SOME REMARKS ON THE HISTORY OF LINEAR ALGEBRA*

C. T. Chong
National University of Singapore

The origins of the concepts of a determinant and a matrix, as well as an understanding of their basic properties, are historically closely connected. Both concepts came from the study of systems of linear equations. Already Leibniz (1646–1716) had considered patterns of coefficients in such systems and represented them with pairs of numbers. Around 1729 Maclaurin presented the solution of simultaneous linear equations in two, three, and four unknowns. This was published in his book *Treatise of Algebra* published posthumously in 1748. The rule he gave was the one given by Cramer (1750), who studied the coefficients of the general conic

$$A + By + Cx + Dy^2 + Exy + x^2 = 0$$

pasisng through five given points. Cramer gave the solution in terms of the ratios of determinants, precisely what is known today as *Cramer's rule*. In 1764, Bezout showed that the vanishing of the coefficient determinant is the condition that non-zero solutions exist.

Vandermonde (1772) was the first to give an exposition of the theory of determinants (i.e. apart from the solution of linear equations although such applications were also made by

---

*Text of talk given at the Workshop on Linear Algebra and its teaching from 9 - 10 September 1985 at the National University of Singapore.

him). He gave a rule for expanding a determinant by using second order minors and their complementary minors. Also in 1772, Laplace expanded Vandermonde's rule by using a set of minors of r rows and the complementary minors.

Although determinants and matrices received a great deal of attention in the 19th century and many papers were written on these subjects they do not constitute great innovations in mathematics. Nonetheless, the idea of the determinant found applications not only in systems of linear equations, but also in the simultaneous solution of equations of higher degree (known as elimination theory), in the transformation of coordinates, in the change of variables in multiple integrals, in the solution of systems of differential equations arising in planetary motion, and in the reduction of quadratic forms to standard forms.

The word *determinant* was already used by Gauss (1777–1855) for the discriminant of the quadratic form

$$ax^2 + 2bxy + cy^2$$

in his number theoretic investigations. The word was later applied by Cauchy (1789–1857) to the determinants that had already appeared in the 18th century work. In his 1815 paper he introduced the idea of arranging the elements in a square array, and used the double subscript notation. A third order determinant would then appear as

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}$$

The two vertical lines as we know it today were introduced by Cayley in 1841. In the 1815 paper Cauchy gave the first systematic and almost modern treatment of determinants. One of the major results was the multiplication theorem for determinants. This had been obtained by Lagrange in 1773 for third order determinants.

The mathematician James Joseph Sylvester (1814–1897) worked in the theory of determinants over a period of 50 years. He studied at Cambridge University and was a professor at the University of Virginia from 1841 to 1845. He then returned to London and spent the next ten years as an actuary and a lawyer. In 1876 he went to the United States again, this time teaching at Johns Hopkins Universtiy. In 1884 he returned to England and at the age of seventy became a professor at Oxford University. One of Sylvester's major accomplishments was an improved method of eliminating x from two polynomials (called the dialytic method; 1840). For example, given

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0$$
$$b_0x^2 + b_1x + b_2 = 0$$

he formed the determinant

$$\begin{vmatrix} a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & 0 \\ 0 & 0 & b_2 & b_1 & b_2 \end{vmatrix}$$

The vanishing of the determinant is the necessary and sufficient condition for the two equations to have a common root.

Jacobi applied the method of determinants to the study of the change of variables in multiple integrals. He also studied determinants whose entries are functions (1841). We will not dwell on this.

There is an intimate connection between determinants and quadratic forms. Already in the 18th century, the problem of transforming equations of the conic sections and quadric surfaces to simpler forms, by choosing appropriate coordinate axes, was already known. Indeed it was known that

$$\sum a_{ij}x_i x_j$$

for i, j not exceeding n, can always be reduced to a sum of r squares

$$y_1^2 + \ldots + y_s^2 - y_{s+1}^2 - \ldots - y_r^2$$

where r is the rank of the coefficient matrix (in modern terminology), by a linear transformation of the form

$$x_i = \sum_j b_{ij}y_j \qquad (i = 1,\ldots,n).$$

Sylvester (1852) stated that the number s of positive terms and r–s negative ones is always the same no matter what real transformation is used. This is known as his law of inertia of quadratic forms in n variables. Regarding the law as self-evident, he never proved it. The law was later rediscovered and proved by Jacobi in 1857.

The further study of the reduction of quadratic forms involves the notion of the characteristic equation of a quadratic form or of a matrix. A quadratic form in three variables is written as

$$Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz.$$

Associated with it is the matrix

$$H = \begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix}$$

The characteristic equation of the form or of the matrix is then $\det(H - tI) = 0$. The values t which satisfy this equation are called *characteristic roots*. From the values of t the lengths of the principal axes are easily obtained.

The notion of the characteristic equation appears implicitly in the work of Euler (1748) when he dealt with the above problem. The notion of characteristic equation first appeared explicitly in Lagrange's work on systems of linear differential equations (1762) in his study of the motions of the six planets known in his day, and in Laplace's work in the same area (1775).

The term characteristic equation is due to Cauchy (1840). He continued the study of the problem of reduction of quadratic forms. One of the results which he obtained, in modern terminology, states that any real symmetric matrix of any order has real characteristic roots. The study of the reduction of quadratic forms and the theory of bilinear forms was later done by Weierstrass (1858, 1868). He gave a general method of reducing two quadratic forms simultaneously to sums of squares.

New results on determinants were obtained throughout the 19th century.

Matrices

One could say that the subject of matrices was well developed before it was created. It was apparent from the immense amount of work on determinants that the array of numbers for determinants could be studied and manipulated for various purposes whether or not the value of the determinant came into question. It was clear then that the array itself merited independent study. The word was first used by Sylvester (1850) when he wished to refer to a rectangular array of numbers. Cayley (1855) insisted that logically the idea of a matrix preceded that of the determinant. Historically, however, the order was just the reverse. In any case he was the first one to single out the matrix itself and the first to publish a series of articles on the subject. Thus he is generally credited with being the creator of the theory of matrices.

Cayley (1821–1895) studied mathematics at Cambridge but turned to law and spent 15 years in that profession. During this period he managed to devote considerable time to mathematics and published close to 200 papers. It was during this period too that he began his long friendship and collaboration with Sylvester. In 1863 he was appointed to a professorship at Cambridge where he remained till his death, except for the year 1882 spent at Johns Hopkins University at the invitation of Sylvester. Together with Sylvester, he was the founder of the theory of invariants.

Matrices were introduced by Cayley as a convenient way of expressing transformations

$$x' = ax + by$$
$$y' = cx + dy$$

The use of matrices simplified the notations involved. He also introduced the basic operations of addition and multiplication of matrices. In his 1885 paper *A memoir on the theory of matrices* he also gave an expression of the inverse of a matrix in terms of the determinant and cofactors as we know it today. He also announced in the same article what is now called the *Cayley-Hamilton Theorem* : If M is a square matrix and $\det(M - xI) = 0$ is the characteristic equation of M, then when x is replaced by M, the resulting matrix is the zero matrix. For example, if M is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then $\det(M - xI)$ is $x^2 - (a + d)x + (ad - bc)$, and when M is substituted for x, we get the zero matrix. Cayley stated that he had verified this for 3 by 3 matrices and that no further proof was necessary. Hamilton's association with the theorem rests on the fact that in introducing his *Lectures on Quaternions* (1853) a certain linear transformation of variables was involved. He proved that the matrix of this linear transformation satisfied the characteristic equation of that matarix, though he did not think formally in terms of matrices.

The question of the minimal polynomial satisfied by the matrix M was first raised by Frobenius in 1878. He stated that it is formed from the factors of the characteristic polynomial and is unique. The uniqueness was proved by Hensel in 1904. In his 1878 paper Frobenius also gave the first general proof of the Cayley-Hamilton Theorem. The first formal definition of an orthogonal matrix was given by Frobenius in 1878 (although the term was previously used by Hermite). Finally Jordan showed in his book published in 1870 that every matrix can be transformed into a similar one that is called the *Jordan canonical form*.

Vectors

By the year 1800 mathematicians were using freely the various types of real numbers and even complex numbers, but the precise definitions of these various types of real numbers and even complex numbers were not available nor was there any logical justification for the mathematical operations (such as addition and multiplication) used on them. The greatest concern seemed to be caused by the fact that letters were manipulated as though they had the properties of the integers, yet the results of these operations were valid when numbers were substituted for the letters. By the middle of the 19th century the mathematical community generally accepted the following axioms :

1. Equal quantities added to the third yield equal quanitites.
2. a + (b + c) = (a + b) + c
3. a + b = b + a
4. Equals added to equals give equals
5. a(bc) = (ab)c
6. ab = ba
7. a(b + c) = ab + ac .

These axioms constituted the *Principle of the Permanence of Form*.

The notion of a vector was already used by Aristotle to represent forces. He was already aware of the parallelogram law. Gauss and others had also introduced the geometric representation of complex numbers. In 1837 Hamilton suggested that complex numbers be expressed as ordered pairs of real numbers satisfying the conditions that if a + bi and c + di are represented as (a,b) and (c,d), then

$$(a,b) + (c,d) = (a + c,\ b + d) \text{ (similarly for } -)$$

$$(a,b) \cdot (c,d) = (ac - bd,\ ad + bc)$$

$$(a,b) / (c,d) = ((ac + bd)/(c^2 + d^2), (bc - ad)/(c^2 + d^2)).$$

The use of complex numbers to represent vectors in the plane became well known around the 1830's. However, the utility of complex numbers is limited. If several forces act on a body, these forces need not lie in a plane. To treat these forces algerbraically a 3-dimensional analogue of complex numbers is needed. One could use Cartesian coordinates (x,y,z) of a point to represent a vector but there were no operations with the triples of numbers to represent the operations with vectors. These operations would seem to have to include addition, product, subtraction, division, and must be associative, commutative, distributive. There began the search for 3-dimensional complex number and its algebra. Gauss made some inroads into this but, as usual, never published them.

The creation of a useful spatial analogue of complex numbers is due to William Hamilton (1805-1865). At the age of five Hamilton could read Latin, Greek and Hebrew. At age 22 he was appointed professor of astronomy at Trinity College in Dublin.

He made many important contributions to optics, astronomy, dynamics and others. His most important mathematical work was the discovery of quaternions (1843). Two major concessions were made in connection with the discovery : the new numbers contained 4 components (not three), and that the commutativity law did not hold.

A quaternion is a number of the form

$$a + bi + cj + dk ,$$

where i, j, k obey certain special rules

$$jk = i,\ kj = -i,\ ki = j\ ,\ ik = -j,\ ij = k,\ ji = -k,$$

$$i^2 = j^2 = k^2 = -1\ .$$

The 'a' part is called the scalar part, while the others are called the vector part. Addition is obtained componentwise, while multiplication is defined via the above rules.

Hamilton had unbounded enthusiasm for his new creation. He believed that it was as important as the creation of the calculus and that it would be the key instrument in mathematical physics. He himself made some applications to geometry, optics and mechanics. But the physicists ignored quaternions and continued to work with Cartesian coordinates. Nevertheless, his work did lead to an algebra and analysis of vectors that physicists adopted, principally via the work of Gibbs. In any case, Hamilton's discovery proved to be of immeasurable importance for algebra. It pointed out the possibility of building up useful systems of numbers which may not obey the commutativity law or other laws of real numbers. More generally it led to the theory of linear associative algebras.

There is a way to advance algebra as far beyond what Vieta and Descartes have left us as they carried it beyond the ancients... . We need an analysis which is distinctly geometrical or linear, and which will explain situation directly as algebra expresses magnitude directly.

—*Leibniz*

From Pythagoras to the mid 19th century, the fundamental problem of geometry was to relate numbers to geometry. It played a key role in the creation of field theory (via the classical construction problems), and quite differently, in the creation of linear algebra. The integration of real numbers into geometry began with Descartes and Fermat in the 1630's. From the point of view of analysis, which focuses on functions, this was entirely satisfactory. But from the point of view of geometry, the method of attaching numbers to geometric entities is too clumsy, and in view of Euclid, the choice of origin and axes irrelevant. This brings us to Hermann Grassmann, the creator of linear algebra.

Grassmann was born in Stettin in 1809, lived there most of his life, and died in 1877. He was one of twelve children, and after marrying at the age of 40, produced 11 children. He spent three years in Berlin studying theology and philology. He had no university training, nor did he ever hold a university post. His life was spent as a school teacher. His major mathematical works are *A theory of tides* written in 1840 (a thesis unpublished) in the hope of raising his status as a teacher; a book known briefly as the *Ausdehnungslehre* (Theory of extension) published in 1844 and almost totally ignored, though it was drawn to the attention of Möbius, Gauss, Kummer, Cauchy and others; and the *Ausdehnungslehre* which appeared in 1862, a new book on the same subject. He made important contributions in mathematics and physics. He also wrote textbooks on languages and edited a political journal. It is unfortunate that his achievements in mathematics were not recognized in his time, and it took almost a century for the importance of his work to become clearly visible.

Grassmann looked upon geometry as being applied mathematics. In his view, there is a part of mathematics, linear algebra, that is applicable to a part of the physical world. Geometry acts as the link between these two, and therefore does not belong to mathematics. One may say that Grassmann invented linear algebra and showed how properly to apply it in geometry. From the beginning, Grassman regarded linear algebra as a formal theory independent of any interpretation. He wanted to distinguish it from its application in geometry. In the first *Ausdehnungslehre*, however, the algebra is interspersed with its geometric interpretation, indicating how he came upon the ideas. The 1862 *Ausdehnungslehre*, on the other hand, gives the full mathematical development of the theory before any application.

Grassmann began with the introduction of a collection of units $e_1, e_2, \ldots$ and considered formal linear combinations of the form $\Sigma\, a_i e_i$, where the a_i are numbers. He defined addition and scalar multiplication of these entities, and proved formally the linear space properties for these operations. He then developed the theory of linear independence in a way surprisingly similar to modern presentation. He also defined the notions of subspace, independence, span, dimension, join and meet of subspaces, and projections of elements onto subspaces. A proof of the invariance of dimension under change of basis was also given. He also showed that any finite set has an independent subset of the same span, and that any independent set extends to a basis. The formula

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

was also obtained by him.

In 1855 Grassmann published a paper in which he defined a product of elements of a linear space by

$$(\Sigma a_i e_i)(\Sigma b_j e_j) = \Sigma a_i b_j e_i e_j$$

and proved its distributivity. Here he allowed the scalars to be complex numbers. If the $e_i e_j$ are themselves linear combinations of the e_i's we have here the concept of an algebra. Later he observed that the algebra of quaternions would be a special case of this. His declared that the objective of publishing this paper was to claim priority for some results that had been published by Cauchy. The story goes like this : In 1847 Grassmann had wanted to send a copy of the *Ausdehnungslehre* to Saint-Vernant to show that he had anticipated some of the latter's ideas on vector addition and multiplication, but not knowing the address, he sent the book to Cauchy with a request to forward it. This was not done. Six years later Cauchy published a paper proving some results which Grassmann had obtained earlier. On reading this, Grassmann said, "I recalled at a glance that the principles which are there established and the results which are proved were exactly the same as those which I published in 1844, and of which at the same time I gave numerous applications to algebraic analysis, geometry, mechanics and other branches of physics." An investigating committee of three members of the French Academy, including Cauchy himself, never came to a decision on the question of priority.

Grassmann introduced two kinds of products : the inner product where $e_i e_j = 1$ if $i = j$, and equal to 0 otherwise, and the outer product (or exterior multiplication) where $[e_i e_j] = -[e_j e_i]$, and $[e_i e_i] = 0$. From these product operations one could then derive the usual formulas on lengths of vectors, and those involving sine and cosine, such as areas formed by two vectors and volumes formed by three vectors.

Grassmann also did extensive work on linear transformations. He showed that given a linear transformation, the whole space can be decomposed into the direct sum of invariant subspaces obtained in terms of the characteristic roots and their multiplicities. This results is sometimes called the primary decomposition theorem.

On his own work, Grassmann had this to say :

> *I remain completely confident that the labour I have expended on the science presented here and which has demanded a significant part of my life as well as the most strenuous application of my powers, will not be lost. It is true that I am aware that the form which I have given the science is imperfect and must remain imperfect. But I feel obliged to state that even if this work should remain unused for another 17 years or longer, without entering into the actual development of science, still that time will come when it will be brought forth from the dust of oblivion and when ideas now dormant will bring forth fruit. I know that if I also fail to gather around me a circle of scholars, whom I could stimulate to develop and enrich them further, yet there will come a time when these ideas, perhaps in a new form, will*

arise anew and will enter into a living communication with contemporary developments. For truth is eternal and divine.

Grassmann's achievements in linear algebra were noted by Gibbs and by the Italian mathematician G. Peano (1858–1932), who was the first to introduce axioms for the theory of vector spaces. This was given in the book published in 1888 : *Geometrical Calculus, according to the Ausdehnungslehre of H. Grassmann, preceded by the operations of deductive logic.* The book was a reworking of Grassman's *Ausdehnungslehre* and Peano made no claim to originality for the ideas contained in it. There is no doubt that the extreme clarity of his presentation, in contrast to the notorious difficulty of reading Grassmann's work, helped to spread Grassmann's ideas and made them more popular. This was indeed Peano's objective in publishing the book, as he stated in the forword.

I shall end with a quotation from Lord Kelvin :

> *Quaternions came from Hamilton after his really good work had been done; and though beautifully ingenious, have been an unmixed evil to those who have touched them in any way... . Vector is a useless survival, or offshoot from quaternions, and has never been of the slightest use to any creature.*

History has a way of showing that remarks should never be lightly made, even by great men.