Investigating the Foundations of Physical Law
2 Mathematical ideas and methods

Quaternions and octonions

There is a mathematics, vital to us, which is spoken of with respect in some quarters, is recognised as the parent of several other branches of mathematics very much in current use, is used every day by software engineers to represent motion in space and time, and appears in a disguised form in the treatment of relativity and quantum mechanics, and yet is regarded with contempt by many who don't know anything about it. Those who don't know their history are condemned to repeat it, and the people who repeat what they have learned from their tutors about the intrinsic worthlessness of this mathematics usually have no idea that they are contributing to the perpetuation of what amounts to an episode of 'cultural genocide' begun more than a hundred years ago.

For physics, the prevalence of this attitude has had disastrous consequences because it has kept us away from the branch of mathematics that is most necessary to understanding its foundations, the only one that has ever given us a handle on the meaning of 3-dimensional space. It is, of course, absurd to claim that any branch of pure mathematics is *intrinsically* worthless and that it will never find valid applications. How can anyone know? In any case, the current massive use in the commercially-driven software industry disproves the claim many of thousands of times over. Yet, it has meant that this branch of mathematics is very little taught in physics courses, making the intrinsic lack of utility a self-fulfilling prophecy. This is, quite frankly, the biggest missed opportunity in the history of the subject.

It was back in 1843 when Sir William Rowan Hamilton realised, for the first time, that he had discovered the meaning of 3-dimensionality. He was trying to extend the idea of complex numbers to the next level. He reasoned that this could be done by adding a third axis to the well-known Argand diagram. On the Argand diagram, the real numbers, based on unit 1, are on the horizontal or x-axis, the imaginary numbers, based on unit $i = \sqrt{(-1)}$, are on the vertical or y-axis. Hamilton proposed drawing a third, z-axis perpendicular to the other two, which couldn't contain real numbers, as these are all on the x-axis, but could conceivably contain another set of imaginary numbers based on a unit $j = \sqrt{(-1)}$ which is different from i. The only problem with this is that it doesn't work algebraically. We can find products of numbers with units 1 and 1, 1 and i, 1 and j, i and i, and j and j, but we can't find a product of numbers with units i and j. It can't have real units, because this would equate i and j; it can't have unit i, because this would equate 1 and j; and it can't have unit j because this would equate 1 and i. The system doesn't exhibit *closure*.

After many years of struggle, he found a solution, but this meant violating one of the cardinal principles of algebra as then known: the principle of commutativity. In effect this meant that if you took the product *ba*, that produced exactly the same answer as taking the product *ab*. What Hamilton did was to remove the real axis entirely and have 3 imaginary axes, with units ***i***, ***j***, ***k***, all different from each other but all equating to √(–1), and following a rotation cycle, so that

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = \boldsymbol{ijk} = -1$$
$$\boldsymbol{ij} = \boldsymbol{k}$$
$$\boldsymbol{ki} = \boldsymbol{j}$$
$$\boldsymbol{jk} = \boldsymbol{i}$$

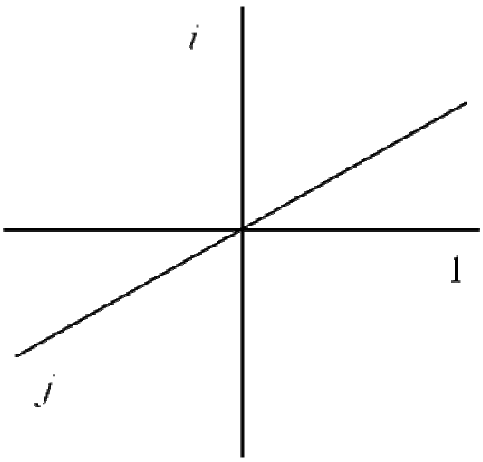
This works well but it requires a reversal of sign when we reverse the order of multiplication. That is:

$$\boldsymbol{ji} = -\boldsymbol{k}$$
$$\boldsymbol{ik} = -\boldsymbol{j}$$
$$\boldsymbol{kj} = -\boldsymbol{i}$$

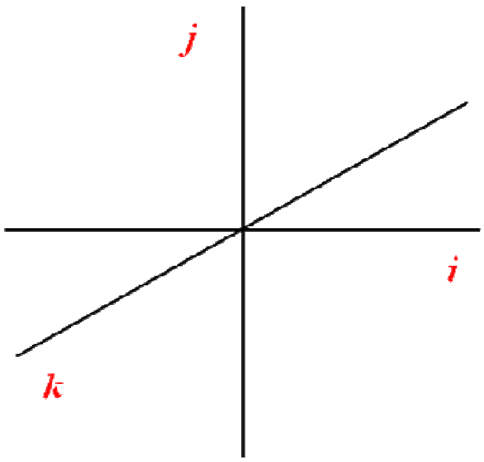
The units are *anticommutative*. The reason can be seen immediately. Let us, for example, take the product ***ij ji***. We multiply the ***j*** units first, and their product is –1. The remaining term is – ***ii***, which clearly equals 1. So

$$\boldsymbol{ij}\,\boldsymbol{ji} = -\,\boldsymbol{ii} = 1$$

But this can only be true if ***ji*** = –***k***. Accepting this had to be done, Hamilton now had a closed algebra, with four basic units 1, ***i***, ***j***, ***k***, which he called *quaternions*, and which was double the size of ordinary complex algebra with units 1, *i*, and four times the size of real algebra, based on unit 1.



extended Argand diagram



quaternions

The next question he asked was could we increase the size of this algebra and still maintain the same rules, say with one real and four imaginary units: 1, ***i***, ***j***, ***k***, ***l***? And the answer is we can't. You can't get consistency if you extend the number of imaginary units beyond three. Hamilton suspected this was the case; Frobenius proved it in 1878. There is just one exception. You can create a consistent system with one real and *seven* imaginary units, 1, ***i***, ***j***, ***k***, ***e***, ***f***, ***g***, ***h***, called the *octonions*. However, to do this you have to break another algebraic rule, the law of associativity in which, say, $(ab)c$ is always equal to $a(bc)$. So, octonions, unlike quaternions, are antiassociative as well as anticommutative, leading to products equivalent to $(ab)c = -a(bc)$.

The trick can't be repeated at any other level, so we are left with just four so-called division algebras:

Real	norm 1	commutative	associative
Complex	norm –1	commutative	associative
Quaternions	norm –1	anticommutative	associative
Octonions	norm –1	anticommutative	antiassociative

If we are looking for quaternions, as Hamilton did, as the potential explanation for the 3-dimensionality of space, the thing to note is that the '3-ness' isn't the primary cause. It is simply a result of anticommutativity. If we have two axes, ***i*** and ***j***, that are anticommutative with each other, then we cannot draw any other axis that is anticommutative with them, unless it is ***ij***, which we also call ***k***. Anticommutativity forces 3-dimensionality. The strange arbitrariness of the number 3 is explained. Commutative things, of course, can be defined to infinity. If ***i*** and ***j*** were commutative, we could have ***i***, ***j***, ***k***, ***l***, ***m***, etc. without limit. In a sense we can say that anticommutative things 'know' about each other's presence and have to act accordingly; commutative things do not.

Now, just as a complex number is fully represented by the sum of a real part and an imaginary part, say, $x + iy$, so a quaternion number will have a real and three imaginary parts, say, $\boldsymbol{a} = w + \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$, where w, x, y, z are just positive or negative scalars or real numbers. (For convenience, we will always represent quaternion units by ***bold italic*** symbols.) Let us now suppose we have another quaternion, say $\boldsymbol{a}' = w' + \boldsymbol{i}x' + \boldsymbol{j}y' + \boldsymbol{k}z'$. If we take the product of $\boldsymbol{a}$ and $\boldsymbol{a}'$, multiplying it out term by term, we will obtain:

$$\boldsymbol{aa}' = ww' - (xx' + yy' + zz')$$
$$+ \boldsymbol{i}\,(w + w' + yz' - zy') + \boldsymbol{j}\,(w + w' + zx' - xz') + \boldsymbol{k}\,(w + w' + xy' - yx')$$

Hamilton, who first obtained this product, described w as the *scalar* part of the quaternion $w + \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$, and $\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$ as the *vector* part. He also described $ww' - (xx' + yy' + zz')$ as the *scalar product* and $\boldsymbol{i}\,(w + w' + yz' - zy') + \boldsymbol{j}\,(w + w' + zx' -$

xz') + $\boldsymbol{k}$ ($w + w' + xy' - yx'$) as the *vector product*. We may note the similarity to the use of these terms in vector theory today. He also introduced a quaternion differential operator, $\nabla = \boldsymbol{i}\, \partial / \partial x + \boldsymbol{j}\, \partial / \partial y + \boldsymbol{k}\, \partial / \partial z$. He speculated that the real or scalar part of the quaternion represented time and the imaginary or vector part space, and that the quaternion structure showed the long-sought link between them. Hamilton also realised that unit quaternions could be used to represent rotations in 3-dimensional space. He was convinced that he had discovered the reason why space had to be 3-dimensional and that quaternions would be the key to unlocking the secrets of the universe.

It all looked very promising, and Maxwell provided alternative quaternion treatments of mathematical operations in his famous *Treatise on Electricity and Magnetism* of 1873. But then it all went wrong. In the late nineteenth century, after both Hamilton and Maxwell were dead, people started complaining about the fact that quaternions, when squared in Pythagoras' theorem, produced the wrong sign of product, negative instead of positive. They also disliked the connection between the real and imaginary parts, preferring a structure that had just three real parts and no imaginary part. Ultimately, Gibbs and Heaviside formulated a new *vector* theory, which made the 3-dimensional or vector part real, discarded the fourth component, and recreated the scalar and vector products as two separate operations. Their vector theory was simply a rule book which has been the everyday tool for physicists ever since, though its arbitrariness has always caused problems for students. After all, vector 'algebra' is not an algebra at all as it has no multiplication and its operations do not exhibit closure. Vectors can have two 'products', neither of which results in another vector. The scalar product produces a scalar and the vector product a new type of quantity, a *pseudovector* (or axial vector), which transforms differently to vectors, while the scalar product of a vector and a pseudovector produces yet another type of quantity, a *pseudoscalar*, which is quite different from a scalar.

Unfortunately, the supporters of the new theory decided there wasn't room for both vector theory and quaternions and an intensive vilification campaign led to the annihilation of the quaternionists. Their mathematics was not only less useful than the new vector theory but, in fact, utterly worthless and totally without application. Hamilton's view that quaternions were the key to the universe was one of the most self-deluding ideas ever attained by a great mathematician. His career, so promising at first, had ended in total tragedy. When relativity came along soon afterwards and made the connection between space and time that Hamilton had sought in his mathematics, so solving one of the two main objections to the quaternion theory, no one wanted to know. The connection, they said, was nothing to do with quaternions. It is hard to find another defeat so comprehensive in the history of science.

Yet Hamilton had not only been right all along about the space-time connection, so removing the first objection, he had already produced the mathematics that would have removed the second objection as well. This was in his very first development of the original idea. Since quaternions were quite distinct from ordinary complex numbers, why not combine the two and produce complexified quaternions? Our base set is now 1, i, $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, and, multiplying everything out, we will also generate terms like $i\boldsymbol{i}$, $i\boldsymbol{j}$, $i\boldsymbol{k}$. For reasons that will soon become clear, I also write $i\boldsymbol{i} = \mathbf{i}$, $i\boldsymbol{j} = \mathbf{j}$, $i\boldsymbol{k} = \mathbf{k}$. If we take the products of these terms, we can write:

$$(i\boldsymbol{i})^2 = (i\boldsymbol{j})^2 = (i\boldsymbol{k})^2 = -i(i\boldsymbol{i})\,(i\boldsymbol{j})\,(i\boldsymbol{k}) = 1$$
$$(i\boldsymbol{i})\,(i\boldsymbol{j}) = i(i\boldsymbol{k})$$
$$(i\boldsymbol{k})\,(i\boldsymbol{i}) = i(i\boldsymbol{j})$$
$$(i\boldsymbol{j})\,(i\boldsymbol{k}) = i(i\boldsymbol{i})$$

The complexified quaternion units $i\boldsymbol{i} = \mathbf{i}$, $i\boldsymbol{j} = \mathbf{j}$, $i\boldsymbol{k} = \mathbf{k}$ are, of course, anticommutative in exactly the same way as ordinary quaternion units, but we now notice an extra feature, the i term outside the bracket that has appeared on the right-hand side of the equations. This becomes clearer if we write the equations in our alternative, more compactified, notation:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -i\mathbf{ijk} = 1$$
$$\mathbf{ij} = i\mathbf{k}$$
$$\mathbf{ki} = i\mathbf{j}$$
$$\mathbf{jk} = i\mathbf{i}$$

The remarkable thing now is that these objects have the properties that we require of vectors. In particular, they square to positive values. But they have something else in addition, an extra property whose meaning didn't emerge until well into the twentieth century. They also incorporate *spin*, that is, the mysterious property introduced by quantum mechanics. We can recognize them as being isomorphic to the Pauli matrices, originally introduced into nonrelativistic quantum mechanics to incorporate the experimentally-discovered concept of spin.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Hestenes, later in the twentieth century, termed **i**, **j** and **k** as the units of a *multivariate vector* algebra. In general, multivariate vectors **a** and **b** followed a full multiplication rule, which incorporated both scalar and vector products in the same way as quaternions:

$$\mathbf{ab} = \mathbf{a.b} + i\,\mathbf{a} \times \mathbf{b}$$

He showed that if we used the full product $\nabla\nabla\psi$ for a multivariate vector ∇ (basically, Hamilton's own definition of the symbol!) instead of the scalar product $\nabla.\nabla\psi$ for an ordinary vector ∇, we could obtain spin ½ for an electron in a magnetic field from the nonrelativistic *Schrödinger equation*. Though the first explanation of spin ½ came from the relativistic Dirac equation, the effect is nothing to do with relativity. It comes from properties deep within 3-dimensionality. The reason why Dirac first obtained it is because he effectively included these properties in the extra algebra he needed to make his equation linear.

In fact, all physical vectors are really multivariate vectors and not ordinary vectors at all, and generally in these lectures when I use the word 'vector' I will use the multivariate definition unless there is a specific reason to do otherwise. I will also always use **bold** symbols for vectors to distinguish them from the ***bold italics*** used to represent quaternions and the *italics* used for ordinary complex numbers. Unlike ordinary vector algebra, multivariate algebra is a real algebra. It has closure and a genuine product. It also makes sense of such things as pseudovectors and pseudoscalars, which appear arbitrarily in ordinary vector algebra. As we can see from its multiplication rules, multivariate vector algebra generates such terms as *i***i**, *i***j**, *i***k**, which are not vector units, and also *i*, which is not an ordinary scalar. Let's take a simple example. Imagine we have a rectangle with sides **a** and **b**. To find the area, we take the product **ab** = **a.b** + *i* **a** × **b**. Since **a** and **b** are orthogonal, the first term on the right-hand side disappears, leaving us with an imaginary vector in a direction perpendicular to **a** and **b**. The area is an imaginary or *pseudovector*, say *i***A**. If we then suppose that the rectangle is the base of a solid body with height **c** in the direction of this pseudovector, then the volume will be the product *i***Ac** = *i***A.c** + *i* *i***A** × **c**. This time, since the vector and pseudovector are parallel, it is the second term which disappears, leaving the product as an imaginary scalar. So volume (the 'triple product' in fact, as well as in name) is a pseudoscalar.

Because we require the pseudovector and pseudoscalar terms as well as vectors and scalars, vector algebra is a larger algebra than quaternion algebra. In fact, quaternions can be seen as a subalgebra of vectors, composed of the pseudovectors and scalars. Remarkably, pseudovectors, which include such concepts as torque and angular momentum, are identical in principle to quaternions, and we can (with appropriate sign adjustments) switch from quaternion to vector representations and vice versa simply by multiplying the units by *i*. In addition, the combination of vector and pseudoscalar units, (**i**, **j**, **k**) and *i*, in the vector algebra brings us to the recognition that vector algebra in this form also incorporates the 4-vector algebra of relativity, which requires the units (**i**, **j**, **k**, *i*) and is equivalent to the complexified version of a quaternion (with units ***i***, ***j***, ***k***, 1).

It is amazing how powerful quaternions look in this new context, and we are only just beginning! Was Hamilton also right about quaternions unlocking the secrets of the universe? I think he probably was, and I hope we will see why as we proceed through the lectures.

Clifford algebra

The most remarkable extension of quaternions comes with the algebra invented by Clifford in the 1870s. In fact it was Clifford who first wrote down the expression **ab** = **a.b** + *i* **a** × **b**. Clifford algebra is one of the most powerful tools ever offered to the physicist, and this is no coincidence, since, in many ways, it seems to be the mathematical code built deep in the structure of physics. It is still massively under-used, though Dirac recognised early on that the algebra he had devised for his equation for relativistic quantum mechanics was, in fact, a Clifford algebra.

Clifford algebra (also called geometrical algebra) unites real, complex numbers, quaternions and vectors into a single system of infinite potential complexity. It is structured on defining a system with *m* units which are square roots of 1 (norm 1) and *n* which are square roots of –1 (norm –1), where *m* and *n* are integers of any size. We write this as *Cl*(*m*, *n*) or *G*(*m*, *n*). However, this is not necessarily a unique specification for it is often possible to produce the same algebra with a quite different specification of *m* and *n*.

One way of building up Clifford algebras is to use commuting sets of quaternions, which may also be complexified. I could for instance build up a Clifford algebra using three sets of quaternion units, each of which commutes with the others, although the units within the quaternion sets anticommute. I could then create a higher Clifford algebra by complexifying it.

The basic Clifford algebra is the Clifford algebra of 3-dimensional space, which we have already described in detail under its later name of multivariate vector algebra. In its full specification it has 8 basic units, each of which can be + or –.

i **j** **k**	vector		
*i***i** *i***j** *i***k**	bivector	pseudovector	quaternion
i	trivector	pseudoscalar	complex
1	scalar		

It has 3 subalgebras: bivector / pseudovector / quaternion, composed of:

*i***i** *i***j** *i***k**	bivector	pseudovector	quaternion
1	scalar		

trivector / pseudoscalar / complex, composed of:

i	trivector	pseudoscalar	complex
1	scalar		

and scalar, with just a single unit:

1

Here, we use the term 'bivector' for the product of two vectors and 'trivector' for the product of three. It is important for us to recognise that we could specify the entire algebra, either by using either **i j k** or a combination of its 3 subalgebras. This will become very significant in our work.

A particularly interesting algebra emerges if we combine this algebra with an identical algebra of 3-dimensional space to which this is commutative.

i j k	vector		
*i***i** *i***j** *i***k**	bivector	pseudovector	quaternion
i	trivector	pseudoscalar	complex
1	scalar		

When we multiply these two algebras by each other, term by term, we produce an algebra that has 64 basic units, which are + and – versions of:

i	**j**	**k**	*i***i**	*i***j**	*i***k**	*i*	1
i	**j**	**k**	*i***i**	**ii**	*i***k**		
ii	**ij**	**ik**	*i***ii**	**iii**	*i***ik**		
ji	**jj**	**jk**	*i***ji**	**iji**	*i***jk**		
ki	**kj**	**kk**	*i***ki**	**iki**	*i***kk**		

Since vectors are complexified quaternions and quaternions are complexified vectors, we obtain an identical algebra if we use complexified double quaternions:

i	***j***	***k***	*i**i***	*i**j***	*i**k***	*i*	1
i	***j***	***k***	*i**i***	*i**i***	*i**k***		
ii	***ij***	***ik***	*i**ii***	*i**ii***	*i**ik***		
ji	***jj***	***jk***	*i**ji***	*i**ji***	*i**jk***		
ki	***kj***	***kk***	*i**ki***	*i**ki***	*i**kk***		

Yet another variation can be found using a combination of vectors (blue) and quaternions (red):

i	**j**	**k**	*i***i**	*i***j**	*i***k**	*i*	1
i	***j***	***k***	*i**i***	*i**i***	*i**k***		
i*i*	**i*j***	**i*k***	*i***i*i***	**ii*i***	*i***i*k***		
j*i*	**j*j***	**j*k***	*i***j*i***	**ij*i***	*i***j*k***		
k*i*	**k*j***	**k*k***	*i***k*i***	**ik*i***	*i***k*k***		

One of the remarkable things about Clifford algebra is that the same algebra can have different dimensionalities simultaneously. It depends on how you distribute between m and n. So two square roots of –1 can become a single square root of 1 and vice versa. And this can be done multiples of times. So, you can have a 10-dimensionality in one perspective simultaneously with a 3-dimensionality in another. Ultimately, as Hamilton recognised at that moment of discovery in 1843, 3-dimensionality is special and is due to a single, simple property. Any higher dimensionality can always be resolved into structures based on 3, and even complex numbers can be represented as incomplete quaternion sets.

Groups

Groups are an important aspect of the mathematics of symmetry, and, since symmetry is going to be a major aspect of our foundational approach, we can expect groups to play a significant part. A finite or infinite number of elements form a group if they contain:

(1) a binary operation (e.g. multiplication, permutation) between any two elements to produce another
(2) an identity element, so that the binary operation between the identity element and another element produces that element
(3) an inverse to each element, so that the binary combination of any element and its inverse produces the identity element
(4) closure, so that the binary operation between any two elements always produces an element within the group

As an illustration, we can take the simplest group C_2, which is of order 2 and so has only two elements, which we could, for example, take as 1 and –1. Here the binary operation is multiplication. The identity element is 1. Each element is its own inverse, and we have closure because each binary operation between 1 and –1 only ever produces either 1 or –1. Complex numbers give us a group of order 4 (C_4) made out of the base units 1, –1, i and $-i$. Quaternions (Q) give us a group of order 8, from the base units 1, –1, ***i***, –***i***, ***j***, –***j***, ***k***, –***k***. Multivariate vectors, which complexify the quaternion algebra, are a group of order 16. In all these cases negative as well as positive units are required for closure. Octonions (O), however, are not a group because their multiplication (see the Appendix) is not associative. For any group, finite or infinite, it is possible to produce all the elements of the group using a finite number of elements as *generators*. For a finite group, this may be less than the number of elements in the group. Often, this can be done in many different ways.

C_2 and C_4 are examples of *cyclic* groups. There is a cyclic group at every order, and how they operate can be illustrated by writing down the multiplication table for C_3.

*	1	a	a^2
1	1	a	a^2
a	a	a^2	1
a^2	a^2	1	a

A cyclic group of order n, has elements, I, a, a^2, ... a^{n-1}, with a^n reverting to the identity. So, if we wanted to express the complex number units as a cyclic group of order 4, then I, a, a^2, a^3 would become 1, i, –1, –i, in that order. Now, there is a second group of order 4 which will be very important to us, which is called D_2 or the Klein-4 group. It is called D_2 or dihedral 2 because it is the group of rotations of the rectangle (identity, and rotations along 3 different axes), and it is the only other group than C_4 of order 4. It can be generated by a 'double' algebra known as H_4, which is made up of 4 units, made up of two commutative sets of quaternions, 1, ***i***, ***j***, ***k***, and 1, ***i***, ***j***, ***k***. Here, the red and blue units multiply commutatively with each other in the ordinary way, but the red units anticommute with *each other*, as do the blue units. The H_4 algebra units can be constructed as 1, ***ii***, ***jj***, ***kk***, and you can see that the units ***ii***, ***jj***, ***kk*** have now become commutative with other, unlike the units of their parent systems. So ***iijj*** = ***jjii***, etc. The H_4 algebra units become a group with the multiplication table:

*	1	***ii***	***jj***	***kk***
1	1	***ii***	***jj***	***kk***
ii	***ii***	1	***kk***	***jj***
jj	***jj***	***kk***	1	***ii***
kk	***kk***	***jj***	***ii***	1

It's effectively the same thing as using quaternions but ignoring the negative signs.

Another way of generating the same table is by creating elements made up of variations of three 'components', ± x, ± y and ± z:

A	x	y	z
B	–x	–y	z
C	x	–y	–z
D	–x	y	–z

The binary operation (*) doesn't have to be anything to do with multiplication. So we can make up our own rules as long as they are rigidly followed and lead to closure. Let's say we have:

$$x * x = -x * -x = x$$
$$x * -x = -x * x = -x$$
$$x * y = y * -x = 0$$

and similarly for *y* and *z*. Then we will have a group table:

*	A	B	C	D
A	A	B	C	D
B	B	A	D	C
C	C	D	A	B
D	D	C	B	A

This version of the group will become even more important to us than the one illustrating the H_4 algebra, and I hope to show later that, overall, there is no more important group in the whole of physics.

We can also arrange for a dual version of this group, with assignments such as:

A*	$-x$	y	z
B*	x	$-y$	z
C*	$-x$	$-y$	$-z$
D*	x	y	$-z$

Another very significant group for us is the group of order 64 represented by the commutative product of two vector algebras or of a vector algebra with a quaternion algebra. If we write out the complete set of 64 terms for the vector-quaternion combination, we find that, apart from the scalar and pseudoscalar units, something very interesting emerges.

1	i				-1	$-i$			
$\boldsymbol{i}\mathbf{i}$	$\boldsymbol{i}\mathbf{j}$	$\boldsymbol{i}\mathbf{k}$	$i\boldsymbol{k}$	$\boldsymbol{j}$	$-\boldsymbol{i}\mathbf{i}$	$-\boldsymbol{i}\mathbf{j}$	$-\boldsymbol{i}\mathbf{k}$	$-i\boldsymbol{k}$	$-\boldsymbol{j}$
$\boldsymbol{j}\mathbf{i}$	$\boldsymbol{j}\mathbf{j}$	$\boldsymbol{j}\mathbf{k}$	$i\boldsymbol{i}$	$\boldsymbol{k}$	$-\boldsymbol{j}\mathbf{i}$	$-\boldsymbol{j}\mathbf{j}$	$-\boldsymbol{j}\mathbf{k}$	$-i\boldsymbol{i}$	$-\boldsymbol{k}$
$\boldsymbol{k}\mathbf{i}$	$\boldsymbol{k}\mathbf{j}$	$\boldsymbol{k}\mathbf{k}$	$i\boldsymbol{j}$	$\boldsymbol{i}$	$-\boldsymbol{k}\mathbf{i}$	$-\boldsymbol{k}\mathbf{j}$	$-\boldsymbol{k}\mathbf{k}$	$-i\boldsymbol{j}$	$-\boldsymbol{i}$
$i\mathbf{i}\boldsymbol{i}$	$i\mathbf{i}\boldsymbol{j}$	$i\mathbf{i}\boldsymbol{k}$	$i\mathbf{k}$	$\mathbf{j}$	$-i\mathbf{i}\boldsymbol{i}$	$-i\mathbf{i}\boldsymbol{j}$	$-i\mathbf{i}\boldsymbol{k}$	$-i\mathbf{k}$	$-\mathbf{j}$
$i\mathbf{j}\boldsymbol{i}$	$i\mathbf{j}\boldsymbol{j}$	$i\mathbf{j}\boldsymbol{k}$	$i\mathbf{i}$	$\mathbf{k}$	$-i\mathbf{j}\boldsymbol{i}$	$-i\mathbf{j}\boldsymbol{j}$	$-i\mathbf{j}\boldsymbol{k}$	$-i\mathbf{i}$	$-\mathbf{k}$
$i\mathbf{k}\boldsymbol{i}$	$i\mathbf{k}\boldsymbol{j}$	$i\mathbf{k}\boldsymbol{k}$	$i\mathbf{j}$	$\mathbf{i}$	$-i\mathbf{k}\boldsymbol{i}$	$-i\mathbf{k}\boldsymbol{j}$	$-i\mathbf{k}\boldsymbol{k}$	$-i\mathbf{j}$	$-\mathbf{i}$

Apart from the 4 units of ordinary complex algebra, 1, i, −1, $-i$, the other sixty units arrange themselves in 12 groups of 5. Remarkably, *any* of these groups of 5 will generate the *entire group*, and this is in fact the minimum number of generators. Significantly, they all have the same overall structure. We take the 8 base units provided by a 4-vector ($\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, i) combined with a quaternion ($\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, 1), and break the symmetry of one of the two 3-dimensional structures to create the minimum structure for the generators. From the perfect symmetry of

$i \quad \mathbf{i} \quad \mathbf{j} \quad \mathbf{k} \quad 1 \quad \boldsymbol{i} \quad \boldsymbol{j} \quad \boldsymbol{k}$

we rearrange to produce:

$i \quad \mathbf{i} \quad \mathbf{j} \quad \mathbf{k} \quad 1$
$\boldsymbol{k} \quad \boldsymbol{i} \quad \boldsymbol{j}$

and finally:

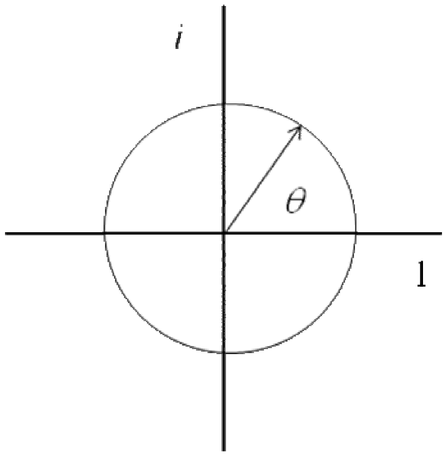
$i\boldsymbol{k} \quad \boldsymbol{i}\mathbf{i} \quad \boldsymbol{i}\mathbf{j} \quad \boldsymbol{i}\mathbf{k} \quad \boldsymbol{j}$

Here, the symmetry of the ***red quaternion operators*** is broken in that they are attached to different types of object: pseudoscalar i, vector **i j k**, and scalar 1. The symmetry of the vector operators, however, is retained because each of **i j k** is associated with the same object. We can imagine that, if we wanted to make the minimum package containing two 3-dimensional structures, the symmetry of one would be broken in this way.

Exactly the same result would have followed if we had used double vectors (or the units of two 'spaces'), where the 5 generators would have been

$i\mathbf{k} \quad \mathbf{ii} \quad \mathbf{ij} \quad \mathbf{ik} \quad \mathbf{j}$

In addition to finite groups there are also infinite or Lie groups, with an infinite number of elements and a finite number of generators. Lie groups are often generated from the rotational aspects of finite groups. The unitary groups, $U(n)$, are defined as the groups of $n \times n$ unitary matrices, and so have a complex determinant with norm 1. A complex square matrix, such as the ones which define $U(n)$ groups, is unitary if the product of the matrix U and its conjugate transpose U^* is the identity matrix I. $U(n)$ groups have n^2 generators. The simplest of these, $U(1)$, can be illustrated by drawing a circle of radius 1 on the Argand diagram centred round the origin. The group then consists of all the complex numbers, $x = iy$, on the circumference of this circle. If we imagine a radius vector drawn from the origin to any point on the circumference, the length of the vector will remain at unity regardless of the angle that the vector makes with the x or y axis, which is equivalent to an arbitrary phase term θ between 0 and 2π. In fundamental physics, the $U(1)$ group is significant as the one connected with the inverse-square law of force, which is derived from the spherical symmetry of space round a point source, and is particularly associated with the electric interaction, which has no other component. In the quantum version of this interaction, the single generator of the group becomes the photon, the boson which mediates the interaction.



$U(1)$ *symmetry*

Of particular significance in physics are the special unitary groups of degree n, or $SU(n)$, which are groups of $n \times n$ unitary matrices, like $U(n)$ groups, but this time with the extra specification that the matrices have determinant 1. $SU(n)$ groups have $n^2 - 1$ generators, the extra constraint of determinant 1 reducing the degrees of freedom, and hence the number of generators, by 1. $SU(2)$ (which is isomorphic to the group of unit quaternions and can be mapped smoothly onto the 3-sphere) is significant as the symmetry group of the weak interaction in the Standard Model, the 3 generators being equivalent to the 3 weakly interacting bosons, W^+, W^- and W^0. In the Standard Model the W^0 mixes with the $U(1)$ field to produce the Z^0 and the photon. It is also the symmetry group for fermion spin. $SU(3)$ is the symmetry group for the strong interaction, with 8 generators, which can be identified with the gluons or bosons mediating the interaction.

A more advanced discussion of Lie groups would show that those groups of most interest in physics, including $SU(2)$, $SU(3)$, and the exceptional groups F_4, E_6, E_7 and E_8, can be derived from the symmetries associated with the real numbers, complex numbers, quaternions and octonions, and are ultimately an expression of the special nature of 3-dimensional space, glimpsed in Hamilton's discovery of quaternions.

Nilpotents and idempotents

Clifford algebra allows us to create some unusual algebraic objects, ones which are square roots of zero, which we call *nilpotents*, and ones which are square roots of themselves, which we call *idempotents*. In fact, nilpotents are more strictly defined as objects for which *any* finitely repeated operation will yield zero, but we will here be concerned only with the case where the operation is squaring. Trivially, of course, 0 is both idempotent and nilpotent, while 1 is idempotent. However, we can also create nonzero nilpotents and nonunit idempotents. The main reason for this is the incorporation of terms that are anticommuting, which then removes all the cross terms in the product by mutual cancellation. Pythagoras' theorem can be structured entirely in nilpotent form. In a simple example, a Pythagorean triplet could be written in the form:

$$(ik5 + i4 + j3)^2 = 5^2 - 4^2 - 3^2 = 0.$$

$(ik5 + i4 + j3)$ is now a nilpotent, or square root of zero. We could even make it a little more complicated by making the 4 a multivariate vector, commutative to the quaternions in the equation:

$$(ik5 + \mathbf{i4} + j3)^2 = 0.$$

To create an idempotent, we can premultiply $(ik5 + \mathbf{i4} + j3)$ by $\boldsymbol{k}$ and a scaling factor a to be determined later. So we try:

$$a\boldsymbol{k}\ (ik5 + \mathbf{i4} + j3)\ a\boldsymbol{k}\ (ik5 + \mathbf{i4} + j3).$$

We can multiply this all out term by term, but there is also a shortcut, because

$$a\boldsymbol{k}\ (ik5 + \mathbf{i4} + j3)\ \boldsymbol{k} = a\ (-ik5 + \mathbf{i4} + j3)$$

and $$(-ik5 + \mathbf{i4} + j3) = (-2ik5 + ik5 + \mathbf{i4} + j3)$$

which means that

$$a\boldsymbol{k}(ik5 + \mathbf{i4} + j3)\ a\boldsymbol{k}(ik5 + \mathbf{i4} + j3) = a\boldsymbol{k}(-2ik5)\ a(ik5 + \mathbf{i4} + j3) = 10ia^2\ \boldsymbol{k}\ (ik5 + \mathbf{i4} + j3)$$

So we make $a = 1 / 10i$ and $\boldsymbol{k}\ (ik5 + \mathbf{i4} + j3) / 10i$ becomes an idempotent, or square root of itself.

Standard and non-standard analysis

There are one or two aspects of mathematics that will be important to us but that we don't need to cover in detail. One concerns real numbers and also calculus and the properties of space. Real numbers begin with the natural numbers, 1, 2, 3, etc., which are quickly generalised to the integers, which may be negative as well as positive. The next stage is to take fractions of integers and define these as rational numbers. Then come algebraic numbers, which are solutions of algebraic equations, that is equations in which a polynomial expression in one variable is equated to 0, for example $x^2 - 2 = 0$. Finally, there are transcendental numbers, such as π and e, which are not generated by algebraic equations, but can usually be expressed only in terms of an infinite series. All of these combined are the real numbers.

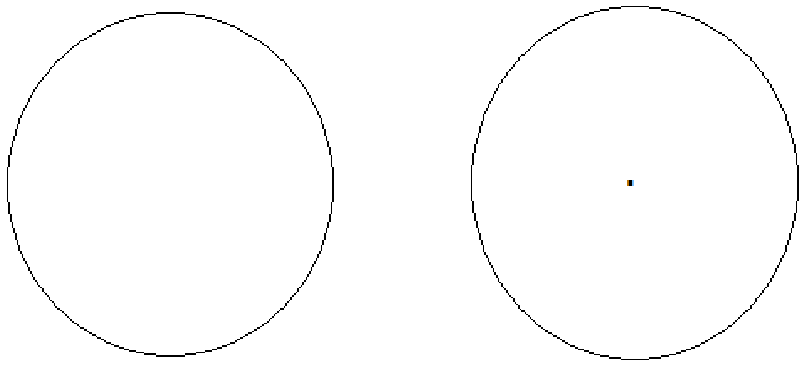
Most of us are familiar with the Cantor argument by which rational and algebraic numbers can be counted by being put into a one-to-one relation with the integers, but the transcendental numbers (and hence the real numbers as a whole) cannot. In the Cantor argument, between any two numbers that can be counted in this way there are an infinite number of real numbers that cannot. The real numbers, therefore, are uncountable or non-denumerable. This argument is valid and most people think it is uniquely so, in that the opposite construction of countable real numbers is false. This, however, is not the case. Even though the Cantor argument is valid in that real numbers can be defined in this way, it is *not uniquely true*. Real numbers can be constructed in such a way that they can be put into a one-to-one correspondence with the integers and so can be made denumerable in the same way as all other numbers. This is because to construct a real number requires an algorithmic process, and algorithmic processes can be counted. So, if we think of real numbers as simply 'there' in nature, then they cannot be counted. If we think of them as always the result of a construction, then they can.

The procedure for counting real numbers in this way was first generated by Skolem in 1934 and is called 'non-standard arithmetic' (as opposed to 'standard arithmetic', based on the Cantor argument). It relates to the Löwenheim-Skolem theorem, in which any consistent finite, formal theory has a denumerable model, in which the elements of its domain are in a one-to-one correspondence with the positive integers. It was subsequently applied to the real numbers in the construction of space in non-Archimedean geometry, a fact which will be specially important to us. A parallel development also applies to calculus in which non-standard analysis aligns itself with non-standard arithmetic in the same way as standard analysis aligns itself with the arithmetic based on the Cantor continuum.

Many people will remember being taught differentiation using infinitesimals. You would draw a graph of a function and draw a line between two points on the graph with horizontal and vertical separations δx and δy, then use the function of the graph to find an expression for δy / δx. You would then make the two points approach each and say that at an infinitesimal distance apart you could cancel any term with δx in it, leaving you with an expression for the differential dy / dx. This was a very effective procedure, used by Newton and others in the seventeenth century and cast into the form familiar to us by Leibniz, but it was always considered nonrigorous. To do differentiation 'properly' meant 'taking the limit', a process introduced by Newton and perfected by Cauchy, in which there were no such cancellations. However, in the twentieth century, Abraham Robinson developed a new procedure in which infinitesimals could be made as rigorous as limits (incidentally, using a nilpotent-type structure). It was found that this 'non-standard' procedure and the standard analysis based on limits produced exactly the same results, and neither method was always superior to the other in proving mathematical theorems. Sometimes one method was superior, sometimes the other. Remarkably the methods of standard and non-standard analysis and arithmetic, and Archimedean and non-Archimedean geometry, are completely dual and cannot be distinguished by any known mathematical theorem or methodology.

Topology

Topology has many insights that are valuable in physics. Here, we are concerned only with one. This is the distinction between simply- and multiply- connected spaces. A simply-connected space is one without singularities. A multiply-connected space is one with a topological singularity. Now, if we imagine 'parallel-transporting' a vector round a closed circuit in the simply-connected space, that is, moving the vector round the circuit in such a way that it is always at a tangent to the path, then, on returning to the starting position it will be pointing in the original direction. If the space is multiply-connected, however, it will, on its return, be pointing in the *opposite direction*. That is, there will be a phase change of 180^o or π radian. To return to the start pointing in its original direction, it will have to do a *double circuit*. This is an illustration of a general phenomenon, known as the Berry phase or geometric phase, and, although it is normally described in geometric terms, it can always be mapped onto a topological representation.



simply-connected space *multiply-connected space*

The topology of $SU(n)$ groups, notably, is simply-connected, while that of $U(n)$ groups is not.

Key numbers in duality, anticommutativity and symmetry-breaking

The final mathematical topic is the simplest, but it is one of the most important. At this stage it is simply introduced as a comment to be followed up at a later period. A few key integers turn up regularly in physics and sometimes in other areas of science (particularly, in biology). Some of these have a very primitive origin in the deepest levels of physics, though this has never been recognised by those who think they need 'sophisticated' explanations. The most important of these are 2, 3 and 5. Generally, where 2 occurs it can be tracked down to duality, where 3 occurs it comes from anticommutativity, and where 5 occurs it represents symmetry-breaking. 5 always comes from complexity, but the other two numbers come from a genuinely primitive level. It is astonishing that nearly all the group structures that are important in physics, including such seemingly complicated ones as E_8, can be interpreted in terms of these numbers, and the individual introduction of the numbers tracked down to their roots.

Appendix

Complex numbers as a group:

*	1	–1	*i*	–*i*
1	1	–1	*i*	–*i*
–1	–1	1	–*i*	*i*
i	*i*	–*i*	–1	1
–*i*	–*i*	*i*	1	–1

Quaternions as a group:

*	1	***i***	***j***	***k***	–1	–***i***	–***j***	–***k***
1	1	***i***	***j***	***k***	–1	–***i***	–***j***	–***k***
i	***i***	–1	***k***	–***j***	–***i***	1	–***k***	***j***
j	***j***	–***k***	–1	***i***	–***j***	***k***	1	–***i***
k	***k***	***j***	–***i***	–1	–***k***	–***j***	***i***	1
–1	–1	–***i***	–***j***	–***k***	1	***i***	***j***	***k***
–***i***	–***i***	1	–***k***	***j***	***i***	–1	***k***	–***j***
–***j***	–***j***	***k***	1	–***i***	***j***	–***k***	–1	***i***
–***k***	–***k***	–***j***	***i***	1	***k***	***j***	–***i***	–1

Octonion multiplication table:

*	1	***i***	***j***	***k***	***e***	***f***	***g***	***h***
1	1	***i***	***j***	***k***	***e***	***f***	***g***	***h***
i	***i***	–1	***k***	–***j***	***f***	–***e***	–***h***	***g***
j	***j***	–***k***	–1	***i***	***g***	***h***	–***e***	–***f***
k	***k***	***j***	–***i***	–1	***h***	–***g***	***f***	–***e***
e	***e***	–***f***	–***g***	–***h***	–1	***i***	***j***	***k***
f	***f***	***e***	–***h***	***g***	–***i***	–1	–***k***	***j***
g	***g***	***h***	***e***	–***f***	–***j***	***k***	–1	–***i***
h	***h***	–***g***	***f***	***e***	–***k***	–***j***	***i***	–1

Peter Rowlands
Physics Department, University of Liverpool
15 March 2013