A Way of Thinking about Quaternions

Ted Szylowiec

1 Introduction

Classic textbooks[1] on quaternions usually begin by introducing additive and associative properties of quaternion vectors. Then they present many problems and examples from elementary geometry that are solved by applying these properties. Although the introductory problems and examples presented in these old books are fascinating in their own right and likely unfamiliar to today's students, the method of solution by quaternion vectors is hardly different from solutions which can be obtained by modern linear algebra.

The multiplicative properties of quaternion vectors are developed later in the books—sometimes much later—upon which the reader may be alarmed to discover that p times q is not the same as q times p in the calculus of quaternions. Heaviside, Graves and De Morgan objected to this non-commutativity, suggesting that Hamilton's calculus was flawed because of it. But non-commutativity was not an arbitrary choice on the part of Hamilton. Quaternions were not an invention but a discovery. And they turned out to be non-commutative. It is the multiplicative aspect of quaternion calculus that distinguishes it. Concerning this Tait writes in chapter II of *Elementary Treatise on Quaternions*:

> We now come to the consideration of question in which the Calculus of Quaternions differs entirely from any previous mathematical methods; and here we shall get an idea of what a Quaternion is, and when it derives its name. These questions are fundamentally involved in the novel use of the symbols of multiplication and division.

In light of this, perhaps it is best to approach the study of quaternions the other way around—starting with their multiplicative properties. We will depart from the usual pedagogical development of the classic quaternion books and instead introduce non-commutativity immediately, even before introducing quaternions or even vectors. Geometry is essentially non-commutative, so this development seems logical. We will see that non-commutativity is a feature of Hamilton's calculus, not a defect as claimed by Heaviside. Algebraic computations with quaternions give geometrical answers in remarkable ways precisely because quaternion multiplication is non-commutative.

[1] W. R. Hamilton, *Elements of Quaternions*. P. G. Tait, *An Elementary Treatise on Quaternions*. P. Kelland and P. G. Tait, *An Introduction to Quaternions*.

2 Three propositions

Let's dispense with the requirement that quantities should be commutative and agree that α, β, γ may or may not be commutative. That is, $\alpha\beta$ may or may not be the same as $\beta\alpha$.

We are at once led to the natural idea of geometrically interpreting three different algebraic cases regarding these quantities. First, where α and β commute: $\alpha\beta = \beta\alpha$. Second, where α and β anti-commute: $\alpha\beta = -\beta\alpha$. And the third, where α and β neither commute nor anti-commute: $\alpha\beta \neq \beta\alpha$ and $\alpha\beta \neq -\beta\alpha$.

At this point we don't know exactly what these α and β are, but since we will be associating them with orientation of lines, we might as well call them *vectors*.

We can study the nature of the product $\alpha\beta$ by defining two operators, $\mathbf{S}$ and $\mathbf{V}$, in terms of the commutator $\alpha\beta - \beta\alpha$ and anti-commutator $\alpha\beta + \beta\alpha$ of α and β:

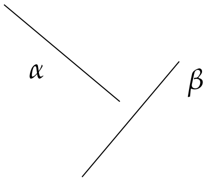
$$\mathbf{S}\alpha\beta = \frac{\alpha\beta + \beta\alpha}{2}$$
$$\mathbf{V}\alpha\beta = \frac{\alpha\beta - \beta\alpha}{2}. \quad (1)$$

The three cases can be phrased in terms of $\mathbf{S}$ and $\mathbf{V}$ as three propositions.

Let's propose that if α and β commute then they represent—in some sense—parallel lines.

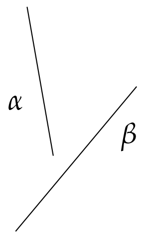
In this case the commutator $\alpha\beta - \beta\alpha$ vanishes while the anti-commutator $\alpha\beta + \beta\alpha$ does not vanish. Using the operators defined above, we can say that if $\alpha \parallel \beta$ then $\mathbf{S}\alpha\beta \neq 0$ and $\mathbf{V}\alpha\beta = 0$.

Next, let us propose that if α and β anti-commute then they represent, in some way, perpendicular lines.



Here, $\alpha\beta + \beta\alpha$ vanishes while $\alpha\beta - \beta\alpha$ does not. Using the operators in (1), we can say that if $\alpha \perp \beta$ then $\mathbf{S}\alpha\beta = 0$ while $\mathbf{V}\alpha\beta \neq 0$.

Finally, let's propose that if α and β neither commute nor anti-commute, then they represent lines which are neither parallel nor perpendicular.



i.e., lines that are inclined obliquely or acutely. In this case neither $\alpha\beta + \beta\alpha$ nor $\alpha\beta - \beta\alpha$ vanish. Using our operators, we can say that, for this case, $\mathbf{S}\alpha\beta \neq 0$ and $\mathbf{V}\alpha\beta \neq 0$.

Arrange these propositions in a table for easy reference:

Commutativity	Product $\alpha\beta$	Geometry
α, β commute	$\mathbf{S}\alpha\beta \neq 0$, $\mathbf{V}\alpha\beta = 0$	$\alpha \parallel \beta$
α, β anti-commute	$\mathbf{S}\alpha\beta = 0$, $\mathbf{V}\alpha\beta \neq 0$	$\alpha \perp \beta$
neither	$\mathbf{S}\alpha\beta \neq 0$, $\mathbf{V}\alpha\beta \neq 0$	neither

3 Decomposition of $\alpha\beta$

When $\alpha \parallel \beta$ we have $\alpha\beta = \mathbf{S}\alpha\beta$. And when $\alpha \perp \beta$ we have $\alpha\beta = \mathbf{V}\alpha\beta$. In any other orientation of α and β we have $\alpha\beta = \mathbf{S}\alpha\beta + \mathbf{V}\alpha\beta$. These are easy to verify. We will verify the last one:

$$\begin{aligned}\mathbf{S}\alpha\beta + \mathbf{V}\alpha\beta &= \frac{\alpha\beta + \beta\alpha + \alpha\beta - \beta\alpha}{2} \\ &= \alpha\beta. \end{aligned} \tag{2}$$

The relation $\alpha\beta = \mathbf{S}\alpha\beta + \mathbf{V}\alpha\beta$ tells us that the product of any two vectors can be decomposed into the sum of an $\mathbf{S}$–part and a $\mathbf{V}$–part.

From the definitions (1) it can be easily shown that $\mathbf{S}$ and $\mathbf{V}$ are distributive over addition and homogeneous with respect to multiplication by numbers m and n:

$$\begin{aligned}\mathbf{S}(m\alpha\beta + n\gamma\delta) &= m\mathbf{S}\alpha\beta + n\mathbf{S}\gamma\delta \\ \mathbf{V}(m\alpha\beta + n\gamma\delta) &= m\mathbf{V}\alpha\beta + n\mathbf{V}\gamma\delta\end{aligned}$$

With the help of these algebraic properties, it's easy to establish other interesting results concerning $\mathbf{S}$ and $\mathbf{V}$. For example: that $\mathbf{S}$ annihiliates the $\mathbf{V}$–part of $\alpha\beta$, while $\mathbf{V}$ annihilates the $\mathbf{S}$–part of $\alpha\beta$:

$$\begin{aligned}\mathbf{S}\,.\,\mathbf{V}\alpha\beta = \mathbf{S}\,.\,\frac{\alpha\beta - \beta\alpha}{2} &= \frac{1}{2}(\mathbf{S}\alpha\beta - \mathbf{S}\beta\alpha) \\ &= \frac{1}{4}(\alpha\beta + \beta\alpha - \beta\alpha - \alpha\beta) = 0.\end{aligned}$$

Likewise it can be shown that $\mathbf{V}\,.\,\mathbf{S}\alpha\beta = 0$. Here we used the *point* or *stop* symbol, '$.$', which Hamilton and Tait use to make expressions involving $\mathbf{S}$ and $\mathbf{V}$ less ambiguous without resorting to writing a lot of parentheses. It means that whatever is to the left of the point is applied to the expression on the right of the point. In the same manner, which is left as an exercise, it can be shown that $\mathbf{S}\,.\,\mathbf{S}\alpha\beta = \mathbf{S}\alpha\beta$ and $\mathbf{V}\,.\,\mathbf{V}\alpha\beta = \mathbf{V}\alpha\beta$. Combining these properties of $\mathbf{S}$ and $\mathbf{V}$ with the decomposition (2) we see that $\mathbf{S}$ is an operator that, when applied to a product of vectors $\alpha\beta$, returns the $\mathbf{S}$–part of the product and annihilates the $\mathbf{V}$–part:

$$\begin{aligned}\mathbf{S}\,.\,\alpha\beta &= \mathbf{S}\,.\,(\mathbf{S}\alpha + \mathbf{V}\alpha\beta) = \mathbf{S}\,.\,\mathbf{S}\alpha\beta + \mathbf{S}\,.\,\mathbf{V}\alpha\beta \\ &= \mathbf{S}\alpha\beta\end{aligned}$$

while $\mathbf{V}$ is an operator that, when applied to $\alpha\beta$, preserves the $\mathbf{V}$–part of the product and annihilates the $\mathbf{S}$–part:

$$\begin{aligned}\mathbf{V}\,.\,\alpha\beta &= \mathbf{V}\,.\,(\mathbf{S}\alpha + \mathbf{V}\alpha\beta) = \mathbf{V}\,.\,\mathbf{S}\alpha\beta + \mathbf{V}\,.\,\mathbf{V}\alpha\beta \\ &= \mathbf{V}\alpha\beta.\end{aligned}$$

4 Interpretation of $\mathbf{S}\alpha\beta$ and $\mathbf{V}\alpha\beta$

Let α and β be parallel. By the propositions of part 2, $\alpha\beta - \beta\alpha = 0$. If γ is any vector perpendicular to α, it will be perpendicular to β as well, and γ will anti-commute with α and anti-commute with β. We have

$$\begin{aligned}\gamma\,.\,\alpha\beta - \alpha\beta\,.\,\gamma &= \gamma\alpha\beta \ - \alpha\beta\gamma \\ &= -\alpha\gamma\beta + \alpha\gamma\beta \\ &= 0.\end{aligned}$$

Therefore $\alpha\beta$ commutes with γ, which is only possible if it is either a vector parallel to γ, or a scalar. But if we choose *another* vector, δ, also perpendicular to α but not perpendicular to γ, we find that $\alpha\beta$

commutes with δ as well. A vector cannot be parallel to both γ and δ. Therefore $\alpha\beta$ must be either 0 or a scalar. We have determined that if two vectors are parallel, *their product is either 0 or a scalar*.

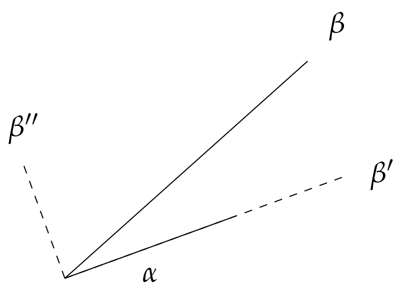
Let α be perpendicular to β. By the propositions of part 2, $\alpha\beta + \beta\alpha = 0$. If γ is any vector parallel to α, it will be perpendicular to β, and γ will commute with α and anti-commute with β:

$$\begin{aligned}\gamma \,.\, \alpha\beta + \alpha\beta \,.\, \gamma &= \gamma\alpha\beta + \alpha\beta\gamma \\ &= \alpha\gamma\beta - \alpha\gamma\beta \\ &= 0.\end{aligned}$$

Thus, γ anti-commutes with the product $\alpha\beta$. By the propositions of part 2, $\alpha\beta$ must be perpendicular to γ. Since a scalar other than 0 cannot be perpendicular to anything, $\alpha\beta$ must be either 0 or a vector. We have determined that if two vectors are perpendicular, *their prodict is either 0 or a vector*.

From the propositions of part 2, $\mathbf{S}\alpha\beta = \alpha\beta$ when $\alpha \parallel \beta$ and $\mathbf{V}\alpha\beta = \alpha\beta$ when $\alpha \perp \beta$. Combined with the results above we have the following: when $\alpha \parallel \beta$, $\mathbf{S}\alpha\beta$ either vanishes or *is a scalar*, and when $\alpha \perp \beta$, $\mathbf{V}\alpha\beta$ either vanishes *or is a vector*.

We now examine the general case where α and β are neither parallel nor perpendicular. Given some vector α, the vector β can be written as the sum of two vectors, $\beta' + \beta''$, where β' is parallel to α and β'' is perpendicular to α.



We can write the product $\alpha\beta$ as $\alpha(\beta' + \beta'')$. For the $\mathbf{S}$-part of $\alpha\beta$ we obtain

$$\begin{aligned}\mathbf{S}\alpha\beta &= \mathbf{S} \,.\, \alpha(\beta' + \beta'') \\ &= \mathbf{S}\alpha\beta' + \mathbf{S}\alpha\beta'' \\ &= \mathbf{S}\alpha\beta' \\ &= \alpha\beta'.\end{aligned}$$

But as we have shown in the previous paragraphs, $\alpha\beta'$ must either vanish or be a scalar. Therefore $\mathbf{S}\alpha\beta$ either vanishes or is a scalar. Now for the $\mathbf{V}$-part:

$$\begin{aligned}\mathbf{V}\alpha\beta &= \mathbf{V} \,.\, \alpha(\beta' + \beta'') \\ &= \mathbf{V}\alpha\beta' + \mathbf{V}\alpha\beta'' \\ &= \mathbf{V}\alpha\beta'' \\ &= \alpha\beta''.\end{aligned}$$

Since have previously shown that $\alpha\beta''$ must either vanish or be a vector, $\mathbf{V}\alpha\beta$ either vanishes or is a vector.

We are now in a position to interpret the general nature of the product of two vectors $\alpha\beta$ and the nature of the $\mathbf{S}$ and $\mathbf{V}$ parts in the decomposition (2),

$$\alpha\beta = \mathbf{S}\alpha\beta + \mathbf{V}\alpha\beta.$$

The arguments of the preceeding paragraphs show that the $\mathbf{S}$-part is always a scalar and the $\mathbf{V}$-part is always a vector. This is true even when one of the two vanishes, as happens when $\alpha \parallel \beta$ or $\alpha \perp \beta$, beause *there is no difference, in this calculus, between the 0 scalar and the 0 vector*. You could say that 0 is the only thing whose $\mathbf{S}$-part is equal to its $\mathbf{V}$-part. Another way to see that 0 is both a scalar and a vector is to consider that if every vector is a linear combination of Hamilton's imaginaries,

$$\gamma = xi + yj + zk,$$

then setting $x = y = z = 0$ gives the zero vector.

The justification for Hamilton's choice of symbols $\mathbf{S}$ and $\mathbf{V}$ is now clear. Hamilton calls $\mathbf{S}\alpha\beta$ the *scalar of* $\alpha\beta$ and $\mathbf{V}\alpha\beta$ the *vector of* $\alpha\beta$. Note that Hamilton's use of the terms scalar and vector is different from the way we usually use them in linear algebra, or even the way we have been using them in this article. In Hamilton's context, *the scalar of* and *the vector of* refer to operators that act on objects, while we usually use the terms *scalar* and *vector* to signify the objects themselves.

Thus the product of two vectors is either a scalar, $\alpha\beta = w$, or a vector, $\alpha\beta = \gamma$, or neither: $\alpha\beta = w + \gamma$, that is, a scalar plus a vector. The latter is what Hamilton calls *a quaternion.* Scalars and vectors are degenerate kinds of quaternions. A scalar is a quaternion with no vector part. A vector is a quaternion with no scalar part. Although α and β are vectors and thus have no scalar parts, the product $\alpha\beta$ has, in the general case described above, both a scalar part and a vector part. Therefore, in general, the product of two vectors is not a vector, but a quaternion.

It's hard to imagine a more natural progression of ideas leading up to this. From simple propositions concerning non-commutative multiplication of vectors, we are led to the notion of quaternion, i.e., scalar plus vector.