CAMILLE JORDAN 1875

Essay on the Geometry of n Dimensions

Translated with commentary by
G. W. Stewart

In 1875 the French mathematician Camille Jordan published a paper on geometry in n dimensions with the aim of generalizing results obtained for two and three dimensions. Although he largely succeeded, his paper is not much cited, principally because his approach of defining planes as the solution space of linear equations was superseded by vector space methods. Yet there is much to be learned from this paper. Here we give a partial translation from the original French along with a commentary.

NOTE TO THE READER

February 29, 2016

This work is a preliminary posting of my partial translations of a paper on geometry in n dimensions by Camille Jordan along with my comments on the papers. A PDF file may be obtained at

`http://www.umiacs.umd.edu/~stewart/Jord75.pdf`

I would appreciate any comments and errata, which will be used to update this document. For the date of the latest changes go the above URL and look at the date on this page. Pete Stewart

stewart@umiacs.umd.edu

Commentary on

Camille Jordan's

Essay on the Geometry of n Dimensions

G. W. Stewart

In 1875 Camille Jordan [13] published a paper on n-dimensional geometry in the *Bulletin de la Société Mathématique (tome 3).*[1] Let Jordan tell us why.

> It is well known that Descarte's merger of analysis and geometry has proved equally fruitful for each of these two disciplines. On the one hand, geometers have learned from their contact with analysis to give their investigations an unprecedented generality. Analysts, for their part, have found a powerful resource in the images of geometry, as much for discovering theorems as for presenting them in a simple, impressive form.
>
> This resource vanishes when one turns to the consideration of functions of more than three variables. Moreover, the theory of these functions is, comparatively speaking, poorly developed. It appears that the time has come to fill this gap by generalizing the results already obtained for the case of three variables. A large number of mathematicians have considered this topic in more or less specialized ways. But I am not aware of any general work on this subject.

It is no exaggeration to call this work a foray by a great mathematician into largely unexplored territory, and Jordan's geometric intuition is something to behold. One of the high points of his work is his unearthing of what today we call the canonical angles between subspaces. Another is his anticipation of Lie groups. He uses many of the ideas of modern vector space theory, such as the standard definition of linear independence and the Steinitz exchange theorem. Moreover, he describes a general orthogonalization procedure that includes the widely used Gram–Schmidt method. Yet the paper seems not to have attracted much attention. Although it is sporadically referenced today for its introduction of canonical angles, it is not cited in the historical surveys of Dorier [9] and Moore [16]. Nor is it mentioned in the biographies the Dictionary of Scientific Biography [8] or the McTutor History of Mathematics [17]. Moreover, it seems to have had little influence on Jordan's successors. So far as I can tell, the modern theory of finite dimensional liner spaces was developed entirely without reference to Jordan's work.

One reason for this is, perhaps, that the paper presents many difficulties for its reader, then and today. First of all, Jordan's planes are defined as the locus of points satisfying a system of (typically inhomogeneous) equations — what we would call a translated subspace or a linear manifold. This creates an uncomfortable inversion between

[1]The introductory section of Jordan's paper appeared word-for-word in 1872 the Comptes Rendus de l'Académie des Sciences, Paris [11], which suggests that the paper was by that time complete.

the dimensions of his planes and their numbering. What Jordan calls a 1-plane, or simply a plane, is defined by single equation and has dimension $n - 1$. On the other hand, our standard two dimensional plane is Jordan's $(n-2)$-plane, whose points satisfy $n-2$ independent linear equations. It takes a great deal of effort to pass back and forth between these two points of view.

In addition, Jordan establishes his results by manipulating the defining equations of his planes, which obscures the physical geometry of the planes themselves. For example, the intersection of two planes corresponds to the union of their equations. Again, whereas in linear algebra subspaces are built up as linear combinations of basis vectors, Jordan regards his planes as being the intersection of 1-planes.[2]

Another difficulty is Jordan's widespread use of elimination to prove his results. This technique, which is related to what today we would call block Gaussian elimination, consists of starting with a set of linear equations and obtaining a smaller set of equations in a reduced number of variables. This is, of course, perfectly legitimate. But Jordan often leaves the details unclear or fails to show that the required elimination can be effected.

Finally, Jordan's exposition is not easy going. At a low level he does not always arrange his materials well, and he is frequently obscure. Typos abound. At a higher level, he often does not provide enough detail to understand the statements of his theorems or their proofs. In my view the paper has an improvisatory flavor — Jordan seems to be developing his subject as he writes. It is reasonable to conjecture that had he rewritten the paper, it would it would have been more accessible

The comments above should not be allowed to obscure the significance of Jordan's results. He his profound understanding of his subject leads him to deep theorems — especially the nature of parallelism and the concept of canonical angles between planes. Later he introduces the Lie group of orhogonal matrices. If his rigor is not up to our standards, all his results, perhaps with minor modifications, remain nonetheless true.

This commentary is written to supplement my translation of Jordan's paper. The translation itself stops at the point where Jordan has introduced the canonical angles between planes.[3] For reasons given above, the reader can expect both the original and the translation to be hard going — hence the commentary. In preparing it I have had to compromise between two extremes.

---

[2] J. Dieudonné, in his introduction to the third volume of Jordan's collected works [7], has this to say. "Up to 1880 the elementary notions about n-dimensional space had not been widely disseminated, and in this work Jordan wishes to show how the classical results from the geometry of two and three dimension may be generalized. He confines himself to real Euclidian geometry and uses only the methods of analytic geometry. Thus he defines linear manifolds by their equations (or alternatively as the intersections of hyperplanes) and not (in contrast with Möbius or Grassmann) as being generated by points or vectors. This [approach] obviously complicates his manipulations."

[3] See Dieudonné's introduction [7] for commentary in modern terms of the material not appearing in this translation.

On the one hand, a blow-by-blow description of Jordan's development in his own terms will not be of much help to the modern reader. On the other hand, to recast and prove his results in the terms of vector spaces would be anachronistic and would trivialize his accomplishments. The compromise adopted here is to stick close to Jordan's proofs, but recast his manipulations in modern terms. For example, we will use matrices freely along the terminology and techniques of modern linear algebra, e.g., null spaces.

Two points need further amplification. First, in this overview we will make free use of matrix inverses. Although Cayley had introduced matrices in the middle of the nineteenth century [4, 5], at the time of Jordan's paper they had not caught on. Nonetheless, mathematicians of the time had ways of getting at the elements of the inverse matrix. Jordan himself gives an example in his treatment of changes of coordinate systems.

Second, many of the matrix reductions, canonical or otherwise, that we now write as transformations of the matrix in question (e.g., similarity transformations), were discovered in the nineteenth century as reductions of linear or bilinear forms by changes of variables [14, p. 804 ff.]. Since the correspondence is rather obvious, I will not hesitate to use the modern approach in this commentary.

A striking aspect of Jordan's essay is the constructive approach that informs his proofs. Behind a typical proof lies an algorithm, however meagerly sketched. One way of understanding what he is doing is to make these algorithms explicit, which I will do from time to time in this commentary.

Jordan's paper consists major divisions identified by Roman numerals and titles. The entire paper is divided into numbered sections, the numbers running consecutively without regard to the major divisions. This organization will be followed here, with frequent merger of sections.

I. Definitions — Parallelism

1. Jordan starts off with basic definitions, beginning with the underlying n-dimensional space.

> We define the position of a point in an n-dimensional space by the n coordinates $x_1, \ldots, x_n$.

Note that Jordan's space is not equipped with the usual apparatus of a vector space — addition, multiplication by a scalar, etc. Later he will introduce the Euclidean distance between two points and subsequently transformations of the coordinates. But his theorems and proofs are all cast in terms of planes, which he defines as follows.

> One linear equation in these coordinates defines a *plane.* Two simultaneous linear equations that are distinct and not incompatible define a *biplane;* k equations, a *k-plane;* $n-1$ equations define a *line;* n equations define a point. By the generic term *multi-plane* we will understand any of the above geometric entities.

Although it does not become clear until later, by linear equation Jordan means a (generally) inhomogeneous equation. Typically he will write the equation of a k-plane as

$$\begin{aligned} A_1 &= a_{11}x_1 + \cdots + a_{1n}x_n + \alpha_1 = 0, \\ &\cdots \\ A_k &= a_{k1}x_1 + \cdots + a_{kn}x_n + \alpha_n = 0, \end{aligned} \tag{1}$$

or for short $A_1 = \cdots = A_k = 0$. The abbreviated version conceals the difference between the terms dependent on x and the constants independent of x. In general we will write the defining equations of a k-plane in the form

$$Ax = a, \tag{2}$$

where A is a $k \times n$ matrix, x is an n-vector, and a is a k-vector.

It will be instructive to investigate these definitions in the light of today's mathematical language. In his definition of plane Jordan's uses the terms distinct and compatible only for 2-planes. What he requires for his subsequent definitions is that the equations have a nonempty set of solutions and that the removal of any equation will change that set. In today's parlance this means that the rows of A in (2) are linearly independent.

The term k-plane suggests that k has something to do with the dimensionality of its solution set. Specifically, the set of solutions of (2) has the following form

$$x = x_0 + z_A, \quad z_A \in \mathcal{N}(A) \tag{3}$$

where x_0 is a particular solution of (2) [e.g., $A^*(AA^*)^{-1}a$] and $\mathcal{N}(A)$ is the null space of A; i.e., $\{x; Ax = 0\}$. Thus the Jordan plane generated by (2) is just a translated subspace

$$x_0 + \mathcal{N}(A).$$

The definition (2) has important consequences. First, since the k rows of A are linearly independent, the dimension of its null space is $n - k$. In other words a k-plane has dimension $n - k$. In particular, what Jordan calls a plane (i.e., a 1-plane) has dimension $n - 1$. As we shall soon see, the Jordan's planes are sort of dual to vectors in a vector space.

Second, a plane is a point set, not a set of equations.[4] As long as that point set is maintained, A and a in (2) can be changed. In particular if L is nonsingular the equation

$$LAx = La \tag{4}$$

defines the same plane as does (2). Note for later reference that (4) implies that we can replace A with a matrix having orthogonal rows; e.g., by $(AA^*)^{-\frac{1}{2}}A$. Jordan will

[4] However, Jordan often identifies planes and their equations and will on occasions refer to, e.g., *the multi-plane* (1).

later describe an orthogonalization method of his own that is related to the well-known Gram–Schmidt orthogonalization method.

Third, in (3) the vector x_0, may be replaced by $x_0 + z_0$, where $z_0 \in \mathcal{N}(A)$, without changing the plane. Informally, sliding a plane along itself does not produce a new plane.

Finally, Jordan's multi-planes can move about $\mathbb{R}^n$. This movement can be controlled by varying the right-hand side a of (2). Note that if $a \neq a'$, then the intersection of the planes defined by the equations $Ax = a$ and $Ax = a'$ is the empty set. Thus the planes defined by A are either identical or nonintersecting (i.e., parallel).

2. Jordan now introduces the concept of a generating plane:

> Let
>
> $$A_1 = 0, \ldots, A_k = 0 \tag{J1}$$
>
> be the equations of a k-plane P_k. If these equations are combined linearly, we get an infinite number of equations of the form
>
> $$\lambda_1 A_1 + \cdots + \lambda_k A_k = 0.$$
>
> The various planes represented by these formulas clearly have P_k as their common intersection, and for short we will call them the *generating planes* of P_k.
>
> ...
>
> It is clear that in place of the equations (J1), we can define P_k by the equations of any k generators
>
> $$\lambda_1 A_1 + \cdots + \lambda_k A_k = 0, \quad , \lambda_1' A_1 + \cdots + \lambda_k' A_k = 0, \ldots,$$
>
> provided that the determinant of the coefficients λ is not zero.

In other words, the generating planes are of the form $\ell^* A = \ell^* a$. It takes only k independent generating planes to define their plane. Here Jordan uses the fact that if $(A^* \ B^*)$ has full column rank then the plane generated by the equation

$$\begin{pmatrix} A \\ B \end{pmatrix} x = \begin{pmatrix} a \\ b \end{pmatrix}$$

is the intersection of the planes $Ax = a$ and $Bx = b$.

Generating planes play a role complementary to vectors in linear spaces. Jordan builds his multi-planes from the top down by intersecting generating planes. This should be contrasted with the technique of building up subspaces from a basis of vectors.

Finally, Jordan notes that the points of a k-plane can be represented as a function of $n - k$ variables. He shows that by adjoining equations to (2) to form a nonsingular equation of the form

$$\begin{pmatrix} A_{11} & A_{12} \\ L_{11} & L_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a \\ \ell \end{pmatrix}. \tag{5}$$

This equation may then be solved for x_1 and x_2 in terms of L. Although Jordan does not give the details, it is a good illustration of the method of elimination, which Jordan frequently uses to establish his results. We will take time out to show how it works.

Assume that A_{11} is nonsingular. If we multiply (5) by the matrix

$$\begin{pmatrix} I & 0 \\ -L_{21}A_{11}^{-1} & I \end{pmatrix},$$

we get

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & L_{22} - L_{21}A_{11}^{-1}A_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a \\ \ell - L_{21}A_{11}^{-1}a \end{pmatrix}.$$

If, in addition, $L_{22} - L_{21}A_{11}^{-1}A_{12}$ is nonsingular, then we can write

$$x_2 = (L_{22} - L_{21}A_{11}^{-1}A_{12})^{-1}(\ell - L_{21}A_{11}^{-1}a)$$

which exhibits x_2 as a function of ℓ. Similarly, x_1 is a function of x_2 and hence of ℓ:

$$x_1 = A_{11}^{-1}(a - A_{12}x_2). \tag{6}$$

Thus we have parameterized the vectors x belonging to the plane defined by (2) as a function of the $(n{-}k)$-vector ℓ.

There are three comments to make about the above development

- The particular procedure given here is called block Gaussian elimination with back substitution. It is one of several variants of elimination. What they all have in common is that they take a linear system in a set of variables and produced a smaller system in a subset of the variables. Hence the name elimination. They also provide a means for finding the values of the eliminated variables once the smaller system has been solved. We will see other examples a little later.

- If A_{11} is singular, the process breaks down. In this case there is a cure: rearrange the columns of A so that A_{11} is nonsingular, a strategy that now goes under the name of pivoting. In this case it can always done because the rows of A are independent. It should be noted that Jordan does not mention that pivoting may be necessary, but later he sometimes suggests the equivalent by requiring changes in the ordering of his equations or their terms.

- Even after pivoting, there is no guarantee that the matrix $L_{22} - L_{21}A_{11}^{-1}A_{12}$ (sometimes called a Schur complement) is nonsingular. It may be possible to choose L_{11} and L_{12} so that this is always true. For example, if we choose $L_{21} = 0$ and $L_{22} = I$, then $L_{22} - L_{21}A_{11}^{-1}A_{12} = I$, so that $x_2 = \ell$ and $x_1 = A_{11}^{-1}(a - A_{12}x_2)$. In other words, if A_{11} is nonsingular, the last $k - n$ components of x may be chosen arbitrarily and the first k can be recovered from (6).

3. Jordan now turns to the interpolation problem of how planes may be determined by a choice of points in $\mathbb{R}^n$.

> *A k-plane is determined by $n-k+1$ points.* Specifically, consider an arbitrary plane that is forced to contain the $n-k+1$ points. This condition gives $n-k+1$ linear equations in the $n+1$ coefficients of the plane. If $n-k+1$ coefficients are eliminated using the these conditional equations, k arbitrary coefficients remain in the equation of the plane. Therefore, the general equation of any plane that pass through the $n-k+1$ given points has the form
>
> $$\lambda_1 A_1 + \cdots + \lambda_k A_k = 0, \tag{J2}$$
>
> and the k-plane $A_1 = \cdots = A_k = 0$ is the common intersection of these planes passing through the $n-k+1$ given points.

Let's see how this plays out in matrix terms. Let X denote the $n{\times}(k{+}1)$ matrix whose columns consist of the coordinates of the points determining the plane. Write a single generating plane in the form $a^*x - \alpha = 0$; The $n-k+1$ coefficients are given by the equation $a^*X - \alpha\mathbf{e}^* = 0$, where $\mathbf{e}$ is the vector of dimension $n-k+1$ consisting of all ones. If we set

$$\hat{X} = \begin{pmatrix} X_1 \\ X_2 \\ \mathbf{e}^* \end{pmatrix},$$

where X_1 is of order $n-k+1$, and partition $(a^* \ \ \alpha) = (a_1^* \ \ a_2^* \ \ \alpha)$ conformally, then

$$a_1^* X_1 + a_2^* X_2 - \alpha\mathbf{e}^* = 0.$$

Hence if X_1 is nonsingular

$$a_1^* = -(a_2^* X_2 - \alpha\mathbf{e}^*) X_1^{-1}. \tag{7}$$

Thus we may we may select $(a_2^* \ \ \alpha)$ and fill out a by computing a_1 from (7). In fact, since the dimension of $(a_2^* \ \ \alpha)$ is k, we can select k independent samples and obtain k generating planes for the k-plane containing the points $x_1, \ldots, x_{n-k+1}$.

All this depends on X_1 being nonsingular. Otherwise, we can use row pivoting to make it so. Jordan allows the possibility by not specifying which of the $n-k+1$ of the variables are to be eliminated. He has thus shown that the ability to carry out the elimination is a sufficient condition for the theorem. However, he does not show that the condition is necessary. It might be that there are points that determine a plane of the appropriate dimension for which the elimination cannot be carried through. Fortunately, it it can be shown that any such plane cannot be unique unless $\hat{X}$ is of full rank.

4. In §§4–12 Jordan defines a notion of parallelism between two planes and treats its consequences. Essentially, two hyperplanes are parallel if they have some parallel generating planes. Since generating planes are 1-planes, Jordan begins his discussion of parallelism by considering what it means for two 1-planes

$$a^*x = \alpha \quad \text{and} \quad b^*x = \beta, \qquad a, b \neq 0, \tag{8}$$

to be parallel. He treats three cases.

1. If the components of a are not proportional to those of b—i.e., if a and b are linearly independent—then (8) has a solution x. Hence, the two planes intersect at x and cannot be parallel.
2. If $a = \mu b$ for some $\mu \neq 0$ and $\alpha \neq \mu\beta$, then (8) has no solution; i.e., the two planes do not intersect and hence are parallel.
3. If $a = \mu b$ and $\alpha = \mu\beta$ then the two planes coincide and are by convention parallel.

To complete the picture here is a fourth case that Jordan will consider later. Specifically, suppose that a and b are orthogonal: that is,

$$a^*b = 0. \tag{9}$$

Then then the two planes are at right angles. For example, suppose that $\alpha = \beta = 0$, a is the lies along the x-axis in three-dimensional space, and b lies along the y-axis. Then the two resulting planes are the (y, z)-plane and the (x, z)–plane. Jordan will generalize this notion and say that the resulting planes are perpendicular.

5. Jordan now turns to a general definition of parallelism. He begins enigmatically.

> Let P_k and P_l be two arbitrary multi-planes. If from among the generating planes of P_k there are some that that are parallel to generating planes of P_l, then they generate a multi-plane.

Specifically, Jordan considers the set $\mathcal{P}_k$ of all generating planes of P_k that are parallel to one of the generating planes of P_l and the set $\mathcal{P}_l$ of all generating planes of P_l that are parallel to one of the generating planes of P_k. Since the equations of two such planes are of the form $a^*x = \alpha$ and $b^*x = \beta$, where the vectors a and b are proportional, we can write the two planes in the form $c^*x = \hat{\alpha}$ and $c^*x = \hat{\beta}$. It follows that the correspondence between generating planes in P_k that are parallel to generating planes in P_l is one-one.

From $\mathcal{P}_k$ Jordan draws ρ equations $C_1 = 0, \ldots, C_\rho = 0$ that satisfy the following conditions:

1. They are mutually independent; i.e., they satisfy no linear identity of the form

$$\lambda_1 C_1 + \cdots + \lambda_\rho C_\rho = 0.$$[5]

2. There is no generating plane of P_k that is independent of the chosen planes and is parallel to a generating plane of P_l.

This set of equations, which we will write $Cx = a_1$, generates a plane which Jordan calls P_ρ. In addition there is a corresponding plane P'_ρ whose generators are taken from the $\mathcal{P}_l$ and can be written in the from $Cx = b_1$. Jordan then says that P_k and P_l have a common parallelism of order ρ.

This definition leads directly to equivalents to the three cases that derive from equation (8). To see this, let us write the equations for P_k and P_l in the form

$$\begin{pmatrix} C \\ A_2 \end{pmatrix} x = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} C \\ B_2 \end{pmatrix} x = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{10}$$

where A_2, a_2 are chosen from the equations for the generating planes of P_k and similarly for B_2 and b_2.[6] Now first, if C is is an empty matrix, the rows $A = A_2$ and $B = B_2$ are linearly independent, and the the equations always have a solution, no matter what their right hand sides. In other words, the planes P_k and P_l always have a point of intersection and are therefore not parallel. Second if C is not empty and $a_1 \neq b_1$, planes defined by the the the two systems are inconsistent and have no point of intersection; i.e., the two planes are parallel. Third, if $a_1 = b_1$ then the two planes intersect in a $(k+l-\rho)$-plane. In Jordan's terms, P_ρ and P'_ρ are identical.

When $k + l > n$ it would seem possible that $k + l - \rho$ could also be greater than n, in which case the partitioning in (10) would make no sense. However, if $k + l > n$ the union of the spaces spanned by the vectors in A^* and B^* cannot fit into n-space unless the spaces have at least $k + l - n$ independent vectors in common — vectors that will contribute to C. Thus, ρ cannot be not less than $k + l - n$, and hence $k + l - \rho$ cannot be greater than n.

6, 7, 8, 9. Jordan devotes these sections to finding

> conditions that must be satisfied by the coefficients of the equations of P_k and P_l for these planes to have a mutual parallelism of order ρ and for them to be contained in the same ρ-plane.

---

[5] It is worth noting that this conditions for equations is essentially the same as Frobenius' definition of linear independent set of vectors [10, p.232] Also see [9, p.258].

[6] Later (§12) Jordan will introduce these equations expanded from C. The existence of A_2 and B_2 requires the use of some variant of the Steinitz exchange lemma: see [9, p.250–251]. Whether or not Jordan regarded the lemma as obvious is not easily decided.

Jordan begins with the defining equations of his multi-planes P_k and P_l: namely, $Ax = a$ and $Bx = b$, where it is assumed that A and B have full row rank. For a linear combination of the equations, say $\ell^* Ax = \ell^* a$ and $m^* Bx = m^* b$ to generate parallel planes, we must have $\ell^* A = m^* B$ (up to a scaling factor which we may assume to be one). Equivalently we must have

$$(B^* \;\; A^*) \begin{pmatrix} m \\ -\ell \end{pmatrix} = 0. \tag{11}$$

Thus the problem becomes one of determining a maximal set of independent null vectors of $(B^* \;\; A^*)$. The cardinality of this set is the order of parallelism ρ of P_k and P_l.

The solution to this problem involves elimination. In §7, Jordan assumes that $k+l \leq n$. He then eliminates m followed by the elimination of $k-\rho$ components of ℓ. This allows the remaining ρ components of ℓ to be chosen arbitrarily, after which the remaining components of ℓ and the components of m may be determined.

The following algorithm implements these ideas using Gaussian elimination with partial and complete pivoting. The pivoting strategies ensure that m is eliminated first and that zeros end up in the proper place.[7]

> Perform Gaussian elimination on the matrix $(B^* \;\; A^*)$, using partial pivoting on B^* and then switching to complete pivoting to process A^*.

The result of this algorithm may be summed up in the following equation:

$$T(B^* \;\; A^* E) = \begin{matrix} \\ l \\ k-\rho \\ n-k-l+\rho \end{matrix} \begin{matrix} \begin{matrix} l & k-\rho & \rho \end{matrix} \\ \begin{pmatrix} \hat{B}_{11} & \hat{A}_{12} & \hat{A}_{13} \\ 0 & \hat{A}_{22} & \hat{A}_{23} \\ 0 & 0 & 0 \end{pmatrix} \end{matrix}, \tag{12}$$

where

1. T is a nonsingular matrix of order n. Because it is nonsingular, the null vectors of the right-hand side of (12) are the same as those of $(B^* \;\; A^* E)$.

2. E is a permutation matrix of order k that reflects the reordering of the columns of A^* due to complete pivoting. Its inverse must be applied to the right-hand side to get the null vectors of $(B^* \;\; A^*)$. To keep things simple, we will assume that $E = I$.

[7] Briefly, Gaussian elimination on a matrix of order n consists in successively using diagonal elements to zero out (eliminate) the subdiagonal elements of a matrix. If, at step k, the diagonal element a_{kk} is zero, the elimination may fail. Pivoting consists replacing a_{kk} with another, nonzero element. In partial pivoting the pivot is chosen to be the largest element (in magnitude) from the pivot column $a_{kk}, \ldots, a_{nk}$. In complete pivoting the choice is made from the elements a_{ij} $(i, j = k, \ldots, n)$.

3. $\hat{B}_{11}$ is a nonsingular upper triangular.

4. $\hat{A}_{22}$ is a nonsingular upper triangular matrix.

We may obtain a complete set of independent null vectors by solving the system

$$\begin{pmatrix} \hat{B}_{11} & \hat{A}_{12} & \hat{A}_{13} \\ 0 & \hat{A}_{22} & \hat{A}_{23} \end{pmatrix} \begin{pmatrix} M \\ -L_2 \\ I_\rho \end{pmatrix} = 0.$$

This gives

$$L_2 = \hat{A}_{22}^{-1}\hat{A}_{23} \quad \text{and} \quad M = \hat{B}_{11}^{-1}(\hat{A}_{12}L_2 - \hat{A}_{13}).$$

Thus the full matrix L is given by

$$L = \begin{pmatrix} \hat{A}_{22}^{-1}\hat{A}_{23} \\ I_\rho \end{pmatrix}. \tag{13}$$

The ith row of L^* is $(\ell_i \;\; \mathbf{e}_{\rho-i+1})$. On comparing this with (11) we see that the coefficients of ith generating plane of P'_k consists of the $(\rho-i+1)$th generating plane of P_k and a linear combination of the first $k-\rho$ generating planes. The right-hand side of P'_k consists of the same linear combination of the right-hand sides of the generating planes of P_k. This is essentially the result that Jordan announces in (J10) and (J11).

In §8 Jordan briefly treats the case $k + l > n$. In this case, as noted above, at least $k+l-n$ of $(A^* \;\; B^*)$ must be linearly dependent on the others. Removing those columns reduces the case to where $k + l \leq n$, which can be treated as above.

The algorithm given above also works when $k > n$ and gives the following result:

$$T(B^* \;\; A^*E) = \begin{matrix} & \\ l \\ n-l-\sigma \\ \sigma \end{matrix} \begin{matrix} \begin{matrix} l & n-l-\sigma & \sigma & k+l-n \end{matrix} \\ \begin{pmatrix} \hat{B}_{11} & \hat{A}_{12} & \hat{A}_{13} & \hat{A}_{14} \\ 0 & \hat{A}_{22} & \hat{A}_{23} & \hat{A}_{24} \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix},$$

where $0 \leq \sigma \leq n - l$ and $\hat{A}_{22}$ is nonsingular and upper triangular. The null vectors of this matrix can be obtained in the form

$$\begin{pmatrix} \begin{pmatrix} \hat{B}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \hat{A}_{13} & \hat{A}_{14} \\ \hat{A}_{23} & \hat{A}_{24} \end{pmatrix} \\ -I_{k+l+\sigma-n} \end{pmatrix}. \tag{14}$$

Thus the order of parallelism is $\rho = k + l + \sigma - n$.

Here are some comments on these results.

- Jordan does not mention pivoting explicitly. But he suggests it by saying that there are choices of variables to eliminate and then assumes that their indices are conveniently numbered; e.g. at the beginning or end of a sequence of variables. For example:

> Then equation (J7) allows the determination of $\mu_1, \ldots, \mu_l$ and $k-\rho$ of the quantities λ, **say** $\lambda_{\rho+1}, \ldots, \lambda_k$, as functions of the ρ free parameters $\lambda_1, \ldots, \lambda_\rho$ [emphasis added].

• The above quotation suggests that Jordan requires a knowledge of ρ to carry through his argument. This is consistent with his statement that he wants to find conditions under which P_k and P_l have parallelism of order ρ. In the Gaussian elimination algorithm, on the other hand, the value ρ is evident in the final form (12). All you have to do is count the number of rows in A_{22} and subtract it from k.

• The Gaussian elimination algorithm and (14) can be implemented to run on a computer and determine the coefficients of the generating planes of P'_k and P'_l. The actual computation, however, must be performed in finite precision arithmetic with rounding error, and the result is that quantities that should be zero will be in general nonzero, which is to say that no parallelism will be detected. Thus for the algorithm to work, these offending quantities must be detected and set to zero. Unfortunately, this process is fraught with difficulties only too well known to numerical analysts.

10, 11, 12. Jordan makes the following statements

> *If two multi-planes P_k and P_l have no parallelism, they intersect in a multi-plane P_{k+l}.*
>
> *If P_k and P_l are parallel of order ρ and do not have the same ρ-plane, then they do not intersect.*
>
> *If P_k and P_l have the same ρ-plane, then they intersect in a $(k+l-\rho)$-plane.*

These three statements follow easily from the form (10).

13. Jordan states:

> A k-plane P_k sliding over an l-plane while remaining parallel to itself it produces a multi-plane.

Here Jordan seems to be going around Robin Hood's barn to prove what he has established previously. For if P_k and P_l intersect then either they must have no parallelism or they must lie in the same ρ-plane. In the first case they intersect in an $(k+l)$-plane [§10]. In the second they intersect in a $(k+l-\rho)$-plane [§12].

II. Distance and Perpendicularity

Jordan now proceeds to the metric geometry of his planes. Here is an outline of his results.

1. Jordan begins by defining what we now call the Euclidean distance in n-space.

2. He then shows how to compute the point x in a k-plane P_k that is nearest to an arbitrary point y. He calls x the projection of y onto x.

3. He uses projections to define the mutual perpendicularity between two planes P_k and P_l and gives an elegant characterization of perpendicularity.

4. He then shows that any k-plane can be written as the intersection of k mutually perpendicular planes.

5. He finally establishes what is now called the Pythagorean equality for projections.

14. Jordan begins with some definitions.

> *The distance between two points* whose coordinates are respectively $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ is defined by the formula
>
> $$\Delta = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$
>
> The *distance between a point p and a multi-plane* is its distance to the point of the multi-plane that is nearest it. This point q is the *projection* of the point p onto the multi-plane.
>
> The *distance between two multi-planes* that do not intersect is the distance between the nearest neighbors of their points.
>
> The *projection of a multi-plane onto another* is the locus of the projections of its points.

Note that Jordan's distance is defined only between two points; the points themselves do not have associated magnitudes, unlike vectors in a vector space.

15. Jordan now sets out to determine the coordinates of the projection x of a point y onto the plane P_k whose equations are $Ax = a$. His argument goes as follows. The differential of the distance $(x-y)^*(x-y)$ with respect to x is $2(x-y)dx$. Hence, setting the differential to zero, we get

$$(x - y)dx = 0. \tag{15}$$

But x is constrained to lie on A_k. Hence

$$A\,dx = 0. \tag{16}$$

This means that (15) must be a linear combination of the rows of (16). Hence for some k-vector ℓ

$$x - y = A^*\ell. \tag{17}$$

Now if we eliminate the variables ℓ, we end up with $n - k$ equations for x. Conjoining these with the equations

$$Ax = a,$$

we get n equations in n unknowns for x.

It is worth while to see how the elimination plays out in matrix terms. Write the equation $x - y = A^*\ell$ in the partitioned form

$$\begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix} = \begin{pmatrix} A_1^* \\ A_2^* \end{pmatrix} \ell,$$

where A_1^* is assumed to be nonsingular. If we premultiply this equation by

$$\begin{pmatrix} I_k & 0 \\ -A_2^* A_1^{-*} & I_{n-k} \end{pmatrix},$$

we get

$$\begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 - A_2^* A_1^{-*}(x_1 - y_1) \end{pmatrix} = \begin{pmatrix} A_1^* \\ 0 \end{pmatrix} \ell,$$

or

$$\begin{pmatrix} I_{n-k} & -A_2^* A_1^{-*} \end{pmatrix} \begin{pmatrix} x_2 - y_2 \\ x_1 - y_1 \end{pmatrix} = 0,$$

and finally

$$\begin{pmatrix} I_{n-k} & -A_2^* A_1^{-*} \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} = y_2 - A_2^* A_1^{-*} y_1.$$

If we now append the system $Ax = a$, we get

$$\begin{pmatrix} I_{n-k} & -A_2^* A_1^{-*} \\ A_2 & A_1 \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} y_2 - A_2^* A_1^{-*} y_1 \\ a \end{pmatrix}. \tag{18}$$

If we multiply the first row of this equation by A_2 and subtract it from the second row, we get the equation

$$(A_1 + A_2 A_2 A_1^{-*}) x_1 = a - A_2 y_2 + A_2 A_2^{-*} A_1^{-*} y_1 \tag{19}$$

for x_1. Then x_2 may be obtained from the equation

$$x_2 = y_2 + A_2^* A_1^{-*}(x_1 - y_1).$$

There are three comments to be made on this derivation.

- In practice the matrix A_1^* may be singular, in this case, one can pivot by interchanging rows of (17) so that the new A_1^* is nonsingular. This is always possible because A^* is assumed to be of full rank.

- Jordan does not show that the augmented the system (18) is nonsingular. From (19) we see that it is true if and only if $A_1 + A_2 A_2 A_1^{-*}$ is nonsingular. This matrix is nonsingular if and only if $A_1A_1^* + A_2A_2^* = AA^*$ is nonsingular, which is true because A has full row rank.

• Jordan overlooks the possibility of determining x directly from (17). The derivation is simple. On multiplying (17) by A, we get

$$a - Ay = AA^*\ell.$$

Since the rows of A are independent, AA^* is nonsingular and

$$\ell = (AA^*)^{-1}a - (AA^*)^{-1}Ay. \tag{20}$$

Finally substituting this expression in (17), we get

$$x = A^*(AA^*)^{-1}a + [I - A^*(AA^*)^{-1}A]y \tag{21}$$

It is worth noting that (21) is in the form (3). Specifically, $A^*(AA^*)^{-1}a$ is a particular solution of $Ax = a$ and $[I - A^*(AA^*)^{-1}A]y$ lies in the null space of A.

16. Jordan now introduces the notion of perpendicularity between planes.

> In general, an l-plane P_l will be said to be *perpendicular to a k-plane* P_k if given two planes P_l' and P_k' that are parallel to P_l and P_k respectively and pass through an arbitrary point q, the projection of each point of P_l' onto P_k' lies in intersection of P_l' and P_k'.

It is clear that the concept of perpendicular planes is not the same as that of orthogonal subspaces, since the latter intersect only at the origin. The model to look to is the coordinate planes of Cartesian 3-space. Here, for example, the xy-plane and the xz-plane have a nontrivial intersection, namely, the x-axis. However, all projections from one plane onto the other lie on the x-axis.

17, 18, 19, 20. In these sections Jordan establishes a condition for perpendicularity of two planes P_k and P_l and shows that perpendicularity is a reciprocal relation, something that is not obvious from his definition. The result, which uses an extension of the representation (10), is simple and elegant.

Let the equations for P_k and P_l be written in the form

$$\begin{pmatrix} C \\ A_2 \end{pmatrix} x = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} C \\ B_2 \end{pmatrix} y = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{22}$$

where

$$CA_2^* = 0 \quad \text{and} \quad CB_2^* = 0. \tag{23}$$

Then a necessary condition for P_l to be perpendicular to P_k is that

$$B_2 A_2^* = 0. \tag{24}$$

The representation (22) differs from (10) by the orthogonality hypotheses (23). Jordan provides the algorithmic wherewithal to obtain the required orthogonality in §§20–21, which we will discuss separately.

Jordan gives three proofs of his result. The first treats the case where C is empty; i.e., when P_k and P_l have no parallelism. The second treats the general case, but does not explicitly give the condition (24). The third proof also treats the general case, and that is the one that will be described here here.

First note that the definition of perpendicularity is stated in terms of two planes P_k' and P_l' that have a common point q. Without loss of generality, we may assume that $q = 0$, which amounts to setting a_1, a_2, b_1, and b_2 to zero in (22). To keep the notation simple we will assume that $P_k = P_k'$ and $P_l = P_l'$. Thus our equations become

$$\begin{pmatrix} C \\ A_2 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} C \\ B_2 \end{pmatrix} y = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{25}$$

Suppose now that $y \in P_l$, so that y satisfies (25). Let x be the projection of y onto P_k, so that x satisfies (25). Then x and y satisfy (17), which we may write in the partitioned form

$$x - y = (C^* \;\; A_2^*) \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix}. \tag{26}$$

By the definition of perpendicularity, x also must be in P_l; i.e.,

$$\begin{pmatrix} C \\ B_2 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

From (25) It follows that

$$\begin{pmatrix} C \\ B_2 \end{pmatrix} (x - y) = 0. \tag{27}$$

If the value of $x - y$ in (26) is substituted into (27), the result is

$$\begin{pmatrix} C \\ B_2 \end{pmatrix} (C^* \;\; A_2^*) \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} = \begin{pmatrix} CC^* & 0 \\ 0 & B_2 A_2^* \end{pmatrix} \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

where the zeros in the matrix come from (23). It follows that $\ell_1 = 0$ and

$$B_2 A_2^* \ell_2 = 0. \tag{28}$$

At this point we have essentially followed Jordan's development. However, Jordan goes on to claim that "all the values of $\lambda_1, \ldots, \lambda_k$ that satisfy these ρ equations [i.e., $CC^*\ell_1 = 0$] must satisfy as an identity all the other equations of the system." The assertion is not obvious, since ℓ_2 depends on y. What is needed is the assertion that as y varies over P_l, the vector ℓ_2 varies over *all* of $\mathbb{R}^{k-\rho}$. This, along with (28), would imply that $B_2A_2^* = 0$.

In fact, this assertion is true. But even granted the assertion, it only establishes necessity of the conditions for perpendicularity. Jordan goes on to state that, since $B_2A_2^* = 0$ implies $A_2B_2^* = 0$, P_k is orthogonal to P_l. But this is true only if the condition $A_2B_2^* = 0$ is sufficient for P_k to be orthogonal to P_l, something Jordan does not prove.

It is possible to establish sufficiency by algebraic manipulations. Necessity appears to require additional analytic methods — specifically the invocation of the singular value decomposition or an equivalent. In an appendix to this commentary I will give the proofs of both necessity and sufficiency. As it turns out, Jordan introduced the singular value decomposition in a paper that appeared in the year before the present paper [12, 19][8] and hence would have understood these proofs.

—— ◇ ——

We will now turn to Jordan's method of orthogonalization, which appears in §20. The problem is to adjust A_2 and B_2 in (22) so that they satisfy (23).

Thus Jordan considers the problem of finding a linear combinations the rows of C and A_2 that is orthogonal to the rows of C. Since the planes generated by C and A_2 have no parallelism, we must find row vector $p^* = (p_C^*\ p_A^*)$, where p_C has ρ components and p_A has $k - \rho$ components such that

$$(p_C^*\ p_A^*) \begin{pmatrix} C \\ A_2 \end{pmatrix} C^* = 0. \tag{29}$$

It is easily seen that

$$p_C^* = -p_A^* A_2 C^* (CC^*)^{-1}. \tag{30}$$

(Jordan arrives at essentially the same result by invoking elimination). If we choose $k - \rho$ independent vectors for p_A (e.g., unit vectors), then we obtain $k - \rho$ independent vectors p whose planes are all perpendicular to the generating planes of C and may replace A_2.

It should be noted that if CC^* is a diagonal matrix (i.e., the rows of C are orthogonal) and choices for the vectors p_A are the unit vectors, then Jordan's algorithm is equivalent to a single step of the modified Gram–Schmidt algorithm. We will return to this point in a moment.

21. Jordan states:

[8] However, see footnote 1

> Any k-plane P_k can be represented in an infinite number of ways as the intersection of k rectangular planes. [Note the distinction between a k-plane and k planes.]

Jordan does not define the term rectangular, but he means that the k planes are mutually perpendicular. As we have seen, perpendicularity in 1-planes amounts to orthogonality of the vectors of coefficients of the planes. It follows that if the equation for P_k is $Ax = a$, and we replace it by $SA = Sb$, where $SA(SA)^*$ is a diagonal matrix, then the k rows of $SA = Sb$ are equations of the required 1-planes.

Jordan proceeds by successive orthogonalization using the algorithm he sketched in §20. Specifically, he chooses a plane generator A_1 of P_k. He then orthogonalizes rows of A against A_1 to produce a $(k-1)$-plane P_{k-1}. From P_{k-1} he draws another plane generator A_2. By construction A_2 is orthogonal to A_1. The next step is to find a $(k-2)$-plane P_{k-2} in P_k that is perpendicular to A_1 and A_2. From P_{k-2}, he draws a plane generator A_3 that is necessarily orthogonal to A_1 and A_2. The procedure continues by finding a $(k-3)$-plane in P_k that is perpendicular to A_1, A_2, and A_3, from which he which draws plane generator A_4, and so on. The 1-planes defined by $A_1, \ldots, A_k$ are the ones required by his assertion.

Jordan is vague about he how he finds the planes P_{k-1}, P_{k-2}, etc. A likely possibility is to take C in (29) from, say for example, the generators A_1 and A_2, and A from the generators of P_{k-2}. One then orthogonalizes as above to get the generators of P_{k-3}. Because of the orthogonality of A_1 and A_2, CC^* is diagonal, and the whole procedure reduces to a scaled version of the modified Gram–Schmidt algorithm.

Yet it would be wrong to assert that Jordan anticipated the Gram–Schmidt algorithm. He never specifically mentions the choices in the preceding paragraph. The best that can be said is that he presents a constructive orthogonalization procedure which, in a particular incarnation, becomes the modified Gram–Schmidt algorithm.[9]

22, 23. Jordan states:

> Let p be an arbitrary point, q be its projection onto a multi-plane P_k, and r be an arbitrary point of P_k. Then the distances between the three points p, q, and r satisfy the relation
> $$\overline{pr}^2 = \overline{pq}^2 + \overline{qr}^2.$$

This is the well-known generalization of the Pythagorean theorem. Jordan proves it using his characterization of projection in §15.

Jordan next considers the projection of a projection.

> Let P_{k+l} be a multi-plane contained in P_k. The projection [r] of p onto P_{k+l} falls on the same point s as the projection of q [onto P_{k+l}]. [Here q is the projection of p onto P_k. The notation here comes from his statement of the Pythagorean theorem above.]

[9]For more on the history of the Gram–Schmidt algorithm see [15].

Specifically, he shows that if r is an arbitrary point in P_{k+l}, then

$$\overline{pr}^2 = \overline{ps}^2 + \overline{sr}^2. \tag{31}$$

and concludes that his assertion is true. However, he implicitly uses the result that if (31) holds for all r in P_{k+l} then s is the projection of p onto l_{k+l}, which is true but requires proof.

III. Change of coordinates

The chief contribution of this section is the introduction of changes in coordinate systems and expressions for the distance between two points in terms of the new coordinates. However, Jordan begins with a discussion of lines between parallel planes.

24. Jordan states:

> Two parallel lines D and D' extending between two parallel planes P and P' have the same length.

The proof is a straightforward manipulation of the Jordan's definitions.

At the end of this section he concludes:

> Three parallel planes divide two arbitrary lines proportionally.

25, 26, 27. Jordan now turns to the problem defining new coordinate systems. While we would define a coordinate system by introducing a set of n independent vectors, Jordan defines his coordinate system in terms of n independent planes.

He begins with the following statement that would seem to belong in the last section.

> *The locus of points whose distance along a given direction from a fixed plane P is constant will necessarily be a plane parallel to P.*

He then defines his new coordinate system.

> Given a system of n independent planes $P, Q, \ldots$, a point in space is completely determined when one knows its distance from each of these planes, each direction being taken along, say, the direction of the intersection of the $n-1$ other planes. Specifically, [according to the statement quoted above] these distances determine n planes parallel to $P, Q, \ldots$ whose intersection is the point in question.

As an example, consider the 1-planes $x = 0$, $y = 0$, and $z = 0$ in Euclidean 3-space. These planes are commonly called the (y, z)-, (x, y)-, and (x, z)-planes. According to Jordan's definition for any point (u, v, w) the first of his coordinates is the length projection of (u, v, w) onto the (y, z)-plane along the direction of the x axis — i.e., u as might be expected. Similarly the second coordinate is v, and the third is w. These turn out to be the same as the original coordinates because of the orthonormality of the original coordinate axes.

By distance Jordan presumably means a signed magnitude, but he does not specify the sign. However, a plane P defined by $a^*x = \alpha$ has two sides, one the set of points y for which $\alpha - a^*y > 0$ and the other the set for which $\alpha - a^*y < 0$. (Note that if $\alpha - a^*y = 0$, then $y \in P$). Thus it is reasonable to take the sign of the direction of the point y to the plane to be the sign of, say, $\alpha - a^*y$.

Jordan then turns to the problem of computing of computing the new coordinates for an arbitrary point and then the distance between two points in terms of their coordinates. At the heart of his derivation is the solution of linear equations, which he effects by means of determinants, essentially expressing the inverse of a matrix in terms of its adjugate matrix. Here we will translate his development into the language of matrices. To distinguish between vectors formed from Jordan's scalars, we will denote all vectors by an overline; e.g., $\overline{\xi}$.

Jordan writes for the equations of the planes $P, Q, \ldots$ in the form $\overline{a}_i^*\overline{x} = \alpha_i$ $(i = 1, \ldots n)$. Their intersection is the point $\overline{x}$ satisfying $A\overline{x} = \overline{\alpha}$, where $A = (\overline{a}_1^* \ \cdots \ \overline{a}_n^*)^*$. He then turns to the problem of determining the coordinates of an arbitrary point $\overline{\xi}$. He treats them a coordinate at a time.

Specifically, partition

$$A = \begin{pmatrix} \overline{a}_1^* \\ A_2 \end{pmatrix}, \quad \text{and} \quad \overline{\alpha} = \begin{pmatrix} \alpha_1 \\ \overline{\alpha}_2 \end{pmatrix}$$

Then the locus of points parallel to the last $n-1$ planes must satisfy the equation

$$A_2(\overline{x} - \overline{\xi}\ ) = 0.$$

Moreover, the point where $\overline{x}$ intersects the first plane must satisfy

$$\overline{a}_1^*(\overline{x} - \overline{\xi}) = \overline{\alpha}_1 - a_1^*\overline{\xi}.$$

These two equations may be combined into one by writing

$$A(\overline{x} - \overline{\xi}) = \overline{\mathbf{e}}_1(\alpha_1 - \overline{a}_1^*\overline{\xi}),$$

where $\overline{\mathbf{e}}_1$ is the vector whose first component is one and whose remaining components are zero. This equation has the solution

$$\overline{x} - \overline{\xi} = A^{-1}\overline{\mathbf{e}}_1(\alpha_1 - \overline{a}_1^*\overline{\xi}) = \overline{a}_1^{(-1)}(\alpha_1 - \overline{a}_1^*\overline{\xi}),$$

where $\overline{a}_1^{(-1)}$ is the first column of the inverse of A. It follows that the directed distance from ξ to x is given by

$$X_1 = \pm\|\overline{x} - \overline{\xi}\| = \|\overline{a}_1^{(-1)}\|(\alpha_1 - \overline{a}_1^*\overline{\xi}),$$

where $\|\cdot\|$ is the Euclidian norm. The remaining components of the vector of new components may be found in the same way and the result summarized in the equation

$$\overline{X} = W(\overline{\alpha} - A\overline{\xi}), \qquad W = \operatorname{diag}(\|\overline{a}_1^{(-1)}\|, \ldots, \|\overline{a}_n^{(-1)}\|). \tag{32}$$

Jordan uses his equivalent of this formula to compute the distance between two points whose new components are $\overline{X}$ and $\overline{X}'$. Specifically, (32) can be inverted to give the following expressions

$$\overline{\xi} = A^{-1}W^{-1}\overline{X} + A^{-1}\overline{\alpha} \quad \text{and} \quad \overline{\xi}' = A^{-1}W^{-1}\overline{X}' + A^{-1}\overline{\alpha}.$$

Hence

$$\begin{aligned}\|\overline{\xi} - \overline{\xi}'\|^2 &= (\overline{X} - \overline{X}')^*(W^{-1}A^{-*}A^{-1}W^{-1})(\overline{X} - \overline{X}') \\ &= (\overline{X} - \overline{X}')(\overline{X} - \overline{X}') + (\overline{X} - \overline{X}')B(\overline{X} - \overline{X}'),\end{aligned}$$

where $B = W - I$. Because because the diagonal elements of $W^{-1}A^{-*}A^{-1}W^{-1}$ are one, the diagonal elements of B are zero and its off-diagonal elements are those of $W^{-1}A^{-*}A^{-1}W^{-1}$.

Finally, Jordan observes that if A is orthogonal then $B = 0$ and $\|\overline{X} - \overline{X}'\| = \|\overline{\xi} - \overline{\xi}'\|$. In other words, if the planes defining the new coordinate system are mutually perpendicular, then the standard formula for the distance between two points remains unchanged.

It may be of interest to see how Jordan writes the elements of A^{-1}. He effectively uses the formula

$$A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)},$$

where $\operatorname{adj}(A)$ is the adjugate matrix of A—i.e. the transpose of the matrix of cofactors of the elements of A— and $\det(A)$ is the determinant of A. He rather cleverly writes the (i,j)-cofactor in the form

$$c_{ij} = \frac{d\det(A)}{d\,a_{ij}},$$

which can be easily verified by expanding $\det(A)$ along its ith row or jth column and then differentiating.

Invariant Angles

In this section—one of the high points of the essay—Jordan defines what we now call the canonical angles between subspaces. His angles are defined for multi-planes, but planes or subspaces they are the same angles.

28, 29. Jordan begins with some definitions from geometry. Two geometric figures are said to be *congruent* whenever they can be represented by the same equations with respect to systems of orthogonal coordinates suitably chosen for each. They are *identical* if one can pass from the coordinate system of one to the coordinate system of the other by an orthogonal substitution with determinant one. In this case they differ only in their positions in space. The are *reflective* if the determinant of the substitution is minus one.[10]

Jordan goes on to show that any k-plane is both identical and reflective to itself.

On might wonder why the sign of the determinant determines whether two figures are identical or reflective. The reason is that for any orthogonal matrix U of order n there is an orthogonal matrix V with positive determinant such that

$$V^{*}U = \mathrm{diag}(I_{n-1} \ \pm 1) \tag{33}$$

Equivalently, $V = U$. The plane V is the product of rotations with positive determinant in the planes. Geometrically this says that given to congruent figures, one may be obtained from the other by a sequence of rotations in the in planes defined by pairs of coordinate axes, followed, if the minus sign obtains obtains in (33), by a reflection in the $(n-1)$ dimensional plane orthogonal to the nth coordinate axis.

30, 31, 32. In these sections, Jordan derives the angle between two planes P and Q in two ways. He assumes that the planes intersect and that one of the points of intersection is the origin.

The first and simplest derivation is to introduce a coordinate system consisting of the plane P and of another plane perpendicular to the first that contains the intersection of both. The remainder of the the coordinate system is an arbitrary set of planes containing the origin and perpendicular to P and Q.

In such a coordinate system the equations of P and Q are

$$\begin{gathered} x_1 = 0, \\ ax_1 + bx_2 = 0. \end{gathered}$$

If we normalize a and b so that their sum of squares as one, then we can regard the normalized a as a cosine and b as a sine:

$$\begin{gathered} x_1 = 0, \\ x_1 \cos\alpha + x_2 \sin\alpha = 0. \end{gathered} \tag{34}$$

Thus the angle α characterizes the relation of the subspaces P and Q. As Jordan puts it:

> Therefore, systems of two planes are not all the same but differ among themselves by a characteristic element.

[10] The French words in the paper corresponding the congruent, identical, and reflective are respectively *pareil, egál,* and *symetŕique.*

Note that if we were to interchange P and Q, the coordinate system would change, and it is not immediately clear that the angle α so obtained is the same as the original. Jordan, therefore, turns to an approach that works with the actual equations of P and Q. The idea is simple, even though the execution is tedious. Let q be a point of Q, and let p be its projection onto P. Let D^2 be the square of the distance between p and q. Similarly let r be the projection of q onto the intersection of P and Q, and let Δ^2 be the square of its length. Then D^2/Δ^2 is the square of the sine of the angle α between $\overline{qr}$ and $\overline{rp}$. Since distance is invariant under orthogonal substitutions, P and Q have the same value of $\sin^2\alpha$ in any rectangular coordinate system.

Jordan concludes by noting that $\sin^2\alpha$ and not $\sin\alpha$ is the invariant, since sign of the latter can change under orthogonal substitutions.

33, 34, 35, 36, 37. Neither of the above approaches will work in the general case of a k-plane P_k and an l-plane P_l, since the two planes will typically have more than one canonical angle. Jordan's approach to this problem is to peel off perpendicular multi-planes from P_k and P_l until all that is left is a set of multi-planes from which the canonical angles can be extracted. Unfortunately for those of us who do not share Jordan's geometric insight, the meaning of his constructions is not readily grasped. We will therefore introduce a canonical form for the matrices A and B defining P_k and P_l that may used to illustrated Jordan's constructions.

We will consider the case where $k \le l$, $k+l \le n$, and $0 \in P_k \cap P_l$. By row operations on A and B and an orthogonal change of the variables x, the matrices A and B for P_k and P_l may be reduced to the forms given below

$$
A^* = \begin{array}{c} \\ \rho \\ r \\ t \\ \rho \\ r \\ t \\ l{-}k \\ q \end{array}
\!\!\begin{array}{c} \begin{array}{ccc} \rho & r & t \end{array} \\
\begin{pmatrix}
I & 0 & 0 \\
0 & I & 0 \\
0 & 0 & I \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}\end{array}, \qquad
A_\perp^* = \begin{array}{c} \\ \rho \\ r \\ t \\ \rho \\ r \\ t \\ l{-}k \\ q \end{array}
\!\!\begin{array}{c} \begin{array}{ccccc} \rho & r & t & l{-}k & q \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 \\
0 & 0 & I & 0 & 0 \\
0 & 0 & 0 & I & 0 \\
0 & 0 & 0 & 0 & I
\end{pmatrix}\end{array}, \qquad (35)
$$

,

$$
B^* = \begin{array}{c} \\ \rho \\ r \\ t \\ \rho \\ r \\ t \\ l{-}k \\ q \end{array}
\begin{array}{c} \begin{array}{cccc} \rho & r & t & l{-}k \end{array} \\
\begin{pmatrix} I & 0 & 0 & 0 \\ 0 & \Gamma & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \Sigma & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array},
\qquad
B_\perp^* = \begin{array}{c} \\ \rho \\ r \\ t \\ \rho \\ r \\ t \\ l{-}k \\ q \end{array}
\begin{array}{c} \begin{array}{cccc} \rho & r & t & q \end{array} \\
\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\Sigma & 0 & 0 \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \\ 0 & \Gamma & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \end{array}. \qquad (36)
$$

The columns of the matrices $A_\perp^*$ and $B_\perp^*$ span the orthogonal complements of the column spaces of A and B. (This representation is adapted from a more general representation in [20])

A guided tour of this representation will be useful.

- The first columns of A^* and B^* express the parallelism of P_k and P_l. They correspond to the rows labeled C in the earlier part of the paper.

- The matrices Γ and Σ have the form

$$\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_r) \quad (1 > \gamma_1 \geq \cdots \geq \gamma_r > 0), \tag{37}$$

$$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) \quad (0 < \sigma_1 \leq \cdots \leq \sigma_r < 1), \tag{38}$$

where

$$\Gamma^2 + \Sigma^2 = I.$$

The γ_i and σ_i are the sines and cosines of Jordan's invariant angles. Note that these angles are nonzero and oblique. The third column of A^* contains vectors that are orthogonal to B^*. The third and fourth columns of B^* contain vectors that are orthogonal to A^*.

- The dimensions in the representation are constrained as follows:

  1. $\rho + r + t = k$,

  2. $0 \leq \rho \leq n$,

  3. $0 \leq t \leq n$,

  4. $q = n - (k + l)$.

As we have noted, Jordan excludes the possibility of angles equal to zero or $\pi/2$. There are, however, natural candidates for such angles.

The identity matrix in the first row ρ of column ρ in B^* can be regarded as containing cosines with angle zero, while the sines of these angles are in are contained in the zero matrix in the second row ρ in column ρ. Similarly, the identity in the second row t of column t can be regarded as sines of angles $\pi/2$, while their cosines are contained in the zero matrix in the first row t and the column t.

The question what to do with these angles is not easily resolved. Jordan has decided that neither belongs among his canonical angles. But this has decision has its problems. Consider the angles in column t. They are ephemeral in the sense that arbitrarily small perturbations of A^* and B^* can cause them them to deviate from $\pi/2$, in which case they must be placed among the canonical angles. Otherwise put, the number of canonical angles is sensitive to arbitrarily small perturbations. Note that the potential angles of $\pi/2$ in column $l-k$ are structural in the sense that if $l > k$ then $\mathcal{R}(B)$ *must* have $k-l$ vectors perpendicular to $\mathcal{R}(B)$, and no perturbation that leaves the equations of P_k and P_l respectively independent can cause such vectors to vanish.

If $k+l \leq n$, as we are assuming here, then all the angles in column ρ are ephemeral, and once again they may deviate from zero under arbitrarily small perturbations. Since ρ is the order of parallelism of P_k and P_l, this says that if $l+k \leq n$, then parallelism itself is an ephemeral property. On the other hand if $k+l > n$, then structural vectors appear in column ρ, since there must be at least $k+l-n$ vectors common to both P_k and P_l. In this case parallelism has at least a partial stability.

One can only conjecture why Jordan chose to reject the angles 0 and $\pi/2$. I will give three possibilities.

1. In writing his essay Jordan was wearing his hat as a geometer and considered his planes as objects that were not subject to perturbations.

2. Jordan wanted to give prominence to the number, ρ which is the order of parallelism.

3. Jordan's method for extracting canonical angles will not work if there are angles of 0 and $\pi/2$. For more on this see the comments on §32.

Or perhaps the truth is some combination of the three. In any event, in §49 Jordan is forced to allow zero and $\pi/2$ as invariant angles when establishing the number of invariant angles.

As we have noted earlier, Jordan works entirely with the equations of his multiplanes, and largely ignores the structure of the planes themselves. This structure can be seen in the orthogonal complements in (35) and (36). Briefly, the columns of $B_\perp^*$ labeled ρ and q represent the common subspaces of P_k and P_l [see (3)]. The column labeled r

of $B_\perp^*$ contain the invariant angles, which are the same as those of the equations. The column labeled t in $A_\perp^*$ is orthogonal to the same column of $B_\perp^*$. Finally the column labeled $l-k$ in $A_\perp^*$ is orthogonal to all the columns of $B_\perp^*$.

We are now in a position to understand Jordan's strategy for defining canonical angles. If we denote by A_r^* and B_r^* the second columns of A^* and B^* in (35) and (36), then $A_r^* B_r = \Gamma$. Thus, as observed above, the problem of finding canonical angles can be reduced to stripping off the unwanted columns of A^* and B^* to leave only A_r^* and B_r^*, which contain the canonical angles. Jordan gives a general prescription for accomplishing this. We will now follow his development, illustrating it by showing how it applies to the canonical representations (35) and (36).

Jordan, as might be expected, proceeds by defining a sequence of perpendicular planes along with perpendicular plane generators. We give his definitions here followed by the order of the planes, their defining equations, and finally the names of the functions defining the perpendicular plane generators. Jordan assumes that P_k and P_l have a common point π, which we will take to be zero. He denotes by P_{n-k} the $(n-k)$-plane perpendicular to P_k. This is the plane $A_\perp^* x = 0$ in (35). Similarly, for the plane P_{n-l} is the plane $B_\perp^* x = 0$. For brevity we will call the "ρ-plane" containing two planes the "smallest covering plane." Note that "smallest" refers to the number of equations, not the dimension of the plane.

1. P_ρ is the smallest covering plane for P_k and P_l.

$$\rho = \rho, \quad \begin{pmatrix} I_\rho & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$x_1, \ldots, x_\rho$$

2. P_σ is the smallest covering plane for P_{n-k} and P_{n-l}.

$$\sigma = \rho + n - k - l, \quad \begin{pmatrix} 0 & 0 & 0 & I_\rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_{n-k-l} \end{pmatrix}$$
$$y_1, \ldots, y_\sigma$$

3. P_τ is the smallest covering plane for P_{n-k} and P_l.

$$\tau = t + l - k, \quad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & I_t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & I_{l-k} & 0 \end{pmatrix}$$
$$z_1, \ldots, z_\tau$$

4. P_υ is the smallest covering plane for P_k and P_{n-l}.

$$\upsilon = t, \quad \begin{pmatrix} 0 & 0 & I_t & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$u_1, \ldots, u_\upsilon$$

5. P_α is the plane whose generators are those of P_k that are perpendicular to P_ρ and P_υ.

$$\alpha = r, \qquad \underset{v_1,\ldots,v_\alpha}{\begin{pmatrix} 0 & I_r & 0 & 0 & 0 & 0 & 0 \end{pmatrix}}$$

6. P_β is the plane whose generators are those of P_{n-k} that are perpendicular to P_σ and P_τ.

$$\beta = r, \qquad \underset{w_1,\ldots,w_\beta}{\begin{pmatrix} 0 & 0 & 0 & I_r & 0 & 0 & 0 \end{pmatrix}}$$

7. P_γ is the plane whose generators are those of P_l that are perpendicular to P_ρ and P_τ.

$$\gamma = r, \qquad \begin{pmatrix} 0 & \Gamma & 0 & \Sigma & 0 & 0 & 0 \end{pmatrix}$$

From this we see that P_α (or P_β) and P_γ contain enough information to recover the canonical angles. It should be stressed that the example given here is constrained by the hypotheses $k + l \leq n$. If $k + l > n$, the results would be somewhat different but would still reveal the canonical angles.

At this point Jordan turns to the problem of extracting the canonical angles from P_α and P_γ. In our illustration α, β, and γ are equal, as they must be. But Jordan has to first establish the equality of these dimensions.

The first six multi-planes Jordan constructs above form an orthogonal system spanning the entire space of points. If we make a change of variables corresponding to this system, the individual planes correspond to nonintersecting sets of the variable $x_1, \ldots, x_n$, and we may assume that the variables corresponding to each of the multi-planes are contiguous. In particular, if we label the sets corresponding to P_α and P_β as x_α and x_β, the we can write the equations for P_α, P_β and P_γ in the form

$$\begin{aligned} (I\ \ 0)\begin{pmatrix} x_\alpha \\ x_\beta \end{pmatrix} &= 0, \\ (0\ \ I)\begin{pmatrix} x_\alpha \\ x_\beta \end{pmatrix} &= 0, \end{aligned} \tag{39}$$

and

$$(A\ \ B)\begin{pmatrix} x_\alpha \\ x_\beta \end{pmatrix} = 0. \tag{40}$$

It should be stressed that initially Jordan does not make this change of variables, though he will do so later. It is also worth noting that he *does* make such a change in the example leading to (34).

The problem then is to manipulate these two equation to extract Γ and Σ. Here are the rules of the game.

1. Any transformation must preserve the rank of $(A\ \ B)$. Specifically, we may form the product $M(A\ \ B)$ provided that M is nonsingular.

2. Any transformation must preserve the relation (40) along with perpendicularity of the equations. Thus we may write

$$(AU\ \ BV)\begin{pmatrix} U^* x_\alpha \\ V^* x_\beta \end{pmatrix} = 0.$$

   where U and V are orthogonal.

Jordan now shows that A is nonsingular. On multiplying (40) by A^{-1}, we get

$$(I\ \ C)\begin{pmatrix} x_\alpha \\ x_\beta \end{pmatrix} = 0, \tag{41}$$

where $C = A^{-1}B$.[11] Now Jordan, citing an important paper by Cauchy [3], constructs an orthogonal matrix F such that $F^*(C^*C)F = G^2 = \operatorname{diag}(g_1^2, \ldots, g_\alpha^2)$, where the g_i are positive. It is easily shown that if $D = CFG^{-1}$ then $D^*D = I$ and $D^*CF = G$. Now

$$(I\ \ CF)\begin{pmatrix} x_\alpha \\ F^* x_\beta \end{pmatrix} = 0,$$

and hence on premultiplying by D^* we get

$$(D^*\ \ D^*CF)\begin{pmatrix} x_\alpha \\ F^* x_\beta \end{pmatrix} = (D^*\ \ G)\begin{pmatrix} x_\alpha \\ F^* x_\beta \end{pmatrix} = 0.$$

Finally

$$(D^*D\ \ G)\begin{pmatrix} D^* x_\alpha \\ F^* x_\beta \end{pmatrix} = 0,$$

or

$$(I\ \ G)\begin{pmatrix} D^* x_\alpha \\ F^* x_\beta \end{pmatrix} \equiv (I\ \ G)\begin{pmatrix} x'_\alpha \\ x'_\beta \end{pmatrix} = 0.$$

If we define Θ by $G = \tan(\Theta)$ and premultiply $\Gamma = \cos(\Theta)$, we get

$$(\Gamma\ \ \Sigma)\begin{pmatrix} x'_\alpha \\ x'_\beta \end{pmatrix} = 0. \tag{42}$$

The equation (42) specifically exhibits the cosines and sines of the invariant angles between P_k and P_l. If we wish to preserve the identity matrices in (39), observe that in the case of P_α

$$(D\ \ 0)\begin{pmatrix} x'_\alpha \\ x'_\beta \end{pmatrix} = 0,$$

[11] Note that A, B, and C are not the matrices in (22).

and on premultiplying by D^* we get

$$(I \;\; 0)\begin{pmatrix} x'_\alpha \\ x'_\beta \end{pmatrix} = 0. \tag{43}$$

A similar maneuver applies to P_β.

There are three things to note about this procedure.

- The procedure does not work if there is a canonical angle of $\pi/2$. This is most easily seen from the point of view of our running example. In this case $A = \Gamma$ and $B = \Sigma$. Since $\cos(\pi/2) = 0$, the diagonal matrix Γ must have a zero on its diagonal, and hence is singular. Thus we cannot form $C = A^{-1}B$ in (41), and the process breaks down.

We can cure this problem by interchanging A and B so that the equation for P_γ becomes

$$(B \;\; A)\begin{pmatrix} x_\beta \\ x_\alpha \end{pmatrix} = 0.$$

But if there is a canonical angle of zero then B will also be singular. Thus if P_k and P_l have both angles of zero and $\pi/2$, the process breaks down completely.

It should be added that if there are eigenvalues very near zero and $\pi/2$ then Jordan's procedure theoretically works but will suffer from numerical problems.

- In diagonalizing C (i.e., $D^*CF = G$) Jordan has computed the very useful singular value decomposition. However, it is not numerically the most stable way of proceeding. Jordan gave a better characterization of this decomposition in the year before the paper treated here appeared. (Recall, however, that an abstract of the present paper appeared in 1872 [11].) For more on the singular value decomposition see [19].

- Jordan's procedure can be adapted to calculate the important CS decomposition [18]]. But the procedure suffers from the problems mentioned above.

38,39,40,41,42. In these sections Jordan establishes the orthogonal invariance of his canonical angles. He gives two proofs, of which we will only consider the first.

Jordan's strategy is to take a geometric construction that is obviously orthogonally invariant and express the canonical angles as a function of this construction. For his first proof he considers the angle between the two generating planes for P_α and P_γ.

Specifically let

$$(\ell^* \;\; 0)\begin{pmatrix} x'_\alpha \\ x'_\beta \end{pmatrix} = \ell^* x'_\alpha = 0$$

be a generating plane of P_α as defined in (43). Similarly let

$$m^*(\Gamma \;\; \Sigma)\begin{pmatrix} x'_\alpha \\ x'_\beta \end{pmatrix} = 0.$$

be a generating plane of P_γ as defined in (42). Then the angle θ between these planes is given by

$$\cos^2\theta = \frac{(m^*\Gamma\ell)^2}{\|m\|^2\|l\|^2}.$$

The values of the local maxima and minima of $\cos^2\theta$ regarded as a function of ℓ and m are clearly orthogonal invariants.

Jordan goes on to show by a rather complicated argument that the maxima are attained when $m_i = n_i = \mathbf{e}_i$, where $\mathbf{e}_i$ is the ith column of the identity matrix of order α. The maximum value at that point is γ_i^2. Thus, the γ_i^2, which are the squares cosines of the angles between P_α and P_γ, are invariant under orthogonal substitutions.

Jordan summarizes these results as follows.

> Let two α-planes P_α and P_γ have only a single point in common. If we seek pairs of their generating planes whose angles are [locally] maximal or minimal, we will get two corresponding systems of real perpendicular planes $A_1, \ldots, A_\alpha$ and $A'_1, \ldots, A'_\alpha$. The desired maxima and minima are none other than the angles of the multi-planes P_α and P_γ.

It should be noted that Jordan implicitly assumes that the angles are distinct. If there is a multiple angle, then there is a multiplicity of corresponding planes. Selecting perpendicular pairs from this collection is a nontrivial problem.

Jordan concludes by noting once again that that the it is the squares of the trigonometric functions of the angles that are the invariants and not the functions themselves.

43,44,45,46,47. Jordan has defined the geometric relations of congruence, equality, and symmetry (§28). He now considers the relation of the planes P_α, P_γ in the canonical form of a plane P and the planes P'_α, P'_γ of a plane P' He first observes, that if $2\alpha < n$ and P_α and P'_α have the same invariants then the P_α and P'_α are both identical and reflective. For they one can be transformed to the other by an orthogonal transformation. But the determinant of this transformation can be changed by changing the sign of a coordinate that does not occur in the canonical equations.

The sign-changing option is not available when $2\alpha = n$, and therefore the two planes cannot be identical and reflective at the same time. Jordan now treats the problem of determining which occurs. He begins by transforming the equations of the of the α- and γ-planes into a common system of coordinate planes, so that the equations for the pair P_α and P_γ have the equations

$$v_r = 0 \quad (r = 1, 2, \ldots, \alpha)$$

and

$$A_r = a_{r1}v_1 + \cdots a_{r\alpha}v_\alpha + b_{r1}w_1 + \cdots b_{r\alpha}w_\alpha = 0 \quad (r = 1, 2, \ldots, \alpha)$$

while the pair P_α and P_γ have the equations

$$v_r = 0 \quad (r = 1, 2, \ldots, \alpha)$$

and

$$\mathcal{A}_r = A_{r1}v_1 + \cdots A_{rn}v_n + B_{r1}w_1 + \cdots B_{rn}w_n \quad (r = 1, 2, \ldots, \alpha).$$

He tacitly assumes that the signs of the determinants of the orthogonal substitutions that result in these equations has been noted. He then states that

> *they are identical if the product of the two determinants*
>
> $$\Delta_1 = \begin{vmatrix} A_{11} & \cdots & A_{1\alpha} \\ \vdots & & \vdots \\ A_{\alpha 1} & \cdots & A_{\alpha\alpha} \end{vmatrix} \quad \text{and} \quad \Delta_2 = \begin{vmatrix} B_{11} & \cdots & B_{1\alpha} \\ \vdots & & \vdots \\ B_{\alpha 1} & \cdots & B_{\alpha\alpha} \end{vmatrix}$$
>
> *has the same sign as the product as the product of the determinants*
>
> $$\delta_1 = \begin{vmatrix} a_{11} & \cdots & a_{1\alpha} \\ \vdots & & \vdots \\ a_{\alpha 1} & \cdots & a_{\alpha\alpha} \end{vmatrix} \quad \text{and} \quad \delta_2 = \begin{vmatrix} b_{11} & \cdots & b_{1\alpha} \\ \vdots & & \vdots \\ b_{\alpha 1} & \cdots & b_{\alpha\alpha} \end{vmatrix}.$$
>
> *In the opposite case, they are reflective.*

Jordan's proof of this fact amounts to repeating the reduction in §35 of the planes P_α and P_γ to canonical form by transformations of positive determinant so that $g_2, \ldots, g_\alpha$ are positive, and the sign of g_1 is the same as the sign as $\delta_1\delta_2$. He then does the same thing for P'_α and P'_γ, where the sign of g'_1 has the same sign as $\Delta_1\Delta_2$. But if g_1 and g_2 have the same sign then their canonical forms are identical and the planes are identical. Otherwise, the canonical forms differ only in the sign of g_1 and g'_1, and the planes are reflective.

49. In this section Jordan shows that the planes P_{n-k} and P_{n-l} perpendicular to P_k and P_l have the same canonical invariants. He establishes this fact by exhibiting the specific generating planes for the P_{n-k} and P_{n-l} corresponding to the generating planes for the canonical forms of P_k and P_l.

A curious feature is that the angles in P_{n-k} and P_{n-l} are shifted. This is an artifact of the necessity to keep the α-planes of the complements orthogonal to those of the original. For example, one of the plane planes of P_α is

$$v'_1 \cos\theta_1 + w'_1 \sin\theta_1$$

and the corresponding plane of P_{n-k} is

$$-v'_1 \sin\theta_1 + w'_1 \cos\theta_1$$

which is duly orthogonal. But it is not in canonical form. It can be brought into canonical form by shifting θ_1 by $\frac{\pi}{2}$. This shift, however, does not alter the fact that θ_1 is an invariant of P_{n-k}.

This result answers a question that must occur to the present day reader: what do the canonical angles of the equations of two planes have to do with the planes themselves? As we have seen the the equations of P_k and P_l generate a planes that are perpendicular to their equations. Hence the planes themselves have the same invariants as the equations.

49. Jordan states:

> A system formed from a k-plane P_k and an l-plane P_l has in general α invariants, where α is the smallest of the four numbers k, l, $n-k$, and $n-l$.

There is a puzzle in this statement. The invariant angles are constrained to lie strictly between zero and $\pi/2$. However, if $k \leq l$, it is possible for P_k to be contained entirely in P_l, in which case the only possible invariant angle is zero. Or P_k could be orthogonal to P_l so that an invariant angle can be only be $\pi/2$. In either event (or in the case where there is a mixture of angles zero and $\pi/2$), P_k and P_l, have strictly speaking, have invariant angles.

The answer to the puzzle lies in the weasel words "in general." A little reflections will show that a line P_k that is also in P_l, can be moved out of P_l by an arbitrarily small perturbation of P_k. Likewise any line in P_k that is perpendicular to P_l can be made to loose its perpendicularity by an arbitrarily small perturbation in P_k. Thus unless P_k and P_l have some special relation, one can expect in general that the system will have k invariant angles.

This situation is illustrated in the example based on the decompositions (35) and (36). Specifically P_ρ is the plane common to P_k and P_l, while P_υ is the plane in P_k that is perpendicular to P_l. The rest of P_k is contained in the plane γ. In the particular case (as opposed to the general case) where $\alpha \neq k$, Jordan says,

> we can express the general case by saying that among the k angles of P_k and P_l, there are ρ that are equal to zero and υ that are equal to $\frac{\pi}{2}$.

Also see the comments on page 25.

50. Jordan begins:

> We have seen (§33) that the inquiry into the angles between two arbitrary multi-planes can immediately be reduced to an inquiry into the angles between two alpha planes, where α is at most equal to $\frac{n}{2}$. This last inquiry can be resolved by reducing the two α-planes to their canonical form, as we have done in §34 and the following. But one can also treat the problem directly.

Jordan continues with equations for two α-planes P and Q:

$$Ax = 0 \quad \text{and} \quad Bx = 0,$$

where A and B have the dimensions $\alpha \times n$. The coefficients of a generating plane of P can be written in the form $\ell^* A = 0$, and likewise $m^* B = 0$ is a generating plane of Q. The angle between these two planes is given by

$$s^2 = \cos^2\theta = \frac{(\ell^* A B^* m)^2}{(\ell^* A A^* \ell)(m^* B B^* m)}. \tag{44}$$

The problem is then to find the stationary points of s^2 (§38 ff.). (Actually, Jordan states that the problem is "to find the minimum of" s^2.)

Although Jordan does not mention it, the square in the numerator of (44) creates problems, since any nonzero ℓ and m for which $s = 0$ becomes a stationary point of s^2. In his development, however, he effectively cancels s from s^2. We will not go into the details here.

The result of Jordan's subsequent manipulations is the equation

$$\begin{pmatrix} 0 & AB^* \\ BA^* & 0 \end{pmatrix} \begin{pmatrix} \ell \\ m \end{pmatrix} = s \begin{pmatrix} AA^* & 0 \\ 0 & BB^* \end{pmatrix} \begin{pmatrix} \ell \\ m \end{pmatrix}. \tag{45}$$

This is a symmetric generalized eigenvalue problem in which the matrix on the right is positive definite. Therefore, the problem has real 2α real eigenvalues. The eigenvalues come in pairs of σ and $-\sigma$; for if ℓ and m satisfy (45), ℓ and $-m$ satisfy the same equation with s replaced by $-s$. (Jordan himself states that it "contains pairs of powers of s.")

It worth noting that if A and B are orthonormal, so that $AA^* = BB^* = I$, then the problem reduces to the ordinary eigenvalue problem

$$\begin{pmatrix} 0 & AB^* \\ BA^* & 0 \end{pmatrix} \begin{pmatrix} \ell \\ m \end{pmatrix} = \sigma \begin{pmatrix} \ell \\ m \end{pmatrix}.$$

The eigenvalues of this matrix are the singular values of AB^* and their negatives. This proves the well known fact that if A^* and B^* are orthonormal bases for two subspaces then the eigenvalues AB^* are the cosines of the canonical angles between the subspaces.

Appendix: A proof of Jordan's conditions for perpendicularity.

Here we will give a proof of the sufficiency and necessity of the condition (24) — i.e, $B_2 A_2^* = 0$ — for P_k and P_l to be perpendicular to one another. Applying Jordan's orthogonal procedure described in §21, we may assume that A_2, B_2, and C are orthonormal. The proof is based on the alternate characterization given in (21) of the projection of a vector onto a plane.

For necessity we will need the following lemma.

Let M be a $k\times l$ matrix. Let u_1 and v_1 be vectors such that $u_1^*u_1 = v_1^*v_1 = 1$

$$\gamma = u_1^* M v_1 = \max_{u^*u=v^*v=1} u^* M v^*. \tag{46}$$

Let $U = (u_1\ \ U_2)$ and $V = (v_1\ \ V_2)$ be orthogonal matrices. Then

$$U^* M V = \begin{pmatrix} \gamma & 0 \\ 0 & \hat{M} \end{pmatrix}. \tag{47}$$

This lemma appears in a paper, mentioned above [12], that Jordan published in 1874 — a year before the present paper — in which he shows how a general matrix may be diagonalized by two sided orthogonal transformations. Jordan goes on to apply this lemma to $\hat{M}$, and so on until M has been reduced to diagonal form. The result is what today is called the singular value decomposition, which was also described by Beltrami in 1873 [2] (the discoveries were independent). It should be noted that Jordan does not work in terms of transformations of matrices. Instead, as would be natural in the nineteenth century, he shows how a bilinear form may be diagonalized by orthogonal changes of its variables.

We will assume that the defining equations for P_k and P_l are in the form (22). Without loss of generality, we may assume that $CC^* = I_\rho$, $A_2A_2^* = I_{k-\rho}$ and $B_2B_2^* = I_{l-\rho}$. Then according to (20), the projection of a point y onto P_k is

$$\left[\begin{pmatrix} C \\ A_2 \end{pmatrix} (C^*\ \ A_2^*)\right]^{-1} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} - \left\{ I - (C^*\ \ A_2^*) \begin{pmatrix} C \\ A_2 \end{pmatrix} \left[\begin{pmatrix} C \\ A_2 \end{pmatrix} (C^*\ \ A_2^*)\right]^{-1} \begin{pmatrix} C \\ A_2 \end{pmatrix} \right\} y.$$

But since a_1 and a_2 are zero and since by orthonormality

$$\left[\begin{pmatrix} C \\ A_2 \end{pmatrix} (C^*\ \ A_2^*)\right]^{-1} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

the formula for the projection reduces to

$$(I - C^*C - A_2^*A_2)y \equiv Q_k y.$$

Similarly the projection of y on P_l is given by

$$(I - C^*C - B_2^*B_2)y \equiv Q_l y.$$

If P_l is to be perpendicular to P_k, a vector z in P_l when projected onto P_k must also lie in P_l. Now for any vector y, $Q_l y$ lies in P_l, and moreover any vector in P_l can be so represented. If we project this vector onto P_k we get the vector Q_kQ_ly. But by perpendicularity, this projection must also lie in P_l. Hence $Q_l Q_k Q_l y = Q_k Q_l y$ Since this last equation must hold for all y, it follows that $Q_l Q_k Q_l = Q_k Q_l$. In terms of the above definitions of Q_k and Q_k, for P_l to be perpendicular to P_k we must have

$$(I - C^*C - B_2^*B_2)(I - C^*C - A_2^*A_2)(I - C^*C - B_2^*B_2) \\ = (I - C^*C - B_2^*B_2)(I - C^*C - A_2^*A_2) \quad (48)$$

We will now simplify this equation. First if we write it in the form

$$(I-C^*C-B_2^*B_2)(I-C^*C-A_2^*A_2)-(I-C^*C-B_2^*B_2)(I-C^*C-A_2^*A_2)(C^*C+B_2^*B_2) \\ = (I - C^*C - B_2^*B_2)(I - C^*C - A_2^*A_2),$$

then we get

$$(I - C^*C - B_2^*B_2)(I - C^*C - A_2^*A_2)(C^*C + B_2^*B_2) = 0$$

Further term-by-term simplification gives (recall that $C^*A_2 = 0$ and $C^*B_2 = 0$)

$$A_2^*A_2B_2^*B_2 = B_2^*B_2A_2^*A_2B_2^*B_2. \quad (49)$$

Now (49) is fully equivalent to (48). Hence if $B_2^*A_2 = 0$, equality holds in (48), and P_l is perpendicular to P_k. Thus $B_2^*A_2 = 0$ is a sufficient condition for P_l to be perpendicular to P_k.

To prove the necessity of the condition, premultiply (49) by A_2 and postmultiply by B_2^* to get

$$A_2B_2^* = (A_2B_2^*)(A_2B_2^*)^*(A_2B_2^*). \quad (50)$$

Set $M = A_2B_2^*$ so that equation (50) becomes $M = MM^*M$. Now if $M \neq 0$, Then by (47) we have $UMV^* = UMV^*VM^*U^*UVM^*$ or

$$\mathrm{diag}(\gamma, \hat{M}) = \mathrm{diag}(\gamma, \hat{M})\mathrm{diag}(\gamma, \hat{M}^*)\mathrm{diag}(\gamma, \hat{M}),$$

where $\gamma \neq 0$ (since $M \neq 0$). In particular, for perpendicularity we must have

$$\gamma = \gamma^3, \quad (51)$$

which is possible only if $\gamma = 1$. But if $\gamma = 1$, then by (46) we must have $(A_2u_1)^*(B_2v_1) = 1$. By orthonormality of $A_2^* = B_2^*$, this implies $A_2^*u_1 = B_2^*v_1$. But by hypotheses no nontrivial linear combinations of the rows of A_2 and B_2 can be the same. Hence $0 < \gamma < 1$, and (51) cannot be satisfied. This shows that the condition $A_2^*B_2 = 0$ is necessary for P_l to be perpendicular to P_k.

Finally, we note that if we interchange B_2 and A in (48) we end up with the condition $A_2^*B_2 = 0$. Which is equivalent to $B_2^*A_2 = 0$. Thus P_k is perpendicular to P_l if and only if P_l is perpendicular to P_k, which, as noted above, is not obvious from the definition of perpendicularity.

AFTERWORD

My main interest in Jordan's paper was to see how he constructed the canonical angles between two subspaces or, as Jordan would have it, the angles between two multi-planes. Consequently, at this point about halfway through Jordan's paper, the translation comes to an end. But Jordan goes on to treat several interesting topics that he mentioned in his introduction. Among these is an determinantal expression for the distance between two multi-planes. It is worth noting that the distance can be computed directly from the representation of two multi-planes that Jordan uses to define his canonical angles between the multi-plane; see [20]. Jordan also gives a canonical form for an orthogonal matrix which consists of a block diagonal matrix of 2×2 plane rotations, followed by a diagonal element of ±1 if the order of the matrix is odd. This decomposition can be computed from the real Schur form of the matrix, for which there is off-the-shelf software [1]. Finally, Jordan introduces the Lie group of orthogonal matrices.

Essay on geometry in n dimensions: by M. Camille Jordan

Session: May 12, 1875

Translated by G. W. Stewart

It is well known that Descarte's merger of analysis and geometry has proved equally fruitful for each of the two disciplines. On the one hand, geometers have learned from their contact with analysis to give their investigations an unprecedented generality. Analysts, for their part, have found a powerful resource in the images of geometry, as much for discovering theorems as for presenting them in a simple, impressive form.

This resource vanishes when one turns to the consideration of functions of more than three variables. Moreover, the theory of these functions is, comparatively speaking, poorly developed. It appears that the time has come to fill this gap by generalizing the results already obtained for the case of three variables. A large number of mathematicians have considered this topic in more or less specialized ways. But we are not aware of any general work on this subject.*

In this essay we propose to show how the formulas for straight lines and planes may be generalized to cover linear functions of an arbitrary number of variables. The study of these elementary topics must naturally precede any investigations concerning functions of higher degrees.

Although these investigations will be purely algebraic, we thought it useful to adopt at the outset certain expressions from geometry. Thus we will take a *point* in n-dimensional space to be defined by the values of the n *coordinates* $x_1, \ldots, x_n$. A single linear equation in these coordinates will define a plane; k simultaneous linear equations will define a *k-plane*; $n-1$ equations, a line. The *distance* between two points will be $\sqrt{(x_1 - x_1')^2 + \cdots}$; etc.

Given these definitions, in section I of this memoir we will treat the various degrees of parallelism that can exist between two multi-planes. In the second section, we will give conditions for perpendicularity, and in the third the formulas for transforming coordinates.

The following sections include results of greater interest. Sections IV and V are devoted to the study of relations that can exist between two multi-planes independent of the choice of axes (with the coordinates remaining rectangular). The main results are summarized in the following propositions.

1. *A system formed from a k-plane and an l-plane passing through a common point in the space has ρ distinct invariants, where ρ is the smallest of the numbers k, l, $n-k$, and $n-l$. One can regard these invariants as defining the angles between the two planes.*

*It seems that a only a part of this subject can be regarded as nearly complete: namely the curvature of surfaces. (See the thesis of M. Morin, 1867 and the memoirs of M. Sophus Lie, *Goettinger Nachrichten, 1871.*)

2. *The various planes perpendicular to P_k and P_l form by their respective intersections an $(n-k)$-plane P_{n-k} and an $(n-l)$-plane P_{n-l} that have the same angles as P_k and P_l.*

3. *If P_k and P_l do not have a point in common, they will have another invariant, namely, the shortest distance between them. This invariant can be written as a fraction whose numerator and denominator are sums of squares of determinants.*

In section VI we give a system of formulas that connect the mutual angles of several multi-planes consisting of n arbitrary planes (all meeting at the same point). For $n = 3$ these formulas reduce to those of spherical trigonometry. We will unify these formulas by a consideration of the determinant of the quadratic form that gives the distance between two points (when the n planes taken as the coordinate planes).

In section VII we will show how an orthogonal change of variables having determinant one can be reduced by a change in the rectangular axes to a simple canonical form that depends on $\frac{n}{2}$ invariants if n is even and $\frac{n-1}{2}$ if n is odd. We will give partial differential equations that are satisfied by these invariants. From these investigations, we will deduce, among other things, generalizations of the following theorems.

Any plane motion may be reduced to a rotation about a point.

Any movement in space is a helicoidal movement.

From this we further obtain a generalization of the law of reciprocity noted by M. Chasles, which serves as the basis for his elegant investigations of the movement of solid bodies.

We conclude by giving the laws of composition of infinitesimal movements in four-dimensional space. The result we obtain is summarized in the following theorem.

A rotation R about a point in a four-dimensional space can be represented by two lines A B in a three-dimensional space of suitable lengths and directions. Two rotations R_1 and R_2 that are represented respectively by the lines A_1 and B_1 and A_2 and B_2 will have a combined rotation represented by a line A depending on A_1 and A_2 and a line B depending on B_1 and B_2 (the lines are combined according the parallelogram rule).

I. Definitions — Parallelism

1. We define the position of a point in an n-dimensional space by n coordinates $x_1, \ldots, x_n$.

One linear equation in these coordinates defines a *plane.* Two simultaneous linear equations that are distinct and not incompatible define a *biplane; k* equations, a *k-plane; $n-1$* equations define a *line; n* equations define a point.

By the generic term *multi-plane* we will understand any of the above geometric entities.

2. Let

$$A_1 = 0, \ldots, A_k = 0 \tag{52}$$

be the equations of a k-plane P_k. If these equations are combined linearly, we get an infinite number of equations of the form

$$\lambda_1 A_1 + \cdots + \lambda_k A_k = 0.$$

The planes represented by these equations clearly have P_k as their common intersection, and for short we will call them the *generating planes* of P_k. Their intersections, taken $2, 3, \ldots, k-1$ at a time, will give an infinite number of biplanes, triplanes, etc. that contain P_k. We may call them the *generating biplanes, triplanes, etc.* of P_k.

It is clear that in place of the equations (52), we can define P_k by the equations of any k generators

$$\lambda_1 A_1 + \cdots + \lambda_k A_k = 0, \quad , \lambda'_1 A_1 + \cdots + \lambda'_k A_k = 0, \ldots,$$

provided that the determinant of the coefficients λ is not zero.

On the other hand, let k' be any integer greater than k but not greater than n. If we adjoin $k' - k$ new equations $A_{k+1} = 0, \ldots, A_{k'} = 0$ to the equations (52) that determine P_k, then the set of these equations (assumed distinct and not incompatible) determines a k'-plane that lies entirely in P_k. We say that the k'-planes obtained in this manner are *the k'-planes of P_k.*

Finally we note that the standard coordinates of any k-plane

$$A_1 = 0, \ldots, A_k = 0$$

can be expressed as a linear function of $n - k$ independent auxiliary variables. Specifically, it is sufficient to set

$$A_{k+1} = \lambda_1, \ldots, A_n = \lambda_{n-k},$$

where $A_{k+1}, \ldots, A_n$ are arbitrary linear functions of $x_1, \ldots, x_n$. We will then have a system of n equations which allow the these coordinates to be expressed as a function of the new variables λ.

3. *In general a plane is determined by n points.* In fact, the general equation of a plane contains $n+1$ coefficients, whose ratios are determined by n linear equations that are obtained by successively substituting the values of the coordinates of the n given points into the equation.

A k-plane is determined by $n-k+1$ points. Specifically, consider an arbitrary plane that is forced to contain the $n - k + 1$ points. This condition gives $n - k + 1$ linear equations in the $n + 1$ coefficients of the plane. If $n - k + 1$ coefficients are eliminated using the these conditional equations, k arbitrary coefficients remain in the equation of the plane. Therefore, the general equation of any plane that pass through the $n - k + 1$ given points has the form

$$\lambda_1 A_1 + \cdots + \lambda_k A_k = 0, \tag{53}$$

and the k-plane $A_1 = \cdots = A_k = 0$ is the common intersection of these planes passing through the $n - k + 1$ given points. Moreover, this k-plane is the only one to have this property. For if a k-plane P_k contains the given points, its generating planes necessarily contain them, and they will be of the form (53). Thus P_k is the same as the k-plane $A_1 = \cdots = A_k = 0$, which is the common intersection of these planes.

4. *In general two planes given by the equations*

$$A = a_1x_1 + \cdots + a_nx_n + \alpha = 0,$$
$$B = b_1x_1 + \cdots + b_nx_n + \beta = 0$$

will intersect in a biplane. But an exception occurs when

$$\frac{a_1}{b_1} = \cdots = \frac{a_n}{b_n}.$$

For in this case the two equations $A = 0$ and $B = 0$ are incompatible. We then say that the two planes are *parallel.* Finally, if we have

$$\frac{a_1}{b_1} = \cdots = \frac{a_n}{b_n} = \frac{\alpha}{\beta},$$

the two equations $A = 0$ and $B = 0$ cannot generate more than a single plane, and the two planes are not only parallel but coincide.

5. *Let P_k and P_l be two arbitrary multi-planes. If from among the generating planes of P_k there are any that that are parallel to some generating planes of P_l, then they generate a multi-plane.*

Specifically, let

$$C_1 = 0, \ldots, C_\rho = 0$$

be the generating planes that are parallel to generating planes of P_l and are chosen in such a way that:

1. They are mutually independent; i.e., they satisfy no linear identity of the form[12]

$$\lambda_1 C_1 + \cdots + \lambda_\rho C_\rho = 0.$$

2. There is no generating plane of P_k that is independent of the chosen planes and is parallel to a generating plane of P_l.

By definition P_l will have planes

$$C_1 = \delta_1, \ldots, C_\rho = \delta_\rho,$$

[12]This definition of independence for equations is essentially the same as the usual definition of linear independent vectors, also stated in 1876 by Frobenius [10, p.236]. Also see [9, p.232].

among its generating planes that are parallel to their respective predecessors. The multi-plane P_k will have all the planes

$$\lambda_1 C_1 + \cdots + \lambda_\rho C_\rho = 0 \tag{54}$$

among its generating planes. These planes generate the ρ-plane

$$C_1 = 0, \ldots, C_\rho = 0,$$

which we will denote by P_ρ. For its part, P_l will have among its generating planes all planes of the form

$$\lambda_1 C_1 + \cdots + \lambda_\rho C_\rho = \lambda_1 \delta_1 + \cdots + \lambda_\rho \delta_\rho \tag{55}$$

which are parallel to their predecessors that generate the ρ-plane

$$C_1 = 0, \ldots, C_\rho = 0,$$

and which we will denote by P'_ρ.

Therefore, for a generating plane of P_k to be parallel to generating plane of P_l it is not only necessary but also sufficient that it be a generating plane of P_ρ. Conversely, for a generating plane of P_l to be parallel to one of those of P_k is is necessary and sufficient that it be one of the generating planes of P'_ρ. We express this relation by saying that the two multi-planes P_k and P_l have a common *parallelism of order* ρ.

Speaking absolutely, we will say that P_k is *parallel* to P_l if all its generating planes are parallel to those of P_l. For this to be true it is obviously necessary that $k \le l$. If $k = l$, it is clear that P_l will also be parallel to P_k.

If $\delta_1 = 0, \ldots, \delta_\rho = 0$ [In the text, $\lambda_1 = 0, \ldots, \lambda_\rho = 0$] simultaneously, the various planes of P_ρ will not only be parallel to those of P'_ρ, but they will be the same, and P_l and P_k will be contained in the same ρ-plane P_ρ.

Otherwise, if, say, δ_1 is different from zero, then for the planes (54) and (55) to coincide we must have the relation

$$\lambda_1 \delta_1 + \cdots + \lambda_\rho \delta_\rho = 0.$$

Solving this relation for λ_1 and substituting it into (54), we get we get the following relation for the generating planes common to P_k and P_l:

$$\lambda_2 \left(C_2 - \frac{\delta_2}{\delta_1} C_1 \right) + \cdots + \lambda_\rho \left(C_\rho - \frac{\delta_\rho}{\delta_1} C_1 \right) = 0.$$

From this we see that these planes are none other than the generating planes of the (ρ−1)-plane

$$C_2 - \frac{\delta_2}{\delta_1} C_1 = 0 \cdots C_\rho - \frac{\delta_\rho}{\delta_1} C_1 = 0$$

that contain P_l and P_k.

6. Let us now look for conditions that must be satisfied by the coefficients of the equations of P_k and P_l for these planes to have a mutual parallelism of order ρ and for them to be or not to be contained in the same ρ-plane.

Let

$$\begin{aligned} A_1 &= a_{11}x_1 + \cdots + a_{1n}x_n + \alpha_1 = 0, \\ &\cdots \\ A_k &= a_{k1}x_1 + \cdots + a_{kn}x_n + \alpha_k = 0, \end{aligned} \tag{56}$$

be the equations of P_k and let

$$\begin{aligned} B_1 &= b_{11}x_1 + \cdots + b_{1n}x_n + \beta_1 = 0, \\ &\cdots \\ B_l &= b_{l1}x_1 + \cdots + b_{ln}x_n + \beta_k = 0, \end{aligned} \tag{57}$$

be those of P_l. For a generating plane of P_k to be parallel to a generating plane of P_l, the following identity [in x] must hold:

$$\lambda_1 A_1 + \cdots + \lambda_k A_k = \mu_1 B_1 + \cdots + \mu_l B_l + \text{constant}.$$

If we equate separately the coefficients of the variables to zero, we get[13]

$$\begin{aligned} \lambda_1 a_{11} + \cdots + \lambda_k a_{k1} &= \mu_1 b_{11} + \cdots + \mu_l b_{l1}, \\ &\cdots \\ \lambda_1 a_{1n} + \cdots + \lambda_k a_{kn} &= \mu_1 b_{1n} + \cdots + \mu_l b_{ln}. \end{aligned} \tag{58}$$

If we want the above generating planes to be not only parallel but also coincident, then the constant must be zero, and hence we get the additional equation

$$\lambda_1 \alpha_1 + \cdots + \lambda_k \alpha_k = \mu_1 \beta_1 + \cdots + \mu_l \beta_l. \tag{59}$$

7. Given these results, suppose first of all that $k + l \leq n$. Then the number of parameters λ and μ will be less than or equal to the number of equation in (58). Therefore, if these equations are distinct, which is the typical case, then one cannot satisfy them except by setting all the parameters to zero. Hence in this case P_k and P_l will not in general have any parallelism.

[13]This argument, which, it seems, derives (58) by setting $x_1, \ldots, x_n$ successively to one while holding the other x's zero, will not work, since it ignores the constant terms in the equation above (58). An alternative is to note that, according to Article 5, the coefficients of the x_i of two parallel planes be proportional. If, without loss of generality we assume they are equal, then (58) follows directly. Moreover, for the planes to coincide their constant terms must be equal, which implies (59).

But this conclusion does not hold if the coefficients a, b have been determined in such a way that the equations (58) reduce to a number p of distinct equations that is less than $k+l$. Moreover, it is easy to see that that p cannot be less than l. Specifically, since the equations (57), which define P_l, are assumed to be distinct and mutually consistent, at least one of the the determinants obtained by taking l columns from the tableau

$$\begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{l1} & \cdots & b_{ln} \end{vmatrix}$$

must be nonzero: say that

$$\Lambda = \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \vdots & & \vdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix} \neq 0.$$

With this choice, the first l equations of (58) are distinct and allow $\mu_1, \ldots, \mu_l$ to be determined as functions of $\lambda_1, \ldots, \lambda_k$; for the determinant of the coefficients multiplying $\mu_1, \ldots, \mu_l$ is equal to Λ, which is nonzero.

Therefore, let $p = k + l - \rho$ and $\rho \leq k$. Then equation (58) allows the determination of $\mu_1, \ldots, \mu_l$ and $k - \rho$ of the quantities λ, say $\lambda_{\rho+1}, \ldots, \lambda_k$, as functions of the ρ free parameters $\lambda_1, \ldots, \lambda_\rho$.

Let

$$\begin{gathered} \lambda_{\rho+1} = m_{\rho+1}\lambda_1 + \cdots + n_{\rho+1}\lambda_\rho, \\ \cdots \\ \lambda_k = m_k\lambda_1 + \cdots + n_k\lambda_k. \end{gathered} \tag{60}$$

If we substitute these values in the expression

$$\lambda_1 A_1 + \cdots + \lambda_k A_k$$

for the plane generators of P_k and for short set

$$\begin{gathered} C_1 = A_1 + m_{\rho+1}A_{\rho+1} + \cdots + m_k A_k, \\ \cdots \\ C_\rho = A_\rho + n_{\rho+1}A_{\rho+1} + \cdots + n_k A_k, \end{gathered} \tag{61}$$

then the equation

$$\lambda_1 C_1 + \cdots + \lambda_\rho C_\rho = 0 \tag{62}$$

defines the generating planes of P_k that are parallel to those of P_l. Moreover, the planes $C_1 = 0, \cdots, C_\rho = 0$ are mutually independent. For if there were an identity of the form

$$\nu_1 C_1 + \cdots + \nu_\rho C_\rho = 0,$$

then by replacing the values of $C_1, \dots, C_\rho$ by the values from equation (61), we would obtain a new identity of the form

$$\nu_1 A_1 + \cdots + \nu_\rho A_\rho + \nu_{\rho+1} A_{\rho+1} + \cdots = 0.$$

This is a contradiction, since the equations $A_1 = 0, \dots, A_k = 0$ were assumed to be distinct.

The planes defined by equation (62) are therefore none other than precisely the generating planes of the ρ-plane $C_1 = 0, \dots, C_\rho = 0$, and P_k and P_l have parallelism of order ρ. Moreover, P_k and P_l will lie in the same ρ-plane provided (59) is a consequence of (58).

8. Suppose now that $k + l > n$. In the general case where the equations in (58) all are all distinct, we will have a parallelism of order $k + l - n$. But if these equations can be reduced to p distinct equations, the parallelism is of order $k + l - p$. Finally, P_k and P_l will lie in the same ρ-plane provided (59) is a consequence of (58).

9. Consequently, to write down the conditions for a parallelism of any order, we need only note that the equations (58) reduce to p distinct equations. But it is well known that for this to be true it is necessary and sufficient that 1) at least one of the minors of degree p formed from the coefficients of these equations be nonzero, and 2) all the minors of degree $p + 1$ vanish.

In addition, for P_k and P_l to be in a single ρ-plane it is necessary that when equation (59) is adjoined to the system of equations (58) the minors of order $p + 1$ containing coefficients of the new equation vanish along with the original minors.

10. *If two multi-planes P_k and P_l have no parallelism, they intersect in a multi-plane P_{k+l}.*

Specifically, by hypothesis, the equation

$$\lambda_1 A_1 + \cdots + \lambda_k A_k = \mu_1 B_1 + \cdots + \mu_l B_l + \text{constant}$$

cannot be satisfied. Equivalently, neither can the equation

$$\lambda_1(A_1 - \alpha_1) + \cdots + \lambda_k(A_k - \alpha_k) = \mu_1(B_1 - \beta_1) + \cdots + \mu_l(B_l - \beta_l)$$

be satisfied. Therefore the $k + l$ functions $A_1 - \alpha_1, \dots, A_k - \alpha_k$, $B_1 - \beta_1, \dots, B_l - \beta_l$ are distinct, and the equations

$$A_1 = 0, \dots, A_k = 0, B_1 = 0, \dots, B_l = 0,$$

which can be written in the form

$$A_1 - \alpha_1 = -\alpha_1, \dots, B_l - \beta_l = -\beta_l,$$

are distinct and mutually compatible, and therefore determine a $k + l$ plane.

11. *If P_k and P_l are parallel of order ρ and do not have the same ρ-plane, then they do not intersect.*

Specifically, P_k and P_l have, by hypothesis, generating planes $C_1 = 0$ and $C_1 = \delta_1$ that are parallel but not coincident. The points of intersection have to satisfy these two equations, which are incompatible. Thus if P_k and P_l are parallel of order ρ and intersect, then they have the same ρ-plane.

12. *If P_k and P_l have the same ρ-plane, then they intersect in a $(k+l-\rho)$-plane.*

Specifically, let

$$C_1 = \cdots = C_\rho = 0$$

be the ρ-plane formed from the common generating planes of P_k and P_l. One can replace the equations

$$B_1 = \cdots = B_l = 0,$$

which defines P_l by the equivalent equations

$$C_1 = 0, \ldots, C_\rho = 0, B_{\rho+1} = 0, \ldots, B_l = 0,$$

in which the first ρ equations are solely combinations of the equations $A_1 = 0, \ldots, A_k = 0$, which define P_k.[14] Thus the number of distinct distinct equations satisfied by the intersection of P_k and P_l are reduced to $k + l - \rho$.

13. *A k-plane P_k sliding over an l-plane while remaining parallel to itself it produces a multi-plane.*

Suppose the two multi-planes are defined by equations (56) and (57). Then the equations of a k-plane passing through the point $\xi_1, \ldots, \xi_n$ are

$$\begin{aligned} A_1 &= a_{11}\xi_1 + \cdots + a_{1n}\xi_n + \alpha_1 [= 0], \\ &\cdots \\ A_k &= a_{k1}\xi_1 + \cdots + a_{kn}\xi_n + \alpha_n [= 0]. \end{aligned} \tag{63}$$

But if the point $\xi_1, \ldots, \xi_n$ is to belong to P_l, then

$$\begin{aligned} B_1 &= b_{11}\xi_1 + \cdots + b_{1n}\xi_n + \beta_1 = 0, \\ &\cdots \\ B_l &= b_{l1}\xi_1 + \cdots + b_{ln}\xi_n + \beta_n = 0. \end{aligned} \tag{64}$$

Therefore, on eliminating $\xi_1, \ldots, \xi_n$ in (63) and (64), we obtain equations — linear equations — for the desired locus.

[14] This is an example of what is now called the Steinitz exchange lemma.

To fix our ideas, suppose that P_k and P_l are in the same ρ-plane P_ρ. We may assume that the planes $C_1 = 0, \ldots, C_\rho = 0$ are among the planes chosen to define P_k and P_l. Hence we may assume that

$$A_1 = B_1 = C_1, \ldots, A_\rho = B_\rho = C_\rho.$$

But then the right-hand sides of the first ρ equations in (63) are zero because of (64). Hence the equations

$$A_1 = 0, \ldots, A_\rho = 0$$

are among the equations for the desired locus, and the other equations may be obtained by eliminating the n variables ξ among the last $k - \rho$ equations of the system (63) and the last l equations of the system (64). This elimination gives $l + k - \rho$ [$l + k - \rho - n$ in the original] equations that suffice to determine the desired locus.

II. Distance and Perpendicularity

14. *The distance between two points* whose coordinates are respectively $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ is defined by the formula

$$\Delta = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

The *distance between a point p and a multi-plane* is its distance to the point of the multi-plane that is nearest it. This point q is the *projection* of the point p onto the multi-plane.

The *distance between two multi-planes* that do not intersect is the distance between the nearest neighbors of their points.

The *projection of a multi-plane onto another* is the locus of the projections of its points.

15. We will now try to find the coordinates $x_1, \ldots, x_n$ of the projection q of a point p whose coordinates are $y_1, \ldots, y_n$ onto the multi-plane P_k whose equations are

$$\begin{gathered} a_{11}x_1 + \cdots + a_{1n}x_n + \alpha_1 = 0, \\ \cdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n + \alpha_n = 0. \end{gathered} \tag{65}$$

Since the point q is in P_k, its coordinates satisfy (65). On the other hand, the expression

$$\Delta^2 = (x_1 - y_1)^2 + \cdots + (x_n - y_n)^2$$

must be smaller than for neighboring point q' lying in P_k and having coordinates $x_1 + dx_1, \ldots, x_n + dx_n$.

Setting the differential of Δ^2 to zero, we get the equation

$$(x_1 - y_1)dx_1 + \cdots + (x_n - y_n)dx_n = 0. \tag{66}$$

Moreover, since q' lies in P_k, the differentials $dx_1, \dots, dx_n$ are not arbitrary but must satisfy

$$\begin{gathered} a_{11}dx_1 + \cdots + a_{1n}dx_n = 0, \\ \cdots \\ a_{k1}dx_1 + \cdots + a_{kn}dx_n = 0. \end{gathered} \tag{67}$$

Since equation (66) must be satisfied whenever (67) is, it is a linear combination of the latter. Hence,

$$\begin{gathered} x_1 - y_1 = \lambda_1 a_{11} + \cdots + \lambda_k a_{k1}, \\ \cdots \\ x_n - y_n = \lambda_1 a_{1n} + \cdots + \lambda_k a_{kn}, \end{gathered} \tag{68}$$

where $\lambda_1, \dots, \lambda_k$ are suitable multipliers. If we eliminate $\lambda_1, \dots \lambda_k$ from these equations, there remain $n - k$ distinct equations

$$C_1 = 0, \dots, C_{n-k} = 0 \tag{69}$$

among the $x_1, \dots, x_n$, which along with the equations (65) completely determine these quantities.

16. Suppose now that instead of being unknowns to be determined $x_1, \dots, x_n$ are the coordinates of a given point of P_k. The equations (68) or (69), in which $y_1, \dots, y_n$ are regarded as variables, represent the locus of points in space whose projection falls on the point $x_1, \dots, x_n$. Hence this locus is a $(n-k)$-plane, say P_{n-k}. We will say that that P_{n-k} *is the is the $(n-k)$-plane perpendicular to P_k over the point $x_1, \dots, x_n$.*

In general, an l-plane P_l will be said to be *perpendicular to a k-plane* P_k if given two planes P'_l and P'_k that are parallel to P_l and P_k respectively and pass through an arbitrary point q, the projections of each point of P'_l onto P'_k lies in intersection of P'_l and P'_k.

17. We will now try and establish conditions for perpendicularity. As before, let be the equations of P'_k be

$$\begin{gathered} A_1 = a_{11}x_1 + \cdots + a_{1n}x_n + \alpha_1 = 0, \\ \cdots \\ A_k = a_{k1}x_1 + \cdots + a_{kn}x_n + \alpha_n = 0, \end{gathered} \tag{70}$$

and let those of P'_l be

$$\begin{gathered} B_1 = b_{11}x_1 + \cdots + b_{1n}x_n + \beta_1 = 0, \\ \cdots \\ B_l = b_{l1}x_1 + \cdots + b_{ln}x_n + \beta_n = 0. \end{gathered} \tag{71}$$

An arbitrary point of P'_l, say $y_1, \dots, y_n$ is connected with its projection by the relations (68). For P'_l to be perpendicular to P'_k it is by definition necessary that $x_1, \dots, x_n$ satisfy the relations (71) along with $y_1, \dots, y_n$. Hence, we have the equations

$$\begin{aligned} b_{11}(x_1 - y_1) + \cdots + b_{1n}(x_n - y_n) = 0, \\ \cdots \\ b_{l1}(x_1 - y_1) + \cdots + b_{ln}(x_n - y_n) = 0. \end{aligned} \tag{72}$$

[For P_l to be perpendicular to P_k,] these equations must be a consequence of (68) whenever $x_1, \dots, x_n$ satisfies (70) and $y_1, \dots, y_n$ satisfies (71).

18. First suppose that P_k and P_l have no parallelism; i.e., P'_k and P'_l have no common generating plane. The equations (70), which are satisfied by $x_1, \dots, x_n$ and the equations

$$\begin{aligned} B'_1 = b_{11}y_1 + \cdots + b_{1n}y_n + \beta_1 = 0 \\ \cdots \\ B'_l = b_{l1}y_1 + \cdots + b_{ln}y_n + \beta_n = 0 \end{aligned} \tag{73}$$

do not furnish any relation among the quantities $x_1 - y_1, \dots, x_n - y_n$.

Specifically, suppose we can deduce a relation of the form

$$\mu_1 A_1 + \cdots + \mu_k A_k + \nu_1 B'_1 + \cdots + \nu_l B'_l = c_1(x_1 - y_1) + \cdots + c_n(x_n - y_n) = 0.$$

If $x_1, \dots, x_n$ and $y_1, \dots, y_n$ are interchanged, this relation gives the identity

$$\mu_1 A_1 + \cdots + \mu_k A_k + \nu_1 B_1 + \cdots + \nu_l B_l = 0.$$

Hence P'_l and P'_k have a common generating plane

$$\mu_1 A_1 + \cdots + \mu_k A_k = -(\nu_1 B_1 + \cdots + \nu_l B_l) = 0,$$

which is contrary to hypothesis.

Therefore the equations (72) have to be deduced from the equations (68) alone. If the values of $x_1 - y_1, \dots, x_n - y_n$ in the equations (68) are substituted into the equations (72) and then the coefficients of the indeterminants $\lambda_1, \dots, \lambda_k$ are set to zero,[15] we get the system

$$\begin{aligned} b_{11}a_{11} + \cdots + b_{1n}a_{1n} = 0 \\ \cdots \\ b_{r1}a_{s1} + \cdots + b_{rn}a_{sn} = 0 \\ \cdots \end{aligned} \tag{74}$$

[15]What Jordan means is that for each value of j the coefficients other than λ_j are set to zero while λ_j remains nonzero. Since the λ_j are functions of y, this requires that as y varies of $\mathbb{R}^k$ the vectors $(\lambda_1, \dots, \lambda_k)$ span $\mathbb{R}^n$. The assertion is true, but Jordan does not prove it.

Each of these kl equations taken alone says that one of the planes $B_1, \dots, B_l$ is perpendicular to one of the planes $A_1, \dots, A_l$. Moreover their symmetric form shows that *if P_l is perpendicular to P_k, then P_k is likewise perpendicular to P_l.*[16]

19. More generally, suppose that P_k and P_l have a parallelism of order ρ; i.e., P'_k and P'_l are contained in the same ρ-plane.

Let

$$C_1 = \mu_{11}A_1 + \dots + \mu_{1k}A_k = \nu_{11}B_1 + \dots + \nu_{1l}B_l = c_{11}x_1 + \dots + c_{1n}x_n + \gamma_1 = 0,$$
$$\dots$$
$$C_\rho = \mu_{\rho 1}A_1 + \dots + \mu_{\rho k}A_k = \nu_{\rho 1}B_1 + \dots + \nu_{\rho l}B_l = c_{\rho 1}x_1 + \dots + c_{\rho n}x_n + \gamma_\rho = 0$$

be the equations of the ρ-plane P'_ρ. Let

$$\nu_{11}B_1 + \dots + \nu_{1l}B_l = C_1 - \delta_1 = 0,$$
$$\dots$$
$$\nu_{\rho 1}B_1 + \dots + \nu_{\rho l}B_l = C_\rho - \delta_\rho = 0$$

be the generating planes corresponding to P_ρ. Equations (70) and (73) give the following equations:

$$\mu_{11}A_1 + \dots + \mu_{1k}A_k - \nu_{11}B'_1 - \dots - \nu_{1l}B'_l = c_{11}(x_1 - y_1) + \dots + c_{1n}(x_n - y_n) = 0,$$
$$\dots$$
$$\mu_{\rho 1}A_1 + \dots + \mu_{\rho k}A_k - \nu_{\rho 1}B'_1 - \dots - \nu_{\rho l}B'_l = c_{\rho 1}(x_1 - y_1) + \dots + c_{\rho n}(x_n - y_n) = 0.$$

Substituting the values of $x_1 - y_1, \dots, x_n - y_n$ from equation (68) into the above equations, we get the following conditional equations in the parameters $\lambda_1, \dots, \lambda_k$:

$$D_1 = 0, \dots, D_\rho = 0 \tag{75}$$

In order to obtain the conditions for perpendicularity, we must first substitute the values of $x_1 - y_1, \dots, x_n - y_n$ from equation (68) into (72), then eliminate the parameters $\lambda_1, \dots, \lambda_k$ with the help of equation (75), and finally set the coefficients of the remaining parameters to zero. Then each of equations (72) will decompose into $k - \rho$ distinct equations. Each of these l systems of $k - \rho$ equations express the fact that the point $x_1, \dots, x_n$ lies respectively in the planes $B_1, \dots, B_l$.

But these systems of conditions are not distinct. In fact, just after the point $x_1, \dots, x_n$ has been processed by the planes $C_1 = 0, \dots, C_\rho = 0$, it is sufficient to complete the process by a suitable choice of ρ planes from the planes $B_1, \dots, B_l$. This reduces the number of distinct conditions necessary for perpendicularity to $(l - \rho)(k - \rho)$.[17]

[16] Note that Jordan has proved the necessity of the conditions (74) for A to be orthogonal. He does not prove the sufficiency of these conditions. Yet sufficiency is what he needs to show that the perpendicularity of A to B implies the perpendicularity of B to A.

[17] This is a another example of Jordan's use of the Steinitz exchange lemma.

20. Once again we have the following: *If P_l is perpendicular to P_k, then conversely P_k is perpendicular to P_l.* To make this reciprocity clear, we may suppose that the planes $C_1, \ldots, C_\rho$ have been chosen to be among the planes defining P'_k and P'_l.

In addition, the remaining planes that define these two multi-planes may be chosen so that they are perpendicular to the preceding planes. Specifically, let

$$\pi_1 A_1 + \cdots + \pi_k A_k = 0$$

be one of the generating planes of P'_k. It will be perpendicular to the planes $C_1 = 0, \ldots, C_\rho = 0$ if the equations

$$(\pi_1 a_{i1} + \cdots + \pi_k a_{1k})c_{r1} + \cdots + (\pi_1 a_{1n} + \cdots + \pi_k a_{kn})c_{rn} = 0$$

are satisfied for $r = 1, \ldots, \rho$. These ρ equations determine ρ of the constants π, say $\pi_{k-\rho+1}, \ldots, \pi_k$ as a function of the others.

For example, let

$$\begin{aligned} \pi_{k-\rho+1} &= m_1 \pi_i + \cdots + n_1 \pi_{k-\rho}, \\ \pi_k &= m_\rho \pi_1 + \cdots + n_\rho \pi_{k-\rho}. \end{aligned}$$

The desired planes are given by the formula

$$\pi_1(A_1 + m_1 A_{k-\rho+1} + \cdots + m_\rho A_k) + \cdots + \pi_\rho(A_\rho + n_1 A_{k-\rho+1} + \cdots + n_\rho A_k) = 0.$$

This formula represents a $(k-\rho)$-plane from which we may freely choose $(k-\rho)$ planes to define P'_k [i.e., by varying $\pi_1, \ldots, \pi_\rho$].[18]

Therefore let

$$\begin{aligned} A_1 = C_1 &= c_{11}x_1 + \cdots + c_{1n}x_n + \gamma_1 = 0, \\ &\cdots \\ A_\rho = C_\rho &= c_{\rho 1}x_1 + \cdots + c_{\rho n}x_n + \gamma_\rho = 0, \\ A_{\rho+1} &= A_{\rho+1,1}x_1 + \cdots A_{\rho+1,n}x_n + \alpha_{\rho+1} = 0, \\ &\cdots \\ A_k &= A_{k,1}x_1 + \cdots + A_{k,n}x_n + \alpha_k = 0, \end{aligned} \tag{76}$$

be the planes that define P'_k chosen as described above. We then have a sequence of conditional equations of the form

$$c_{r1}a_{s1} + \cdots + c_{rn}a_{sn} = 0 \tag{77}$$

[18] When $\rho = 1$, Jordan's algorithm amounts to using the Gram–Schmidt method to orthogonalize the remaining generating planes of A to C_1.

Likewise, we can define P_l' by means of the planes

$$\begin{aligned} &C_1 = 0, \cdots, C_l = 0 \\ &B_{\rho+1} = B_{\rho+1,1}x_1 + \cdots + B_{\rho+1,n}x_n + \beta_{\rho+1} = 0, \\ &\cdots \\ &B_l = B_{l,1}x_1 + \cdots B_{l,n}x_n + \beta_k = 0, \end{aligned} \tag{78}$$

which have been chosen so that

$$c_{r1}b_{s1} + \cdots + c_{rn}b_{sn} = 0. \tag{79}$$

If we compare the equations (76) and (78) to the equations (70) and (71), we get

$$a_{11} = b_{11} = c_{11}, \ldots, a_{\rho n} = b_{\rho n} = c_{\rho n}. \tag{80}$$

Let us now substituted the values of (68) into the equations (72). If we take into account the equations (80), (77), and (79) and for short set

$$c_{r1}c_{s1} + \cdots + c_{rn}c_{sn} = K_{rs},$$
$$a_{r1}b_{s1} + \cdots + a_{rn}b_{sn} = L_{rs},$$

then we get

$$\begin{aligned} \lambda_1 K_{11} + \cdots + \lambda_\rho K_{1\rho} &= 0, \\ &\cdots \\ \lambda_1 K_{\rho 1} + \cdots + \lambda_\rho K_{\rho\rho} &= 0, \\ \lambda_{\rho+1} L_{\rho+1,1} + \cdots + \lambda_k L_{\rho+1,k} &= 0, \\ &\cdots \\ \lambda_{\rho+1} L_{l,1} + \cdots + \lambda_k L_{lk} &= 0. \end{aligned}$$

The first ρ equations of this system are the result of substituting the values (68) into the equations $C_1 = 0, \ldots, C_\rho = 0$. Moreover, all the values of $\lambda_1, \ldots, \lambda_k$ that satisfy these ρ equations must satisfy as an identity all the other equations of the system. This gives the relations

$$L_{\rho+1,1} = 0, \ldots, L_{lk} = 0.$$

These relations are clearly symmetric with respect to a and b, and they say that each of the planes $A_{\rho+1}, \ldots, A_k$ is perpendicular to each of the planes $B_{\rho+1}, \ldots, B_l$.[19]

21. *Any k-plane P_k can be construed as the intersection of k rectangular planes in an infinite number of ways.*[20]

[19] This derivation suffers from the same problems as the as the derivation in §18. Namely, the values of $\lambda_1, \ldots, \lambda_k$ must by shown to have a sufficiency of independence, and the conditions $L_i = 0$ ($i = \rho + 1, \ldots k$) must be shown to be sufficient.

[20] By "rectangular" Jordan means "mutually perpendicular."

Specifically, let $A_1 = 0$ be an arbitrary generating plane of P_k. We have seen that P_k contains a $(k-1)$-plane P_{k-1} that is perpendicular to A_1.[21] Let A_2 be an arbitrary generating plane of P_{k-1}. It will be perpendicular to A_1. Moreover, P_k contain a $(k-2)$-plane perpendicular to the biplane (A_1, A_2). Let A_3 be one of its plane generators. It will be perpendicular to A_1 and A_2. On can find in P_k a new plane P_{k-2} perpendicular to the three proceeding planes. And so on.

If we set $k = n$ in the above proposition, we see that *we can cause an infinite number of systems of rectangular planes to pass through an arbitrary point in the space.*

22. *Let p be an arbitrary point, q be its projection onto a multi-plane P_k, and r be an arbitrary point of P_k. Then the distances between the three points p, q, and r satisfy the relation*

$$\overline{pr}^2 = \overline{pq}^2 + \overline{qr}^2.$$

Specifically, let $y_1, \ldots, y_n$, $x_1, \ldots, x_n$, and $\xi_1, \ldots, \xi_n$ be the coordinates of p, q, and r, and suppose that P_k is defined by (65). Then

$$\begin{aligned}
\overline{pr}^2 &= (y_1 - \xi_1)^2 + \cdots + (y_n - \xi_n)^2 \\
&= (y_1 - x_1 + x_1 - \xi_1)^2 + \cdots + (y_n - x_n + x_n - \xi_n)^2 \\
&= \overline{pq}^2 + \overline{qr}^2 + 2\{(y_1 - x_1)(x_1 - \xi_1) + \cdots + (y_n - x_n)(x_n - \xi_n)\}
\end{aligned}$$

But if we replace $y_1 - x_1, \ldots, y_n - x_n$ by their values in (68), the term between the braces becomes

$$\lambda_1\{a_{11}(x_1 - \xi_1) + \cdots + a_{1n}(x_n - \xi_n)\} + \cdots + \lambda_k\{a_{k1}(x_1 - \xi_1) + \cdots + a_{kn}(x_n - \xi_n)\}$$

In this expression the multipliers of $\lambda_1, \ldots, \lambda_k$ are clearly zero since $x_1, \ldots, x_n$ and $\xi_1, \ldots, \xi_n$ satisfy (65).

23. *Let P_{k+l} be a multi-plane contained in P_k. The projection of p onto P_{k+l} falls on the same point s as the projection of q.*

In fact, let be r an arbitrary point of P_{k+l}. Then we have from the above

$$\overline{pr}^2 = \overline{pq}^2 + \overline{qr}^2.$$

Moreover, since s is the projection of q on P_{k+l},

$$\overline{qr}^2 = \overline{qs}^2 + \overline{rs}^2.$$

Finally,

$$\overline{ps}^2 = \overline{pq}^2 + \overline{qs}^2.$$

[21] By "contains" Jordan means that there are $k-1$ of the plane generators of A that generate P_{k-1} with the desired properties. These may be obtained by the orthogonalization algorithm described in §20. As a point set, P_{k-1} contains P_k.

All this implies that

$$\overline{pr}^2 = \overline{ps}^2 + \overline{sr}^2.$$

which shows that the point s is the projection of p onto P_{k+1}.[22]

III. Changes of coordinates

24. *Two parallel lines D and D' extending between two parallel planes P and P' have the same length.*

Specifically, let the equations of the line D be

$$\begin{aligned} A_1 &= a_{11}x_1 + \cdots + a_{1n}x_n + \alpha_1 = 0, \\ &\cdots \\ A_{n-1} &= a_{n-1,1}x_1 + \cdots + a_{n-1,n}x_n + \alpha_{n-1} = 0 \end{aligned} \tag{81}$$

and let the equation of the plane P be

$$B = b_1x_1 + \cdots + b_nx_n + \beta = 0. \tag{82}$$

Let the equations of the line D' be

$$A_1 = \alpha'_1, \ldots, A_{n-1} = \alpha'_{n-1} \tag{83}$$

and let the equation of the plane P' be

$$B = \beta'. \tag{84}$$

The coordinates $\xi_1, \ldots, \xi_n$ of the intersection of D with P satisfy equations (81) and (82). The coordinates $\eta_1, \ldots, \eta_n$ of the point of intersection of D with P' satisfy equations (81) and (84). Hence

$$\begin{aligned} & a_{11}(\eta_1 - \xi_1) + \cdots + a_{1n}(\eta_n - \xi_n) = 0, \\ & \cdots \\ & a_{n-1,1}(\eta_1 - \xi_1) + \cdots + a_{n-1,n}(\eta_n - \xi_n) = 0, \\ & b_1(\eta_1 - \xi_1) + \cdots + b_n(\eta_n - \xi_n) + \beta' = 0, \end{aligned} \tag{85}$$

The coordinates $\xi'_1, \ldots, \xi'_n$ and $\eta'_1, \ldots, \eta'_n$ of the points of intersection of D' with P and P' satisfy the same equations. Hence

$$\eta'_1 - \xi'_1 = \eta_1 - \xi_1, \ldots, \eta'_n - \xi'_n = \eta_n - \xi_{n1},$$

[22] Jordan assumes here that if the last equality holds for all $r \in P_{k+l}$ then s is the projection of p onto P_{k+1}, which is true but requires proof.

whence

$$(\eta_1' - \xi_1')^2 + \cdots + (\eta_n' - \xi_n')^2 = (\eta_1 - \xi_1)^2 + \cdots + (\eta_n - \xi_n)^2,$$

QED.

In addition, we note that the values of $\eta_1 - \xi_1, \ldots, \eta_n - \xi_n$ obtained from (85) vary in proportion to the constant β'. From this we have the following conclusion.

Three parallel planes divide two arbitrary lines proportionally.

25. *The locus of points whose distance along a given direction from a fixed plane P is constant will necessarily be a plane parallel to P.*

Given a system of n independent planes $P, Q, \ldots$, the position of a point in space is completely determined when one knows its distance from each of these planes, each direction being taken along, for example, the direction of the intersection of the $n-1$ other planes. Specifically, [according to the statement italicized above] these distances determine n planes parallel to $P, Q, \ldots$ whose intersection is the point in question.

The above distances, $X_1, \ldots, X_n$, form a new system of coordinates which can be used instead of $x_1, \ldots, x_n$ to define the points in space.

26. The new coordinates are related to the old by linear equations. Specifically, let

$$\begin{gathered} a_{11}x_1 + \cdots + a_{1n}x_n + \alpha_1 = 0, \\ \cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n + \alpha_n = 0, \end{gathered} \tag{86}$$

be the equations of the planes $P, Q, \ldots$. Let $\xi_1, \ldots, \xi_n$ be the coordinates of an arbitrary point. The line parallel to the intersection the planes $Q, \ldots$ passing through this point has the equations

$$a_{r1}(x_1 - \xi_1) + \cdots + a_{rn}(x_n - \xi_n) = 0 \qquad (r = 2, \ldots, n). \tag{87}$$

In addition, the coordinates $x_1, \ldots, x_n$ of the intersection of this parallel line with P satisfy the equation P or, what amounts to the same thing,

$$a_{11}(x_1 - \xi_1) + \cdots + a_{1n}(x_n - \xi_n) = a_{11}\xi_1 + \ldots + a_{1n}\xi_n - \alpha_1.^{23} \tag{88}$$

[23]The right-hand side of this equation should be $-(a_{11}\xi_1 + \ldots + a_{1n}\xi_n + \alpha_1)$. This means that the definition of L below and should be modified, as should the right-hand side of (89).

For brevity set

$$a_{11}\xi_1 + \ldots + a_{1n}\xi_n - \alpha_1 = L,$$

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \Delta,$$

$$\frac{d\Delta}{da_{rs}} = b_{rs},$$

$$\sqrt{b_{r1}^2 + \cdots + b_{rn}^2} = M_r.$$

Then the following can be derived from (87) and (88):

$$x_1 - \xi_1 = \frac{b_{11}L}{\Delta}, \ldots, x_n - \xi_n = \frac{b_{1n}L}{\Delta},$$

$$X_1 = [\pm]\sqrt{(x_1 - \xi_1)^2 + \cdots + (x_n - \xi_n)^2} = \frac{M_1}{\Delta}L, = \frac{M_1}{\Delta}(a_{11}\xi_1 + \cdots + a_{1n}\xi_n - \alpha_i).$$

In the same way, we find for all values of r that

$$X_r = \frac{M_r}{\Delta}(a_{r1}\xi_1 + \cdots + a_{rn}\xi_n - \alpha_r). \quad (89)$$

Inverting these equalities we get

$$\xi_r = \frac{b_{1r}}{M_1}X_1 + \cdots + \frac{b_{nr}}{M_n}X_n + \frac{b_{1r}\alpha_1 + \cdots + b_{nr}\alpha_n}{\Delta}. \quad (90)$$

27. It is easy to express the distance between two points as a function of the new coordinates. Specifically, let $\xi_1', \ldots, \xi_n'$ and $X_1', \ldots, X_n'$ be the coordinates of the second point. Since they satisfy (87) and (90), we have

$$\xi_r - \xi_r' = \frac{b_{1r}}{M_1}(X_1 - X_1') + \cdots + \frac{b_{nr}}{M_n}(X_n - X_n').$$

The distance D between the two points is given by the formula

$$D^2 = \sum_r (\xi_r - \xi_r')^2 = \sum_r (X_r - X_r')^2 + 2\sum_{rs} \frac{b_{r1}b_{s1} + \cdots + b_{rn}b_{sn}}{M_r M_s}(X_r - X_r')(X_s - X_s').$$

This formula simplifies when the coordinate new planes are orthogonal. Specifically, by definition,

$$a_{r1}a_{s1} + \cdots a_{rn}a_{sn} = 0 \quad \text{if } r \neq s. \quad (91)$$

Moreover, without any change to the system of new planes we can multiply the equation of each by a constant so that

$$a_{r1}^2 + \cdots + a_{rn}^2 = 1.$$

But then $a_{11}, \ldots, a_{nn}$ will be the coefficients of an orthogonal substitution. From the properties of such substitutions we have

$$\begin{gathered} a_{r1}a_{s1} + \cdots + a_{rn}a_{sn} = 0, \qquad a_{1r}a_{sr} + \cdots a_{nr}a_{ns} = 0, \\ \Delta = \pm 1, \quad \frac{b_{rs}}{\delta} = a_{sr}, \quad M_r = \sqrt{\Delta^2(a_{1r}^2 + \cdots + a_{nr}^2)} = 1, \\ \frac{b_{r1}b_{s1} + \cdots + b_{rn}b_{sn}}{M_r M_s} = \Delta^2(a_{1r}a_{1s} + \cdots + a_{nr}a_{ns}) = 0, \end{gathered} \tag{92}$$

so that D^2 reduces to a sum of squares.

Conversely, if the equations (92) are satisfied, the equations (90) represents an orthogonal substitution, and the inverse substitution defined by (89) will also be orthogonal. Hence the equations (91) will be satisfied, and the new coordinate planes will therefore be mutually perpendicular.

IV. Invariant Angles

28. The transformation of coordinates in the space of n dimensions offers the same benefits as in ordinary geometry. It allows one to change the equations representing a single figure and to simplify them by a suitable choice of the variables that appear in the transformations.

Two arbitrary figures in space are said to be *congruent* whenever they can be represented by the same equations with respect to systems of orthogonal coordinates suitably chosen for each.

If the orthogonal systems by which the two figures are represented are such that one can pass from one to the other by a orthogonal substitution of determinant 1 [*coefficient* in the original], the two figures are said to be *identical,* and they differ only in their position in space. They are *reflective* if the determinant of the substitution is equal to -1.

29. *All k-planes* are simultaneously identical and reflective. Specifically, we have seen that any k-plane can be regarded as resulting of k intersecting planes that are mutually perpendicular. From arbitrary point of P_k we construct the P_{n-k} plane that is perpendicular to it, which will be the intersection of $n-k$ mutually perpendicular planes that will also be perpendicular to the preceding planes. If we take these n planes so defined as coordinate planes, the equations of P_k assume the form

$$x_1 = 0, \ldots, x_k = 0.$$

Since the equations of an arbitrary k-plane can be reduced to the same form, this k-plane will be identical or reflective to P_k. But at the same time it will be identical and reflective. For P_k is itself reflective, since its equations do not change when x is changed to $-x$ (an operation that is an orthogonal substitution of determinant -1).

30. We will now consider a system of two intersecting planes P and Q. We will take one of the points of intersection as the origin and the plane P for the plane x_1 [$= 0$]. For the plane x_2 [$= 0$] we will take a plane perpendicular [rectangular] to the first plane and passing through the biplane (P, Q). For the other coordinate planes we will take $n-2$ rectangular planes associated with the $(n-2)$-plane perpendicular to (P, Q). The equations of the two planes then will have the form

$$x_1 = 0,$$
$$ax_1 + bx_2 = 0.$$

If we divide the second equation by $\sqrt{a^2+b^2}$, which is permissible, and set $\frac{a}{\sqrt{a^2+b^2}} = \cos\alpha$ and $\frac{b}{\sqrt{a^2+b^2}} = \sin\alpha$, we have

$$\begin{aligned} x_1 &= 0, \\ x_1\cos\alpha + x_2\sin\alpha &= 0. \end{aligned} \tag{93}$$

Thus the canonical form to which we can reduce the equations of a system of two planes contains an angle α as a parameter. Therefore, systems of two planes are not the same but differ among themselves by a characteristic element.

31. That this must be the case can be seen *a priori* [i.e., without changing of coordinates] Consider the equations of two planes in their general form:

$$P = a_1x_1 + \cdots a_2x_n + \alpha = 0,$$
$$Q = b_1x_1 + \cdots b_2x_n + \beta = 0.$$

Let $y_1, \ldots, y_n$ be the coordinates of an arbitrary point q of Q. It projection $x_1, \ldots, x_n$ onto P is given by the relations

$$a_1(x_1 - y_1) = \lambda a_1, \ldots, a_n(x_n - y_n) = \lambda a_n$$

where λ is a constant to be determined from the condition that $x_1, \ldots, x_n$ satisfy the equation $P = 0$. Equivalently,

$$a_1(x_1 - y_1) + \cdots + a_n(x_n - y_n) = L$$

where, for short, $L = a_1y_1 + \cdots + a_ny_n - \alpha$.

It follows that

$$\lambda = \frac{L}{a_1^2 + \cdots + a_n^2},$$

and the distance between the two points $y_1, \dots, y_n$ and $x_1, \dots, x_n$ is given by the formula

$$D^2 = \lambda^2(a_1^2 + \cdots + a_n^2) = \frac{L^2}{a_1^2 + \cdots + a_n^2}. \tag{94}$$

We will now determine the distance Δ from q to the intersection of the planes P, Q. The projection $z_1, \dots, z_n$ of the point p on its intersection is given by the relations

$$z_1 - y_1 = \lambda a_1 + \mu b_1, \dots, z_n - y_n = \lambda a_n + \mu b_n,$$

the parameters λ, μ being determined by the condition that $z_1, \dots, z_n$ must satisfy the relations $P = 0$, $Q = 0$. Equivalently,

$$a_1(z_1 - y_1) + \cdots + a_n(z_n - y_n) = L,$$
$$b_1(z_1 - y_1) + \cdots + b_n(z_n - y_n) = 0.$$

Hence

$$\lambda \Sigma_r a_r^2 + \mu \Sigma_r a_r b_r = L, \quad \lambda \Sigma_r a_r b_r + \mu \Sigma_r b_r^2 = 0,$$

or

$$\lambda = \frac{\Sigma b_r^2 \cdot L}{\Sigma a_r^2 \Sigma b_r^2 - (\Sigma a_r b_r)^2}, \quad \mu = -\frac{\Sigma a_r b_r \cdot L}{\Sigma a_r^2 \Sigma b_r^2 - (\Sigma a_r b_r)^2}.$$

Substituting these values in the expression

$$\Delta^2 = (\lambda a_1 + \mu b_1)^2 + \cdots + (\lambda a_n + \mu b_n)^2$$

and canceling the common factor $\Sigma_r a_r^2 \Sigma_r b_r^2 - (\Sigma_r a_r b_r)^2$, we get

$$\Delta^2 = \frac{\Sigma_r b_r^2 \cdot L}{\Sigma_r a_r^2 \Sigma_r b_r^2 - (\Sigma_r a_r b_r)^2}. \tag{95}$$

Dividing this equation into (94), we get

$$\frac{D^2}{\Delta^2} = \frac{\Sigma_r a_r^2 \Sigma_r b_r^2 - (\Sigma_r a_r b_r)^2}{\Sigma_r a_r^2 b_r^2} \tag{96}$$

The ratio of $\frac{D^2}{\Delta^2}$ is therefore independent of the position of the point q in plane Q and is therefore equal to a constant K.

Let us now replace the current coordinate by another system, also rectangular. Then $\frac{D^2}{\Delta^2} = K'$, where K' is the function analogous to K formed with the new coefficients. But the distances D and Δ are not changed under orthogonal substitutions, and hence $K = K'$. Thus the function K is invariant under all orthogonal substitutions, and two systems of two planes cannot be congruent if they have different values of this invariant.

If the two planes are reduced to the canonical form (93), then

$$K = \sin^2 \alpha$$

The quantity

$$1 - K = \frac{(\Sigma a_r b_r)^2}{\Sigma a_r^2 \Sigma b_r^2} = \cos^2 \alpha$$

is a second invariant, which, like the first, is reflective with respect to the coefficients of the two planes.

The angle α may be called *the angle of the two planes.*

32. It should be noted that although $\sin^2 \alpha$ and $\cos^2 \alpha$ are invariants, the same is not true of the coefficients $\sin \alpha$ and $\cos \alpha$ in the canonical equation for the plane Q. Specifically, changing the sign of one of the coordinates x_1, x_2, or both, will change the signs of the corresponding coefficients.

33. Let us now turn to the consideration of two multi-planes P_k and P_l having a common point π. In order to treat the problem in full generality, we will assume the following.

1. The multi-planes P_k and P_l lie in a common ρ-plane P_ρ (without lying in a common $(\rho + 1)$-plane).
2. The multi-planes P_{n-k} and P_{n-l} erected at the point π and perpendicular P_k and P_l lie in a common σ-plane P_σ (without lying in a common $(\sigma + 1)$-plane).
3. P_{n-k} and P_l are lie in a common τ-plane P_τ(without lying in a common $(\tau + 1)$-plane).
4. P_{n-l} and P_k are lie in a common υ-plane P_υ (without lying in a common $(\upsilon + 1)$-plane).

We take the following for coordinate planes:

1. ρ rectangular planes

   $$x_1 = 0, \ldots, x_\rho = 0$$

   chosen from the plane generators of P_ρ;
2. σ rectangular planes

   $$y_1 = 0, \ldots, y_\sigma = 0$$

   chosen from the plane generators of P_σ;
3. τ rectangular planes

   $$z_1 = 0, \ldots, z_\tau = 0$$

   chosen from the plane generators of P_τ;

4. υ rectangular planes
$$u_1 = 0, \dots, u_\upsilon = 0$$
chosen from the plane generators of P_υ;

5. $k - \rho - \upsilon = \alpha$ rectangular planes
$$v_1 = 0, \dots, v_\tau = 0$$
chosen arbitrarily from the generators of the multi-plane P_α that are in P_k and are perpendicular to P_ρ and P_υ.

6. $n - k - \sigma - \tau = \beta$ rectangular planes
$$w_1 = 0, \dots, w_\beta = 0$$
chosen from the generators of the multi-plane P_β that are in P_{n-k} and perpendicular P_σ and P_τ.

It is clear that coordinate planes defined above are all mutually perpendicular and that the equation of P_k take the form

$$\begin{gathered} x_1 = 0, \dots, x_\rho = 0 \\ u_1 = 0, \dots, u_\rho = 0 \\ v_1 = 0, \dots, v_\alpha = 0 \end{gathered}$$

The multi-plane P_l is the result of intersecting the planes

$$\begin{gathered} x_1 = 0, \dots, x_\rho = 0 \\ z_1 = 0, \dots, z_\rho = 0 \end{gathered}$$

with a $(l - \rho - \tau = \gamma)$-plane P_γ perpendicular to them. Moreover, since P_γ is perpendicular to P_{n-l}, they will be perpendicular to the planes

$$y_1 = 0, \dots, y_\sigma = 0, \quad z_1 = 0, \dots, z_\mu = 0,$$

which are contained in P_{n-l}. Therefore, the equations of its generating planes have the form

$$\begin{gathered} A_1 = a_{11}v_1 + \dots + a_{1\alpha}v_\alpha + b_{11}w_1 + \dots + b_{1\beta}v_\beta = 0, \\ \dots \\ A_\gamma = a_{\gamma 1}v_1 + \dots + a_{\gamma\alpha}v_\alpha + b_{\gamma 1}w_1 + \dots + b_{\gamma\beta}v_\beta = 0. \end{gathered} \tag{97}$$

Thus the comparison of two multi-planes P_k and P_l reduces to the comparison of the γ-plane P_γ with the α-plane P_α, which is the intersection of the planes

$$v_1 = 0, \dots, v_\alpha = 0.$$

Let us now consider how to carry out this comparison.

34. In the first place, it is easy to see that *the three integers α, β, γ are equal.*

Specifically, if γ were greater than α, we could eliminate $v_1, \ldots, v_\alpha$ in (97) to obtain a generating plane of P_γ of the form

$$c_1 w_1 + \cdots + c_\beta w_\beta = 0. \tag{98}$$

This plane will be perpendicular to P_α and hence to P_k. On adjoining this equation those defining P_τ, we get a $(\tau + 1)$-plane containing both P_{n-k} and P_l, contrary to our assumption [about the maximality of τ].

If γ were less than α, then we could find a $(\alpha{-}\gamma)$-plane in P_α that is perpendicular to P_γ. For the general equation

$$\lambda_1 v_1 + \cdots + \lambda_\alpha v_\alpha$$

of the generating plains of P_α contains α parameters $\lambda_1, \ldots, \lambda_\alpha$, and there are only γ conditions for perpendicularity. Hence $\alpha - \gamma$ parameters remain free. If the equations of this $(\alpha{-}\gamma)$-plane are adjoined to those of P_υ, we obtain a $(\nu{+}\alpha{-}\gamma)$-plane that contains both P_k and P_{n-l}, which is contrary to hypothesis. It follows that $\alpha = \gamma$.

If γ were greater than β, we could eliminate $w_1, \ldots, w_\beta$ in (98) so as to obtain a generating plane of P_γ whose equation is of the form

$$c_1 v_1 + \cdots + c_\alpha v_\alpha = 0.$$

This plane is among the generating planes of P_k. On adjoining one of these planes to those of P_ρ, we get a $(\rho{+}1)$-plane that contains both P_k and P_ρ, which is contrary to hypothesis.

If γ were less than β, we could find an $(\beta{-}\gamma)$-plane in P_β that is perpendicular to P_γ. On adjoining the equations of this $(\beta{-}\gamma)$-plane to those of P_σ, we have a $(\sigma{+}\beta{-}\gamma)$-plane that contains both P_{n-k} and P_{n-l}, which is contrary to hypothesis.

Therefore, $\gamma = \beta = \alpha$, and we can write (97) in the form

$$A_r = a_{r1} v_1 + \cdots + a_{r\alpha} v_\alpha + b_{r1} w_1 + \cdots + b_{r\alpha} v_\alpha = 0 \quad (r = 1, 2, \ldots, \alpha). \tag{99}$$

35. In the second place, *the α functions of the form*

$$B_r = a_{r1} v_1 + \cdots + a_{r\alpha} v_\alpha \tag{100}$$

are all distinct. For if they were related by a linear equation, [say],

$$k_1 B_1 + \cdots + k_\alpha B_\alpha = 0,$$

then P_γ would have a generating plane

$$k_1A_1 + \cdots + k_\alpha A_\alpha = 0,$$

whose equation reduces to the form (100), a result we have shown to be impossible [§34].

Since the functions B_r are distinct, we can solve the equations (97) for $v_1, \ldots, v_\alpha$ and get inverse relations of the form

$$v_r = e_{r1}B_1 + \cdots + e_{r\alpha}B_\alpha \quad (r = 1, 2, \ldots, \alpha).$$

Given this, it is clear that the planes

$$C_r = e_{r1}A_1 + \cdots + e_{r\alpha}A_\alpha = 0 \tag{101}$$

reduce to the form

$$C_r = v_r + b'_{r1}w_1 + \cdots + b'_{r\alpha}w_\alpha = 0 \quad (r = 1, 2, \ldots, \alpha). \tag{102}$$

We may thus regard P_γ as no longer being determined by the intersection of the planes

$$A_1 = 0, \ldots, A_\alpha = 0,$$

but instead by the intersection of the planes

$$C_1 = 0, \ldots, C_\alpha = 0,$$

whose equations have a simpler form.

36. To obtain a further simplification, consider the function

$$\sum_\rho (b'_{1\rho}w_1 + \cdots + b'_{\alpha\alpha}w_\alpha)^2 = \Phi(w_1, \ldots, w_\alpha).$$

We can find an orthogonal substitution with real coefficients

$$\left| \begin{array}{ll} w_1 & f_{11}w_1 + \cdots + f_{\alpha 1}w_\alpha \\ & \quad\cdots \\ w_1 & f_{\alpha 1}w_1 + \cdots + f_{\alpha\alpha}w_\alpha \end{array} \right| \tag{103}$$

that, when applied to this function, causes the mixed terms to vanish [i.e., terms of the form $b'_{i\rho}b'_{j\rho}w_iw_j$ $(i \neq j)$. Jordan calls them *les rectangles.*] (Cauchy, *Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires, etd.; Exercices de mathematics, t. IV.*). Once these coefficients have been determined, we replace the coordinate planes

$$v_1 = 0, \ldots, v_\alpha = 0$$

by new rectangular planes

$$v'_1 = 0, \ldots, v'_\alpha = 0 \tag{104}$$

given by the relations

$$v'_r = f_{r1}v_1 + \cdots + f_{r\alpha}v_\alpha \quad (r = 1, 2, \ldots, \alpha). \tag{105}$$

On the other hand, in place of the equations

$$C_r = 0 \quad (r = 1, 2, \ldots, \alpha)$$

used to define P_γ, we will use the equivalent system formed from the following equations:

$$A'_r = f_{r1}C_1 + \cdots + f_{r\alpha}C_\alpha = v'_r + b''_{r1}w_1 + \cdots b''_{r\alpha}w_\alpha. \quad (r = 1, 2, \ldots, \alpha), \tag{106}$$

where for short we have set

$$b''_{r\rho} = f_{r1}b'_{1\rho} + \cdots + f_{r\alpha}b'_{\alpha\rho} \qquad \begin{pmatrix} r = 1, 2, \ldots, \alpha \\ \rho = 1, 2, \ldots, \alpha \end{pmatrix}$$

We then have

$$\begin{aligned} \Sigma_\rho(b''_{1\rho}w_1, \ldots, b''_{\alpha\rho}w_\alpha)^2 &= \Sigma_\rho[\Sigma_r b'_{r\rho}(f_{r1}w_1 + \cdots + f_{r\alpha}w_\alpha)]^2 \\ &= \Phi(f_{11}w_1 + \cdots + f_{\alpha 1}w_\alpha + \cdots + f_{1\alpha}w_1 + \cdots + f_{\alpha\alpha}w_\alpha). \end{aligned}$$

Therefore, this expression does not have any cross terms in the variables owing to the condition

$$\Sigma_\rho b''_{r\rho}b''_{s\rho} = 0, \qquad r \neq s, \tag{107}$$

which says that the planes $A'_1, \ldots, A'_\alpha$ are mutually rectangular.

37. Furthermore, let

$$\Sigma_\rho b''^2_{r\rho} = g_r^2 \qquad (r = 1, 2, \ldots, \alpha). \tag{108}$$

When the relations (107) and (108) are divided by $g_r g_s$ and g_r^2 respectively, they show that the quantities

$$\frac{b''_{11}}{g_1}, \ldots, \frac{b''_{r\rho}}{g_r}, \ldots, \frac{b''_{\alpha\alpha}}{g_\alpha}$$

are the coefficients of an orthogonal substitution. If we take

$$w'_r = \frac{b''_{r1}}{g_1}w_1 + \cdots + \frac{b''_{r\alpha}}{g_r}w_\alpha \qquad (\alpha = 1, 2, \ldots, r) \tag{109}$$

in place of $w_1, \ldots, w_\alpha$, then P_α will still be determined by the equations (104), while the equations for P_γ reduce to the simple form

$$A'_r = v'_r + g_r w'_r \qquad (r = 1, \ldots, \alpha), \tag{110}$$

in which no more than α parameters $g_1, \ldots, g_\alpha$ appear.

If we set

$$g_r = \tan\theta_r,$$

and multiply the equations (110) by $\cos\theta_r$, they take the form

$$A'_r = v'_r \cos\theta_r + w'_r \sin\theta_r = 0. \tag{111}$$

We will say that $\theta_1, \ldots \theta_\alpha$ are *the angles of the two α-planes P_α and P_γ.*

38. The quantities $cos^2\theta_r$ and $sin^2\theta_r$ are orthogonal invariants. On can show this in two ways.

First, let

$$\lambda_1 v'_1 + \cdots + \lambda_\alpha v'_\alpha,$$

be an arbitrary plane generator of P_α and

$$\mu_1(v'_1 \cos\theta_1 + w'_1 \sin\theta_1) + \cdots + \mu_\alpha(v'_\alpha \cos\theta_\alpha + w'_\alpha \sin\theta_\alpha)$$

be arbitrary plane generator of P_γ. The angle between these planes is given by the formula

$$\cos^2\theta = \frac{(\lambda_1\mu_1\cos\theta_1 + \cdots + \lambda_\alpha\mu_\alpha\cos\theta_\alpha)^2}{(\mu_1^2 + \cdots + \mu_\alpha^2)(\lambda_1^2 + \cdots + \lambda_\alpha^2)} = \frac{N^2}{ML}. \tag{112}$$

If one varies the indeterminants λ and μ, the [local] maxima and minima this expressions will clearly be invariants.

First of all suppose that only the λ are varied. If the λ's are determined so that $\frac{N^2}{ML}$ [$\frac{M}{N}$ in the text] has a maximum or minimum value t^2, then we must have

$$t^2 = \frac{N^2}{ML} = \frac{N^2 + d.N^2}{ML + MdL}$$

whatever the variations $d\lambda_1, \ldots, d\lambda_n$. From this it follows that

$$d.N^2 = t^2 MdL.$$

Hence

$$\frac{dN^2}{\lambda_1} = t^2 M\frac{dL}{d\lambda_1}, \ \ldots, \ \frac{dN^2}{\lambda_\alpha} = t^2 M\frac{dL}{d\lambda_\alpha},$$

or

$$N\mu_1\cos\theta_1 = t^2 M\lambda_1, \ldots, N\mu_\alpha\cos\theta_\alpha = t^2 M\lambda_\alpha, \tag{113}$$

On subtracting the values of λ from these equations, substituting them into (112), and simplifying, we get

$$t^2 = \cos^2\theta = \frac{\mu_1^2\cos^2\theta_1 + \cdots + \mu_\alpha^2\cos^2\theta_\alpha}{\mu_1^2 + \cdots + \mu_\alpha^2} = \frac{R}{S}. \tag{114}$$

39. Let us now take the quantities $\mu_1, \ldots, \mu_n$ as the variables. The variational law for t^2 that follows from (116) may be represented geometrically. Specifically, since t^2 depends only on the ratios of the quantities $\mu_1, \ldots, \mu_\alpha$, we may assume that they may be constrained to vary in such a way that S is always equal to one. That done, in an α-dimensional space construct a line[situated at the origin] of length $r = \frac{1}{t^2}$ and making angles $\mu_1, \ldots, \mu_\alpha$ with the coordinate axes. The locus of the ends of these lines will be the ellipsoid

$$1 = X_1^2 \cos^2 \theta_1 + \cdots + X_\alpha^2 \cos^2 \theta_\alpha.$$

40. It remains to obtain the maxima and minima of t^2. Let s^2 be one of them. We have

$$s^2 = \frac{R}{S} = \frac{R + dR}{S + dS} = \frac{dR}{dS},$$

from which we have

$$\frac{dR}{d\mu_1} = s^2 \frac{dS}{d\mu_1}, \ldots, \frac{dR}{d\mu_\alpha} = s^2 \frac{dS}{d\mu_\alpha},$$

or

$$\mu_1 \cos^2 \theta_1 = \mu_1 s^2, \ldots, \mu_\alpha \cos^2 \theta_\alpha = \mu_\alpha s^2, \tag{115}$$

These equations will be satisfied if we set

$$s^2 = \cos^2 \theta_\rho, \quad \mu_1 = \ldots = \mu_{\rho-1} = \mu_{\rho+1} = \ldots = \mu_\alpha = 0.$$

Then the equations (113) give the corresponding values of the unknowns λ:

$$\lambda_1 = \ldots = \lambda_{\rho-1} = \lambda_{\rho+1} = \ldots = \lambda_\alpha = 0. \tag{116}$$

Therefore, there exist α distinct maxima and minima corresponding respectively to the angles between the planes $A_1, \ldots, A_\alpha$ and $A'_1, \ldots, A'_\alpha$. We thus established the following theorem.

Let two α-planes P_α and P_γ that have only a single point in common be given. If we seek pairs of their generating planes whose angles are maximal or minimal, we will get two corresponding systems systems of real perpendicular planes $A_1, \ldots, A_\alpha$ and $A'_1, \ldots, A'_\alpha$. The desired maxima and minima are none other than the angles of the multi-planes P_α and P_γ.

41. For the second way, let $v'_1, \ldots, v'_\alpha$ and $w'_1, \ldots, w'_\alpha$ be the coordinates of an arbitrary point in P_γ. Its distance h from the multi-plane defined by the equations

$$v'_1 = \cdots = v'_\alpha = 0$$

is evidently given by

$$h^2 = v'^2_1 + \cdots + v'^2_\alpha.$$

On the other hand, its distance k from the origin of the coordinates is given by the formula

$$k^2 = v_1'^2 + \cdots + v_\alpha'^2 + 2_1'^2 + \cdots + w_\alpha'^2.$$

Hence, on taking the equations (60) into account,

$$k^2 = v_1'^2 + \cdots + v_\alpha'^2 + v_1'^2 \cot^2\theta_1 + \cdots + v_\alpha'^2 \cot^2\theta_\alpha = \frac{v_1'^2}{\sin^2\theta_1} + \cdots + \frac{v_\alpha'^2}{\sin^2\theta_\alpha}.$$

Given this, the maxima and minima of $\frac{h^2}{k^2}$ are evidently invariants. Let one of them be u^2. Then as before

$$u^2 = \frac{h^2}{k^2} = \frac{h^2 + d.h^2}{k^2 + d.k^2},$$

whence

$$\frac{d.h^2}{dv_1'} = u^2 \frac{d.k^2}{dv_1'}, \ldots, \frac{d.h^2}{dv_\alpha'} = u^2 \frac{d.k^2}{dv_\alpha'}.$$

These equations may be satisfied by setting

$$u^2 = \sin^2\theta_\rho,\ v_1' = \cdots = v_{\rho-1}' = v_{\rho+1}' = \cdots = v_\alpha' = 0.$$

Thus we have α distinct invariants, which are the squares of the sines of the angles $\theta_1, \ldots, \theta_\alpha$.

42. We note, once again, that the quantities $\cos\theta_r$, $\sin\theta_r$, and $\tan\theta_r = g_r$ are not invariant, since one can change their signs simply by changing the sign of one of the coordinates. Only their squares are invariant.

43. Consider two arbitrary systems of two α-plains P_α, P_γ and P_α', P_γ'. If they have the same invariants, then they are congruent; for we have just seen that they can be reduced by orthogonal substitutions to a single canonical form. But it will be useful to distinguish the case where the two systems are identical from the case where they are reflective.

If $n > 2\alpha$, then the two systems are simultaneously identical and reflective. For each of them is reflective to itself; its equation does not change under the orthogonal substitution of determinant -1 that one obtains by changing the sign of one of the coordinates that does not occur in the canonical equations.

Such is no longer the case when $n = 2\alpha$: one has either identity or reflectivity. Let us see how the two cases may be distinguished.

44. As before, let

$$v_r = 0 \quad (r = 1, 2, \ldots, \alpha)$$

and

$$A_r = a_{r1}v_1 + \cdots a_{r\alpha}v_\alpha + b_{r1}w_1 + \cdots b_{r\alpha}w_\alpha = 0 \quad (r = 1, 2, \ldots, \alpha)$$

be the equations that define P_α and P_γ. Likewise, we can write the equations that define P'_α and P'_γ in the form

$$V_r = 0 \quad (r = 1, 2, \dots, \alpha)$$

and

$$A_{r1}V_1 + \cdots A_{r\alpha}V_\alpha + B_{r1}W_1 + \cdots B_{r\alpha}W_\alpha = 0 \quad (r = 1, 2, \dots, \alpha)$$

where, $V_1 = 0, \dots, W_\alpha = 0$ are rectangular planes.

The substitution

$$|v_1, \dots, w_\alpha \; V_1, \dots, W_\alpha|.$$

is orthogonal, and since its determinant is equal to ± 1, it is clear that the system P'_α, P'_γ will be equal or symmetric to the system of two multi-planes P_α, P_γ defined by

$$v_r = 0 \quad (r = 1, 2, \dots, \alpha)$$

and

$$\mathcal{A}_r = A_{r1}v_1 + \cdots A_{rn}v_n + B_{r1}w_1 + \cdots B_{rn}w_n \quad (r = 1, 2, \dots, \alpha).$$

The question therefore reduces to determining when the new system is identical to the system P_α and P_γ and when, on the other hand, the two are reflective.

45. We are going to show that *they are identical if the product of the two determinants*

$$\Delta_1 = \begin{vmatrix} A_{11} & \cdots & A_{1\alpha} \\ \vdots & & \vdots \\ A_{\alpha 1} & \cdots & A_{\alpha\alpha} \end{vmatrix} \quad \text{and} \quad \Delta_2 = \begin{vmatrix} B_{11} & \cdots & B_{1\alpha} \\ \vdots & & \vdots \\ B_{\alpha 1} & \cdots & B_{\alpha\alpha} \end{vmatrix}$$

has the same sign as the product as the product of the determinants

$$\delta_1 = \begin{vmatrix} a_{11} & \cdots & a_{1\alpha} \\ \vdots & & \vdots \\ a_{\alpha 1} & \cdots & a_{\alpha\alpha} \end{vmatrix} \quad \text{and} \quad \delta_2 = \begin{vmatrix} b_{11} & \cdots & b_{1\alpha} \\ \vdots & & \vdots \\ b_{\alpha 1} & \cdots & b_{\alpha\alpha} \end{vmatrix},$$

In the opposite case, they are reflective.

46. To define P_γ we will take the the planes

$$C_r = 0$$

defined by the relations (50) and (51) instead of the the planes

$$A = 0.$$

In this new system of planes the determinants corresponding to δ_1 and δ_2 are

$$\delta_1' = \epsilon\delta_1 = 1, \quad \delta_2' = \epsilon\delta_2 = \frac{\delta_2}{\delta_1} = \begin{vmatrix} b_{11}' & \cdots & b_{1\alpha}' \\ \vdots & & \vdots \\ b_{\alpha 1}' & \cdots & b_{\alpha\alpha}' \end{vmatrix}$$

where ϵ denotes the determinant

$$\begin{vmatrix} e_{11}' & \cdots & e_{1\alpha}' \\ \vdots & & \vdots \\ e_{\alpha 1}' & \cdots & e_{\alpha\alpha}' \end{vmatrix}$$

reciprocal to δ. From this it is seen that δ_2' has the same sign as $\delta_1\delta_2$.

Instead of $v_1, \dots, v_\alpha$, let us now take as coordinates the quantities $v_1', \dots, v_\alpha'$ determined by the relations (105). Let us also define P_γ by the intersection of the planes A_r' given by the formula (106). The determinant

$$\delta_2'' = \begin{vmatrix} b_{11}'' & \cdots & b_{1\alpha}'' \\ \vdots & & \vdots \\ b_{\alpha 1}'' & \cdots & b_{\alpha\alpha}'' \end{vmatrix}$$

given by the formula (106) is clearly equal to $\delta_2\psi$, where ψ is the determinate of the coefficients $f_{11}, \dots, f_{\alpha\alpha}$. This determinant is equal to ± 1, since the substitution (103) is orthogonal. We may assume that it is equal to one. For in order to change its sign it is sufficient to change all the signs in $f_{11}, \dots, f_{1\alpha}$. This can be done without changing the fact that the substitution (103) causes the off-diagonal elements in Φ to vanish. Under this supposition we have

$$\delta_2'' = \delta_2.$$

Given this, we clearly have

$$\delta_2'' = g_1 \cdots g_\alpha \psi', \tag{117}$$

where ψ' is the determinant of the equations (109). Since these relations are orthogonal, we have

$$\psi' = \pm 1.$$

Moreover, the signs of the $g_1, \dots, g_\alpha$ are arbitrary. Hence we may assume that $g_2, \dots, g_\alpha$ are positive and determine the sign of g_1 so that ψ' is equal to one. From (117) we see that g_1 has the same sign as δ_2'' or the product $\delta_1\delta_2$.

We can therefore give P_α and P_γ the forms

$$v_1' = 0, \dots, v_\alpha' = 0, \tag{118}$$

and

$$v'_1 + g_1 w'_1 = 0, \ldots v'_\alpha + g_\alpha w'_\alpha = 0 \tag{119}$$

by orthogonal transformations of determinant one.

47. Proceeding is the same way with the system P'_α and P'_γ, which by hypothesis has the same invariants $g_1^2, \ldots, g_\alpha^2$, we can by orthogonal transformations of determinant one transform their equations in to the same canonical form as above, provided $\Delta_1\Delta_2$ have the same sign as $\delta_1\delta_2$. If $\Delta_1\Delta_2$ has a different than $\delta_1\delta_2$, then we can reduce it to a form that differs from the above by the sign of g_1. In the first case, the two systems are clearly identical. In the second case they are reflective, since to make their canonical forms the same, we have only to change the sign of a single coordinate. We have thus established our proposition.

48. We have seen above that one can determine a system of coordinates $x_1, \ldots, x_\rho$, $y_1, \ldots, y_\sigma$, $z_1, \ldots, z_\tau$, $u_1, \ldots, u_\upsilon$, $v'_1, \ldots, v'_\alpha$, and $w'_1, \ldots, w'_\alpha$ such that P_k and P_l are respectively determined by the equations

$$\begin{aligned} &x_1 = 0, \ldots, x_\rho = 0 \\ &u_1 = 0, \ldots, u_\upsilon = 0 \\ &v'_1 = 0, \ldots, v'_\alpha = 0 \end{aligned} \tag{120}$$

and by the equations

$$\begin{aligned} &y_1 = 0, \ldots, y_\sigma = 0 \\ &z_1 = 0, \ldots, z_\tau = 0 \\ &v'_1 \cos\theta_1 + w'_1 \sin\theta_1, \ldots, v'_\alpha \cos\theta_\alpha + w'_\alpha \sin\theta_\alpha. \end{aligned} \tag{121}$$

The multi-planes P_{n-k} and P_{n-l} taken perpendicularly from the origin of their predecessors will have the equations

$$\begin{aligned} &y_1 = 0, \ldots, y_\rho = 0 \\ &z_1 = 0, \ldots, z_\upsilon = 0 \\ &w'_1 = 0, \ldots, w'_\alpha = 0 \end{aligned} \tag{122}$$

and

$$\begin{aligned} &y_1 = 0, \ldots, y_\sigma = 0 \\ &u_1 = 0, \ldots, u_\upsilon = 0 \\ &-v'_1 \sin\theta_1 + w'_1 \cos\theta_1, \ldots, -v'_\alpha \sin\theta_\alpha + w'_\alpha \cos\theta_\alpha. \end{aligned} \tag{123}$$

A comparison of these equations shows:

1. *P_{n-k} subtends with P_k α angles equal to $\frac{\pi}{2} + \theta_1, \ldots \frac{\pi}{2} + \theta_\alpha$*

2. *P_{n-k} and P_{n-l} subtend between themselves α angles equal to $\pi + \theta_1, \ldots, \pi + \theta_\alpha$*

It follows that P_{n-k} *and* P_{n-l} *have the same invariants as* P_k *and* P_l.

49. *A system formed from a* k*-plane* P_k *and an* l*-plane* P_l *has in general* α *invariants, where* α *is the smallest of the four numbers* k, l, $n-k$, *and* $n-l$.

Specifically, let us assume that the coefficients of P_l and P_k have no special relation to one another. Suppose, first of all that k is the smallest of the four numbers above. Then we have $k+l \leq n$ and $k+n-l \leq n$. Then p_k in general has no generating plane in common with P_l or P_{n-l}. Hence according to §33,

$$\rho = 0, \upsilon = 0, \alpha = k - \rho - \upsilon = k.$$

In the particular case that P_k has ρ generating planes in common with P_l and υ generating planes in common with P_{n-l}, we can express the general case by saying that among the k angles of P_k and P_l, there are ρ that are equal to zero and υ that are equal to $\frac{\pi}{2}$.

On the other hand, let $n-k \leq k, l, n-l$. Then, as above, the multi-planes P_{n-k} and P_{n-l} will have in general $n-k$ invariant angles, and P_k and P_l, which are respectively perpendicular to P_{n-k} and P_{n-l}, will have the same invariants (see §48).

50. We have seen (§33) that the inquiry into the angles between two arbitrary multi-planes can immediately be reduced to an inquiry into the angles between two alpha planes, where α is at most equal to $\frac{n}{2}$. This last inquiry can be resolved by reducing the two α-planes to their canonical form, as we have done in §34 and the following. But one can also treat the problem directly.

Specifically, let

$$a_{11}x_1 + \cdots + a_{1n}x_n = 0,$$
$$\cdots \quad \cdots \quad \cdots$$
$$a_{\alpha 1}x_1 + \cdots + a_{\alpha n}x_n = 0,$$

and

$$b_{11}x_1 + \cdots + b_{1n}x_n = 0,$$
$$\cdots \quad \cdots \quad \cdots$$
$$b_{\alpha 1}x_1 + \cdots + b_{\alpha n}x_n = 0,$$

be the equations of the two α-planes — call them P and Q.

By setting

$$A_1 = a_{11}\lambda_1 + \cdots + a_{\alpha 1}\lambda_\alpha, \ \ldots, \ A_n = a_{1n}\lambda_1 + \cdots + a_{\alpha n}\lambda_\alpha \tag{124}$$

we may write the general equation of a generating plane of P as

$$A_1x_1 + \cdots + A_nx_n = 0.$$

Likewise, on setting

$$B_1 = b_{11}\lambda_1 + \cdots + b_{\alpha 1}\lambda_\alpha, \ \ldots, \ B_n = b_{1n}\lambda_1 + \cdots + b_{\alpha n}\lambda_\alpha \tag{125}$$

we can write a generating plane of Q as

$$B_1x_1 + \cdots + B_nx_n = 0.$$

The angle ϕ between two of the planes is given by the expression

$$\cos^2\phi = \frac{(\sum A_\rho B_\rho)^2}{\sum A_\rho^2 B_\rho^2} = \frac{M}{N}. \tag{126}$$

It remains to find the minimum of this expression.

Let s^2 be the required minimum. The values corresponding to the unknowns — $\lambda_1, \ldots, \lambda_n$ on the one hand and $\mu_1, \ldots, \mu_n$ on the other — are determined only up to their ratios, because $\frac{M}{N}$ is homogeneous and of degree zero with respect to each of these systems of variables. One can, therefore, determine the λ's and the μ's in such a way that in addition to the condition

$$\frac{M}{N} = s^2$$

it satisfies the auxiliary conditions

$$\sum A^2 = 1 \quad \text{and} \quad \sum B^2 = 1,$$

whence

$$N = \sum A^2 \sum B^2 = 1 \quad \text{and} \quad s = \sqrt{\frac{M}{N}} = \sqrt{M} = \sum A_\rho B_\rho.$$

All this given, let us perturb the variables λ, μ by infinitely small increments. Then according to the conditions for a minimum, we have

$$s^2 = \frac{M}{N} = \frac{M + dM}{N + dN}.$$

If we multiply by the denominators and set separately each of the coefficients of $\lambda_1, \ldots, \lambda_\alpha$ and $\mu_1, \ldots, \mu_\alpha$ from zero [?], we get for $\sigma = 1, \ldots, \alpha$

$$\frac{dM}{d\lambda_\sigma} = s^2 \frac{dN}{d\lambda_\sigma} \tag{127}$$

and

$$\frac{dM}{d\mu_\sigma} = s^2 \frac{dN}{d\mu_\sigma}. \tag{128}$$

But

$$\frac{dM}{d\lambda_\sigma} = 2\sum A_\rho B_\rho \sum a_{\sigma\rho} B_\rho = 2s \sum a_{\sigma\rho} B_\rho,$$
$$\frac{dN}{d\lambda_\sigma} = 2\sum B_\rho^2 \sum a_{\sigma\rho} A_\rho = 2\sum a_{\sigma\rho} A_\rho.$$

If we substitute these values into equation (127) and cancel the common factor s, we get

$$\sum a_{\sigma\rho} B_\rho = s \sum a_{\sigma\rho} A_\rho. \quad (\sigma = 1, 2, \ldots, \alpha). \tag{129}$$

Likewise equation (128) gives

$$\sum b_{\sigma\rho} A_\rho = s \sum b_{\sigma\rho} B_\rho \quad (\sigma = 1, 2, \ldots, \alpha). \tag{130}$$

Equations (124), (125), (129), and (130), form a system of $2n + 2\alpha$ linear equations among the $2n + 2\alpha$ quantities λ, μ, A, and B. If we set the determinant of this system to zero, we get the characteristic equation that determines s. This equation is of degree 2α, but it is easy to see that that it only contains pairs of powers of s.[24]

[24] There are actually only 2α free quantities since $A\rho$ and B_ρ depend linearly on the λ's and μ's. The characteristic equation give ρ values of s corresponding to the cosine of the canonical angles and in addition the corresponding values of $-s$.

References

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. *LAPACK Users' Guide*. SIAM, Philadelphia, third edition, 1999.

[2] E. Beltrami. Sulle funzioni bilineari. *Giornale di Matematiche ad Uso degli Studenti Delle Universita*, 11:98–106, 1873. An English translation by D. Boley is available in Technical Report 90–37, Department of Computer Science, University of Minnesota, Minneapolis, 1990.

[3] A. L. Cauchy. Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes. In *Oeuvres Complétes (II^e Série)*, volume 9. 1829.

[4] A. Cayley. Remarques sur la notation des fonctions algébriques. *Journal für die reine und angewandte Mathematik*, 50:282–285, 1855. Cited and reprinted in [6, v. 2, pp. 185–188].

[5] A. Cayley. A memoir on the theory of matrices. *Philosophical Transactions of the Royal Society of London*, 148:17–37, 1858. Cited and reprinted in [6, v. 2, pp. 475–496].

[6] A. Cayley. *The Collected Mathematical Papers of Arthur Cayley*. Cambridge University Press, Cambridge, 1889–1898. Arthur Cayley and A. R. Forsythe, editors. Thirteen volumes plus index. Reprinted 1963 by Johnson Reprint Corporation, New York.

[7] J. Dieudonné. In J. Dieudonné, editor, *Oeuvres de Camille Jordan, V.3*, pages v–vii. Gauthier–Villars, Paris, 1962.

[8] Jean Dieudonné. Jordan, Camille. In C. C. Gillispe, editor, *Dictionary of Scientific Biography, VII.* Charles Scribner's Sons, New York, 1973.

[9] Jean-Luc Dorrier. A general outline of the genesis of vector space theory. *Historia Mathematica*, 22:277–261, 1995.

[10] G. F. Frobenius. Über das Pfaffsche problem. *Journal für die reine und angewandte Mathematic*, 82:230–315, 1876.

[11] C. Jordan. Essai sur la géométrie à n dimensions. *Comptes Rendus de l'Académie des Sciences, Paris*, 75:1614–1615, 1872.

[12] C. Jordan. Sur la réduction des formes bilinéaires. *Comptes Rendus de l'Académie des Sciences, Paris*, 78:614–617, 1874.

[13] C. Jordan. Essai sur la géométrie à n dimensions. *Bulletin de la Société Mathématique*, 3:103–174, 1875.

[14] M. Kline. *Mathematical Thought from Ancient to Modern Times.* Oxford University Press, New York, 1972.

[15] S. J. Leon, øA. Bj orck, and W. Gander. Gram–Schmidt orthogonalization: 100 years and more. *Numerical Linear Algebra and Applications*, 20:492–532, 2013.

[16] G. H. Moore. The axiomatization of linear algebra: 1875–1940. *Historia Mathematica*, 22:262–303, 1995.

[17] J. J. O'Connor and E. F. Robertson. Marie Ennemond Camille Jordan. In J. J. O'Connor and E. F. Robertson, editors, *The MacTutor History of Mathematics Archive*. `www-history.mcs.st-andrews.ac.uk/index.html`, 2015.

[18] C. C. Paige and M. Wei. History and generality of the CS-decomposition. *Linear Algebra and Its Applicationsn*, 208/209:303–326, 1994.

[19] G. W. Stewart. On the early history of the singular value decomposition. *SIAM Review*, 35:551–566, 1993.

[20] G. W. Stewart. A canonical CS representation of a pair of subspaces. *SIAM Journal of Matrix Analysis and Applications*, 2015. To apppear.