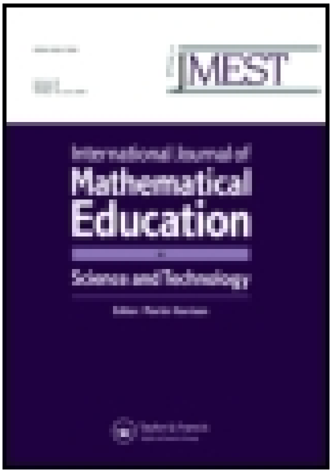
The teaching of vector algebra

Ron Shaw
University of Hull



International Journal of Mathematical Education in Science and Technology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tmes20

The teaching of vector algebra

Ronald Shaw [a]

[a] Department of Applied Mathematics , University of Hull , England

Published online: 09 Jul 2006.

To cite this article: Ronald Shaw (1985) The teaching of vector algebra, International Journal of Mathematical Education in Science and Technology, 16:5, 593-602, DOI: 10.1080/0020739850160504

To link to this article: http://dx.doi.org/10.1080/0020739850160504

INT. J. MATH. EDUC. SCI. TECHNOL., 1985, VOL. 16, NO. 5, 593–602

The teaching of vector algebra

by RONALD SHAW
Department of Applied Mathematics, University of Hull, England

(*Received 4 April 1984*)

To ease the student's passage from vector algebra to general linear algebra the author argues that the (non-metrical) scalar triple product should be defined geometrically in terms of oriented volume and be taught *before* the (metrical) cross product.

1. Introduction

In traditional accounts of vector algebra, the various topics are introduced in the following order:

$$\left.\begin{array}{lll} \text{L.} & \textit{Linear combinations} & a\mathbf{u}+b\mathbf{v} \\ \text{S.} & \textit{Scalar (dot) product} & \mathbf{u}.\mathbf{v} \\ \text{V.} & \textit{Vector (cross) product} & \mathbf{u}\times\mathbf{v} \\ \text{T.} & \textit{Scalar triple product} & [\mathbf{u},\mathbf{v},\mathbf{w}] \end{array}\right\} \tag{1}$$

Such accounts start out from an assumed geometrical knowledge, of a non-rigorous nature, with vectors being introduced as certain equivalence classes of oriented line segments, and their addition being defined by the triangle law. The set of such geometric vectors will be denoted in this note by E_2 or E_3 according as it arises from the Euclidean plane or from Euclidean three-dimensional space. Similarly $\mathbf{u}.\mathbf{v}$ and (in the case of three dimensions) $\mathbf{u}\times\mathbf{v}$ receive geometric definitions in terms of $uv\cos\theta$ and $uv\sin\theta\mathbf{n}$. However the scalar triple product is defined in terms of the scalar and vector products by

$$[\mathbf{u},\mathbf{v},\mathbf{w}]=\mathbf{u}\times\mathbf{v}.\mathbf{w} \tag{2}$$

In this note I wish to urge that the scalar triple product should also, in a first course on vectors, be defined geometrically—namely in terms of oriented (signed) volume—and that *it should be introduced before the vector product*. Moreover it will probably help to lead in to the trilinear product T by discussing first of all its bilinear analogue in two dimensions, namely:

$$\text{B.} \qquad \textit{Bilinear product (oriented area)} \quad [\mathbf{u},\mathbf{v}] \tag{3}$$

which is defined for vectors $\mathbf{u}$, $\mathbf{v}\in E_2$. I will argue that the traditional order L→S→V→T of the topics (1) is misconceived, and should be replaced by the scheme

$$\begin{tikzcd} & \text{S} \arrow[dr] & \\ \text{L} \arrow[ur] \arrow[dr] & & \text{V} \\ & \text{B, T} \arrow[ur] & \end{tikzcd} \tag{4}$$

My concern with the way the scalar triple product is taught is part of a more general quarrel with the usual accounts of vector algebra, which is that they present the student from the outset with a confusing mish-mash of metrical and non-metrical notions—and thereby seriously jeopardise the chances of a later smooth passage from vector algebra to general linear algebra. (Cf. Dieudonné's remarks in his Introduction to [1].) Now even when dealing with the topic L, traditional accounts so often introduce vectors as having length (and sometimes, in two dimensions, even a bearing, in terms of an angle made with some standard direction)—despite the fact that the metrical notions of length and angle do not at all enter into the later setting up of vector space axioms. Instead one should stress the fact that in order to deal with the relation $\mathbf{v}=c\mathbf{u}$ all that one needs is the ratio of segments *which lie on parallel lines*, and that this ratio is non-metrical in character. (For example, two segments lying on parallel lines can be said to have 'equal length' if they form opposite sides of a parallelogram. Note of course that parallelograms are non-metrical objects, since they can be defined via the non-metrical notion of parallelism; in contrast rhombi and rectangles are metrical objects, since the former require the comparison of lengths along *non-parallel lines* and the latter require the notion of angle, or at least that of orthogonality.)

But more pernicious in this respect is the definition of the scalar triple product by equation (2). For not only does the triple product thereby emerge very late in the day, with its symmetry properties and geometrical significance coming somewhat as a surprise, but *the definition (2) makes* $[\mathbf{u}, \mathbf{v}, \mathbf{w}]$ *appear to be a secondary construct which is of a highly metrical nature*—apparently dependent upon both the dot and cross products (both of which are metrical notions). Yet *in fact the scalar triple product is essentially non-metrical in character*; it exists and is of importance for any three-dimensional vector space, whether or not the space has been equipped with dot and cross products. For, as is well known, $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ is just the $n=3$ case of a determinant function $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$, which latter can be defined, non-metrically, for any n-dimensional vector space.

Determinant functions in n dimensions, and the associated theory of $n \times n$ determinants, are typically studied part way through a first-year undergraduate course on linear algebra, the treatment at this level being algebraic, proceeding rigorously from the axioms for an n-dimensional vector space. However, such a rigorous algebraic approach is not, I believe, appropriate for a first encounter with the scalar triple product. Instead I favour introducing $[\mathbf{u}, \mathbf{v}, \mathbf{w}]$ (and its two-dimensional analogue $[\mathbf{u}, \mathbf{v}]$, $\mathbf{u}$, $\mathbf{v} \in E_2$) in the same intuitive geometric spirit as for the rest of the topics (1).

This geometric approach is outlined in §§2–5. Once geometry has provided us with a 'real life' example of a trilinear skew symmetric function $[\ ,\ ,\ ]: E_3 \times E_3 \times E_3 \to \mathscr{R}$, then—and for most mortals only then—is the time ripe for the rigorous algebraic treatment which I briefly sketch in §6. In short (and with only a slight sense of exaggeration) I claim that 90 per cent of geometric inspiration followed by 10 per cent of algebraic perspiration provides an ideal preparation for a first course of linear algebra. Surely we should not be ashamed of our rich inheritance of geometrical understanding, but instead marvel at the ease with which it inspires us to invent an elegant and powerful algebra of vectors?

2. Area and volume

In particular we should not despise making use of our intuitive notions of area in the plane and of volume in three-dimensional space. Now I agree that these notions are traditionally thought of as amongst the more advanced geometrical concepts, and for this reason I think there has been some reluctance to make use of them in the setting up of vector algebra. But those 'facts' concerning area and volume which are needed in the approach outlined in §§3 and 4 are surprisingly simple, few in number and readily acceptable. We need only to assume that certain 'reasonable' subsets $S, S_1, S_2, \ldots$ of the plane can be assigned an area which satisfies

$$\left.\begin{array}{ll}\text{(i)} & \text{area}(S) \geqslant 0, \\ \text{(ii)} & \text{area}(S_1 \cup S_2) = \text{area}(S_1) + \text{area}(S_2), \text{ if } S_1, S_2 \text{ are disjoint} \\ \text{(iii)} & \text{area is invariant under translations.}\end{array}\right\} \quad (5)$$

To fix area uniquely (and to rule out the trivial solution area $(S)=0$, for all S, of (5)) we need to 'normalize' our area function. Let us (arbitrarily) choose some basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ for E_2 and then specify our unit of area by proclaiming that the parallelogram having adjacent sides $\mathbf{e}_1, \mathbf{e}_2$ has unit area. Let us write this normalization convention as

$$\text{area}(\mathbf{e}_1, \mathbf{e}_2) = 1 \quad (6)$$

Similarly we assume that certain subsets of three-dimensional space can be assigned a volume which satisfies conditions (5′), say completely analogous to (5), and which is fixed by a normalization convention

$$\text{vol}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1 \quad (6')$$

—which last proclaims that a certain parallelepiped, defined in terms of some arbitrarily chosen basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for E_3, has unit volume.

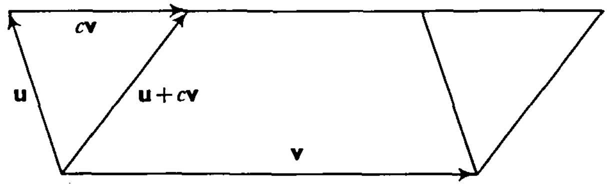
It should be noted that the above properties of area and volume are entirely non-metrical. The ensuing treatment, in §§3 and 4, of the products $[\mathbf{u}, \mathbf{v}]$ and $[\mathbf{u}, \mathbf{v}, \mathbf{w}]$ will consequently be non-metrical, and so will blend in harmoniously with a future rigorous algebraic treatment.

The non-metrical nature of area in the plane, and of volume in space, is not always appreciated. Many people are misled because they have been reared upon results such as

$$\left.\begin{array}{ll}\text{(i)} & \text{area} = \text{base} \times \text{height} \\ \text{(ii)} & \text{volume} = (\text{area of base}) \times \text{height}\end{array}\right. \quad (7)$$

in the case of parallelograms and parallelepipeds, respectively. Certainly the results (7) are stated in a highly metrical language. For example, result (7)(i) involves two lengths along *non*-parallel directions, and uses also the further metrical concept of orthogonality (in order to define 'height'). However all that we need for our purposes are the *following non-metrical versions of the results* (7):

$$\left.\begin{array}{ll}\text{(i)} & \text{parallelograms on the same base whose 'top' sides} \\ & \text{are collinear (see the diagram) have equal area.} \\ \text{(ii)} & \text{parallelepipeds on the same base whose 'top' faces} \\ & \text{are coplanar have equal volumes.}\end{array}\right\} \quad (8)$$



The result (8)(i) can be stated in the form: *the area of a parallelogram is invariant under shearing*. This result is in fact a very easy consequence of properties (5). To see this, note that the two triangles in the diagram have equal area because one is the translate of the other by the vector amount $\mathbf{v}$. Similarly result (8)(ii) can be summarized in the form: *the volume of a parallelogram is invariant under shearing*. This can similarly be seen as a consequence of properties (5′)—for example by considering a sequence of two shears in the two directions defined by the edges forming the base of the parallelepiped. Alternatively it might help to think of a (very large) pack of (very thin) cards. After a shear parallel to the base, the total volume remains unchanged because the volume of each card is unchanged.

3. The product $[\mathbf{u}, \mathbf{v}]$

Let area $(\mathbf{u}, \mathbf{v})$ denote the area of the parallelogram with adjacent sides $\mathbf{u}, \mathbf{v} \in E_2$. Then we define the product $[\mathbf{u}, \mathbf{v}]$ by

$$[\mathbf{u}, \mathbf{v}] = \pm \operatorname{area}(\mathbf{u}, \mathbf{v}) \tag{9}$$

where the sign is fixed as follows. In the case when $\mathbf{u}, \mathbf{v} \in E_2$ are linearly independent, and hence form a basis for E_2, we choose the plus sign or minus sign according as the ordered basis $\{\mathbf{u}, \mathbf{v}\}$ gives rise to the same or opposite orientation for E_2 as the ordered basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ used in our normalization convention (6). In the case when $\mathbf{u}, \mathbf{v}$ are linearly dependent then area $(\mathbf{u}, \mathbf{v}) = 0$ and so $[\mathbf{u}, \mathbf{v}] = 0$ whatever the choice of sign. The important algebraic properties of the product $[\mathbf{u}, \mathbf{v}]$ are as follows:

$$\left.\begin{array}{lll} \text{B.0.} & \textit{Normalization} & [\mathbf{e}_1, \mathbf{e}_2] = +1 \\ \text{B.1.} & \textit{Bilinearity} & \\ \text{B.2.} & \textit{Skew symmetry} & [\mathbf{u}, \mathbf{v}] = -[\mathbf{v}, \mathbf{u}] \end{array}\right\} \tag{10}$$

Properties B.0 and B.2 follow immediately from the definition (9) (granted the normalization agreement (6)). Because of the skew symmetry, to prove bilinearity, i.e. linearity of [,] in each of its arguments, it suffices to prove linearity of [,] in its first argument:

$$\left.\begin{array}{ll} \text{(i)} & [c\mathbf{u}, \mathbf{v}] = c[\mathbf{u}, \mathbf{v}], c \in \mathscr{R} \\ \text{(ii)} & [\mathbf{x}+\mathbf{y}, \mathbf{v}] = [\mathbf{x}, \mathbf{v}] + [\mathbf{y}, \mathbf{v}] \end{array}\right\} \tag{11}$$

Property (11)(i) causes no difficulty: first of all we can prove it for c an integer (easy!), then for c rational and then, by a limiting argument for general real c. (Incidentally the point of the definition (9) of $[\mathbf{u}, \mathbf{v}]$ as *oriented* area was precisely to achieve (11)(i), thus avoiding the $|c|$ which occurs in area $(c\mathbf{u}, \mathbf{v}) = |c|$ area $(\mathbf{u}, \mathbf{v})$ and which stands in the way of bilinearity.)

To prove (11)(ii) let us make use of our previous result concerning the invariance of area under shearing. In our present notation this reads (see the above diagram again):

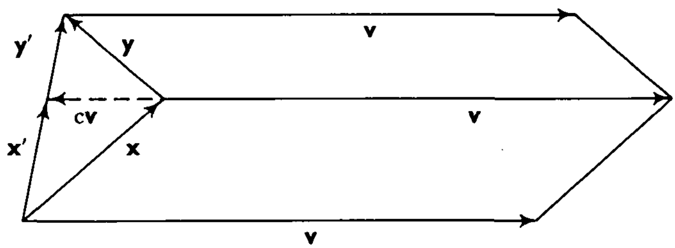
B.3. *Invariance under shearing* $[\mathbf{u}+c\mathbf{v},\mathbf{v}]=[\mathbf{u},\mathbf{v}]$ (12)

This is correct for *oriented* areas, since the ordered bases $\{\mathbf{u},\mathbf{v}\}$ and $\{\mathbf{u}+c\mathbf{v},\mathbf{v}\}$ always define the same orientation for E_2, for all $c\in\mathscr{R}$. Incidentally B.3 is a simple consequence of B.1 and B.2:

$$[\mathbf{u}+c\mathbf{v},v]=[\mathbf{u},\mathbf{v}]+c[\mathbf{v},\mathbf{v}], \text{ by B.1}$$
$$=[\mathbf{u},\mathbf{v}], \text{ by B.2} \tag{13}$$

However, our logic at present is proceeding in the reverse direction, making use of the previously proven B.3 in order to prove B.1.

Now (11)(ii) is a very easy consequence of (5)(ii), and (9) in the case when $\mathbf{x}$ and $\mathbf{y}$ are collinear. But the case of non-collinear $\mathbf{x},\mathbf{y}$ can be reduced to the collinear case by making use of B.3. To see this simply define $c\mathbf{v}$ as in the next diagram so that



$\mathbf{x}'=\mathbf{x}+c\mathbf{v}$ is collinear with $\mathbf{y}'=\mathbf{y}-c\mathbf{v}$. Then we have

$$[\mathbf{x}+\mathbf{y},\mathbf{v}]=[\mathbf{x}'+\mathbf{y}',\mathbf{v}], \quad \text{since } \mathbf{x}+\mathbf{y}=\mathbf{x}'+\mathbf{y}'$$
$$=[\mathbf{x}',\mathbf{v}]+[\mathbf{y}',\mathbf{v}], \quad \text{since } \mathbf{x}',\mathbf{y}' \text{ are collinear}$$
$$=[\mathbf{x},\mathbf{v}]+[\mathbf{y},\mathbf{v}], \quad \text{by B.3} \tag{14}$$

This concludes the proof of the fundamental properties B.0, B.1 and B.2 of the product $[\mathbf{u},\mathbf{v}]$.

A consequence of B.0 and B.2 is that we know the 4 ($=2\times 2$) values $[\mathbf{e}_i,\mathbf{e}_j]$ of the product [,] upon basis vectors:

B.0′. *Values on a basis*

$$[\mathbf{e}_1,\mathbf{e}_2]=1,\ [\mathbf{e}_2,\mathbf{e}_1]=-1,\ [\mathbf{e}_1,\mathbf{e}_1]=0=[\mathbf{e}_2,\mathbf{e}_2] \tag{15}$$

Now any function M which is multilinear in its vector arguments has the great simplifying property that its values $M(\mathbf{u},\mathbf{v},\ldots)$ for general vectors $\mathbf{u},\mathbf{v},\ldots$ are completely determined once its values $M(\mathbf{e}_i,\mathbf{e}_j,\ldots)$ on the vectors of a basis are known. In the particular case of the function [,] of present concern, we deduce from B.1 that

$$[\mathbf{u},\mathbf{v}]=[u_1\mathbf{e}_1+u_2\mathbf{e}_2,v_1\mathbf{e}_1+v_2\mathbf{e}_2]$$
$$=u_1v_1[\mathbf{e}_1,\mathbf{e}_1]+u_1v_2[\mathbf{e}_1,\mathbf{e}_2]$$
$$+u_2v_1[\mathbf{e}_2,\mathbf{e}_1]+u_2v_2[\mathbf{e}_2,\mathbf{e}_2]$$

and hence, from B.0′, obtain the result

$$[\mathbf{u},\mathbf{v}]=u_1v_2-u_2v_1 \tag{16}$$

In this way we have computed the area of a general parallelogram in units of our standard parallelogram:

$$\frac{\text{area}\,(\mathbf{u},\mathbf{v})}{\text{area}\,(\mathbf{e}_1,\mathbf{e}_2)}=|u_1v_2-u_2v_1| \tag{17}$$

It should be observed that *the above derivation of (16) was entirely non-metrical.* In the introduction we noted that the ratio of segments lying in the same line is non-metrical in character. In (17) we have now demonstrated the two-dimensional analogue of this, namely that *the ratio of areas lying in the same plane is entirely non-metrical in character*. (Of course we have only so far considered areas of parallelograms; we can, however, arrive at areas of more complicated regions of the plane by filling up the plane with a finer and finer mesh of parallelograms and thereby squeeze the area of interest between lower and upper bounds which get closer and closer together.) Incidentally a more geometric demonstration of the non-metrical nature of area ratio could have been given, by showing how an arbitrary parallelogram can be transformed into one whose sides are parallel to our standard parallelogram by means of two suitable shears.

Of course the above non-metrical considerations do *not* allow us to compare areas in two *non-parallel* planes lying in E_3. In each plane we can choose a standard parallelogram as unit of area, but we are able to compare the units and so obtain an absolute notion of area only if we make use of the metrical properties of E_3—for example by laying down that every square of side 1 has unit area.

Once we are in a three-dimensional metrical setting I agree that the role of oriented area $[\mathbf{u},\mathbf{v}]$, $\mathbf{u},\mathbf{v}\in E_2$, will be taken over by the vector area $\mathbf{u}\times\mathbf{v}$, $\mathbf{u},\mathbf{v}\in E_3$, with (17) being replaced by area $(\mathbf{u},\mathbf{v})=|\mathbf{u}\times\mathbf{v}|$. Nevertheless I claim that the preliminary teaching of the oriented area $[\mathbf{u},\mathbf{v}]$ (and of the associated 2×2 determinants $u_1v_2-u_2v_1$) is very well worth while because it paves the way for a subsequent painless treatment of oriented volume $[\mathbf{u},\mathbf{v},\mathbf{w}]$ (and of 3×3 determinants)—as outlined in the next section.

4. The scalar triple product $[\mathbf{u},\mathbf{v},\mathbf{w}]$

Let vol$(\mathbf{u},\mathbf{v},\mathbf{w})$ denote the volume of the parallelepiped with concurrent edges $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}\in E_3$. Then we define the scalar triple product $[\mathbf{u},\mathbf{v},\mathbf{w}]$ as the *oriented volume*

$$[\mathbf{u},\mathbf{v},\mathbf{w}]=\pm\text{vol}(\mathbf{u},\mathbf{v},\mathbf{w}) \tag{9'}$$

where the plus or minus sign is chosen according to whether the ordered basis $\{\mathbf{u},\mathbf{v},\mathbf{w}\}$ is right- or left-handed. (In the case when $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are linearly dependent, the parallelepiped is degenerate (planar) and has zero volume, so the choice of sign is immaterial.) In our normalization convention (6′) let us agree to use a right-handed bases $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$. The fundamental algebraic properties of $[\mathbf{u},\mathbf{v},\mathbf{w}]$ are then as follows:

T.0. *Normalization* $[\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3]=+1$

T.1. *Trilinearity*

T.2. *Skew symmetry* $[\mathbf{u},\mathbf{v},\mathbf{w}]=[\mathbf{v},\mathbf{w},\mathbf{u}]=[\mathbf{w},\mathbf{u},\mathbf{v}]$

$$=-[\mathbf{u},\mathbf{w},\mathbf{v}]=-[\mathbf{v},\mathbf{u},\mathbf{w}]=-[\mathbf{w},\mathbf{v},\mathbf{u}] \tag{10'}$$

Properties T.0 and T.2 follow immediately from (9′) and the foregoing normalization agreement. The proof of T.1 proceeds on completely analogous lines to the proof of B.1 given in §3, and involves an appeal to the result in Section 2 concerning the invariance of volume under shearing. This last reads in our present notation:

T.3. *Invariance under shearing* for all $c, d \in \mathscr{R}$

$$[\mathbf{u}+c\mathbf{v}+d\mathbf{w}, \mathbf{v}, \mathbf{w}]=[\mathbf{u}, \mathbf{v}, \mathbf{w}] \tag{12'}$$

The analogue of (14) then goes through, provided that the vector $c\mathbf{v}$ in the previous diagram is replaced by an appropriate linear combination $c\mathbf{v}+d\mathbf{w}$ of the vectors $\mathbf{v}, \mathbf{w}$ defining the bases of the parallelepipeds in question.

Armed with these fundamental algebraic properties, the rest is plain sailing. Properties T.0 and T.2 tell us the $27(=3\times3\times3)$ values $[\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k]$ of the function $[\ ,\ ,\ ]$ upon the vectors of our particular basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$:

T.0′. *Values on a basis*

$$[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]=[\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1]=[\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2]=+1$$

$$[\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2]=[\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3]=[\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1]=-1$$

$$[\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k]=0, \text{ if } i=j \text{ or } j=k \text{ or } k=i \tag{15'}$$

The values for general vectors then follow by trilinearity. In detail, property T.1 entails that

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}]=[\Sigma u_i\mathbf{e}_i, \Sigma v_j\mathbf{e}_j, \Sigma w_k\mathbf{e}_k]$$
$$=\Sigma\Sigma\Sigma u_iv_jw_k[\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k]$$

whence T.0′ yields the result

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}]=u_1v_2w_3+u_2v_3w_1+u_3v_1w_2$$
$$-u_1v_3w_2-u_2v_1w_3-u_3v_2w_1 \tag{16'}$$

In this way we have computed the volume of a general parallelepiped in units of our standard parallelepiped:

$$\frac{\operatorname{vol}(\mathbf{u}, \mathbf{v}, \mathbf{w})}{\operatorname{vol}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)}=|D| \tag{17'}$$

where D denotes the expression (called a 3×3 determinant†) on the right-hand side of (16′). *Our derivation of (16′) was entirely non-metrical, so that we have demonstrated the fact that the ratio of volumes (in three-dimensional space) is a non-metrical notion.* (Of course in a metrical setting we would normally lay down that a cube of side 1 has unit volume—for example by laying down that our normalization convention T.0 should be applied to a right-handed orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.)

At this stage one can profitably, and painlessly, deal with the solution of three linear equations in three unknowns. Let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, be given vectors $\in E_3$ (or given triples $\in\mathscr{R}^3$), and let x, y, z be the 'unknown' real numbers.

Homogeneous equations

The equation

$$x\mathbf{a}+y\mathbf{b}+z\mathbf{c}=\mathbf{0} \tag{18}$$

has a non-trivial solution if and only if $[\mathbf{a}, \mathbf{b}, \mathbf{c}]=0$.

† Observe that the properties T.1, T.2, T.3 already tell us (if we trust our geometrical approach!) that determinants satisfy corresponding properties, say D.1, D.2, D.3, when thought of as functions of their column triples.

Proof

Vol(**a**, **b**, **c**) = 0 if and only if the edges **a**, **b**, **c** are coplanar, that is linearly dependent, that is satisfy (18) for x, y, z not all zero.

Non-homogeneous equations

Suppose that $[\mathbf{a}, \mathbf{b}, \mathbf{c}] \neq 0$. Then the equation

$$x\mathbf{a} + y\mathbf{b} + z\mathbf{c} = \mathbf{d} \tag{19}$$

has the unique solution (*Cramer's Rule*)

$$x = \frac{[\mathbf{d}, \mathbf{b}, \mathbf{c}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}, \quad y = \frac{[\mathbf{a}, \mathbf{d}, \mathbf{c}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}, \quad z = \frac{[\mathbf{a}, \mathbf{b}, \mathbf{d}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]} \tag{20}$$

Proof

(i) *Existence and uniqueness of solution.*

We are given that **a**, **b**, **c** are non-coplanar (and hence form a basis for E_3). Consequently *any* vector $\mathbf{d} \in E_3$ can be expressed in the form (19), its components (x, y, z) being *uniquely* determined.

(ii) *Explicit form of the solution*

$$\begin{aligned}[\mathbf{d}, \mathbf{b}, \mathbf{c}] &= [x\mathbf{a} + y\mathbf{b} + z\mathbf{c}, \mathbf{b}, \mathbf{c}], \text{ from (19)} \\ &= x[\mathbf{a}, \mathbf{b}, \mathbf{c}], \text{ from T.3 and T.1}\end{aligned}$$

Similarly we obtain y and z by considering $[\mathbf{a}, \mathbf{d}, \mathbf{c}]$ and $[\mathbf{a}, \mathbf{b}, \mathbf{d}]$.

5. The vector product $\mathbf{u} \times \mathbf{v}$

It is only *after* we have captured the non-metrical aspects of area and volume in the products $[\mathbf{u}, \mathbf{v}]$ and $[\mathbf{u}, \mathbf{v}, \mathbf{w}]$ that the time is ripe to invoke the metrical aspects involved in the definition of the vector product and its relation to the scalar triple product. So only at this stage should we introduce the notion of vector area and define $\mathbf{u} \times \mathbf{v}$ in terms of the vector area of a parallelogram. It then follows from the metrical result (7)(ii) for the volume of a parallelepiped that the scalar, vector and triple products are related by

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \mathbf{u} \times \mathbf{v} . \mathbf{w} \tag{21}$$

(I am here assuming that the scalar product $\mathbf{u} . \mathbf{v}$ has already been introduced. As suggested in the teaching scheme (4), whether this has been done before or after the introduction of the triple product is of no great importance.)

It follows from its definition that the vector product is skew symmetric: $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$. Of crucial importance (for example in obtaining its component form) is the fact that the vector product is bilinear—and, though possible, it is not quite so easy to give a proof of this directly from the geometrical definition of $\mathbf{u} \times \mathbf{v}$. However, knowing as we do that the scalar triple product exists and is of fundamental importance quite independently of the vector product, why not use the proven trilinearity of the former to prove the bilinearity of the latter? (It would certainly be perverse to go in the other direction!). In fact such a proof follows immediately from (21), upon using the fact that $\mathbf{x} . \mathbf{w} = \mathbf{y} . \mathbf{w}$ for all $\mathbf{w}$ implies that $\mathbf{x} = \mathbf{y}$. Of course one needs to invoke also the bilinearity of the scalar product $\mathbf{u} . \mathbf{v}$.

Incidentally (as the reader may well have gathered already!) I believe that in the teaching of vector algebra no opportunity should be lost to stress the simplicity and importance of multilinearity ideas. Such an opportunity presents itself at this stage, since such ideas lead one to very simple proofs of certain vector identities involving the scalar and vector triple products—see [2].

6. The rigorous algebraic approach

Once our geometric intuition has convinced us that an elegant algebra of vectors 'really exists', the challenge arises to place our 'knowledge' into a rigorous algebraic framework—by setting up appropriate axioms and rigorously deducing theorems. Such theorems will apply to the great variety of situations (often non-geometric in nature) which fit in with the axioms. However, some of them can be interpreted geometrically, and we thereby can reverse our original progression and now use algebra to provide us with rigorous proofs of geometric theorems. For example, the axioms for an inner product space yield the fact that $\mathbf{a}.\mathbf{b}=0$ implies that $(\mathbf{a}+\mathbf{b})^2=\mathbf{a}^2+\mathbf{b}^2$, which amounts to a one-line proof of Pythagoras' theorem.

Let me in this section give only the barest outlines of the algebraic approach (as suitable to be taught in a first-year undergraduate course on linear algebra). Firstly the study of linear combinations of geometric vectors leads one to set up the familiar axioms for a real vector space. Secondly for any three-dimensional real vector space V_3 one *defines* the determinant function (= scalar triple product) to be that unique function [, ,] which satisfies properties T.0, T.1 and T.2 of (10′). (Such a function is unique on account of the algebraic computation which led to equation (16′.) To prove its existence, define a function by (16′) and check that it indeed satisfies T.0, T.1 and T.2).

Next one axiomatizes the notion of an inner product $\mathbf{u}.\mathbf{v}$ by means of the three properties:

$$\left.\begin{array}{ll} \text{S.1.} & \textit{Bilinearity} \\ \text{S.2.} & \textit{Symmetry} \\ \text{S.3.} & \textit{Positive-definiteness} \end{array}\right\} \qquad (22)$$

For a given real vector space there are a host of choices of inner product. For example if $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is *any* basis for V_3 we can define an inner product for V_3 by decreeing that this basis shall be orthonormal: $\mathbf{e}_i.\mathbf{e}_j=\delta_{ij}$ (definition of a bilinear function by its values on basis vectors).

Let V_3 now be equipped with a particular choice of inner product. Then one can prove that an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ exists. Let us fix the triple scalar product for V_3 by the normalization condition that $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ equals $+1$ for this particular orthonormal basis. If $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is any orthonormal basis then it follows from the identity (4) in [2] that $[\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3]^2=1$. (Geometrically speaking we have thus demonstrated algebraically that *every* cube of side 1 has unit volume!) Consequently orthonormal bases fall into two orientation classes, say (i) positive, or 'right-handed', and (ii) negative, or 'left-handed', according to whether

$$\text{(i)}\quad [\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3]=+1, \quad \text{or (ii)}\quad [\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3]=-1 \qquad (23)$$

Finally, relative to a particular choice of inner product for V_3, we can define algebraically an associated vector product $\mathbf{u}\times\mathbf{v}$ as follows. We appeal to the (easily proved) result that any linear function $f: V_3 \to \mathscr{R}$ is of the form $f(\mathbf{w})=\mathbf{f}.\mathbf{w}$ for some unique $\mathbf{f}\in V_3$. It follows from this that for fixed $\mathbf{u}, \mathbf{v}\in V_3$ there exists a unique vector $\mathbf{f}$, depending of course upon $\mathbf{u}, \mathbf{v}$, such that

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}]=\mathbf{f}.\mathbf{w}, \quad \text{for all} \quad \mathbf{w}\in V_3 \qquad (24)$$

Let us write the dependence of $\mathbf{f}$ upon $\mathbf{u}$, $\mathbf{v}$ as $\mathbf{f}=\mathbf{u}\times\mathbf{v}$. Thus we see that (21) now appears as the *definition* of the vector product $\mathbf{u}\times\mathbf{v}$! The algebraic properties of $\mathbf{u}\times\mathbf{v}$ are now easily proved from the corresponding ones of $[\mathbf{u},\mathbf{v},\mathbf{w}]$. For example (23)(i) and (24) yield the result that the vectors of any positive orthonormal basis satisfy

$$\mathbf{f}_2\times\mathbf{f}_3=+\mathbf{f}_1,\qquad \mathbf{f}_3\times\mathbf{f}_1=+\mathbf{f}_2,\qquad \mathbf{f}_1\times\mathbf{f}_2=+\mathbf{f}_3 \tag{25}$$

Observe that the host of possible inner products for V_3 is filtered through the *unique* (granted our normalization agreement) triple product $[\mathbf{u},\mathbf{v},\mathbf{w}]$ *to give rise to a corresponding host of associated vector products.*

Starting out from the algebraic characterizations (22) and (24) of $\mathbf{u}.\mathbf{v}$ and $\mathbf{u}\times\mathbf{v}$, let me conclude by indicating how to make contact with the original geometric definitions uv cos θ and uv sin $\theta\mathbf{n}$. Firstly, the axioms S.1–S.3 imply the Schwarz inequality $|\mathbf{u}.\mathbf{v}|\leqslant uv$ and so allow us to *define* θ by $\cos\theta=\mathbf{u}.\mathbf{v}/(uv)$. Secondly by setting $\mathbf{w}=\mathbf{u}$ in (24) we obtain

$$\mathbf{u}\times\mathbf{v}.\mathbf{u}=[\mathbf{u},\mathbf{v},\mathbf{u}]=0$$

Similarly $\mathbf{u}\times\mathbf{v}.\mathbf{v}=0$, and so $\mathbf{u}\times\mathbf{v}$ is indeed orthogonal to both $\mathbf{u}$ and $\mathbf{v}$. The identity (see (2) of [2])

$$\mathbf{u}\times\mathbf{v}.\mathbf{u}\times\mathbf{v}=u^2v^2-(\mathbf{u}.\mathbf{v})^2$$

now yields $(\mathbf{u}\times\mathbf{v})^2=u^2v^2(1\times\cos^2\theta)=(uv\sin\theta)^2$, so that $\mathbf{u}\times\mathbf{v}$ has the 'correct' magnitude. Finally $\mathbf{u}\times\mathbf{v}$ 'points in the correct sense' since (24) yields

$$[\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}]=(\mathbf{u}\times\mathbf{v})^2\geqslant 0$$

which states, in the case of linearly independent $\mathbf{u}$, $\mathbf{v}$, that $[\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}]$ and $[\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3]$ both define the same orientation for V_3.

References

[1] DIEUDONNÉ, 1969, *Linear Algebra and Geometry* (Kershaw).
[2] SHAW, R., 1985, A new proof of $(\mathbf{a}\times\mathbf{b})\times\mathbf{u}=(\mathbf{a}.\mathbf{u})\mathbf{b}-(\mathbf{b}.\mathbf{a})\mathbf{a}$. *Int. J. Math. Educ. Sci. Technol.* **16,** 561.