = \sum \frac{p!}{i! k!} \alpha_1'^i \alpha_2'^k \varphi_i(x', y').$$

Writing  $\alpha_1' = \alpha_1 a + \alpha_2 c$ ,  $\alpha_2' = \alpha_1 b + \alpha_2 d$  and comparing coefficients of  $\alpha_1^i \alpha_2^k$  we got  $\varphi_i$  as a linear combination of  $\varphi_k$ .

The substratum of the representation  $\Gamma_{m-1} \times \Gamma_{n-1}$  consists of the polynomials  $g$  of order  $m-1$  in  $x$ ,  $n-1$  in  $y$ . They are not polynomials of the type  $f$ , but the product  $[xy]g$  increases the order of  $g$  by 1 in each vector without affecting the representation.

For the final step in the proof we consider an arbitrary polynomial  $f$  homogeneous, of degree  $m$  in  $x$  and  $n$  in  $y$ , and show that the polynomial may be written in the form

$$(7.5) \quad f(x, y) = \sum_{i=0}^{m+n} \alpha_i \varphi_i(x, y) + [xy] g(x, y).$$

We observe that we have  $(m+1)(n+1)$  linear equations to solve for  $mn + m + n + 1 = (m+1)(n+1)$  unknowns. To show that this system has a solution we show that the corresponding homogeneous equations have but a trivial solution.

Put

$$\sum \alpha_i \varphi_i + [xy] g(x, y) = 0.$$

Identifying  $x$  and  $y$ , we have at once  $\alpha_i = 0$ , and hence

$$[xy]g(x, y) \equiv 0. \quad \text{Since } [xy] \neq 0, g(x, y) \equiv 0.$$

Hence the form (7.5) is possible and, indeed, in a unique way.

The development in terms of powers of the determinant  $[xy]$  obtained inductively from (7.5) is the Clebsch-Gordan development.

Formula (7.4) lends itself to repeated application and we obtain

$$(7.6) \quad \Gamma_m \times \Gamma_n = \Gamma_{m+n} + \Gamma_{m+n-2} + \dots + \Gamma_{|m-n|}.$$

This is the Clebsch-Gordan development in the language of representation theory.

Finally, we indicate how explicit expressions for the coefficients in the development might be obtained.

Writing

$$f(x, y) = \mathcal{D}_{xy}^m \varphi_0(y) + [xy] \mathcal{D}_{xy}^{m-1} \varphi_1(y) + [xy]^2 \mathcal{D}_{xy}^{m-2} \varphi_2(y) + \dots$$

we obtain

$$f(y, y) = \frac{(m+n)!}{m!} \varphi_0(y),$$

by identifying  $x$  and  $y$ .

If we apply the  $\Omega$  operator to the whole equation and then identify  $x$  and  $y$ , we obtain an expression for  $\varphi_1(y)$ ; applying  $\Omega^2$  and identifying  $x, y$ , we obtain  $\varphi_2(y)$  etc.

To carry through the details of this process we make first two remarks:

(1)  $\Omega$  and  $\mathcal{D}_{xy}$  commute.

We have

$$\begin{aligned}\Omega \mathcal{D}_{xy} f &= \left( \frac{\partial^2}{\partial x_1 \partial y_2} - \frac{\partial^2}{\partial x_2 \partial y_1} \right) \sum_{i=1}^2 x_i \frac{\partial f}{\partial y_i} \\ &= \sum_{i=1}^2 x_i \frac{\partial}{\partial y_i} \left( \frac{\partial^2 f}{\partial x_1 \partial y_2} \right) - \sum_{i=1}^2 x_i \frac{\partial}{\partial y_i} \left( \frac{\partial^2 f}{\partial x_2 \partial y_1} \right) \\ &= \sum_{i=1}^2 x_i \frac{\partial}{\partial y_i} \left( \frac{\partial^2 f}{\partial x_1 \partial y_2} - \frac{\partial^2 f}{\partial x_2 \partial y_1} \right) = \mathcal{D}_{xy} \Omega f.\end{aligned}$$

$$(2). \quad \Omega \{ [xy]^n \Psi_n(x, y) \} = n(m+n+1-n) [xy]^{n-1} \Psi_n + [xy]^n \Omega \Psi_n,$$

where  $\Psi_n$  is of degrees  $m-r$  in  $x$ ,  $n-r$  in  $y$ .

For by (7.2) we have

$$m(m+1) f - \mathcal{D}_{xy} \mathcal{D}_{yx} f = [xy] \cdot \Omega f,$$

where  $f$  is of degrees  $m$ ,  $n$  in  $x$ ,  $y$  respectively, ( $m > 0$ ). Substituting

$[xy]^n \cdot \Psi_n(x, y)$  for  $f$  in the above, and using (7.1), yields,

$$m(m+1) [xy]^n \Psi_n - [xy]^n \mathcal{D}_{xy} \mathcal{D}_{yx} \Psi_n = [xy] \cdot \Omega \{ [xy]^n \Psi_n \},$$

while applying (7.2) to  $\Psi_n$  we obtain

$$\{ (m-n)(m-n+1) - \mathcal{D}_{xy} \mathcal{D}_{yx} \} \Psi_n = [xy] \Omega \Psi_n.$$

Eliminating  $\mathcal{D}_{xy} \mathcal{D}_{yx} \Psi_n$  from these two expressions, we obtain the desired result,

$$\Omega \{ [xy]^n \Psi_n \} = n(m+n+1-n) [xy]^{n-1} \Psi_n + [xy]^n \Omega \Psi_n.$$

If we apply this formula for the case  $\Psi_n = \mathcal{D}_{xy}^{m-n} \varphi_n(y)$ , with

$\varphi_n$  of degree  $m+n-2r$ , we have

$$\Omega \{ [xy]^n \mathcal{D}_{xy}^{m-n} \varphi_n(y) \} = n(m+n+1-n) [xy]^{n-1} \mathcal{D}_{xy}^{m-n} \varphi_n(y),$$

$$\text{since } \Omega \Psi_n = \Omega \mathcal{D}_{xy}^{m-n} \varphi_n(y) = \mathcal{D}_{xy}^{m-n} \Omega \varphi_n(y) = 0.$$

By iteration, we readily obtain

$$\Omega^n \{ [xy]^n \mathcal{D}_{xy}^{m-n} \varphi_n(y) \} = \frac{n!(m+n+1-n)!}{(m+n+1-2n)!} \mathcal{D}_{xy}^{m-n} \varphi_n(y),$$

while  $\Omega^{n+1} = 0$ .

Whence, for  $f(x, y)$  we get

$$\left( \Omega^n f(x, y) \right)_{x=y} = \frac{n! (m+m+1-n)!}{(m+m+1-2n)(m-n)!} \cdot \varphi_n(y),$$

and the coefficients of  $[x y]^r$  in the Clebsch-Gordan expansion are explicitly given.

### 8. A basis for binary vector invariants. Rumer's Theorem

In this section we study polynomials  $f(x, y, z, \dots)$  of  $h$  binary vectors  $x, y, z, \dots$ , which are invariant under the group of unimodular linear transformations. We suppose  $f$  is of given degrees  $a, b, c, \dots$  in  $x, y, z, \dots$ , respectively. We denote by  $N(a, b, c, \dots)$  the number of linearly independent invariants of  $f$ .

If we consider  $f(x, y, z, \dots)$  as a polynomial in  $x, y$  with coefficients depending on the other arguments  $z, \dots$ , we get, upon applying the Clebsch-Gordan theorem,

$$f(x, y; z, \dots) = \sum a_i(z, \dots) \varphi_i(x, y) + [xy] \cdot g(x, y; z, \dots).$$

Since  $f$  is an invariant,  $\sum a_i \varphi_i$  and  $[xy]g(x, y; z, \dots)$  are invariants, for under transformation

$$f(x, y; z, \dots) = f(x', y'; z', \dots) = \sum a_i(z', \dots) \varphi_i(x', y') + [x'y'] g(x', y'; z', \dots),$$

and both of the first summands are linear combinations of  $\varphi_i(x, y)$ . Since the decomposition is unique, they are identical and hence are invariant. Hence

$$\sum a_i \varphi_i(x, x) \quad \text{is an invariant depending on } h-1 \text{ vectors } x, z, \dots.$$

They are in one-to-one correspondence with the invariants of the type

$$\sum a_i(z, \dots) \varphi_i(x, y). \quad \text{We have the recursive formula}$$

$$N(a, b, c, \dots) = N(a+b, c, \dots) + N(a-1, b-1, c, \dots),$$

and by repeated application of the Clebsch-Gordan expansion,

$$N(a, b, c, \dots) = \sum_{r=0,1,\dots,\min(a,b)} N(a+b-2r, c, \dots),$$

a recursive expression for the number of linearly independent invariants.

By the First Fundamental Theorem for vector invariants, we learned that all such invariants are expressible in terms of the bracket factor  $[xy]$ . If we call  $[xy]^\alpha [xz]^\beta [yz]^\gamma \dots$  a monomial, we know that all invariants are expressible as a linear combination of monomials, with numerical coefficients.

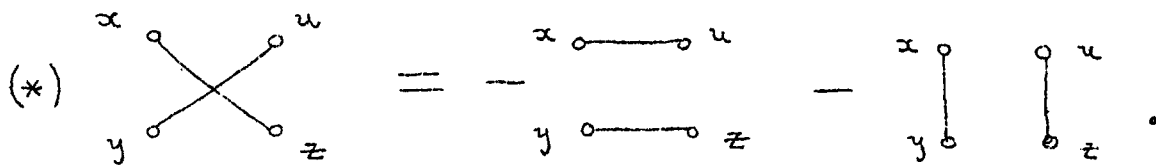
We depict a monomial by means of a diagram in which a point represents a vector and dashes represent bracket factors. We call the dashes "valence dashes". The degree of a monomial in  $x$  is evidently, pictorially, the number of dashes issuing from  $x$ .

Though the monomials form a base for the invariants, linear relations may exist among these basic invariants. The Second Fundamental Theorem assures us that only one fundamental relation exists, namely

$$(8.1) \quad [xu][yz] + [yu][zx] + [zu][xy] = 0.$$

All relations among the basic invariants are, by the Second Fundamental Theorem, algebraic consequences of (8.1). We shall give a new proof of this fact.

Relation (8.1) may be represented diagrammatically as follows:



Thus, the identity (8.1) furnishes a means of "uncrossing" the two dashes appearing on the left-hand side, replacing them by the two diagrams on the right-hand side, each of which is free from crossed dashes.

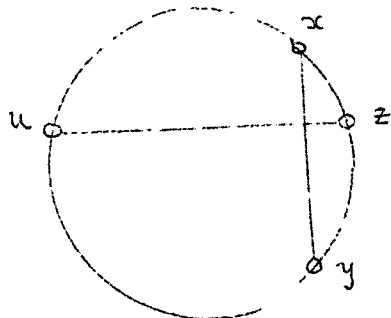
Suppose, now, a certain diagram of arbitrary complexity is given. Is it possible, by means of identity (8.1) to replace each pair of crossed dashes by a combination that is free of crossed dashes, so that the given diagram is changed into a linear combination of diagrams that contain no crossed dashes?

This may indeed be accomplished!

Consider the points of the diagram distributed in a definite order along the circumference of a circle. Either of the circular arcs joining two of these points is called a valence arc. By the length of a valence arc we mean the number of points lying on it, where the two end-points are counted together as 1 point. Thus the minimum length of a valence arc is unity. We suppose the diagram has  $N$  dashes.

Case 1. Suppose a valence arc  $A$  joining  $x$  and  $y$  is of minimum length 1. If, now, the dash  $x \text{---} y$  is omitted, we have a diagram containing  $N-1$  dashes. This diagram may be "uncrossed"; then, by adding the dash  $x \text{---} y$  to the uncrossed diagram again we arrive at the uncrossing of the original diagram by a number of steps of type (\*).

Case 2. Shortest valence arc has length exceeding 1. Since arc  $B$  is

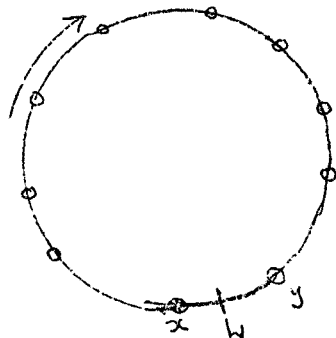


the shortest valence arc, no dash joining points of this arc can exist except the dash  $x \text{---} y$ . If, now,  $xy$  and  $zu$  be uncrossed through (\*), two diagrams arise in which shorter valence arcs  $xz$  and  $yz$ , respectively, occur. By induction with

respect to the length of the shortest valence arc our proof is thus finished.

The number of non-crossing diagrams satisfies the same recurrence formula as the number of linearly independent invariants.

Suppose there are  $d$  dashes joining  $x$  and  $y$ ;  $d \leq \min(a, b)$ . If we let



$x$  and  $y$  coalesce to form a single point  $L$ , the number of dashes issuing from  $L$  is, then,  $a + b - 2d$ . Conversely,  $d$  being given as a number  $\leq \min(a, b)$ , the distribution of the  $a-d$ ,  $b-d$  dashes issuing from  $x$  and  $y$ , respective-

ly, is uniquely determined, as to their endpoints by assigning the first  $a$ -d of the points (in the clockwise direction) to be joined to  $x$ , the other  $b$ -d to be joined to  $y$ . This is of necessity required by the condition of non-crossing.

The result of this section was first announced by the Russian physicist Rumer in 1931; it may be stated thus:

The monomials of non-crossing diagrams of given degrees  $a, b, \dots$  form a base for all invariants of those degrees.

Our proof depends on the first main theorem for binary vector invariants, but it contains a new proof for the second main theorem since the base consisting of linearly-independent members is arrived at by the sole use of relation (\*).

#### 9. The Burnside-Frobenius-Schur Theorem, in elementary disguise

The Clebsch-Gordan theorem states

$$\Gamma_a \times \Gamma_b = \sum_{\nu} \Gamma_{\nu}, \quad (a \geq 0, b \geq 0),$$

where the summation is extended to  $|a-b| \leq \nu \leq a+b$ ,  $\nu$  differing from  $a+b$  by an even integer. We may formulate this by the conditions

$$\begin{aligned} a &\geq 0, & \nu &\leq a + b, \\ b &\geq 0, & a &\leq b + \nu, & a + b + \nu &= \text{even integer}, \\ \nu &\geq 0, & b &\leq \nu + a, \end{aligned}$$

Three non-negative numbers  $a, b, \nu$  satisfying the above relations will be said to form an even triangle.

Repeated application of the Clebsch-Gordan formula gives

$$(9.1) \quad \Gamma_a \times \Gamma_b \times \Gamma_c \times \dots = \sum_{\nu} m_{\nu} \Gamma_{\nu},$$

where we suppose the left-hand member to be composed of  $\nu$  factors. We write  $n_{\nu} = n_{\nu}(a, b, c, \dots)$ . What is the significance of (9.1)? The polynomials  $f(x, y, z, \dots)$  of degrees  $a, b, c, \dots$ , in  $x, y, z, \dots$ , respectively, forming

the substratum of the left-hand side, constitute a linear set of dimensionality  $(a+1)(b+1)(c+1) \dots$ . We may choose a basis for this manifold, consisting of certain sets  $f_0, f_1, \dots, f_v$ , each of a certain valence  $v$ . The functions  $f_i$  span a certain linear, invariant, manifold whose law of transformation is  $\Gamma_v$ ; that is, the  $f_i$  transform like  $f_i \sim (-1)^i \binom{v}{i} x_2^i x_1^{v-i}$ . Introduce, now, a "void vector"  $\chi$ : a binary vector transforming cogrediently, and form

$$\sum f_i \chi_1^i \chi_2^{v-i} = F(x, y, z, \dots, \chi), \text{ where } F \text{ is of degree } v \text{ in } \chi.$$

$F(x, y, z, \dots, \chi)$  is an invariant, for

$$f_i l_1^i l_2^{v-i} \sim (-1)^i \binom{v}{i} (l_1 x_2)^i (l_2 x_1)^{v-i} = (x_1 l_2 - l_1 x_2)^v.$$

If, now, we have several sets  $(f_0, \dots, f_v)$ ,  $(f'_0, \dots, f'_v)$ ,  $(f''_0, \dots, f''_v)$ ,

and we write

$$\begin{aligned} \sum f'_i l_1^i l_2^{v-i} &= F' \\ \sum f''_i l_1^i l_2^{v-i} &= F'' \end{aligned},$$

the invariants  $F, F', F''$  are certainly linearly independent. The converse, however, is not trivial; that is, we wish to show that if  $F, F', F''$  are linearly independent invariants, and  $G, G'$  are linearly independent, then the coefficients are linearly independent. This is, indeed, a consequence of the Burnside-

Frobenius-Schur Theorem in the event that  $\Gamma_m$  is irreducible ( $\Gamma_m$  and  $\bar{\Gamma}_m$  are clearly not equivalent for  $n \neq m$ , since the degrees differ). Without concerning ourselves here about the question of irreducibility, we can establish the desired conclusion directly by showing that the total number  $N(a, b, \dots, v)$  of linearly independent invariants of degree  $v$  equals  $n_v(a, b, c, \dots)$ . We deduce this result by an application of the Clebsch-Gordan theorem. We have

$$\Gamma_b \times \Gamma_c \times \dots = \sum_u m_u(b, c, \dots) \Gamma_u,$$

and

$$\begin{aligned} \Gamma_a \times \Gamma_b \times \Gamma_c \times \dots &= \sum_u m_u(b, c, \dots) (\Gamma_a \times \Gamma_u) \\ &= \sum_u m_u(b, c, \dots) \sum_v \Gamma_v, \end{aligned}$$

where  $a, u, v$  form an even triangle. Thence,

$$\Gamma_a \times \Gamma_b \times \Gamma_c \times \dots = \sum m_v(a, b, c, \dots) \Gamma_v, \quad \text{and}$$

$$m_v(a, b, c, \dots) = \sum_u m_u(b, c, \dots),$$

$u$  forming an even triangle with  $a$  and  $v$ .

The asymmetry of the index and the argument is only apparent, for applying the above for  $v = 0$ ,

$$n_0(a, b, c, \dots) = n_a(b, c, \dots)$$

since the only value of  $u$  that forms an even triangle with  $a$  and  $0$  is  $u = a$ .

( $n_0(a, b, c, \dots)$  is the number of times the identical representation  $\Gamma_0$  occurs).

We write

$$m(v, a | b, c, \dots) = \sum_u m(u | b, c, \dots),$$

(where the summation extends over  $a+v, a+v-2, \dots, |a-v|$ ) which coincides with the recursive formula for  $N$  with equal anchorage. Hence

$$n_v(a, b, c, \dots) = N(a, b, c, \dots, v).$$

We can, then, determine a base for all polynomials. Take an arbitrary integer  $v$  and determine all the  $N(v)$  invariants  $F(x, y, z, \dots, \lambda)$  of degree  $v$  in  $\lambda$ . Write each such invariant in the form  $\sum f_i \lambda_1^{i_1} \lambda_2^{i_2} \dots$ . The  $N(v)$  invariants  $F$  are linearly independent and, by the above, the coefficients  $f_i$  are linearly independent. Thus, we obtain a base for all polynomials of preassigned degree from a base depending on one more. (This is usually proved by means of the Burnside-Frobenius-Schur Theorem.)

10. Sketch of the quantum-mechanical approach. Angular momentum and spin.

The close relation existing between the laws by which atoms combine to build molecules, and the laws governing the formation and mutual relationship of binary invariants was exhibited as a formal analogy by Sylvester. This connection between two branches of science that superficially are so remote has, since Sylvester's time, been established as a true physical theory by means of quantum-mechanics and the introduction of electronic spin. We shall give, in this section, a sketch of the physical basis of Sylvester's analogy.

We are here concerned not with the structure of molecules, but with the forces acting between atoms far apart from each other. The nuclei are considered as fixed centers at distance  $r$  much larger than the atomic dimensions.

The state of a single electron is represented in quantum mechanics by a "wave function"  $\psi(x, y, z)$  depending on its position  $q = (x, y, z)$  in space.  $|\psi(q)|^2$  measures the relative probabilities of the electron's presence in volume elements of equal size at different places  $q$ . The adoption of a scalar  $\psi$  is only preliminary; it might have to be replaced by a wave function with several components, - in the same way as the electro-magnetic field strength, the wave function of a photon, has, according to Maxwell's theory, six components. If the state of a single electron is described by a function  $\psi(q)$  of an argument  $q$ , then the state of a system of  $n$  electrons is described by a function

$\psi(q_1, \dots, q_n)$  depending on  $n$  such arguments, the  $i^{\text{th}}$  one of which  $q_i$  refers to the  $i^{\text{th}}$  electron. The most essential feature is this: that a state is represented by a vector in a complex vector space (of infinitely many dimensions). The phenomenon of interference requires the possibility of linear superposition and of shift of phases, i.e. addition and multiplication by complex numbers, and entities allowing these operations form exactly what the mathematician calls a (complex) vector space. The connection with the observational data is of a statistical na-

ture: e.g. the wave field determines the probability of finding an electron at this or that place. Its mathematical expression rests on the fact that the representing vector space bears a Hermitian metric; each vector has a squared absolute value; in the functional space of the  $\Psi(q)$  this is, for instance, the integral  $\|\Psi\|^2 = \int |\Psi(x, y, z)|^2 dx dy dz$ , extending over the whole real space.

A physical quantity is represented by a linear Hermitian operator, e.g. the coordinate  $x$  by the operator multiplying  $\Psi(x, y, x)$  by  $x$ . The eigenvalues of the operator are the possible values of that physical quantity, the corresponding eigenfunctions representing those states in which these values are taken on with certainty. The energy operator  $H$  determines how the state  $\Psi$  changes in time according to the fundamental dynamical law

$$\frac{\hbar}{i} \frac{d\Psi}{dt} = H \Psi ,$$

( $\hbar = h/2\pi$ ,  $h$  is Planck's constant. We choose the units so that  $\hbar = 1$ .)

The eigen-states of  $H$  are the "energy-", "quantum-" or "stationary states".

If two parts of a physical system are combined into a whole, the representing vector space of the whole is the "cross product" of the two partial spaces; this is indicated by the way in which  $n$  electrons and their representing spaces combine to form an  $n$ -electron-system.

Two states of a system which differ merely by orientation in space and hence change into each other by a rotation (an orthogonal transformation  $s$  of coordinates  $x, y, z$ ) are not different in any physical sense and must hence be represented by two vectors  $\Psi$  arising from each other by a unitary transformation  $U(s)$  corresponding to  $s$ :  $s \rightarrow U(s)$  is a representation of the group of rotations in our multidimensional vector space. In particular, to the infinitesimal rotation about the  $x, y, z$ -axes correspond infinitesimal unitary operators

$\frac{i}{\hbar} (M_x, M_y, M_z)$ . The quantities represented by the Hermitian operators

$M_x, M_y, M_z$ , are called the components of angular momentum and  $M^2 = M_x^2 + M_y^2 + M_z^2$  is the square of angular momentum. (This is in agreement with the classical definition of angular momentum in the case of a scalar wave function.) The fact that it is a scalar and hence invariant with respect to rotations is expressed by the equations

$$H M_x - M_x H = 0, \quad \dots, \quad \dots$$

Because these equations express the fact that  $M_x, M_y, M_z$  commute with  $H$ , they can be interpreted the other way around as asserting that angular momentum is constant in time.

$s \rightarrow U(s)$ , as a representation of a compact (closed) group is of necessity decomposable into irreducible parts of finite dimensions. Rotations appear, when the unit sphere is stereographically projected on a complex  $\xi$ -plane, as a linear fractional transformation of the complex variable  $\xi$ , or, after introducing homogeneous variables  $x_+, x_-$  ( $\xi = x_+/x_-$ ), as the homogeneous linear binary transformations:

$$(10.1) \quad \begin{aligned} x_+ &= ax_+^i + bx_-^i, \\ x_- &= cx_+^i + dx_-^i. \end{aligned}$$

They may be normalized as non-homogeneous transformations by the condition  $ad - bc = 1$ . Each  $\Gamma_\nu$  now denotes a representation of the group of rotations; the substratum of  $\Gamma_\nu$  is the totality of all forms of order  $\nu$  of the two variables  $x_+, x_-$ . Such a form may be written either as

$$\sum \varphi_m x_+^i x_-^k, \quad \left( \begin{array}{l} i+k = \nu \\ -i+k = m \end{array}; m = \nu, \nu-2, \dots \right),$$

or as

$$\sum \varphi(\nu_1, \dots, \nu_\nu) x_{\ell_1} \dots x_{\ell_\nu}$$

with symmetric coefficients  $\varphi(\nu_1, \dots, \nu_\nu)$ . It is a fact that the

$\Gamma_\nu, (\nu = 0, 1, \dots)$  form a complete set of inequivalent, irreducible representa-

tions. In particular a quantity of the kind  $\Gamma_1$  is called a spinor: it possesses two components  $\psi_+, \psi_-$ , depending on the Cartesian coordinate system in real space, which transform according to  $\Gamma_1$  under a change of coordinates, i.e. such that  $\psi_+ x_+ + \psi_- x_-$  stays invariant.

Hence the total representation space can be decomposed into subspaces of dimensionality  $\nu + 1$  where the rotation group induces  $\Gamma_\nu$ . [In the "space" of scalar functions  $\psi(x, y, z)$  of one point  $q = (x, y, z)$  this is accomplished by the spherical harmonics: each  $\Gamma_\nu$  with an even  $\nu : \nu = 2\ell$ , appears exactly once.] As  $\Gamma_\nu$  is irreducible, one may say that all vectors in this subspace  $P_\nu$  describe the same fully determined state of our physical system but with the orientation in space left arbitrary. I refer to this as a simple space of inner quantum number  $j = \nu/2$ . To the infinitesimal rotation of angle  $\epsilon$  about the z-axis there corresponds the substitution

$$\delta x_+ = \frac{i\epsilon}{2} x_+, \quad \delta x_- = -\frac{i\epsilon}{2} x_- ;$$

hence in (10.1)

$$\delta \varphi_m = -\epsilon i \cdot \frac{m}{2} \varphi_m.$$

Consequently  $M_z$  is capable of the values

$$\frac{m}{2} = j, j-1, \dots, -j.$$

in our simple state. If one computes  $M^2 = M_x^2 + M_y^2 + M_z^2$  one finds that it is  $j(j+1)$  times the unit matrix; the square of angular momentum hence has a definite value  $j(j+1)$  in  $P_\nu$ .

With respect to angular momentum  $\mathcal{M} = (M_x, M_y, M_z)$ , our whole representation space breaks up into blocks of the kind corresponding to the several distinct values of  $\nu$ . It commutes with the operators induced by the rotation group. Since the several  $\Gamma_\nu$ 's are irreducible and inequivalent, Schur's fundamental lemma -- which here may be confirmed by

$$(10.2) \quad \left\| \begin{array}{ccc} \Gamma_\nu & 0 & 0 \\ 0 & \Gamma_\nu & 0 \\ 0 & 0 & \Gamma_\nu \end{array} \right\|$$

fundamental lemma -- which here may be confirmed by

elementary computation -- tells us that  $H$  breaks up into corresponding blocks of the type

$$(10.3) \quad \begin{vmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{vmatrix}$$

the number  $E_{ik}$  here indicating these multiples of the  $(v+1)$ -rowed unit matrix. By a suitable change of the unitary coordinate system, not affecting the normal form (10.2), the matrix (10.3) can be brought into diagonal form. The effect will be that  $M_z, M^2, H$  are simultaneously in diagonal form, and that the energy  $H$  takes on a definite value  $E_i$  in each simple state. Each energy level or term  $E_i$  is of necessity of a certain degree  $2j+1$  of degeneracy because it corresponds to a certain simple state of inner quantum number  $j$ .  $m/2$  is called the magnetic quantum number; for it describes the splitting up of this  $(2j+1)$ -fold term into  $2j+1$  individual terms under the influence of a homogeneous magnetic field in the  $z$ -direction. (Such a field destroys the spherical symmetry; Zeeman effect.)

What is the meaning of the Clebsch-Gordan formula

$$(10.4) \quad \Gamma_a \times \Gamma_b = \sum \Gamma_\nu, \quad (\nu = a+b, a+b-2, \dots, |a-b|),$$

in the light of this quantum-mechanical interpretation? If two physical systems in a simple state with inner quantum numbers  $j, j'$  [ $j = a/2, j' = b/2$ ] are combined into a whole, the whole is capable of exactly one definite simple state whose inner quantum number  $J (= \nu/2)$  lies within the limit

$$(**) \quad |j - j'| \leq J \leq j + j'.$$

This is in complete accordance with classical mechanics. If the orientation in space of the two parts is left indeterminate (whereas the states in all other respects are completely determined) then the state of the whole is not completely fixed because of the variable mutual orientation of the two parts. It can be described by the angle between the two vectors of partial angular momentum, and this

angle is given if the resultant, the length  $J$  of the total angular momentum, is given;  $J$  varies within the limits (\*\*). Quantum mechanics differs however from classical mechanics in the following three aspects:

- 1)  $j$  is, in classical mechanics, capable of all non-negative values; in quantum mechanics, only of integral and half-integral values  $0, \frac{1}{2}, 1, \dots$ ;
- 2) the square of angular momentum is in the first case  $j^2$ ; in the second  $j(j+1)$ ;
- 3) in classical mechanics the resultant is capable of all values  $J$  within the limits (\*\*); in quantum mechanics only of such values as differ from the limits by an integer.

The vast empirical material of spectroscopy evinces the fact that the wave function of a single electron is not a scalar but a spinor with two components  $\psi_+(q), \psi_-(q)$ . One may write the index as an argument  $i$ ,  $\psi = \psi(q, i)$  capable of the two values  $+1, -1$ . This spin variable  $i$  can be more precisely described as the  $z$ -component of angular spin momentum, measured in the unit  $\hbar/2$ . One could say that the electron in an abstract way consists of two parts, the positional or orbital part represented by the continuous argument  $q$ , and the spin part represented by the discrete argument  $i$ .

In a first approximation when one neglects the interaction of the several electrons in an atom, or a molecule and the interaction of the orbital and the spin parts of each electron, one has an inner quantum number for both parts and for the whole of each electron (they are called: azimuthal, spin and inner quantum numbers, respectively) and a corresponding energy level. Their composition according to formula (10.4) yields the azimuthal, spin and inner quantum numbers of the whole structure. So each energy level of the whole is associated (exactly, with a value of the total inner quantum numbers, and) approximately with values of all the other quantum numbers just mentioned. Thus, (10.4) is the basic formula for the ordering of term spectra according to these quantum numbers.

Analysis of simple states provides an approximate tool for analyzing the energy states.

For many purposes it is permissible to ignore the positional argument  $q$  entirely. The spin state of an  $n$  electron system is then described by a function  $\mathcal{Q}(i_1, \dots, i_n)$  i.e. a vector in a  $2^n$ -dimensional space.  $|\mathcal{Q}(i_1, \dots, i_n)|^2$  is the probability that the first electron has the spin  $i_1$ , the second has the spin  $i_2$ , and so on. The group of rotations induces the representation

$\Gamma_1 \times \Gamma_1 \times \dots \times \Gamma_1$  ( $n$  factors) which according to the Clebsch-Gordan formula is decomposable into terms  $\Gamma_\nu$ , ( $\nu = n, n-2, \dots$ ). Here we operate in an elementary domain; we need have no recourse to such general and profound theorems as that about the irreducibility of  $\Gamma_\nu$  and the completeness of the set of all  $\Gamma_\nu$ , ( $\nu = 0, 1, \dots$ ). The physicists are accustomed to describe the formula

$$\Gamma_1 \times \Gamma_1 = \Gamma_2 + \Gamma_0$$

in a quasi-intuitive way by saying that the spin of the two electrons may be either parallel or anti-parallel; in the first case they enforce, in the second case they destroy or saturate each other. The formula

$$\Gamma_\nu \times \Gamma_1 = \Gamma_{\nu+1} + \Gamma_{\nu-1}$$

allows of an analogous interpretation.

Long before the dawn of the Heisenberg-Schrödinger form of quantum mechanics, N. G. Lewis propounded the idea, supporting it by an enormous amount of chemical material, that chemical bond consisted of some kind of mutual saturation of two electrons, one in each of the two atoms knitted together. Such a bond is of a very peculiar and rather mysterious nature: the "saturation" takes place between two partners only, - and yet shows no polarity. Now the electronic spin is endowed with exactly the same features, as it is capable of two values only and the combination of spins is according to equation (10.1). If this idea is correct, then  $\nu$  is not only the spin part of angular momentum, but also the valence.

As  $v$  is capable of the values  $n, n-2, \dots$ , the valence of a chemical element should not be uniquely determined, but its parity (whether its valence is odd or even) should be. This parity should alternate regularly if one arranges the elements in the periodical system according to their atomic number  $n$ . We are taught by experience that these are indeed the most fundamental facts about valences. The formula (10.4) also is in agreement with the valence picture: two atoms of valences  $a$  and  $b$  respectively can give rise to a molecule with  $v = a + b - 2d$  free valences, where  $d$  is capable of the values  $0, 1, \dots, \min.(a, b)$ ; for  $d$  valence dashes between the atoms absorb  $d$  valences in each atom and this can continue until (with  $d = \min.(a, b)$ ) the valences of one of the atoms are exhausted.

This relationship between valence bond and electron spin appears here as a plausible hypothesis. But in fact no hypothesis is needed: quantum mechanics gives an unambiguous and complete prescription for computing the forces acting between atoms. We make use of three heuristic preliminaries for this purpose: if our idea is correct then the spin space of an atom in which all valences are free or in which no internal saturations occur, is described by a symmetric function  $\Phi(i, \dots, i_n)$  of the  $n$  spins. We shall proceed as if no other than valence electrons were present. This is certainly not correct, but it will isolate the most essential feature. Hence we proceed under the assumption that the wave function of the individual atom depends symmetrically on the spin arguments of its electrons.

Let us now pass to a systematic and quantitative execution of our scheme! The link to binary vectors and their forms has been provided by the spin phenomenon and the transformation law of spinors.

### 11. Symmetry-spin-degeneracy

If the spin is entirely disregarded, the wave function  $\Psi(q_1, \dots, q_n)$  of  $n$  particles depends only upon the positions of the particles. When the particles are indistinguishable from each other, the energy operator  $H$  depends upon them symmetrically.

Consider a number of fixed nuclei and a number of electrons moving about them. The state of the system is described by a wave function  $\Psi(i_1, \dots, i_n)$  depending on  $n$  variables. Instead of the continuous variable  $q$  of position one may choose a discrete variable  $i$  distinguishing the several energy states. The energy operator is symmetric with respect to these particles. To simplify things we replace the infinite range of the variables  $i$  by a finite one ( $i = 1, \dots, v$ ; Hilbert space of finite dimensionality).  $\Psi(i_1, \dots, i_n)$  is then a tensor of rank  $n$  in our  $v$ -dimensional vector space.

The energy is a linear operator  $H: \mathcal{U} \rightarrow \mathcal{U}$ , so that, explicitly

$$\Psi'(i_1, \dots, i_n) = \sum_k H(i_1, \dots, i_n; k_1, \dots, k_n) \Psi(k_1, \dots, k_n);$$

the symmetry of the operator means that  $H$  is unaltered by a permutation applied cogrediently to  $i_1, \dots, i_n$  |  $k_1, \dots, k_n$ .

As a consequence, the tensor space breaks up into sub-spaces that are invariant with respect to all symmetric linear operators. One may impose on a tensor  $\Psi$  various symmetry conditions; the imposition generating certain "symmetry classes". To make the classification of tensors complete one introduces "classes of highest possible symmetry" which are such that it is impossible to cut out of such a class a part by imposing further symmetry conditions upon its members. The class of symmetric tensors is such a class. Different symmetry classes do not intercombine, as was pointed out by Heisenberg in the case  $n = 2$ ; i.e. if the state of our system lies in one of the symmetry classes, it will always stay in that class under whatever dynamical influences. Wigner gave the

complete group-theoretic treatment for arbitrary  $n$ . Instead of the exact problem we consider the approximation characteristic for the perturbation problem. The energy operator consists of two parts:  $H = H_0 + \epsilon W$ , the first  $H_0$  containing the action of the nuclei upon each electron individually, the second  $\epsilon W$  accounting for their comparatively weak interaction.  $\psi'_i(q)$ , ( $i = 1, 2, \dots, n$ ) may be  $n$  energy states of the single electron with the energy levels  $W_i$ . Neglecting first the interaction  $\epsilon W$ ,

$$\psi(q_1, \dots, q_n) = \psi_1(q_1) \cdot \dots \cdot \psi_n(q_n)$$

then describes an energy state of our system with the corresponding energy level  $W = W_1 + \dots + W_n$ , in which the  $i^{\text{th}}$  electron is in the state  $\psi_i$ . For the total energy  $H$  our  $\psi$  is only approximately a stationary state. From the symmetric property of the Hamiltonian  $H$ , the function  $P\psi$  arising from  $\psi$  by an arbitrary permutation of its arguments is an approximate eigen-function of the same energy level. Now under the influence of interaction, the energy level  $W$  of the system breaks up into neighboring energy levels. We assume that no "accidental" degeneracy is superposing our symmetry degeneration; i.e., all other energy levels differ from  $W$  by a higher order of magnitude than  $\epsilon$ . Then one is allowed to disregard the intercombination of the states  $P\psi$  belonging to  $W$  with all other eigenstates, provided one neglects higher powers of  $\epsilon$ : this is the approximation of the perturbation theory. We thus operate in the  $n!$ -dimensional functional space a basis of which is given by the  $n!$  functions  $P\psi$ .

Since  $P\psi$  is a (near)eigen-function, so is  $\sum_P a_P P\psi$ . Writing  $\underline{a} = \sum_P a_P P$  (the general symmetry operator) we have

$$\chi = \sum_P a_P P\psi = \psi \underline{a}.$$

Supposing all  $n!$  functions  $P\psi$  to be linearly independent, we are operating in an  $n!$  vector space of all symmetry operators  $\underline{a}$ .

Its metric is defined by the Hermitian form

$$\int \dots \int \chi \cdot \bar{\chi} dq_1 dq_2 \dots dq_n = \int_{\text{space}} \chi \cdot \bar{\chi} dq.$$

$$\sum_{P,Q} a_P \bar{a}_Q \int P\psi \cdot Q\bar{\psi} dq \text{ depends only on the combination } P^{-1}Q \text{ (or } Q^{-1}P);$$

for

$$\int P\psi \cdot Q\bar{\psi} dq = \int \psi \cdot P^{-1}Q\bar{\psi} dq = G_{P^{-1}Q},$$

if we put  $G_P = \int \psi \cdot P\bar{\psi} dq$ ; and hence

$$\int \chi \cdot \bar{\chi} dq = \sum_{P,Q} a_P \bar{a}_Q \cdot G_{P^{-1}Q}.$$

Since the symmetry of H means commutativity of H with P, we have

$$\int Q\bar{\psi} H P\psi dq = \int Q\bar{\psi} \cdot P H \psi \cdot dq = \int P^{-1}Q\bar{\psi} \cdot H \psi dq,$$

and hence

$$\int \bar{\chi} \cdot H \chi \cdot dq = \sum a_P \bar{a}_Q H_{P^{-1}Q},$$

with

$$H_P = \int H \psi \cdot P\bar{\psi} dq \text{ (exchange energy).}$$

An eigen vector  $\underline{a}$  corresponding to the eigenvalue  $\lambda$  is to satisfy the equation

$$\sum_P a_P H_{P^{-1}Q} = \lambda \sum_P a_P G_{P^{-1}Q}.$$

$b_Q = \sum_P a_P H_{P^{-1}Q}$  is the resultant symmetric operator  $b = aH$ , when first H and then  $\underline{a}$  is performed. We have

$$\underline{a}(\lambda G - H) = 0.$$

To determine  $\lambda$ , put  $|\lambda G - H| = 0$ . The left-hand side of this equation breaks up into several factors (according to the different ways a function may possess a "highest possible degree of symmetry"). For instance, if we start with the symmetric  $\sum_P P\psi$  instead of  $\psi$  then all the  $P\psi$  coincide, the dimensionality  $n!$  reduces to 1.

The Pauli exclusion principle, however, asserts that only one symmetry class is present in nature: the class of anti-symmetric tensors.

But when the existence of spin is taken into account, while its dynamical influence or reaction between orbits and spins is neglected, the idea is experiencing a revival in modified form. This shall be our standpoint as we proceed to weave together the two threads of spin and symmetry.

We now put in the spin, neglecting its dynamical influence. Since the energy operator of our  $n$  electron system is supposed to affect only the positional arguments  $q_\alpha$  of the wave function, an eigen-function  $\Psi(q_1, \dots, q_n)$  will depend on these variables only. But not only can  $\Psi$  be replaced by an arbitrary linear combination  $\underline{a} \Psi$  of the  $n!$  functions  $P\Psi$ , but one has to multiply it by an arbitrary function  $\Phi(i_1, \dots, i_n)$  of the  $n$  spin variables.

$$\Psi(q_1, i_1; q_2, i_2; \dots, q_n, i_n) = \Psi(q_1, \dots, q_n) \cdot \Phi(i_1, \dots, i_n)$$

is an eigen-function of degree of degeneracy  $2^n$ . Each function arising by a permutation performed on the pairs  $q_1, i_1; q_2, i_2; \dots; q_n, i_n$  is an eigen-function. We get thus a symmetry-spin degeneracy of degree  $n!2^n$ . Applying the Pauli principle, we form from the  $n!2^n$  basic eigen-functions such linear combinations as are anti-symmetric in the  $n$  compound arguments  $q_1, i_1; q_2, i_2; \dots; q_n, i_n$ , which are associated with the  $n$  electrons. Putting

$$\delta_P = \begin{cases} +1, & \text{even permutation} \\ -1, & \text{odd permutation,} \end{cases}$$

we accomplish the desired effect by forming

$$(11.1) \quad \chi(q_1, i_1; q_2, i_2; \dots, q_n, i_n) = \sum_P \delta_P \Psi(q_1, \dots, q_n) \cdot P\Phi(i_1, \dots, i_n).$$

$\Psi$  is here a given function, whereas the spin function  $\Phi$  uniquely determining  $\chi$  is quite arbitrary. The linear manifold of these functions  $\chi$  is of dimension  $2^n$  and the perturbation problem is now to be stated in the  $2^n$ -dimensional vector space  $\mathcal{Q}$  rather than in the  $n!$ -dimensional vector space of the symmetry operators  $\underline{a}$ . This situation was first fully recognized by Slater's "determinant" method. Again we assume in the approximation of the perturbation theory that

$\psi(q_1, \dots, q_n)$  is an approximate eigen-function with the approximate energy level  $W$ .

We emphasize that the following basic assumptions are made:

- (1) We suppose the atoms are far apart from each other, i.e., their mutual distances are large compared to the distances of the electron in one and the same atom.
- (2) We have no "accidental" degeneracy; i.e. all the eigen-functions except (11.1) belong to energy levels far off  $W$ . This is a very decisive restriction. It involves the condition that the state of the individual atom is an  $s$ -state described by a spherically symmetric wave function  $\psi$ , for in all other cases we have the degeneracy of order  $2\lambda+1$  in a single atom corresponding to the azimuthal quantum number  $\lambda$  (the state of the atom being described by a  $\psi$  which depends on direction in space like a spherical harmonic of order  $\lambda$ ).
- (3) We neglect all electrons except valence electrons. (This is definitely wrong; the dynamic influence of the non-valence electrons is not negligible, but in a first survey this simplification helps to bring out the essential features.)

## 12. Spin states and binary invariants. Computations of binding energies.

The metric of the  $2^n$ -dimensional vector space  $\mathcal{Q}$  is defined by the

Hermitian form  $\sum_{(i)} \int \chi \cdot \bar{\chi} dq$ . We obtain

$$\sum_{(i)} \int \chi \cdot \bar{\chi} dq = \sum_{(i)} \int \sum_{P, Q} \delta_P \cdot P \psi(q) \cdot P \varphi(i) \cdot \delta_Q Q \bar{\psi}(q) \cdot Q \bar{\varphi}(i) dq.$$

Attaching the sign  $\delta_P$  to our former  $G_P$ :

$$G_P = \delta_P \int \psi(q) P \bar{\psi}(q) dq,$$

we have the Hermitian form in the variables  $\mathcal{Q}(i)$

$$\sum_{(i)} \sum_{P, Q} G_P \cdot \varphi(i) P^{-1} Q \bar{\varphi}(i) \text{ or } n! \sum_{(i)} \sum_P G_P \varphi(i) P \bar{\varphi}$$

For the form  $\int \bar{\chi} \cdot H \chi dq$  we get the same result where  $G$  is replaced by  $H$ :

$$H_P = \delta_P \int H \psi(q) P \bar{\psi}(q) dq.$$

Our eigen-value problem consists in determining those values of  $\lambda$  for which the equations

$$(*) \quad \sum_{\mathcal{P}} (\lambda G_{\mathcal{P}} - H_{\mathcal{P}}) \mathcal{P} \bar{\Phi}(\lambda) = 0 \quad \text{or} \\ (\lambda \underline{G} - \underline{H}) \bar{\Phi}(\lambda) = 0$$

have a non-trivial solution  $\bar{\Phi}(\lambda)$ ;  $\underline{G}$  and  $\underline{H}$  denote the operators

$$\underline{G} = \sum_{\mathcal{P}} G_{\mathcal{P}} \cdot \mathcal{P}, \quad \underline{H} = \sum_{\mathcal{P}} H_{\mathcal{P}} \cdot \mathcal{P}$$

Suppose  $\Psi_1(q_1, \dots, q_a)$  is an exact eigen-function for the first atom containing  $\underline{a}$  electrons. If all electrons are valence electrons according to our third basic assumption, the complete wave function must be of the form

$\Psi_1(q_1, \dots, q_a) \cdot \Phi_1(i_1, \dots, i_a)$  with a spin factor  $\Phi_1$  symmetric in its  $\underline{a}$  arguments;  $\Psi_1$  must then be, according to the Pauli principle, antisymmetric.

For the whole system we put

$\Psi(q_{1,1}, \dots, q_{a,1}, q_{a+1,1}, \dots, q_{m,1}) = \Psi_1(q_{1,1}, \dots, q_{a,1}) \Psi_2(q_{a+1,1}, \dots, q_{a+b,1}) \dots$   
 $\Psi(q_{1,1}, \dots, q_{n,1})$  is an eigen-function of the energy level  $W_1 + W_2 + \dots = W$ , the energy state of the system, when interaction of the atoms is neglected. The indices attached to the functions  $\Psi$  refer to the several atoms of our system, while the indices of the variables  $q$  refer to the electrons. We are denoting by  $a, b, \dots$  the valences of the atoms.

In the expression (11) we have to impose the restriction on

$\Phi(i_1, \dots, i_a, k_1, \dots, k_b, \dots)$  of being symmetric in the  $\underline{a}$  arguments  $i$ , symmetric in the  $b$  arguments  $k$ , of  $\Psi_1, \Psi_2, \dots$ . (In all other cases  $\chi$  would vanish, in consequence of the anti-symmetric nature of  $\Psi_1, \Psi_2, \dots$ ) The linear manifold of such functions  $\Phi$  is of dimensionality  $(a+1) \cdot (b+1) \dots$ . We now consider the form

$$\Phi(x, y, \dots) = \sum_{i, k, \dots} \bar{\Phi}(i_{1,1}, \dots, i_{a,1}, k_{1,1}, \dots, k_{b,1}, \dots) x_{i_{1,1}} \dots x_{i_{a,1}} y_{k_{1,1}} \dots y_{k_{b,1}} \dots$$

of degree  $a$  in the binary vector  $x = (x_+, x_-)$  associated with the first atom, of degree  $b$  in the binary vector  $y$  associated with the second atom, etc. Our eigenvalue-problem

$$(12.1) \quad (\lambda \underline{G} - \underline{H}) \bar{\Phi} = 0$$

is to be solved in the  $(a+1)(b+1) \dots$  dimensional vector space of all these forms  $\Phi(x, y, \dots)$  of degrees  $a, b, \dots$ , in  $x, y, \dots$ , respectively.

We can determine a base for the polynomials  $\bar{\Phi}$  consisting of a number of sections  $\bar{\Phi}_m$  ( $m = v, v-2, \dots, -v$ ), such that the functions of each section transform among each other under the influence of a transformation  $s$ , according to  $\Gamma_v$ . In a state described by any linear combination of the  $v+1$  functions  $\bar{\Phi}_m$ , the total spin momentum of the system is  $v$ . There occur  $n_v$  such sections  $\{\bar{\Phi}_m\}$  of valence  $v$ . Thus, if we introduce the "void" vector  $(\chi_+, \chi_-)$ , the polynomial

$$(12.2) \quad F(x, y, \dots, \chi) = \sum_{\substack{i+k=v \\ -i+k=m}} \bar{\Phi}_m(x, y, \dots) \chi_+^i \chi_-^k,$$

of degrees  $a$  in  $x$ ,  $b$  in  $y$ ,  $\dots$ ,  $v$  in  $\chi$ . is an invariant. We know, vice versa, that we find a base for all  $\bar{\Phi}_m$  when we set up for each  $v$  a complete set of linearly independent invariants  $F$ :  $F_v^{(\alpha)}$ , ( $\alpha = 1, 2, \dots, n_v$ ), and introduce the coefficients  $\bar{\Phi}_m = \bar{\Phi}_{v,m}^{(\alpha)}$ , by (12.2).  $n_v(a, b, \dots)$  is the number of independent invariants  $F(x, y, \dots, \chi)$  of preassigned degrees  $a, b, \dots, v$ .

Since the symmetry operators  $\underline{G}$ ,  $\underline{H}$  obviously change an invariant into an invariant, we can formulate the eigenvalue problem

$$(12.3) \quad (\lambda \underline{G} - \underline{H}) F = 0$$

in the  $n_v$ -dimensional space of all invariants  $F = F_v$  of degree  $v$  in  $\chi$ . If

$$\underline{G} F_v^{(\alpha)} = \sum_{\beta=1}^{n_v} g_{\alpha\beta}^{(v)} F_v^{(\beta)}, \quad \underline{H} F_v^{(\alpha)} = \sum_{\beta=1}^{n_v} h_{\alpha\beta}^{(v)} F_v^{(\beta)}$$

its secular determinant reads

$$(12.4) \quad \left| \lambda g_{\alpha\beta}^{(v)} - h_{\alpha\beta}^{(v)} \right|_{\alpha, \beta = 1, \dots, n_v} = 0.$$

Because  $\underline{G}$  and  $\underline{H}$  operate only on the indices associated with  $x, y, \dots$  and not with  $\nu$ . (12.3) or

$$\sum_{\beta} (\lambda g_{\alpha\beta}^{(\nu)} - h_{\alpha\beta}^{(\nu)}) F_{\nu}^{(\beta)} = 0$$

implies the same relation for each coefficient  $\phi_m$ :

$$(12.5) \quad (\lambda \underline{G} - \underline{H}) \phi_{\nu m} = 0 \quad \text{or} \quad \sum_{\beta} (\lambda g_{\alpha\beta}^{(\nu)} - h_{\alpha\beta}^{(\nu)}) \phi_{\nu m}^{(\beta)} = 0.$$

In this way we succeed in decomposing our problem (12.1) into its several factors (12.5) according to the possible values of valence  $\nu$  and magnetic quantum number  $m$ . The secular determinant  $\Delta$  (12.4) of the part  $\nu m$  is independent of  $m$ ; hence this factor occurs  $(\nu+1)$ -times in the complete determinant, each energy level of valence  $\nu$  remaining  $(\nu+1)$ -fold degenerated even after interaction has been taken into account. This corresponds to the fact that the configuration may be turned around in space without affecting the physical conditions; states are thus treated as alike when they differ only by orientation in space.

In practice one tries to put  $\underline{G}$  and  $\underline{H}$  in better shape.  $H_P$  has the same value  $W_0$  for all permutations  $P$  that interchange electrons within the several atoms; for these permutations  $G_P = 1$ . If  $P$  interchanges one electron in atom  $x$  with one in atom  $y$ , the exchange integral  $H_P$  has the same value  $H_{xy}$  independent of which pair one selects in  $x$  and  $y$ . Compared to  $W_0$  it is of order of magnitude  $e^{-2r}$ . All other terms arising from exchanges of more than two electrons in different atoms are at least of order  $e^{-4r}$ , and hence may be neglected. With this approximation we have

$$\underline{G} = 1 + \sum G_{xy} t_{xy}, \quad \underline{H} = W_0 + \sum H_{xy} t_{xy}.$$

We write  $W_0 + \lambda$  instead of  $\lambda$ , thus denoting by  $\lambda$  the deviation of the energy level from its undisturbed value  $W_0$ .  $\lambda$  being of order  $e^{-2r}$  we have, with the same approximation

$$\begin{aligned} (\lambda + W_0) \underline{G} - \underline{H} &= W_0 (1 + \sum G_{xy} t_{xy}) + \lambda - (W_0 + \sum H_{xy} t_{xy}) \\ &= \lambda - \sum W_{xy} t_{xy}, \quad \text{with} \\ W_{xy} &= H_{xy} - W_0 \cdot G_{xy} \end{aligned}$$

$t_{xy}$  is by definition

$$(\mathcal{D}_{\eta y} \mathcal{D}_{\xi x})_{\substack{\xi=y \\ \eta=x}},$$

where  $\mathcal{D}$  is our old friend, "polarization." When applied to a form  $\Phi(x, y)$  of degrees  $a$  and  $b$  in  $x$  and  $y$  respectively, we find

$$\begin{aligned} \mathcal{D}_{xy}(\mathcal{D}_{yx}\Phi) &= \sum_{l,k} x_l \frac{\partial}{\partial y_k} (y_i \frac{\partial \Phi}{\partial x_i}) = \sum_{l,k} x_l y_k \frac{\partial^2 \Phi}{\partial y_k \partial x_l} + \sum_i x_i \frac{\partial \Phi}{\partial x_i} \\ &= (\mathcal{D}_{\eta y} \mathcal{D}_{\xi x} \Phi)_{\substack{\xi=y \\ \eta=x}} + a \Phi, \end{aligned}$$

hence

$$t_{xy} = \mathcal{D}_{xy} \mathcal{D}_{yx} - a, \quad (= \mathcal{D}_{yx} \mathcal{D}_{xy} - b).$$

We thus finally arrive at this result: When atoms of given valence  $a$ ;  $b$ , ... and given simple energy states combine into a molecule of  $v$  free valences, the spin state of the molecule is described, independently of its orientation in space, by an invariant  $F(x, y, \dots; \lambda')$  depending on the binary vectors  $x, y, \dots, \lambda'$ , and of degrees  $a, b, \dots, v$ , in  $x, y, \dots, \lambda'$ , respectively.

According to the first main theorem, the "monomials" (power-products of the brackets  $[xy], [xz], [yz], \dots$ ) form a base for all invariants, though they are in general not linearly independent. A state of the molecule described by such a monomial and hence representable by a valence diagram, may be called a pure valence state. As a matter of fact, the quantum states do not coincide with the pure valence states but lie between them. They and their corresponding energies  $\lambda$  are to be computed by solving, in the field of all invariants  $F$  of the degrees mentioned, the secular equation

$$(\lambda - \underline{T}) F = 0$$

where

$$\underline{T} = W_{xy} t_{xy} + \dots$$

A base suitable for carrying out this calculation and consisting of linearly independent invariants is procured by Rumer's theorem.

As an illustration, consider a molecule consisting of two atoms  $x, y$  of valences  $a, b$  ( $b \leq a$ ), joined by  $d$  valence dashes;

$$v = a + b - 2d; \quad d = 0, 1, \dots, b.$$

There is only one monomial of the assigned orders

$$F = [xy]^d [x\prime]^{a-d} [y\prime]^{b-d}$$

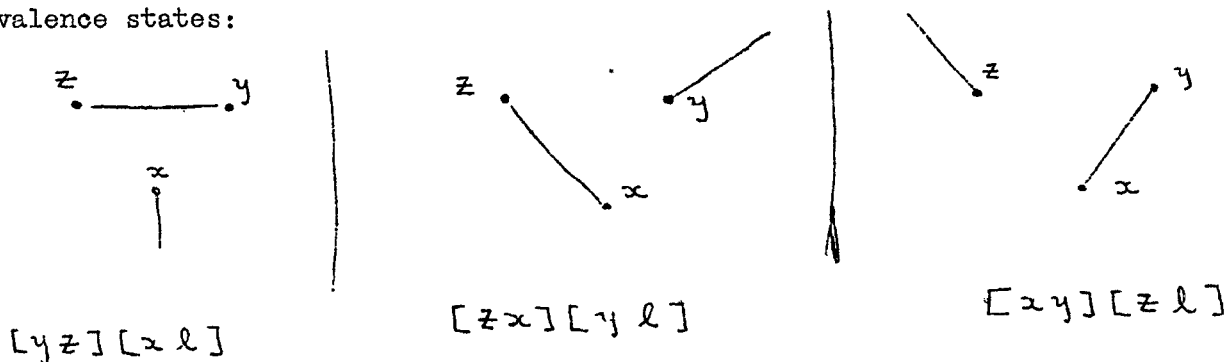
One readily computes

$$t_{xy} F = \{(a-d)(b-d+1) - a\} F, \quad \text{and hence}$$

$$\lambda = \{(a-d)(b-d) - d\} W_{xy}.$$

(This energy state is of multiplicity  $a + b + 1 - 2d$ .) Thus the numerical factor varies from the positive value  $a \cdot b$  to the negative value  $-b$ . If the exchange energy  $W_{xy}$  is positive (computations show this to be true for the hydrogen molecule) the energy  $\lambda$  decreases, the molecule becomes more stable with fastening tie. For  $H_2$  ( $a = b = 1$ ), the two energies in their dependence on the distance  $r$  resemble two reflected exponential curves ( $e^{-2r}$ ) for large values of  $r$ . As both energies become finally positive, the curve corresponding to  $d = 1$  must have a minimum. This accounts for the existence of a stable molecule  $H-H$ .

We remark, finally, that if  $x, y, z$  are three atoms, each of valence 1, and if we consider the states with a free valence  $v = 1$ , we have the three pure valence states:



But it follows from the identity

$$[x\cancel{y}][yz] + [y\cancel{x}][zx] + [z\cancel{y}][xy] = 0,$$

that there exist only two independent states, the energy levels of which are given by

$$\lambda = \pm \sqrt{(W_1^2 + W_2^2 + W_3^2) - (W_2W_3 + W_3W_1 + W_1W_2)}$$

$$[W_1 = W_{yz}, \dots]$$

Indeed, using

$$F_1 = [xz][y\cancel{x}], \quad F_2 = [yz][x\cancel{y}],$$

as a base, and observing that  $t_{xy}$  is the operation of exchanging  $x$  and  $y$ , one finds

$$t_{xy}F_1 = W_{xy} \cdot F_2 = W_{zx}F_1 + W_{yz}(F_1 - F_2) = (W_1 - W_2)F_1 + (W_3 - W_1)F_2,$$

$$t_{xy}F_2 = W_{xy}F_1 = W_{zx}(F_1 - F_2) + W_{yz}F_2 = (W_3 - W_2)F_1 + (W_2 - W_1)F_2.$$

The secular equation for  $\lambda$  is thus

$$\begin{vmatrix} \lambda - (W_1 - W_2) & W_1 - W_3 \\ W_2 - W_3 & \lambda + (W_1 - W_2) \end{vmatrix} = 0$$

or

$$\lambda^2 - \{(W_1^2 + W_2^2 + W_3^2) - (W_2W_3 + W_3W_1 + W_1W_2)\} = 0.$$

Thus the number of different independent states is less than the chemical bond diagrams would indicate, for more than two atomic molecules. This reduction is, strangely enough, in agreement with the facts,

A very difficult part in the application to concrete cases is the computation of the exchange energies; it is an almost insurmountable obstacle unless one is willing to derive their values in a half-empirical way. We know that the combinatorial schemes of valence dashes are far from giving a full account of chemical compounds; they are like a skeleton hardly discernible beneath the living flesh of the chemical facts. So it is in our theory. However, we can by no

means claim that even this theory tells the whole story; the situation is still much more complicated. We pointed out that there are sources for binding forces other than the symmetry-spin-degeneracy here discussed; we neglected the electrons whose spins are mutually saturated within the individual atoms; it is uncertain to what extent our model can be considered an approximation to the ready-made molecule.