Chapter 1

Zero

The bulk of this chapter (**1.2 – 1.5**) is the result of a collaboration with **BERNARD DIAZ**, and incorporates the results of a number of co-authored papers. The aim of the chapter is to show that mathematics, and in particular the mathematics needed for physics, can be generated from the idea of zero. Mathematics is shown to be derivable from a zero totality, without assuming numbers, by structuring it by analogy with a computer rewrite system. Here, we assume that any deviation from the zero state (a 'creation') forces a continued attempt to recover the original zero totality. The fundamental rule has a zero-totality alphabet or set of states, which generates itself by acting on any subalphabet, necessarily generating a new zero-totality alphabet by acting on itself. Then, simply by assuming a totally undefined nonzero state, we produce a conjugate (or 'negative') and a series of complex forms (not yet associated with numbers), which act on each other to produce further complex forms or their conjugates. We can identify operations which we describe as commutative, which can be continued indefinitely, and others which are anticommutative, which are restricted to the two forms and their operation on each other. The requirement of uniqueness forces us to default on the second, with the consequence that we generate an ordinal series of discrete, dimensional systems, which we can use as a number series, and so introduce this concept for the first time. The link between divisibility or discreteness and dimensionality, is now established as the fundamental process by which the idea of discreteness, and, even that of number, actually enters into mathematics. We can now generate an infinite series by the repeated application of three processes, described as conjugation, complexification and dimensionalization, with important consequences for physics. The overall strategy of the chapter is to derive the general process and its important application to mathematics in sections 1.2 to 1.4, before developing the more specifically technical details of the computer rewrite interpretation in section 1.5.

ZERO TO INFINITY - The Foundations of Physics

1.1 An Origin for Everything

What is the most fundamental possible idea? What is the simplest thing we can possibly start with to obtain the universe as we know it today? Is there some point at which we could say that the origin of everything had been established because it could not possibly have been imagined otherwise? If we believe that there really are answers to these questions, then we must imagine that they must come in a form so obvious that no special argument will be needed to present them. It is almost inconceivable, also, that such answers would not also lead to the long-desired ‘unified theory of physics’.

Physics has long been recognised as the most fundamental physical science, perhaps even the most fundamental possible structure for the whole of human knowledge. It certainly gives the appearance of being the most fundamental way possible of accessing knowledge about what we call the ‘natural world’. Our current understanding of physics has been accumulated as the result of a succession of inductive inferences derived from information which is ultimately empirical in origin. We have learnt from experience that certain patterns of thought produce successful outcomes, that relatively simple generalisations appear to have a surprisingly wide currency, and that, as Galileo expressed it: ‘The Book of Nature is written in the language of mathematics.’ The idea that physics may be, in some sense, ‘unified’ in a way which has not so far been discovered has a sound basis in the way that the subject has developed over the last five hundred years.

A ‘unified’ theory, however, means something very particular. It means a theory which derives all its results from a single source, not simply one in which we put all our known results into a single comprehensive package. This is a crucial distinction which is not always made in discussions of the subject; but, in principle, a truly unified theory cannot come from an act of unification; that is, we cannot, for example, create a truly unified theory simply by combining quantum mechanics and general relativity in a new mathematical superstructure. Such attempts have always failed in the past, and will continue to do so in the future, because the concept of unification as combination is invalid. Unification is really about finding *descent from a common origin*. Creating a sophisticated mathematical superstructure will not provide answers to the fundamental questions that we would expect from a truly unified theory, a theory of ‘everything’.

Many physicists believe that we can assume concepts such as space, time and matter to be fundamental and even inexplicable. Many also believe that we must assume that the fundamental ideas will be expressed in a mathematical language that already exists independently of the physical principles we are expressing. Others again are prepared to believe that the unifying mathematical structure, whatever it is, will be a sophisticated one which only a few will be able to truly understand. *I do not believe in any of these things*. If physics is to prove itself the most fundamental possible way of understanding the 'natural world', then it must *explain* space, time and matter, as well as use them, and it must *generate* the mathematical structures it uses; it must also show how all things that are sophisticated arise from things that are much more simple. In principle, all possible complications must be removed from the ultimate starting point. It has to be intrinsically simple and absolutely single.

So, how do we construct such a theory? What would be left after we have removed all the complications? What could be our single and simple starting point? There is only one possible answer: zero, absolutely nothing, or perhaps 'no thing'. This is the only thing in all our experience which has no structure and includes no complication. So, zero must be our starting point. It must also be where we finish, for nothing, as we all know, comes from nothing – *nihil ex nihilo fit*. The idea that the universe may be a totality of 'nothing', in some sense, has been discussed for quite a number of years, especially in the context of the 'big bang' theory of creation, where the total energies of matter (positive) and gravitation (negative) could cancel each other out. This is the kind of reasoning behind the statement of the science-writer, Peter Atkins, that 'the seemingly something is elegantly reorganized nothing, and … the net content of the universe is … nothing'.[1] However, to find the most fundamental possible theory, we must go even further, and state that the universe and everything we can possible experience or conceive is also *conceptually* nothing. No other position is extreme or uncompromising enough to be able to explain *everything*.

1.2 The Genesis of Number

As we have said, many scientists today believe that the concept of nothing has a fundamental role to play in an understanding of the universe and its contents. It is, for example, often stated that the universe itself could have emerged as a quantum fluctuation from an essentially zero state, and that the total zero energy is maintained by the total positive mass energy being equated numerically to the negative gravitational energy generated by this mass through a special

ZERO TO INFINITY - The Foundations of Physics

application of Mach's principle (the idea that the inertia or mass of physical objects is determined by their interactions with the rest of the matter in the universe). However, the concept of nothing may be even more powerful than this, for it may be that the nothingness applies also to the entire conceptual scheme of which matter and the universe are merely components. Nothing is unique among conceptual ideas in being infinitely degenerate (infinitely capable of reinterpretation) and it may be that this infinite degeneracy is the key to an understanding of its special power.

However, since 'nothing comes from nothing', we are left with the question of how we preserve the total nothingness in the presence of the seemingly 'something' which we call 'the universe'. Both mathematics and physics suggest that the answer lies in the concept of *duality*, the idea that to every fundamental concept there exists some kind of partnership or relationship with another that is its 'dual'. As Nicholas Young has noted, 'the idea of duality pervades mathematics',[2] the pairing of positive and negative numbers being only the most obvious example. So we might well expect that the scheme required for preserving the conceptual nothingness of mathematical physics, is essentially a dualistic one. (This will be discussed in the next chapter.) But we still have to identify the origin of such processes, and it is here that we will need the concept of infinite degeneracy, for we have a logical analogy in the *rewrite* systems which are fundamental to computing.

Obviously, we cannot generate a fundamental theory by imagining in advance that computers actually exist or that the universe is structured like a computer, but the 'rewrite' concept is independent of any connection with software or hardware, and can be conceived as a purely abstract process. If we could develop a universal system which endlessly rewrote itself, *superveniently* (that is, without any sense of temporal progression), we would also have a system which was, by definition, infinitely degenerate; and, if at every rewriting, the dualistic principle (or, in the first instance, zero totality) applied, we would have a universal process relevant to mathematics, physics and all such attempts at a fundamental description of ultimate 'reality'.

In a recent paper,[3] Deutsch *et al* state that, 'Though the truths of logic and pure mathematics are objective and independent of any contingent facts or laws of nature, our *knowledge* of these truths depends entirely on our knowledge of the laws of physics.' According to these authors we have been forced by 'recent progress in the theory of computation', 'to abandon the classical view that computation, and hence mathematical proof, are purely logical notions independent of that of computation as a physical process'. Mathematical structures, however autonomous, 'are revealed to us only through the physical world'. We can go further and state that that mathematical structure which is most fundamental in understanding the physical world is also likely to be the structure which is most fundamental to understanding mathematics itself.

The application to physics is particularly important because it is a strong test of the worth of a fundamental idea. Mathematics can be structured on fundamental principles in a large variety of ways, but physics has to survive the test of observation and experiment under many different conditions. Physics thus appears to tell us not only what is necessary to a foundational structure for mathematics, but also what is the most efficient such structure. The mathematical foundations of physics may also be expected to provide a route to understanding the principles important in the foundations of quantum computing. So, how then do we generate mathematics, and, in particular, the mathematics especially relevant to physics, purely from the concept of zero?

Well, we might think that mathematics begins with the natural numbers or positive integers, 1, 2, 3, …, as we are conditioned into believing that counting is an elementary process. So, it has become a standard procedure to derive mathematical structures from the natural numbers, and then progress by successively extending the set to incorporate negative, rational, algebraic, real, and complex numbers, before proceeding to higher algebraic structures involving, say, quaternions, vectors, Grassmann and Clifford algebras, Hilbert spaces, and even higher structures (including transfinite ones). However, to begin mathematics with the integers, though natural to our human perceptions, is to start from a position already beyond the beginning. The integers are loaded with a mass of assumptions about mathematics. They are not fundamentally simple but already contain packaged information about things beyond the integer series itself. This makes them a convenient codification of mathematics, but not a simplified starting-point. The number 1 is not the most obvious initial step from 0 because it contains, for example, the notion of discreteness, as well as ordinality (or ordering). Further, although there are many ways in which mathematics can be generated from foundational elements, physics seems to indicate that discreteness at the foundational level is always associated with dimensionality (see chapter 2), and hence is not a truly primary concept. In addition, there is no obvious route of progression from natural numbers to reals, which then appear as a nonlogical extension. It would seem to be more logical, in terms of rewrite procedures, to consider the real or noncountable 'numbers' as anterior to the integers, though it is even more logical to suppose that even real numbers are not the ultimate starting point.

Suppose, then, that we make no prior mathematical assumptions, in particular any conventional concept of mathematics / numbering. We assume instead that zero totality is the only possible state and that any deviation from this state, such as the assumption of a nonzero state, say R, generates an automatic mechanism for recovering it. Zero totality, however, is also infinitely degenerate; and to express this we propose that, while each state is a zero totality state, it is not a *unique* zero. Our rewriting procedure will then consist of defining a zero totality 'system', otherwise undefined, to contain nothing new within it (symbolized by $\rightarrow$) and a new zero totality outside it (symbolized by $\Rightarrow$). For convenience, though without any assumption as to form, we will describe the totality, as so far understood, as an 'alphabet'. Then an examination – here represented by a 'concatenation' or placing together, with no algebraic significance – of the alphabet with respect to anything other than itself (a 'subalphabet') will always yield the alphabet, because anything less than the totality will generate nothing new. On the other hand, an examination of the alphabet with respect to itself will necessarily generate a new zero totality alphabet, because a finite alphabet cannot represent a unique zero state. To summarise:

(subalphabet) (alphabet) → (alphabet) *i.e. there is nothing new*
(alphabet) (alphabet) ⇒ (new alphabet) *i.e. the zero totality is not unique*

These are the only conditions. However, as we shall see, the maintenance of a total zero state will require duality at all times. All terms are necessarily paired terms, which are, in principle, indistinguishable individually. In addition, the nature of the new alphabet produced by $\Rightarrow$ will always be determined by the need to satisfy $\rightarrow$ in all possible cases. That is, we can't work out what a new alphabet will look like till we have worked out all the ways in which concatenation with its subalphabets will yield only itself.

Suppose, then, that we assume a nonzero R. This will not only be the first subalphabet, but it will also need the zero-creating conjugate, say R^*, to make it a conjugated (zero) alphabet. That is:

$$(R)\,(R) \Rightarrow (R, R^*)$$

We now apply $\rightarrow$ to this alphabet to show that, within itself, it produces nothing new.

$$(R)\,(R, R^*) \rightarrow (R, R^*)\ ;\ (R^*)\,(R, R^*) \rightarrow (R^*, R) \equiv (R, R^*)$$

ZERO TO INFINITY - The Foundations of Physics

We note here that no concept of ‘ordering’ is required by concatenation. So, all possible concatenations have produced only the alphabet itself. From these rules it is easy to show that

$$(R)\,(R) \rightarrow\ (R)\ ;\ (R^*)\,(R) \rightarrow\ (R^*)\ ;\ (R)\,(R^*) \rightarrow (R^*)\ ;\ \ (R^*)\,(R^*) \rightarrow\ (R)$$

But, of course, the zero-totality alphabet (R, R^*) cannot be unique, and concatenation with itself must produce a new conjugated alphabet, whose additional terms must be chosen in such a way that the subalphabets yield nothing new. So we try

$$(R, R^*)\,(R, R^*) \Rightarrow (R, R^*, A, A^*)$$

and find that applying → to this new alphabet means that

$$\begin{aligned}
(R)\,(R, R^*, A, A^*) &\quad\rightarrow\quad (R, R^*, A, A^*) \equiv (R, R^*, A, A^*)\\
(R^*)\,(R, R^*, A, A^*) &\quad\rightarrow\quad (R^*, R, A^*, A) \equiv (R, R^*, A, A^*)\\
(A)\,(R, R^*, A, A^*) &\quad\rightarrow\quad (A, A^*, R^*, R) \equiv (R, R^*, A, A)\\
(A^*)\,(R, R^*, A, A^*) &\quad\rightarrow\quad (A^*, A, R, R^*) \equiv (R, R^*, A, A^*)
\end{aligned}$$

It will be apparent that, to maintain an unchanged alphabet (and to specify that R, R^*, A, A^* remain distinct), we have to arrange that each term cycles into another. From these rules, it would appear that we may also derive

$$(R)\,(A) \rightarrow (A)\ ;\ (R)\,(A^*) \rightarrow (A^*)\ ;\ (R^*)\,(A) \rightarrow (A^*)\ ;\ (R^*)\,(A^*) \rightarrow (A);$$
$$(A)\,(A) \rightarrow (R^*)\ ;\ (A^*)\,(A^*) \rightarrow (R^*)\ ;\ (A)\,(A^*) \rightarrow (R).$$

Of course, absolute duality of terms like A and A^* would seem to imply that we could write down expressions such as

$$(A)\,(A) \rightarrow (R)\ ;\ (A^*)\,(A^*) \rightarrow (R^*)\ ;\ (A)\,(A^*) \rightarrow (R^*)$$

instead of the ones chosen. However, since R and A would now be indistinguishable, this would, in effect, be equivalent to not extending the alphabet.

At the next stage, the process of ensuring that the new alphabet only produces itself when concatenated with its subalphabets (→) requires that we now introduce concatenated *terms*, such as AB, AB^* into the alphabet. So:

$$(R, R^*, A, A^*)\,(R, R^*, A, A^*) \Rightarrow (R, R^*, A, A^*, B, B^*, AB, AB^*).$$

As in all previous cases, if we successively perform the → operation with (R^*), (A), (A^*), (B), (B^*), (AB), (AB^*), or any combination of these that is less than the full alphabet, the answer will be unchanged.

$$(R)\,(R, R^*, A, A^*, B, B^*, AB, AB^*) \rightarrow (R, R^*, A, A^*, B, B^*, AB, AB^*)$$
$$(R^*)\,(R, R^*, A, A^*, B, B^*, AB, AB^*) \rightarrow (R^*, R, A^*, A, B^*, B, AB^*, AB)$$

$$(A)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (A, A*, R*, R, AB, AB*, B, B*)$$
$$(A*)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (A*, A, R, R*, AB*, AB, B*, B)$$
$$(B)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (B, B*, AB, AB*, R*, R, A, A*)$$
$$(B*)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (B*, B, AB*, AB, R, R*, A*, A)$$
$$(AB)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (AB, AB*, B, B*, A, A*, R*, R)$$
$$(AB*)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (AB*, AB, B*, B, A*, A, R, R*)$$

This is, of course, identical in our formalism to:

$$(R)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$
$$(R*)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$
$$(A)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$
$$(A*)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$
$$(B)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$
$$(B*)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$
$$(AB)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$
$$(AB*)\ (R, R*, A, A*, B, B*, AB, AB*) \rightarrow (R, R*, A, A*, B, B*, AB, AB*)$$

However, with both *A* and *B* in the alphabet, we start generating apparent ambiguities, with regard to *AB* and *AB**. Thus, there would seem to be no absolute way of determining the last two terms in the last two expressions. If the concatenated (*AB*), (*AB**), etc., are to be considered as valid terms, then something must disappear when we take (*AB*) (*AB*), under →, as the only new terms allowed are those within the alphabet. In fact, we find that we have two options:

$$(AB)\ (AB) \quad \rightarrow \quad (R) \qquad (\textit{commutative})$$
$$(AB)\ (AB) \quad \rightarrow \quad (R*) \qquad (\textit{anticommutative})$$

with conjugates, of course, producing the appropriately conjugated results.

Both options would appear to be possible, but their effects will be very different. The anticommutative option, will not be repeatable when the alphabet is extended to incorporate new terms, such as (*C*), (*D*), etc., whereas the commutative option, can be repeated indefinitely. The anticommutative option effectively produces a closed ‘cycle’ with components (*A*, *B*, *AB*) and their conjugates, which excludes any further *C*, *D* …-type term of anticommuting with them. All other such terms will commute.

But there is another fundamental difference between the two options. If we choose the commutative option, we have a series of terms, such as *A*, *B*, *C* …, which are completely indistinguishable. In other words, we wouldn't know whether or not we had created something new. We would not be extending the alphabet. If, however, we choose the *anticommutative* option (as we are, in fact, obliged to do if we want to generate something new), each term will always be distinguishable from all the others, because its anticommutative partner must be unique. That is, if *B* is the anticommutative partner to *A*, then *C*, *D*, … etc., are not. So we can always identify a term by its anticommutative partner. Uniqueness is only possible because partnership is involved.

In fact, as we will find, the particular choice that makes both mathematics and physics possible as we know them, is also the most 'efficient', in requiring the minimum of choice or decision-making. That is, we automatically default at the anticommutative option whenever this is possible. The result is that the alphabets generated by ⇒ incorporate a regular series of identically structured closed cycles, each of which commutes with all others. The structure of this is identical to that which is familiar to us as the infinite series of finite (binary) integers of conventional mathematics. The closed cycles form an infinite ordinal sequence, establishing for the first time the meaning of both the number 1 and the binary symbol 1 as it appears in classical Boolean logic. The logical 1 becomes potentially a conjugation state of 0, that is, a subset alphabet defined within the system; and the alphabets structure themselves as an infinite series of binary digits.

The necessity of introducing anticommutativity has, in fact, simultaneously created the concepts of *discreteness* and *dimensionality* (specifically 3-dimensionality). As we will see from physics, there is a key connection between 3-dimensionality and discreteness in physical systems: one cannot exist without the other. Here, 3-dimensionality, i.e. anticommutativity, becomes the ultimate source of discreteness in a zero totality universe, because an anticommutative system has the property of *closure*, whereas a commutative system remains open to infinity. Geometry is logically prior to algebra and arithmetic.

1.3 The Genesis of Algebra

The process we have generated so far is based on the idea that zero totality is a *universal* requirement. It is not intrinsically algebraic, or even mathematical. It has not made the prior assumption that mathematical structures actually exist, and, in principle, it is independent of any concept of observers or physical entities. In addition, although we have used a kind of 'symbolic' representation for *R*, *A*, *B*, *C*, etc., these terms are not symbolic in the conventional sense, and do not, *a priori*, represent algebraic or any other entities.

However, the process has now reached a stage where we have developed an equivalent of numbering, ordinality or counting via a system of integers; and the concept of discreteness has emerged along with dimensionality. To maintain *R*, *A*, *B*, *C*, etc., as non-integral, non-countable and non-discrete, and yet establish their relationship with respect to the new number system which they generate, we describe them as real, non-denumerable and continuous, where these terms imply only the absence of those properties which we associate with the integers. In this sense *R*, which is simply a category of things which are unspecified and undefined, and not even definable as a set, can, by using the new concept of numbering, be re-structured as the set of reals ($\mathfrak{R}$), defined by the Cantor continuum (with cardinality $\aleph_1$, as opposed to the cardinality $\aleph_0$ of the integers). If we regard 1 as a 'unit' within the set, then we can proceed to find in $\mathfrak{R}$ all the number systems that can be constructed as countable with a 1:1 relation with the integers: rational numbers, algebraic numbers, etc., without completely specifying the set, and all of these, of course, become subject to standard arithmetical operations.

At the same time, we can see that the conjugate set R^* can be interpreted, in this re-structuring, as the set of negative reals; while *A* and A^* become the continuous sets based on the complex units i and $-i$, which each square to -1. *B* and B^*, with the necessary introduction of anticommutativity and dimensionality, convert the units of *A*, A^*, *B*, B^* into those of the *quaternions*, $\boldsymbol{i}$, $-\boldsymbol{i}$, $\boldsymbol{j}$, $-\boldsymbol{j}$, with the units of *AB* and AB^* becoming equivalent to $\boldsymbol{ij} = \boldsymbol{k}$ and $-\boldsymbol{ij} = -\boldsymbol{k}$. Further extensions to the alphabet will then be accomplished via an infinite series of new quaternion systems; and a new type of dimensional or constructible 'real' numbers will emerge to represent terms such as *AC*, *BC*, *ABC* (with countable units squaring to 1), which will be equivalent to those of Robinson's non-standard analysis or Skolem's non-standard arithmetic. It will also be possible to develop new types of mathematics by combining different aspects of the overall structure (now definable as a *rewrite* system) in novel ways, as has been the usual procedure in mathematics, and we may conjecture that all branches of mathematics that can conceivably exist may be generated by procedures internal to this structure. In particular, we may note that, once we have the concept of number, and the additional principle of complexity, we may structure John Conway's development of the surreal number system in terms of a rewrite process, with a defined alphabet.

Alternative systems of units will certainly be possible, where they can related by a 1:1 mapping to the overall structure, for example the negative unit (or iso-Minkowskian) system developed by Santilli, with its powerful applications in both physics and pure mathematics.[4] Again, since the whole generating process is defined through an initial lack of specificity, with constraints only applied when zero totality or infinite degeneracy require them, then there is no need to even consider such concepts as associativity in the context of defining a universal rewrite system; however, once numbering is introduced, then we are certainly free to introduce alternative relations between defined and undefined quantities, with particular specificities now introduced – for example, the associativity which we break in defining octonions.

We can, also use the integral ordinal sequence established with dimensionalization to restructure the subset alphabets as a series of finite groups, the order of which doubles at every stage, producing an ordinal binary enumeration. In effect, we use the process to create the concept of finite group, though it is significant that the group concept generally requires the avoidance of specific limitations on the overall system. With the group structure, the 'multiplication' and 'squaring' of elements, in addition to identity and inversion, become operations which are fundamental to the principles of duality and zero totality.

In terms of 'units' (once we have established their existence), we can express the developing structures in the form:

order 2	± 1
order 4	$\pm 1, \pm \boldsymbol{i}_1$
order 8	$\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1$
order 16	$\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1, \pm \boldsymbol{i}_2, \pm \boldsymbol{i}_2\boldsymbol{i}_1, \pm \boldsymbol{i}_2\boldsymbol{j}_1, \pm \boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1$
order 32	$\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1, \pm \boldsymbol{i}_2, \pm \boldsymbol{i}_2\boldsymbol{i}_1, \pm \boldsymbol{i}_2\boldsymbol{j}_1, \pm \boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1,$ $\pm \boldsymbol{j}_2, \pm \boldsymbol{j}_2\boldsymbol{i}_1, \pm \boldsymbol{j}_2\boldsymbol{j}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_1\boldsymbol{j}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_2, \pm \boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{i}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{j}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1$
order 64	$\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1, \pm \boldsymbol{i}_2\boldsymbol{i}_1, \pm \boldsymbol{i}_2\boldsymbol{i}_1, \pm \boldsymbol{i}_2\boldsymbol{j}_1, \pm \boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1,$ $\pm \boldsymbol{j}_2, \pm \boldsymbol{j}_2\boldsymbol{i}_1, \pm \boldsymbol{j}_2\boldsymbol{j}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_1\boldsymbol{j}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_2, \pm \boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{i}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{j}_1, \pm \boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1$ $\pm \boldsymbol{i}_3, \pm \boldsymbol{i}_3\boldsymbol{i}_1, \pm \boldsymbol{i}_3\boldsymbol{j}_1, \pm \boldsymbol{i}_3\boldsymbol{i}_1\boldsymbol{j}_1, \pm \boldsymbol{i}_3\boldsymbol{i}_2, \pm \boldsymbol{i}_3\boldsymbol{i}_2\boldsymbol{i}_1, \pm \boldsymbol{i}_3\boldsymbol{i}_2\boldsymbol{j}_1, \pm \boldsymbol{i}_3\boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1,$ $\pm \boldsymbol{i}_3\boldsymbol{j}_2, \pm \boldsymbol{i}_3\boldsymbol{j}_2\boldsymbol{i}_1, \pm \boldsymbol{i}_3\boldsymbol{j}_2\boldsymbol{j}_1, \pm \boldsymbol{i}_3\boldsymbol{j}_2\boldsymbol{i}_1\boldsymbol{j}_1, \pm \boldsymbol{i}_3\boldsymbol{j}_2\boldsymbol{i}_2, \pm \boldsymbol{i}_3\boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{i}_1,$ $\pm \boldsymbol{i}_3\boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{j}_1, \pm \boldsymbol{i}_3\boldsymbol{j}_2\boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1$

We begin with a unit integer (1) and then imagine finding an infinite series of 'duals' to this unit. We suppose that the dualling process must be carried out with respect to all previous duals, so that the entire set of characters generated becomes the new 'unit', and ensure that the total result is zero at every stage. The first dual then becomes –1, generating a new 'unit' consisting of (1, –1). Following this, we have a series of terms to which we can give symbols such as $\boldsymbol{i}_1$, $\boldsymbol{j}_1$, etc. Usually, of course, $\boldsymbol{i}_1\boldsymbol{j}_1$ would be written $\boldsymbol{k}_1$, but no new independent unit is created by this notation. An alternative expression could be in terms of multiplying factors:

order 2	$(1, -1)$
order 4	$(1, -1) \times (1, \boldsymbol{i}_1)$
order 8	$(1, -1) \times (1, \boldsymbol{i}_1) \times (1, \boldsymbol{j}_1)$
order 16	$(1, -1) \times (1, \boldsymbol{i}_1) \times (1, \boldsymbol{j}_1) \times (1, \boldsymbol{i}_2)$
order 32	$(1, -1) \times (1, \boldsymbol{i}_1) \times (1, \boldsymbol{j}_1) \times (1, \boldsymbol{i}_2) \times (1, \boldsymbol{j}_2)$
order 64	$(1, -1) \times (1, \boldsymbol{i}_1) \times (1, \boldsymbol{j}_1) \times (1, \boldsymbol{i}_2) \times (1, \boldsymbol{j}_2) \times (1, \boldsymbol{i}_3)$,

with the series repeating for an endless succession of indistinguishable $\boldsymbol{i}_n$ and $\boldsymbol{j}_n$ values. To define the character sets as true 'units', we require that the product of any unit with itself, or with any subunit, generates only the unit. So, for example, at order 8, we will have the products:

$$
\begin{aligned}
(\pm 1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \\
(\pm \boldsymbol{i}_1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \\
(\pm \boldsymbol{j}_1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \\
(\pm \boldsymbol{i}_1\boldsymbol{j}_1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \\
(\pm 1, \pm \boldsymbol{i}_1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \\
(\pm 1, \pm \boldsymbol{j}_1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \\
(\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \\
(\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) \times (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1) &= (\pm 1, \pm \boldsymbol{i}_1, \pm \boldsymbol{j}_1, \pm \boldsymbol{i}_1\boldsymbol{j}_1), \text{ etc.}
\end{aligned}
$$

For this to be always true, the terms $\boldsymbol{i}_1, \boldsymbol{j}_1, \boldsymbol{i}_2, \boldsymbol{j}_2, \boldsymbol{i}_3, \boldsymbol{j}_3$, etc. are required to have the properties of imaginary units, or square roots of –1, while the products, such as $\boldsymbol{i}_1\boldsymbol{j}_1$, must be imaginary or real units, that is, square roots of either –1 or +1. The two possibilities lead to entirely different consequences, for we can generate an unlimited number of complex products which are square roots of 1, but, for any complex number, such as $\boldsymbol{i}_1$, there is *only a single complex product* of the form $\boldsymbol{i}_1\boldsymbol{j}_1$, which is itself complex. So, if $\boldsymbol{i}_1\boldsymbol{j}_1$ is complex, then $\boldsymbol{i}_1$, $\boldsymbol{j}_1$, and $\boldsymbol{i}_1\boldsymbol{j}_1$ form a *closed system* – equivalent to the cyclic quaternion system of complex numbers,

$\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$. The system is closed with the dimensionality fixed at 3, and no other option is available. As we have seen, the choice is not arbitrary within a universal system; to create something new, we are obliged to choose the first option as default, generating an infinite number of identically structured closed systems.

Of course, the potentially infinite sequence of $\boldsymbol{i}_n$ values, with commutativity between $\boldsymbol{i}_m$ and $\boldsymbol{i}_n$ or $\boldsymbol{j}_n$ ($m \neq n$), creates the possibility of an infinite-dimensional vector-like algebra, while the anticommutativity between $\boldsymbol{i}_n$ and $\boldsymbol{j}_n$ ensures the finite- and, specifically, three-dimensionality of each of the quaternion systems. The commutativity of $\boldsymbol{i}_m$ and $\boldsymbol{i}_n$ is equivalent to defining $(\boldsymbol{i}_m\boldsymbol{i}_n)^2$ as 1, while the anticommutativity of $\boldsymbol{i}_n$ and $\boldsymbol{j}_n$ defines $(\boldsymbol{i}_n\boldsymbol{j}_n)^2$ as the conjugate, or -1. If $\boldsymbol{i}_1$ forms a quaternion system with $\boldsymbol{j}_1$, and the complex product $\boldsymbol{i}_1\boldsymbol{j}_1$, then no product of $\boldsymbol{i}_1$ with any other complex number of the form $\boldsymbol{i}_1$, $\boldsymbol{i}_2$, $\boldsymbol{i}_3$, $\boldsymbol{i}_4$, … or $\boldsymbol{j}_2$, $\boldsymbol{j}_3$, $\boldsymbol{j}_4$, … will itself be complex. If we take $\boldsymbol{i}_1$, $\boldsymbol{j}_1$, and $\boldsymbol{i}_1\boldsymbol{j}_1$ as a quaternion system ($\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$), then any further complexification, to produce, say, $\boldsymbol{i}_2\boldsymbol{i}_1$, $\boldsymbol{i}_2\boldsymbol{j}_1$, and $\boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1$, will produce a system equivalent to the multivariate vectors, complexified quaternions, or Pauli matrices (**i**, **j**, **k**), which square to positive scalar units. (The mathematical argument is detailed in the appendix to the chapter.)

As a result of this, we can define two separate processes of dualling through complex numbers: the ordinary process of complexification (i.e. multiplying by a complex number of the form $A = \boldsymbol{i}$), which can be continued to infinity (here symbolized by using terms of the form $\boldsymbol{i}_n$); and the restricted process of dimensionalization, which introduces a complementary complex factor of the form $B = \boldsymbol{j}$ (here symbolized by terms of the form $\boldsymbol{j}_n$ for each $\boldsymbol{i}_n$), which applies separately, and uniquely, to every complex operator, converting the $\boldsymbol{i}_n$ into an element of a quaternion set. The processes of *complexification* and *dimensionalization*, with their respective open and closed algebras, become simply alternative forms of duality, along with *conjugation*, or the introduction of alternative signs, + and –. It is notable that there is no such thing, in principle, as a pure complex number, only an incomplete representation of a quaternion set.

While further applications of conjugation have no effect on the structure of the group elements, repeated applications of complexification and dimensionalization take the sequence through the infinite series of quaternionic structures. Since the repeated application of conjugation makes no change to the structure, the series follows the pattern:

order 2	conjugation	×	$(1, -1)$
order 4	complexification	×	$(1, \boldsymbol{i}_1)$
order 8	dimensionalization	×	$(1, \boldsymbol{j}_1)$

order 16	complexification	×	$(1, \boldsymbol{i}_2)$
order 32	dimensionalization	×	$(1, \boldsymbol{j}_2)$
order 64	complexification	×	$(1, \boldsymbol{i}_3)$

We will show later that the three processes, taken together, and repeated indefinitely, provide the entire structure of mathematical duality required for physical application.

There is also a further sense in which such separate processes can also be seen as aspects of a single process of duality. This is evident in the fact that we can switch between the processes in our *representations* of them. For example, we can use the discrete process of dimensionalization to represent the alternative (and intrinsically non-discrete) processes of conjugation and complexification. We can even represent them on the same 2-dimensional (Argand) diagram, where, the x-y plot uses the space to the right and left of the origin to represent + and – values, and the combination of x and y coordinates to represent complex numbers. (It is significant here that we can *only* represent dual processes through discreteness, though this does not mean that the processes themselves must be discrete.) Discreteness also enables us to find more convenient ways of describing the emerging algebra.

1.4 Group Representations

The structure, as so far defined, is closely related to that of Clifford or geometrical algebra, $Cl(m, n) \equiv G(m, n)$, which is structured on m units squaring to 1 and n to –1. It also has elements of Grassmann algebra. It is interesting that these fundamentally dualistic algebras arise as aspects of simply defining a zero totality. The same applies to the group representations, which further emphasize the link with mathematical duality, and are particularly relevant to mathematical physics. Though the discussion is confined to the finite or discrete groups, the incorporation of fundamentally undefined terms as the original basis of the algebra means that it applies to continuous or Lie groups as well.

We may begin with the C_2 group, which can now be represented by 1 and –1; a dual system will extend this to four elements, producing an equivalent to $C_2 \times C_2$, and we choose the only way of extending a group including 1 and –1 to encompass four elements, by making the unknown elements (hitherto represented by the generic A and A^*) acquire the characters that we describe by the algebraic symbols i and $-i$. The group of 1, –1, i, $-i$ is not, of course, $C_2 \times C_2$, or D_2, but C_4.

However, it contains the same *information* as $C_2 \times C_2$, for we can write this information in the form of the complex ordered pairs: 1, i; 1, $-i$; -1, i; -1, $-i$, which *is* of the form $C_2 \times C_2$, and is the only domain in which $\pm i$ can exist.

If we are now required to dual the C_4 group, the most efficient and ordinally-structured way of retaining elements equivalent to 1, -1, i, $-i$ in an extended group of order eight, is by supposing that we can expand i, $-i$ into the necessarily *cyclic* and noncommutative operators $\boldsymbol{i}$, $-\boldsymbol{i}$, $\boldsymbol{j}$, $-\boldsymbol{j}$, $\boldsymbol{k}$, $-\boldsymbol{k}$, which we describe as quaternions. The definition of the quaternion group Q_4, with elements 1, -1, $\boldsymbol{i}$, $-\boldsymbol{i}$, $\boldsymbol{j}$, $-\boldsymbol{j}$, $\boldsymbol{k}$, $-\boldsymbol{k}$, is simply a statement of the fact that the complex C_4 group has been dualistically extended on the basis that $\boldsymbol{ij}$ ($= \boldsymbol{k}$) has the same kind of properties as $\boldsymbol{i}$ and $\boldsymbol{j}$, with $(\boldsymbol{ij})(\boldsymbol{ij}) = -1$. Again, we can represent the same information by a C_2 multiplication, using a group of the form $C_2 \times C_2 \times C_2$. The cyclic nature of the quaternions is significant here, because the eight possible ($C_2 \times C_2 \times C_2$) combinations of $\pm \boldsymbol{i}$, $\pm \boldsymbol{j}$, $\pm \boldsymbol{k}$ become sufficient to generate the entire information produced by the elements of Q_4. In effect, describing a set of operators, such as $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, as 'cyclic' means reducing the amount of independent information they contain by a factor 2, because $\boldsymbol{k}$, for example, arises purely from the product $\boldsymbol{ij}$. It could even be argued that the necessity of maintaining the equivalence of the Q_4 and $C_2 \times C_2 \times C_2$ representations is the determining factor in making the quaternion operators cyclic. In addition, the cyclicity prevents the definition of further complex terms, such as $\boldsymbol{I}$, where $(\boldsymbol{iI})(\boldsymbol{iI}) = -1$, though there are an unlimited number of $\boldsymbol{I}$ terms such that $(\boldsymbol{iI})(\boldsymbol{iI}) = 1$.

The process can be continued further using terms of this kind. We dual Q_4 by complexifying it to the complex quaternion or multivariate 'vector' group 1, -1, i, $-i$, $\boldsymbol{i}$, $-\boldsymbol{i}$, $\boldsymbol{j}$, $-\boldsymbol{j}$, $\boldsymbol{k}$, $-\boldsymbol{k}$, $i\boldsymbol{i}$, $-i\boldsymbol{i}$, $i\boldsymbol{j}$, $-i\boldsymbol{j}$, $i\boldsymbol{k}$, $-i\boldsymbol{k}$, of order 16, which has a related $C_2 \times C_2 \times C_2 \times C_2$ formulation, and which may also be written 1, -1, i, $-i$, $i\mathbf{i}$, $i\mathbf{i}$, $i\mathbf{j}$, $-i\mathbf{j}$, $i\mathbf{k}$, $-i\mathbf{k}$, $\mathbf{i}$, $-\mathbf{i}$, $\mathbf{j}$, $-\mathbf{j}$, $\mathbf{k}$, $-\mathbf{k}$, where a complex quaternion, such as $i\boldsymbol{i}$ becomes the equivalent of the multivariate vector $\mathbf{i}$. (It is significant, here, that a possible alternative dualling of quaternions to octonions, with sixteen components, would fail to maintain the group structure, as octonions are nonassociative.) We then expand the complex terms to a three-dimensional status, to produce a double quaternion group, say 1, -1, $\boldsymbol{I}$, $-\boldsymbol{I}$, $\boldsymbol{J}$, $-\boldsymbol{J}$, $\boldsymbol{K}$, $-\boldsymbol{K}$, $\boldsymbol{i}$, $-\boldsymbol{i}$, $\boldsymbol{j}$, $-\boldsymbol{j}$, $\boldsymbol{k}$, $-\boldsymbol{k}$, of order 32, which has a related $C_2 \times C_2 \times C_2 \times C_2 \times C_2$ formulation. Then we complexify again, to produce a multivariate vector-quaternion group 1, -1, i, $-i$, $i\mathbf{i}$, $-i\mathbf{i}$, $i\mathbf{j}$, $-i\mathbf{j}$, $i\mathbf{k}$, $-i\mathbf{k}$, $\mathbf{i}$, $-\mathbf{i}$, $\mathbf{j}$, $-\mathbf{j}$, $\mathbf{k}$, $-\mathbf{k}$, $\boldsymbol{i}$, $-\boldsymbol{i}$, $\boldsymbol{j}$, $-\boldsymbol{j}$, $\boldsymbol{k}$, $-\boldsymbol{k}$, $i\boldsymbol{i}$, $-i\boldsymbol{i}$, $i\boldsymbol{j}$, $-i\boldsymbol{j}$, $i\boldsymbol{k}$, $-i\boldsymbol{k}$, and 36 real and complex combinations of vectors and quaternions, forming a group of 64, with a related $C_2 \times C_2 \times C_2 \times C_2 \times C_2 \times C_2$ formulation. Because of the reduction of information involved in defining both multivariate vectors and quaternions as cyclic, and in one producing complex, and the other real, products, the $C_2 \times C_2 \times C_2 \times C_2 \times C_2 \times C_2$ formulation can be expressed by the 64 possible combinations of $\pm\,\mathbf{i}$, $\pm\,\mathbf{j}$, $\pm\,\mathbf{k}$, $\pm\,\boldsymbol{i}$, $\pm\boldsymbol{j}$, $\pm\,\boldsymbol{k}$. Further dualling is possible on the same basis, but it is clear that only three fundamental principles are required to continue the dualling to infinity – opposite signs (or equivalent), the distinction between real and imaginary components, and the introduction of cyclic dimensionality – and to establish every conceivable combination of these, that is to establish every type of dualling, requires a group of 64 elements.

C_2	C_2	$\pm\,1$	conjugate
C_4	$C_2 \times C_2$	$\pm\,1, \pm\,i$	complexify
Q_4	$C_2 \times C_2 \times C_2$	$\pm\,1, \pm\,\boldsymbol{i}, \pm\boldsymbol{j}, \pm\,\boldsymbol{k}$	dimensionalize
V_{16}	$C_2 \times C_2 \times C_2 \times C_2$	$\pm\,1, \pm\,i, \pm\,\boldsymbol{i}, \pm\boldsymbol{j}, \pm\,\boldsymbol{k}$	complexify
QQ_{32}	$C_2 \times C_2 \times C_2 \times C_2 \times C_2$	$\pm\,1, \pm\,\boldsymbol{I}, \pm\,\boldsymbol{J}, \pm\,\boldsymbol{K}, \pm\,\boldsymbol{i}, \pm\boldsymbol{j}, \pm$	dimensionalize
VQ_{64}	$C_2 \times C_2 \times C_2 \times C_2 \times C_2 \times C_2$	$\pm\,1, \pm\,i, \pm\,\boldsymbol{I}, \pm\,\boldsymbol{J}, \pm\,\boldsymbol{K}, \pm\,\boldsymbol{i}, \pm\boldsymbol{j}, \pm\,\boldsymbol{k}$	complexify

The process becomes entirely repetitive at the level of V_{16}, while VQ_{64} is what we obtain by combining C_2, C_4, Q_4, and V_{16} as independent elements, establishing conjugation, complexification, dimensionalization and repetition.[5] Beyond this stage, we can consider the sequence proceeding through an infinite series of quaternionic structures by repeated processes of complexification and dimensionalization, creating an algebra of infinite dimensions, whose units are each quaternionic. Repetition necessarily sets in as soon as we establish the principle of closure, and closure, as we shall see, allows us an immediate procedure for returning to zero.

The order 16 group, as we have seen, is of special interest as creating what is effectively a 'real' dimensional structure of the kind observed in normal 3-dimensional vector space. The order 16 group of complex quaternions is notably equivalent to the 'real' dimensional structure of 3-dimensional multivariate vector space, as used in the geometrical algebra of Hestenes and others, and applied by them to the algebra of physical space and time, to generate electron spin as a natural consequence of spatial three-dimensionality.[6,7] The components, $\pm\,1$, $\pm\,\boldsymbol{i}_1$, $\pm\boldsymbol{j}_1$, $\pm\,\boldsymbol{i}_1\boldsymbol{j}_1$, $\pm\,\boldsymbol{i}_2$, $\pm\,\boldsymbol{i}_2\boldsymbol{i}_1$, $\pm\,\boldsymbol{i}_2\boldsymbol{j}_1$, $\pm\,\boldsymbol{i}_2\boldsymbol{i}_1\boldsymbol{j}_1$, can be rearranged and written in the form $\pm\,1$, $\pm\,i$, $\pm\,\mathbf{i}$, $\pm\,\mathbf{j}$, $\pm\,\mathbf{k}$, $\pm\,i\mathbf{i}$, $\pm\,i\mathbf{j}$, $\pm\,i\mathbf{k}$, where $\pm\,1$, $\pm\,i$, become the respective scalar and pseudoscalar, and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, and $i\mathbf{i}$, $i\mathbf{j}$, $i\mathbf{k}$ the respective vector and pseudovector terms of this algebra. In this algebra, which is also isomorphic to that of Pauli matrices, the 'total' product of two multivariate vectors **a** and **b** is of the form $\mathbf{a}.\mathbf{b} + i\,\mathbf{a} \times \mathbf{b}$, and the ‘total’ products of the vector units is of the form $\mathbf{ii} = \mathbf{jj} = \mathbf{kk} = 1$; and $\mathbf{ij} = -\mathbf{ji} = i\mathbf{k}$; $\mathbf{jk} = -\mathbf{kj} = i\mathbf{i}$; and $\mathbf{ki} = -\mathbf{ik} = i\mathbf{j}$. In summary, quaternions follow the multiplication rules:

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

$$\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}$$

$$\boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}$$

$$\boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$$

$$\boldsymbol{ijk} = -1,$$

while those for complexified quaternions are:

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

$$(i\boldsymbol{i})(i\boldsymbol{j}) = -(i\boldsymbol{j})(i\boldsymbol{i}) = i(i\boldsymbol{k})$$

$$(i\boldsymbol{j})(i\boldsymbol{k}) = -(i\boldsymbol{k})(i\boldsymbol{j}) = i(i\boldsymbol{i})$$

$$(i\boldsymbol{k})(i\boldsymbol{i}) = -(i\boldsymbol{i})(i\boldsymbol{k}) = i(i\boldsymbol{j})$$

$$(i\boldsymbol{i})(i\boldsymbol{j})(i\boldsymbol{k}) = i,$$

which are isomorphic to those for multivariate vectors, or Pauli matrices:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = 1$$

$$\mathbf{ij} = -\mathbf{ji} = i\mathbf{k}$$

$$\mathbf{jk} = -\mathbf{kj} = i\mathbf{i}$$

$$\mathbf{ki} = -\mathbf{ik} = i\mathbf{j}$$

$$\mathbf{ijk} = i.$$

The succession, then, allowing for conjugation (±) within each group, becomes:

order 2	real scalar
order 4	complex scalar (real scalar plus pseudoscalar)
order 8	quaternions
order 16	complex quaternions or multivariate 4-vectors
order 32	double quaternions
order 64	complex double quaternions or multivariate vector quaternions

The order 16 group (if we are to retain the maximum indistinguishability by avoiding octonion-type nonassociativity) is also the point at which the extension of the sequence becomes one of repetition. So, a complete specification of an interative generating procedure could be made by using the groups of order 2, 4, 8 and 16. Taken as independent entities, these may be combined minimally in the group of order 64, using the symbols ± 1, $\pm i$, $\pm \boldsymbol{i}$, $\pm \boldsymbol{j}$, $\pm \boldsymbol{k}$, $\pm \mathbf{i}$, $\pm \mathbf{j}$, $\pm \mathbf{k}$, to represent the respective units required by the scalar, pseudoscalar, quaternion and multivariate vector groups. This will take on physical significance when we realize that the algebra of this group is that of the gamma matrices used in the Dirac equation – the quantum equation determining the behaviour of the most fundamental components of matter – and that these matrices may be represented as the terms $\boldsymbol{k}$, $\boldsymbol{ii}\mathbf{i}$, $\boldsymbol{ii}\mathbf{j}$, $\boldsymbol{ii}\mathbf{k}$, $\boldsymbol{ij}$, whose binomial combinations, as we will show, are sufficient to generate the entire group.

1.5 Rewriting Nature

So far, we have concentrated on showing that a requirement of non-unique zero totality automatically generates a structure which has the characteristics of an algebra, including countable numbers, even without the prior assumption of the existence of mathematics. The structure effectively 'rewrites' itself in the way that occurs in the formal rewrite systems used in computing, examples of which are given by von Koch (1905)[8], Chomsky (1956)[9], Naur *et al* (1960)[10], Mandelbrot (1982)[11], Wolfram (1985)[12], and Prusinkiewicz and Lindenmayer (1990)[13], among others, and, as previously suggested, we may include the work of Conway here. The system we have proposed, however, unlike those, is a *universal* one, which effectively rewrites all known rewrite systems, and so extends the applicability and power of the entire concept to any sphere of computing. In this sense, as well as providing a foundation for mathematics, it also provides a foundation for computing, in either its classical or quantum forms.

Now, rewrite systems (or production systems) are synonymous with computing in the sense that most software is written in a language that must be rewritten as symbols for some hardware to interpret. Formal rewrite, or production, systems are pieces of software that take an object usually represented as a string of characters and using a set of *rewrite rules* generate a new string representing an altered state of the object. If required, a second *realisation system* takes the string and produces a visualisation or manifestation of the objects being represented. However, using the concept of zero totality, we can show that we can define a rewrite system from which other rewrite systems may be constructed at a basic level. Of course, while such a system can be encapsulated, for convenience, in a computer program written in a high level language, that program must be recognised as being different from the rewriting mechanism which it represents. It would simply be a way of realising this using existing knowledge, while the originating mechanism would represent ‘computer language’ at a much more fundamental level.

Rewriting in the computing sense always begins with an initial state. In this case, it is a string representation of 0. We begin with the idea that only 0 is unique. Everything that is not 0 is undefined. In rewriting, we start with an argument denying that we have a non-0 starting-point. We assume that we are not entitled to posit anything other than 0, and that we are forced to rewrite when we start from any other position. In the process we observe the significance of the concept of hierarchy, and of the difference between *recursion* (where the entire set of objects and relations is already in existence) and *iteration* (where the objects and relations are successively built up in a process which starts from the simplest).

Traditionally computer rewrite systems involve objects defined in terms of symbols representing characters drawn from a finite alphabet, and a series of states. To move from state to state we apply a finite set of rules – rewrite rules or productions – to a string of the symbols that represents the current state of the complex object. Some stopping mechanism is defined to identify the end of one state and the start of the next (for example we can define that for each symbol or group of symbols in a string, and working in a specific order, we will apply every rule that applies). It is usual in such systems to halt the execution of the entire system if there is no change in the string generated or if the changes are cycling, or after a specified number of iterations. Differing stopping mechanisms determine different families of rewrite systems, and in each family, alternative rules and halting conditions may result in strings representing differing species of object. Allowing new rules to be added dynamically to the existing set and allowing rules to be invoked in a stochastic fashion are means whereby more complexity may be introduced.

To explain how an apparent ‘something’ results from an ultimate nothing we need to show how a universal alphabet that encompasses duality and nothingness can be developed using a universal rewriting system. There are, in fact, two methods by which the elements of this alphabet may be discovered. One of these methods yields an infinite number of subset alphabets each of which has properties that can be exploited, for example using further rewrite systems based on the subset alphabet. As we have seen from the algebraic development, at this stage, a powerful simplification becomes apparent, reducing the procedure to three fundamental processes.

If we relax the rules regarding the finiteness of the characters in the alphabet(s) and the number of states, but continue to assume the rest of the constraints described above, a more universal rewrite system is defined. Such a system has alphabets as its complex objects and subset alphabets (all the symbols so far delivered) as states. For this to remain a rewrite system an initial state (that can be re-written) must exist, and for it to be universal there must be, we may conjecture, a minimum of two rewrite rules (productions). One of these, *create,* delivers a new symbol at each invocation. The term 'symbol' is used here because what is delivered may be a single character of the alphabet, a subset alphabet, or indeed an entire alphabet. The second rule, *conserve*, examines all symbols currently in existence to ensure that no anomalies exist as a consequence of bringing the new one into existence.

With such a minimum universal rewrite system, the initial state (usually called the Δ-state) must contain at least one symbol that we can use to identify that the universe is empty. However, any symbol we choose is immediately (and simultaneously) a symbol, a character of the final alphabet, a subset alphabet and full alphabet in its own right. We may choose, arbitrarily, the single symbol 0 (zero), and set it as the string representing the complex object in the Δ-state {0}. We are obliged to make an arbitrary choice here because we cannot use *create* without the Δ-state – the minimum rewrite system condition for a universal system. If we were to use *conserve* now it would simply return that 0 is unique, fixed, and consistent and no change from the Δ-state would be generated. We must therefore now invoke *create* supplying the Δ-state as parameter, or source, string.

If we presume that *create* is an algorithm with stopping criteria, it returns a result target string containing a new symbol. If the paradigm for the algorithm were recursive, the resulting symbol (E is used here) would represent every character of the alphabet at the first step. To create any refining character, a specific e_x, using the *recursive* paradigm would be impractical because of the implied infinity and storage requirement. We may not use an iterative paradigm at this stage because we would have to supply an upper limit and / or need to identify which of the infinite characters we are creating. Both of these actions require a character not yet in the character set (alphabet) we have so far defined.

The pair of symbols, the string {0, E} is our new object (alphabet) which we now submit to *conserve* which examines every combination of symbols (Table 1.1):

Table 1.1

	0	**E**
0	00	0E
E	E0	EE

We note that 00, the ‘transition’ from 0 to 0, conserves 0. The combination 0E is the transition from 0 to E and is balanced, for all E, by its conjugate partner E0 which is the transition back from E to 0, thereby conserving 0. The combination EE, the transition from every symbol E to every other, is anomalous and must be returned by *conserve* as unexplained or ‘inconsistent’ as it does not appear to conserve 0. However, at infinity, all transitions represented by EE will have been examined, EE will be declared ‘nilpotent’ (i.e. ‘squaring’ to 0) in that it delivers 0, and we will be left with three generic combinations:

(00, 0E, E0)

However, it is impractical to use the recursive version of *conserve* to examine further the elements of E because of the implied infinite number of iterations.

We return to the *create* process and accept that we must postulate symbols Δ_a, Δ_b, … Δ_n drawn from E such that they are in an arbitrary ordinal sequence. We note that there is an infinite number of such sequences because choice of Δ_a is arbitrary. However, we may now use an iterative paradigm for *create* and because n is specified, an iterative (or recursive) *conserve* can be constructed. At the end of each invocation we are now presented with a symmetrical table of transitions that represent the simplest set of properties for the current set of n symbols (Table 1.2).

Table 1.2

	0	$\mathbf{\Delta_a}$	$\mathbf{\Delta_b}$	$\mathbf{\Delta_c}$	**…**	$\mathbf{\Delta_n}$
0	00	$0\Delta_a$	$0\Delta_b$	$0\Delta_c$		$0\Delta_n$
$\mathbf{\Delta_a}$	Δ_a0	$\Delta_a\Delta_a$	$\Delta_a\Delta_b$	$\Delta_a\Delta_c$		$\Delta_a\Delta_n$
$\mathbf{\Delta_b}$	Δ_b0	$\Delta_b\Delta_a$	$\Delta_b\Delta_b$	$\Delta_b\Delta_c$		$\Delta_b\Delta_n$
$\mathbf{\Delta_c}$	Δ_c0	$\Delta_c\Delta_a$	$\Delta_c\Delta_b$	$\Delta_c\Delta_c$		$\Delta_c\Delta_n$
:						
$\mathbf{\Delta_n}$	Δ_n0	$\Delta_n\Delta_a$	$\Delta_n\Delta_b$	$\Delta_n\Delta_c$		$\Delta_n\Delta_n$

The Δ_a row and Δ_a column illustrate the conjugate pair structure observed earlier. The remaining cells of Table 1.2 identify explicitly each Δ symbol to Δ symbol

ZERO TO INFINITY - The Foundations of Physics

transition observed generically in Table 1.1. Off diagonal there are symmetrical conjugate pairs; there are, for example, three such cancelling pairs when $n = b$ and six when $n = c$. The diagonal cells of the table contain transitions from each symbol to itself and do not cancel out in this way.

We now invoke the *conserve* process, noting that it does not define the transition property but merely identifies those novel transition combinations that appear not to conserve 0. When $n = a$, the symbol Δ_a is added to the alphabet and the transition $0\Delta_a$ is introduced. We need $\Delta_a 0$ (and the idea that this is a conjugate form) to conserve 0. However, this leaves the combination $\Delta_a\Delta_a$ unexplained (i.e. novel) and to conserve 0 we must conjecture that, whatever it is, is balanced by whatever is to come – or both are 'nilpotent' in the sense introduced above. To discover this we invoke *create* to add a new symbol to the alphabet which then defines (arbitrarily) the $n = b$ row and column. At $n = b$ (in *conserve*) we continue to require the conjugate explanation for all off-diagonal elements in the table. In addition, we have non-0 to non-0 symbol transitions, each of which has a cancelling conjugate, and which must ultimately yield a symbol already in the alphabet. However, when these transitions are explained we still have $\Delta_b\Delta_b$ as novel, and require the method of explaining the novelty used earlier. We see that at every invocation of *conserve* we define the need for an additional symbol, delivered by *create* – it is inherent that both processes are obligatory. Other processes may now be conjectured within the rewrite system that impart meaning to 'transition' and also to each transition from Δ_n to Δ_n; however, in each case all of what is to come must balance the $\Delta_n\Delta_n$ in the diagonal position. 'Balance' in this explanation assumes that the 00 transition yields 0; however, we could consider it to yield a conjugate of some form. Where this is the case we may consider each newly created diagonal element as 'balancing' that conjugate by delivering the unconjugated form. In each case the new symbol created carries the entire subset alphabet.

The properties and symbols emerge from the application of the two rewrite rules and would have been equally valid for any of the infinite alternative selections. Significantly, since the ultimate aim is to recover the zero state through an infinite series of processes, the emergence should be seen as being of a *supervenient* nature, that is, without temporal connotation. Furthermore, the symbol delivered at each step has all the properties of all the symbols previously delivered, and in a hierarchical and orthogonal fashion.

Finally, we note that the symbol 0, the existence of the Δ-state, and the processes *create* and *conserve* are outside the rewrite system in that they must exist before the system can function. If we can allow these assumptions, we may also presume the existence of some natural machine that will deliver, for a set of appropriate rewrite rules, a corresponding alphabet where the symbols themselves map to specific rules. In terms of the rewrite procedures we have adopted, the assumption of any non-zero category (previously expressed as R) must immediately lead to the return to zero, which, in mathematical terms, becomes equivalent to supposing a 'negative' category or 'conjugate' (R^*) corresponding to the original assumption; the combination must then produce the desired zero state by some process (not yet specified as 'addition'). In terms of Table 1.2, this is the recognition that $\Delta_a\Delta_a$ leads to the creation of the new symbol Δ_b. A process of 'self-referencing', correlation or 'combination', which can be expressed in the form RR (again without assuming a specific mathematical interpretation), has produced an extended categorization, from (R) to (R, $-R$). At this point we have created ordinality (through the existence of an implied + and an explicit –), though not yet counting, as there is no discreteness or anything fixed involved in the procedure. (The same applies in the Dedekind 'cut', used to define real numbers in conventional mathematics; despite its name, this is a definition of ordinality without a prior assumption of discreteness.)

However, discreteness emerges naturally as a result of extending the alphabet, and it is the next application of the create procedure ($\Delta_b\Delta_b \rightarrow \Delta_c$) which leads to the number system as we know it, for now we have an undifferentiated 'set' of possible origins for the 'negative' ordinal category or conjugate. We describe these as complex forms, and each must have its own conjugate. In mathematical terms, the complex category remains completely undefined in respect to the real category, and has no ordinal relation to it. There are infinitely possible or indefinitely possible systems that are represented by the mathematical A, even for a seemingly specified real category. It is only when we express this fact in the next creation stage that we are able to begin to extend ordinality towards enumeration, for this stage leads to what become mathematical 'combinations' of complex categories. We find here that to every conceivable A, e.g. A, B, C, …, there are indefinitely possible (commutative) combinations leading to the original real category (e.g. $AC\,AC = R$), but very definite and restricted (anticommutative) ones leading to the conjugate (e.g. $AB\,AB = -R$).

These alternative possibilities provide the respective foundations for infinite- and finite-dimensional algebras. The infinite-dimensional algebra, as we will see in chapter 3, leads to the infinite Hilbert vector space of quantum mechanics, while the Hamilton or 3-dimensional algebra is responsible for the cyclic system of quaternions. It is the cyclicity of the latter which introduces discreteness or closure, and the concept of ‘unity’, and it is thus no coincidence that discreteness in physics is invariably associated with dimensionality (see chapter 2). As we have seen in 1.3, in our generation of the associated algebraic structure, we can choose the default position of taking the conjugate combination to create a regular ordinal sequence. We now find that only ‘one’ independent A-type concept (say B) is associated with each conceivable A, and we can sequence the terms ordinally by choosing indistinguishability between the As in every conceivable respect. So the sequence, although arbitrary, becomes a series of integral binary enumerations, which we can also apply to ordinality in the real categories. Defining 1, at all, in this way, automatically creates a *dual* system, with –1, etc., equivalent to requiring $1 + 1 = 2$, and generating the Peano idea of ‘successor’. With the reals, integers, and complexity as fundamental aspects of the system, the remaining mathematical number categories (and higher algebras) can be defined by applying the ordinality condition in a variety of ways, as in conventional mathematics. No new principle is required.

In effect, the hierarchical and orthogonal mathematical structure suggested by the rewrite mechanism is the following:

R	undefined	Δ_a
R, R^*	conjugation	Δ_b
R, R^*, A, A^*	complexification	Δ_c
$R, R^*, A, A^*, B, B^*, AB, AB^*$	dimensionalization	Δ_d
$R, R^*, A, A^*, B, B^*, AB, AB^*, C, C^*, AC, AC^*, BC, BC^*, ABC, ABC^*$	repetition	Δ_e

As before, all the character sets are generated by the simple rule that no non-zero character set or alphabet is complete, so every self-referencing yields a new or extended alphabet, which is necessarily self-conjugated; and we can, of course, extend the process to infinity, as in section 1.3. The significance of this system for both classical and quantum computing will become apparent after we have first examined the foundations of physics.

1.6 Quaternions and Vectors

Because of their great importance in physics, it will be convenient at this point to provide a technical appendix, giving a short account of the relationship between quaternions and vectors. Quaternions were finally discovered by William Rowan Hamilton in 1843, but he had actually started work, many years earlier, with the Argand diagram, used to represent complex numbers ($a + bi$), in which the x-axis represented the real system of numbers (a), and the y-axis, drawn orthogonally to it, represented the imaginary numbers (bi). This, he thought, could be seen as representing two-dimensional space, and he sought to extend it to provide a 3-dimensional version using an extension of complex numbers. There is, of course only one set of *real* numbers, but Hamilton reasoned that, with the true nature of imaginary numbers being unknown, there was no reason why there could not be two sets of imaginary numbers, which could be drawn orthogonally to each other, say bi and cj. However, he found very quickly that such a system of 'triplets' would not work, for the system did not exhibit 'closure'; that is, the product ij had no meaning within the system, and had to represent some other quantity.

The breakthrough came when he realised that if he invented a *third* system of imaginary numbers (dk), he could define a system that exhibited closure, though he had to sacrifice the principle of commutativity (the idea that ab always equals ba) to do it. In this system, $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$, the three square roots of –1, are related by the formulae:

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = \boldsymbol{ijk} = -1.$$

However, the products of the units are noncommutative, since

$$\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}$$

and so on. This is obvious from consideration of $\boldsymbol{ij}\,\boldsymbol{ji} = -\boldsymbol{i}^2 = 1$. The full rules become:

$$\begin{gathered} \boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1 \\ \boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k} \\ \boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i} \\ \boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j} \\ \boldsymbol{ijk} = -1, \end{gathered}$$

Hamilton suspected what Frobenius later proved (in 1878) that the system was unique; no other extension of complex algebra was possible that maintained the same algebraic rules. He was convinced also that he had therefore found the true explanation for the three-dimensionality of space. But, as with ordinary complex numbers, the three imaginary parts had to be accompanied by a real part, making a four-part number, which he described as a *quaternion*. But, if the three imaginary parts represented the three dimensions of space, what did the real part mean? Almost immediately, he postulated that it must be time. Now, quaternions were used enthusiastically in the mid-nineteenth century by mathematical physicists like P. G. Tait and James Clerk Maxwell. But, in the end, several factors caused them to be discarded in favour of a new form of vector algebra, promoted strongly by Willard Gibbs and Oliver Heaviside. The first was that quaternions, in normal usage, gave the wrong signs when squared for Pythagorean addition, the answers being negative where positive ones would have been expected; while the second was the fact that quaternions were complicated structures made up of real and imaginary parts, and only the imaginary (or three-dimensional) part was needed in most physical applications. So vector algebra extracted this imaginary part from the complete quaternion and changed it from an imaginary to a real quantity, devising a restricted set of rules for its usage, which were derived ultimately from the more extensive quaternion algebra.

This system proved to be of such immediate utility that, despite the efforts of Tait and others, vector algebra eventually succeeded quaternion algebra in all major applications. However, it has long been established that quaternions were at least the *parent* of vector theory. From the very beginning, Hamilton had realised that quaternions involved a vector part (the imaginary one) and a scalar part (the real one), and he even defined equivalents of the modern vector and scalar products, and the modern vector differential operator, ‘del’, while Maxwell, using quaternions, developed a great deal of modern vector calculus; the terms ‘vector’ and ‘scalar’, used in the modern sense, are also Hamilton’s own. In principle, the vector algebraists did little more, in mathematical terms, than separate out vector and scalar parts, and vector and scalar products, change them over from real to imaginary, or imaginary to real, and devise a neater and more consistent notation, though vector algebra still uses Hamilton’s **i**, **j**, **k** as well as his ∇.

However, quaternions are, in themselves, and not just in their application to the development of vectors, one of the most powerful tools ever presented to the physicist – many people, in fact, use them routinely without realising that it is quaternions that they are actually using. In fact, if they had been used in the correct way early on, they would have pre-empted many later physical developments that came about much more tortuously.[14] For example, Einstein developed, in 1905, a simple kinematical theory, based on only two physical postulates, for explaining a whole series of facts concerning the ‘Electrodynamics of Moving Bodies’, but, within two and a half years, this theory (the special theory of relativity) was, in a way, superseded by an even simpler theory, due to Hermann Minkowski, which was based on mathematical, rather than physical, assumptions.

In classical vector space, Pythagoras' theorem required the length of a line element r to be invariant to arbitrary changes in the components from x, y, z to x', y', z' according to the rule:

$$r^2 = x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2.$$

To this equation, Minkowski added a term representing time, multiplied by a unit-conversion factor numerically equal to the velocity of light. In Einstein's words, written after his conversion to the new scheme: 'Mathematically, we can characterize the generalized Lorentz transformation thus: it expresses x', y', z', t' in terms of linear homogeneous functions of x, y, z, t of such a kind that the relation

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2$$

is satisfied identically.'[15] Minkowski thus replaced the ordinary vector described by three real parts (x, y, z), representing Euclidean space, with a space-time *4-vector* with three real parts and one imaginary (x, y, z, ict); and proceeded to apply the same principle to every part of physics previously described by vectors.

Minkowski's argument, though very striking, was not entirely original. Its mathematics had been almost anticipated by Poincaré, who had effectively introduced the Minkowski metric in his paper of July 1905, without fully realising the relativistic implications, and the idea of four dimensions had had a long prehistory. But the parallel with quaternions and with Hamilton's view of the connection of space and time are obvious, and it was soon pointed out to Minkowski himself that his space-time had a basically quaternion-type structure. A. W. Conway, for example, used quaternions in a simple and elegant way to derive the Lorentz transformations and all the other relativistic equations.[16] The interest in this issue is not just historical, for, even now, it is usually considered mathematically convenient, to write 4-vector line elements in what is effectively a quaternion form,

$$ds^2 = c^2\,dt^2 - dr^2 = c^2dt^2 - dx^2 - dy^2 - dz^2,$$

whose very structure is equivalent to making the time element real and the space elements imaginary.

Many applications of vector theory, of course, don't require the time component, but, even in these cases, a proper attention to quaternion theory shows that what seems arbitrary and awkward in vector theory has a natural and simple explanation in terms of quaternions. Vector theory, as it stands, breaks most of the rules of ordinary algebra – namely, commutativity, associativity, the law of moduli, unambiguous division and closure – and it has two kinds of ‘product’, neither of which is a product in the ordinary algebraic sense. The rules of vector algebra are designed purely to fit physical requirements and seem to make very little mathematical sense. However, they do make sense if seen in a quaternionic context.

The most convenient way of showing this is to use one of the ‘Clifford algebras’, named after William K. Clifford, who discovered its basic principles in the 1870s. The term is generically used to cover all the algebras (also called geometric algebras) which combine such elements as real and complex numbers, vectors and quaternions, but the type which we need for our immediate purposes is the *Cl* (3, 1) algebra, an algebra isomorphic to complex quaternions, which is, in effect, the complete reverse of the *Cl* (1, 3) quaternion algebra. ($Cl\ (m, n) \equiv G\ (m, n)$ means an algebra founded on m square roots of 1 and n square roots of –1.) This is the algebra which Hestenes has called ‘multivariate’ vector algebra, and it effectively subjects 4-vectors to the full quaternionic rules of multiplication.

First of all, we look at quaternion multiplication. If we take two quaternions, with zero real parts:

$$\boldsymbol{a} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$$

and

$$\boldsymbol{a'} = x'\boldsymbol{i} + y'\boldsymbol{j} + z'\boldsymbol{k},$$

and multiply them, we obtain

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

This, as Hamilton originally showed, divides into a scalar product $-(xx' + yy' + zz')$ and a vector product $\boldsymbol{i}\,(yz' - zy') + \boldsymbol{j}\,(zx' - xz') + \boldsymbol{k}\,(xy' - yx')$, and we can recognise, almost immediately, that the first term is virtually identical to the normal scalar product of two vectors (except for the negative sign) and the second term is the same as the normal vector product – except that the unit imaginary quaternions $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ replace the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$. (It is convenient to use the convention of bold italic script for quaternions and bold for vectors.) The important thing about this multiplication, however, is that it obeys the ordinary rules of algebra (except for commutativity). It doesn’t require any special definition or any special rules. In particular, the product is a quaternion, like the multiplicands, and so we have ‘closure’.

In ordinary vector algebra, by contrast, there are two methods of multiplication and neither method exhibits ‘closure’: the dot product of two vectors produces a scalar and the cross product a pseudovector, and both products break other basic algebraic rules. What we really need is a *single* method of multiplication of vectors which exhibits closure and which maintains all normal algebraic properties except commutativity, a method which *combines* dot and cross products. By *direct analogy with quaternions*, it is easily shown that this ‘full’ product of vectors **a** and **a'** is of the form:

$$\mathbf{aa'} = \mathbf{a}.\mathbf{a'} + i\,\mathbf{a} \times \mathbf{a'}.$$

This is the basic rule of the Clifford algebra *Cl* (3,1). Assuming that (as with quaternions) the basic quantity is a combination of vector and scalar, the product exhibits closure and all the other basic algebraic properties except commutativity. The imaginary sign before the cross product is necessary to establish complete symmetry with the quaternion system, and the introduction of imaginary quantities into vector theory gives us an easy explanation of the difference between scalars and pseudoscalars, and vectors and pseudovectors.

Pseudoscalars are simply imaginary scalars, that is, ordinary (real) scalars which have been multiplied by *i*, the square root of –1; and pseudovectors are ordinary (real) vectors multiplied by the same factor. Both types of term arise naturally from the algebra. If we take two orthogonal vectors, **a** and **b**, then the scalar product is 0, and the total product is a pseudovector

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

An example is the area of a rectangle. If **a** and **b** are parallel, on the other hand, then the total product is a real scalar

$$\mathbf{ab} = \mathbf{a}.\mathbf{b}.$$

The product of a pseudovector $i\,\mathbf{a} \times \mathbf{b}$ and parallel real vector **c** will be an imaginary scalar quantity, and, therefore, a pseudoscalar. For example, if **a**, **b** and **c** are the vectors representing the sides of a cube, then the volume is given by the total product

$$\mathbf{abc} = i\,\mathbf{a} \times \mathbf{b}.\mathbf{c}.$$

The product of a pseudovector $i\,\mathbf{a} \times \mathbf{b}$ and a perpendicular real vector **c**, however, which involves the multiplication of two imaginary terms, will be a real quantity, $i\,(i\,\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$ and, therefore, a true vector. An example of this occurs in the formula

$$\mathbf{F} = e\,\mathbf{v} \times \mathbf{B},$$

where we are effectively multiplying a vector $e\,\mathbf{v}$ by an orthogonal pseudovector $\mathbf{B}$. We also obtain a real vector if we multiply a pseudovector like area ($\mathbf{A}$) by a pseudoscalar like pressure (P), as in the formula

$$\mathbf{F} = P\mathbf{A},$$

and we obtain a real scalar, work (PV), if we take the product of the pseudoscalars, pressure and volume. Once again, the imaginary operators are removed by multiplication.

A great advantage of this system is that we have no need to define separate vector and scalar products for the unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. We simply define total products of the form

$$\mathbf{ii} = \mathbf{i.i} + i\,\mathbf{i} \times \mathbf{i} = \mathbf{i.i} = 1,$$

and

$$\mathbf{ij} = \mathbf{i.j} + i\,\mathbf{i} \times \mathbf{j} = i\,\mathbf{k},$$

with

$$\mathbf{ji} = \mathbf{j.i} + i\,\mathbf{j} \times \mathbf{i} = -i\,\mathbf{k}.$$

Products of two vectors are called ‘bivectors’, and products of three vectors are ‘trivectors’. (They are literally ‘triple products’.) For unit vectors $\mathbf{i}$ and $\mathbf{j}$, the bivector (or ‘area’ element) is $i\mathbf{k}$, while the trivector (or ‘volume’ element) is i. The full rules are written:

$$\begin{gathered}
\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = 1 \\
\mathbf{ij} = -\mathbf{ji} = i\mathbf{k} \\
\mathbf{jk} = -\mathbf{kj} = i\mathbf{i} \\
\mathbf{ki} = -\mathbf{ik} = i\mathbf{j} \\
\mathbf{ijk} = i.
\end{gathered} \tag{1.1}$$

Physicists, of course, have long used this algebra, in the form of the so-called ‘Pauli matrices’, which are employed in applying spin to the Schrödinger equation. These are defined as:

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

and follow the rules:

$$\begin{gathered}
\sigma_x\sigma_y = -\ \sigma_y\sigma_x = i\sigma_x \\
\sigma_y\sigma_z = -\ \sigma_z\sigma_y = i\sigma_x \\
\sigma_z\sigma_x = -\ \sigma_x\sigma_z = i\sigma_y.
\end{gathered}$$

and
$$\sigma_x\sigma_x = \sigma_y\sigma_y = \sigma_z\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}.$$

Using the algebraic operators, in fact, makes it completely unnecessary to use any matrices at all, and, by applying the full multiplication properties of vectors to quantum mechanics, we can show that the spin of the electron can be derived from the nonrelativistic Schrödinger equation (by adding a cross product term), just as readily as it can from the relativistic equation of Dirac. So, if the true 'quaternion'-like nature of vector quantities had been recognised at the time, it would have been possible to derive the spin of the electron directly from nonrelativistic quantum mechanics, without having to put it into the equation ad hoc. This also tells us that spin has nothing to do with relativity or 4-vectors; it involves only the vector part of the quaternion, and, as we might expect, analogies have been found in classical physics, where vectors are equally important.

Another way of obtaining the same algebra is to use complexified quaternions, so that
$$\mathbf{i} = i\boldsymbol{i}; \qquad \mathbf{j} = i\boldsymbol{j}; \qquad \mathbf{k} = i\boldsymbol{k}.$$
Then
$$(i\boldsymbol{i})^2 = (i\boldsymbol{j})^2 = (i\boldsymbol{k})^2 = 1$$
$$(i\boldsymbol{i})(i\boldsymbol{j}) = -(i\boldsymbol{j})(i\boldsymbol{i}) = i(i\boldsymbol{k})$$
$$(i\boldsymbol{j})(i\boldsymbol{k}) = -(i\boldsymbol{k})(i\boldsymbol{j}) = i(i\boldsymbol{i})$$
$$(i\boldsymbol{k})(i\boldsymbol{i}) = -(i\boldsymbol{i})(i\boldsymbol{k}) = i(i\boldsymbol{j})$$
$$(i\boldsymbol{i})(i\boldsymbol{j})(i\boldsymbol{k}) = i.$$
These rules are clearly the same as those defined in (1.1).[17] The multivariate vectors are, notably, as anticommutative as the parent quaternions. An anticommutative, associative algebra of this kind is necessarily 3-dimensional, but an infinite vector (Grassmann) algebra can be defined by making the terms commutative.