Representations of a Fundamental Theory

Peter Rowlands

IQ Group and Science Communication Unit, Department of Physics, University of Liverpool, Oliver Lodge Laboratory, Oxford Street, P.O. Box 147, Liverpool, L69 7ZE, UK. e-mail: p.rowlands@liverpool.ac.uk

Abstract. A series of representations of the fundamental theory that physics has a symmetrical or dualistic structure is used to show that both physical and mathematical ideas result from an attempt to maintain a zero total content in any fundamental conceptual scheme for explaining nature.

INTRODUCTION

The fundamental theory to be represented is the idea that physics has its origins in a symmetrical structure which preserves its conceptually zero content. It has become popular in recent years to suppose that the total energy in the universe, counting gravitational energy as negative, may be zero, and that the entire universe may have emerged out of pure nothingness through something like a quantum fluctuation. Perhaps also, if we include the vacuum state, there is some kind of ultimate balance between matter and antimatter. It is not unusual, in fact, to read statements like the one by the well-known chemist and science writer Peter Atkins [1994], who has said of physical matter that 'the seemingly something is elegantly reorganized nothing, and … the net content of the universe is … nothing'. However, it is not just matter and the universe that appear to be nothing, but the entire conceptual scheme of which these are merely components (*nihil ex nihil fit*). Mathematically, only zero is absolutely unique, and this leads one to suppose that an absolute or universal theory must in a sense be, ultimately, a theory about the meaning of nothingness. The idea of conceptual nothingness has been proposed on many occasions by the author as the basis of physics [Rowlands, 1983, 1991, 1999, 2001], in the sense that the perfect symmetry between the only truly fundamental parameters in nature is exactly of this sort, even in algebraic terms. In view of the intrinsically mathematical nature of physical quantities, and the probability that mathematics and physics have the same conceptual origin, it also seems to make sense to describe mathematics in terms of the same totally zero structure. The question then becomes: how do we get something from nothing? Such a question is unlikely to be answered on the basis of an appeal to pure first principles without some initial empirical investigation into the way that physics and mathematics appear to be structured at their foundations. A convenient starting-point is the evidence of duality, or, rather, *something which leads ultimately to the creation of this concept*, as a fundamental principle.

DUALITY AND ITS GROUP STRUCTURE

Duality is fundamental to physics and mathematics. All operations are dual operations. All objects are dual objects. All physical and mathematical theories are dual theories. As Nicholas Young writes in a well-known textbook [1988], 'the idea of duality pervades mathematics', while duality has, from the beginning, been the basis of the author's theories of mathematical physics and philosophy of science [Rowlands, 1983, 1991, 1999, 2001]. In effect, we can't define something without defining also what it is not. Alternatively, we can't characterize 'nature' or 'reality', even to the extent of saying whether it has an independent existence (is ontological) or is a product of our perception (is epistemological). So, the concept even has a fundamental philosophical manifestation. Any attempt to characterize it in one aspect will automatically lead to our discovery of the 'opposite' characterization in another. Every 'probe' will meet with an opposing 'response'.

Why is this? It seems that physics and mathematics, as we have supposed, are attempts at creating something from nothing. Although we assert that we have given nothing a character or aspect, it is, in fact, still nothing, if we take the totality of probe and response. An obvious case is provided by the conservation of linear momentum in an explosion problem; gain of positive momentum in one direction has to be countered by an equal gain of negative momentum in a direction which is precisely opposite. The totality remains zero. The example is a very familiar one, and it is usually treated as an illustration of a very fundamental law of physics, with an independent existence. The law is, in fact, however, just one illustration of a very much more fundamental principle of duality in nature, which can be seen as the ultimate origin of both physics and mathematics. A 'theory of everything' needs first to be a 'theory of nothing'.

How, then, does this principle of duality operate? The answer has to be: in the simplest way possible. It is not possible to imagine any duality simpler than that provided by the C_2 group. We could describe it, in mathematical terms, by the use of the elements 1 and –1, but our starting concept must, in fact, be even simpler than that, and cannot yet assume the discrete numbering associated with the term 'dual'. The concept of 1 may appear to be simple, but it is in reality loaded with information about the meaning of discreteness, in addition to ordinality, which doesn't appear at the most fundamental level. A simpler opposition is that between + and – applied to the unspecified entities which are generically described as the reals. A mathematical-computational approach to this is given by the author, in a paper written with Bernard Diaz [Rowlands and Diaz, 2002].

In effect, the simplest possible thing other than 0 that we can imagine is $\mathfrak{R}$, or can be described as $\mathfrak{R}$, where $\mathfrak{R}$ is a totally unspecified or undifferentiated entity, and its automatic negation or 'conjugation' is the thing we describe as $-\mathfrak{R}$. In principle, as soon as we have $\mathfrak{R}$, we have no option but to take $-\mathfrak{R}$ as well, in order to maintain the zero totality. That is, defining $\mathfrak{R}$, at all, automatically creates what we will eventually call a 'dual' system. Ultimately, we will find that this is equivalent to requiring 1 + 1 = 2, and generating the Peano idea of 'successor', and a natural (binary) numbering system, which avoids the Gödel problem through a zero totality, but this will require the generation of a concept of discreteness which is not a direct feature of $\mathfrak{R}$. We cannot, at this stage, even take $\mathfrak{R} - (-\mathfrak{R})$ to be, say, $2\mathfrak{R}$, until we have defined 2, and the concept of number generally, to exist. The existence of + and – signs can thus be taken as an expression of ordinality, but not yet of a discrete ordinality (as with the Dedekind 'cut', which, despite its name, is a definition of ordinality without a prior assumption of discreteness), and the expression of this process as *duality*, or the group structure as C_2, cannot yet be taken as explicit.

We also have no option but to relate $-\mathfrak{R}$ to $\mathfrak{R}$ in some way other than defining their totality as 0, and the identity $-\mathfrak{R} \times -\mathfrak{R} = \mathfrak{R} \times \mathfrak{R}$ then becomes deeply significant in establishing that the relation between these elements is a *group* relationship, and that the 'multiplication' and 'squaring' of elements, in addition to identity and inversion, are operations which are fundamental to the principle that we will ultimately describe as 'duality', when we have introduced discreteness. Of course, without some concept of enumeration, and no way of identifying $\mathfrak{R}$ more exactly, $-\mathfrak{R} \times -\mathfrak{R} = \mathfrak{R} \times \mathfrak{R}$ is describable only by the generic $-(-\mathfrak{R})$ or $\mathfrak{R}$; there is no defined concept of 'unit', and the × operation is not yet identifiable as 'multiplication'.

But, suppose now that we require a counter concept, even for the $\mathfrak{R}$ and $-\mathfrak{R}$ category, that is, we require a system which avoids 'privileging' these, and privileging one with respect to the other. We may then suppose that conjugate terms must exist which allow us to generate $-\mathfrak{R}$ in the same way as we generate $\mathfrak{R}$ from $-\mathfrak{R}$ (in addition to its generation from $\mathfrak{R}$). In mathematical terms, we describe these as members of the complex set, and each must have its own conjugate. Symbolically, we represent the new terms as $\mathbb{C}$ and $-\mathbb{C}$. However, the new category $\mathbb{C}$ remains undefined in respect to the real category, and has no ordinal relation to it. Consequently, there are infinitely possible or indefinitely possible systems and combinations that are represented by this symbol.

However, when we investigate the combinations of possible $\mathbb{C}$ terms, we find a distinct separation between the infinitely possible combinations leading to the original real category $\mathfrak{R}$, and the very definite *noncommutative* ones leading to the conjugate $-\mathfrak{R}$, where only 'one' independent $\mathbb{C}$-type concept (say $\mathbb{C}'$) is associated with each conceivable $\mathbb{C}$. Thus, we find that the former are infinitely extendible, while the latter are cyclic or enclosed. It is at this point that we can introduce discreteness, and the concept of 'unity', into mathematics. By choosing the default position of assuming indistinguishability between the $\mathbb{C}$s in every conceivable respect (i.e. none is 'privileged'), we can create a regular ordinal sequence, which, although arbitrary in principle, becomes a series of integral binary enumerations, which can also be applied to ordinality in the real categories. The consequent minimising of variation within the generating process additionally allows us to retain (and even create) the powerful notion of group structure. We can represent the generality of the process in the form:

$\mathfrak{R}$	undefined
$\mathfrak{R}, -\mathfrak{R}$	conjugation
$\mathfrak{R}, -\mathfrak{R}, \mathbb{C}, -\mathbb{C}$	complexification
$\mathfrak{R}, -\mathfrak{R}, \mathbb{C}, -\mathbb{C}, \mathbb{C}', -\mathbb{C}', \mathbb{C}\mathbb{C}', -\mathbb{C}\mathbb{C}'$	dimensionalization
$\mathfrak{R}, -\mathfrak{R}, \mathbb{C}, -\mathbb{C}, \mathbb{C}', -\mathbb{C}', \mathbb{C}\mathbb{C}', -\mathbb{C}\mathbb{C}',$ $\mathbb{C}'', -\mathbb{C}'', \mathbb{C}\mathbb{C}'', -\mathbb{C}\mathbb{C}'', \mathbb{C}'\mathbb{C}'', -\mathbb{C}'\mathbb{C}'',$ $\mathbb{C}\mathbb{C}'\mathbb{C}'', -\mathbb{C}\mathbb{C}'\mathbb{C}''$	repetition

The logical operations involved in the sequence can be expressed in a quasi-algebraic form though operations such as × and – are not limited to an algebraic interpretation until we create the concept of integral sequencing via the ordinal series of closed systems:

$\mathfrak{R} \times \mathfrak{R} = -\mathfrak{R} \times -\mathfrak{R} = \mathfrak{R}$	
$\mathfrak{R} \times -\mathfrak{R} = -\mathfrak{R} \times \mathfrak{R} = -\mathfrak{R}$	
$\mathfrak{R} \times \mathbb{C} = \mathbb{C} \times \mathfrak{R} = \mathbb{C}$	
$\mathbb{C} \times \mathbb{C} = -\mathbb{C} \times -\mathbb{C} = -\mathfrak{R}$	
$\mathbb{C} \times -\mathbb{C} = -\mathbb{C} \times \mathbb{C} = \mathfrak{R}$	
$\mathbb{C}' \times \mathbb{C}' = -\mathbb{C}' \times -\mathbb{C}' = -\mathfrak{R}$	
$\mathbb{C}\mathbb{C}' \times \mathbb{C}\mathbb{C}' = -\mathbb{C}\mathbb{C}' \times -\mathbb{C}\mathbb{C}' = -\mathfrak{R}$	closed (anticommutative)
$\mathbb{C}\mathbb{C}'' \times \mathbb{C}\mathbb{C}'' = -\mathbb{C}\mathbb{C}'' \times -\mathbb{C}\mathbb{C}'' = \mathfrak{R}$	unlimited (commutative)

The character sets effectively represent all those, including $\mathcal{R}$, $-\mathcal{R}$ which are generated by operating on themselves:

$$(\mathcal{R}) \times (\mathcal{R}) = (\mathcal{R})$$
$$(\mathcal{R}, -\mathcal{R}) \times (\mathcal{R}, -\mathcal{R}) = (\mathcal{R}, -\mathcal{R})$$
$$(\mathcal{R}, -\mathcal{R}, \mathbb{C}, -\mathbb{C}) \times (\mathcal{R}, -\mathcal{R}, \mathbb{C}, -\mathbb{C}) = (\mathcal{R}, -\mathcal{R}, \mathbb{C}, -\mathbb{C})$$
$$(\mathcal{R}, -\mathcal{R}, \mathbb{C}, -\mathbb{C}, \mathbb{C}', -\mathbb{C}', \mathbb{C}\mathbb{C}', -\mathbb{C}\mathbb{C}') \times (\mathcal{R}, -\mathcal{R}, \mathbb{C}, -\mathbb{C}, \mathbb{C}', -\mathbb{C}', \mathbb{C}\mathbb{C}', -\mathbb{C}\mathbb{C}')$$
$$= (\mathcal{R}, -\mathcal{R}, \mathbb{C}, -\mathbb{C}, \mathbb{C}', -\mathbb{C}', \mathbb{C}\mathbb{C}', -\mathbb{C}\mathbb{C}'), \text{ etc.}$$

The closed sets are those which introduce discreteness through anticommutativity.

We are now in a position to extend the argument using the integral sequence we have created. Beginning 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 $\mathbb{C}$ and –$\mathbb{C}$) 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_8, 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_8. 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_8 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 ***I***, where (***iI***)(***iI***) = –1, though there are an unlimited number of ***I*** terms such that (***iI***)(***iI***) = 1.

The process can be continued further using terms of this kind. We dual Q_8 by complexifying it to the complex quaternion or multivariate 'vector' group 1, –1, *i*, –*i*, ***i***, –***i***, ***j***, –***j***, ***k***, –***k***, *i**i***, –*i**i***, *i**j***, –*i**j***, *i**k***, –*i**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***i**, *i***i**, *i***j**, –*i***j**, *i***k**, –*i***k**, **i**, –**i**, **j**, –**j**, **k**, –**k**, where a complex quaternion, such as *i**i*** becomes the equivalent of the multivariate vector **i** (see Appendix 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, ***I***, –***I***, ***J***, –***J***, ***K***, –***K***, ***i***, –***i***, ***j***, –***j***, ***k***, –***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***i**, –*i***i**, *i***j**, –*i***j**, *i***k**, –*i***k**, **i**, –**i**, **j**, –**j**, **k**, –**k**, ***i***, –***i***, ***j***, –***j***, ***k***, –***k***, *i**i***, –*i**i***, *i**j***, –*i**j***, *i**k***, –*i**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 ± **i**, ± **j**, ± **k**, ± ***i***, ± ***j***, ± ***k***, the algebra of the Dirac gamma matrices. 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	± 1	conjugate
C_4	$C_2 \times C_2$	± 1, ± *i*	complexify
Q_8	$C_2 \times C_2 \times C_2$	±1, ± ***i***, ± ***j***, ± ***k***	dimensionalize
V_{16}	$C_2 \times C_2 \times C_2 \times C_2$	± 1, ± *i*, ± ***i***, ± ***j***, ± ***k***	complexify
QQ_{32}	$C_2 \times C_2 \times C_2 \times C_2 \times C_2$	± 1, ± ***I***, ± ***J***, ± ***K***, ± ***i***, ± ***j***, ± ***k***	dimensionalize
VQ_{64}	$C_2 \times C_2 \times C_2 \times C_2 \times C_2 \times C_2$	± 1, ± *i*, ± ***I***, ± ***J***, ± ***K***, ± ***i***, ± ***j***, ± ***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_8, and V_{16} as independent elements, establishing conjugation, complexification, dimensionalization and repetition. 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 infinite-dimensional Grassmann algebra, 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 process of conjugation, of course, can be repeated, like those of complexification and dimensionalization, but it is defined in such a way that repetition produces no new structure.)

THE PARAMETER GROUP

So far, this sounds purely mathematical. What relevance, then, does it have to physics? The answer is that it is, in fact, purely *physical* in origin. Duality is a *physical* requirement of the description of nature, and not necessarily a requirement of an abstract system of logical thought, though it may well be that such a system cannot be separated from considerations derived from physical requirements. In effect, when we define the dual, we *define the physical*. The words are synonymous. So, we should expect to see manifestations of these structures in physical 'reality', as we ordinarily perceive it.

From purely empirical considerations of physics, it has been possible previously to suggest that it is based on the relationships between only four fundamental parameters: space, time, mass and charge (where charge is a general term for the sources of the electromagnetic, weak and strong interactions) (see Appendix II). Further investigation of these suggest that the most fundamental properties and 'antiproperties' they possess are as follows:

space	nonconserved	real	countable
time	nonconserved	imaginary	noncountable
mass	conserved	real	noncountable
charge	conserved	imaginary	countable

This has the structure of a $C_2 \times C_2$ or D_2 relationship, in which any of space, time, mass or charge may be the group identity element, and each is its own inverse (see Appendix III). It has also been shown that the symmetry is exact, and absolutely unbroken within physics. Especially significant, however, is the fact that countability or discreteness is a necessary requirement for cyclic multidimensionality, for unidimensionality is an obviously necessary property of a continuous or noncountable quantity – it can't have an origin. But, multidimensionality is also a necessary property of discreteness (at least in a nonconserved parameter like space). Discreteness has to have a reference or origin; we can't imagine observing discreteness in space without at least another dimension for reference. However, when we investigate space and charge, we find further that the dimensionality in each case is also *three*-dimensionality and cyclic, just as we require for our dual system. Space, being real, has the properties of a multivariate vector, with the associated pseudoscalar being imaginary time, in the '4-vector' combination; while charge, being imaginary, has the properties of a quaternion, with the associated real scalar being mass. In the case of the quaternions, also, it is significant that three-dimensionality is the only dimensionality which, *mathematically*, preserves the group structure; the mathematical possibility is determined *at the same time* as the physical. With the arguments already presented, we can additionally say that the origin of the *physical* concepts of continuity and discreteness lie in the duality which requires the creation of a cyclic three-dimensionality in our conceptualization of nature. This is, in fact, what we *mean* by continuity and discreteness.

We can see now that two of the distinctions between the parameters, which we have derived inductively from observed physical characteristics (real / imaginary and noncountable / countable), are identical to the C_2 distinctions which extend the original C_2 duality into complexity and cyclic dimensionality. However, even the original C_2 duality (1 / –1) originated from the act of creating 'something from nothing' (1 from 0), the very definition of *nonconservation*, as is the concept of 'successor' which it implies. So, in principle, our group of space, time, mass and charge has all the elements required to extend physical duality to infinity. And, our choice of the distinction between conservation and nonconservation (in effect, incorporating 0 directly as the never-used 'totality', and leaving –1 as implicitly understood rather than explicit) even allows us to simplify a potential $C_2 \times C_2 \times C_2$ structure into the simpler $C_2 \times C_2$ we have used above, with the added bonus that we can represent it as identity element (or single sign of scalar) plus three 'quaternion' terms, thus creating a powerful mapping of the four parameters onto a quaternion space. Alternatively, we could represent mass and charge as 'conjugated' quantities, in the sense that creation of a + value can only be accomplished at the same time as the creation of an equivalent – value. So the group could be written in a form, in which the property / antiproperty distinctions occur as examples of successive applications of the dualling process (nonconjugated → conjugated; real → complex; nondimensional → dimensional):

space	nonconjugated	real	dimensional
time	nonconjugated	complex	nondimensional
mass	conjugated	real	nondimensional
charge	conjugated	complex	dimensional

Mathematically, it is possible to create a dual set or parameters, one form of which is seen in certain versions of the Dirac theory, in which certain characteristics of space and time, and mass and charge are reversed, for example the real / imaginary characteristics.

space*	nonconserved	imaginary	countable
time*	nonconserved	real	noncountable
mass*	conserved	imaginary	noncountable
charge*	conserved	real	countable

The combined group is then extended to $C_2 \times C_2 \times C_2$, with a quaternion representation (Q_8) with both signs of scalar. This extends the mathematical representational space, but is not needed in the physical representational space, and, in the more sophisticated (quantum field) versions of Dirac theory becomes redundant.

There is also a distinction between the representations of the distinctions between the parameter properties (e.g. real / imaginary) by existence / nonexistence conditions, as here; and the explicit representation of these properties by their explicit natures (e.g. vector / quaternion). The minimum representation in the latter case is of the order $C_2 \times C_2 \times C_2 \times C_2 \times C_2 \times C_2$, or the Dirac group. In the former case, there are at least two striking visual representations of the group relations, which bring out the significance of the C_2 distinctions and of the principle of cyclic dimensionality.

COLOUR REPRESENTATION

The four parameters, space, time, mass and charge, are represented by concentric circles, the parameter chosen as the identity element for the group occupying the centre circle. The division of the properties into three components is reflected by the division of the circles into three sectors. The properties (say, Real, Nonconserved, Discrete) are represented in Figure 1, by primary colours (say, Red, Green, Blue), and the ‘antiproperties’ (Imaginary, Conserved, Continuous) by the complementary secondary colours (Cyan, Magenta, Yellow). All of these choices are individually arbitrary (as we see from Appendix III), as is the choice of secondary colours to represent the properties, and primary colours to represent the 'antiproperties' in Figure 2. The division between properties and antiproperties is also a completely free choice. Only the overall pattern is fixed. As configured, with Space selected as the identity element, and the colour representation for the properties selected as indicated, the innermost circle represents Space, the next Charge, the next Mass, and the outermost circle Time. But this will be changed as soon as we redefine any of the colour representations or exchange the status of any or the property-antiproperty pairs. In addition to being an alternative representation of the main group, Figure 2 may also be used, simultaneously with Figure 1, as a representation of the dual group, which can be obtained (for example) by exchanging the status of real and imaginary quantities (as in some versions of the Dirac theory).

Figure 1

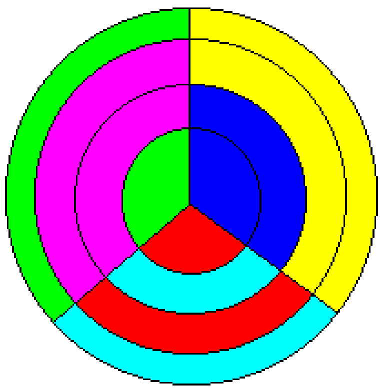
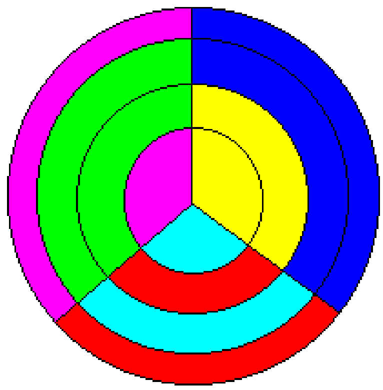


Figure 2



The nature of the fundamental parameter group is demonstrated in this representation by summing up the colour combinations in each of the circles. This results in a white inner circle for the identity element, and a sequence of the three primary colours (Figure 3) or secondary colours (Figure 4), which adds up to a white totality.

Figure 3

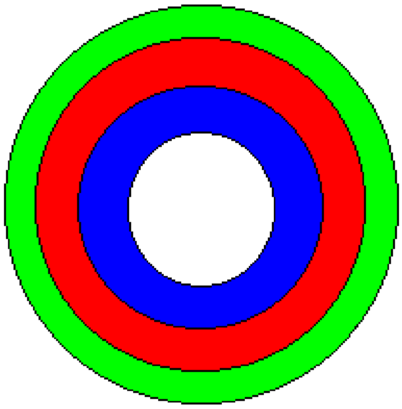
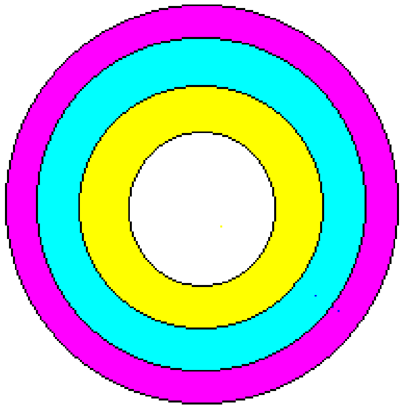
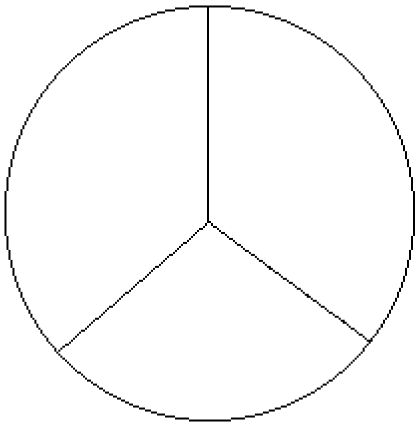


Figure 4



Adding up the property-antiproperty combinations in the sectors also results in a white totality for each sector, as expected (Figure 5).

Figure 5



Clearly, this colour representation derives its effectiveness from the fact that a three-colour system is a simulation of three-dimensionality in an alternative vector or quaternion representation. Such representations are also possible in a more direct form.

3-D (VECTOR) REPRESENTATION

The alternative (vector or quaternion) representation (Figures, 6 and 7) can be seen as a literal interpretation of the *x*, *y* and *z*, used in the tables in Appendix III. The *x*, *y* and *z* directions represent the properties, and the –*x*, –*y*, and –*z* directions the antiproperties (again, according to an arbitrary choice). The four 'Red' lines, in Figure 6 (shown here as continuous), drawn from the origin of these 3-dimensional axes, then represent the four parameters, and the 'Cyan' lines (shown as dotted) those of the dual group. Figure 7 shows the same representation as Figure 6, but without the axes.) The 'Red' lines are reflections of each other in two planes. We can represent these as preserving the sign of the volume element (or identity), if the axes are taken in the same cyclic sense; and so they correspond to the parameter group The Red plus Cyan lines are the reflections of each other in a single plane, and do not preserve the sign of the volume element; and so form the parameter group taken with its dual. The reflection of a line in three planes produces its exact dual.

It will be apparent that the representation of the dualities of the parameter group using either the three real spatial dimensions or the pseudo-dimensions of the three primary colours is a powerful way of bringing out the connections between duality and dimensionality, and the fact that all the individual dualities are, in effect, versions of the same mechanism. It is also a convenient way of showing how the parameter group can be used to represent a kind of 'super-duality' of all the elements, conveniently displayed using the particular duality of dimensionality.

Figure 6

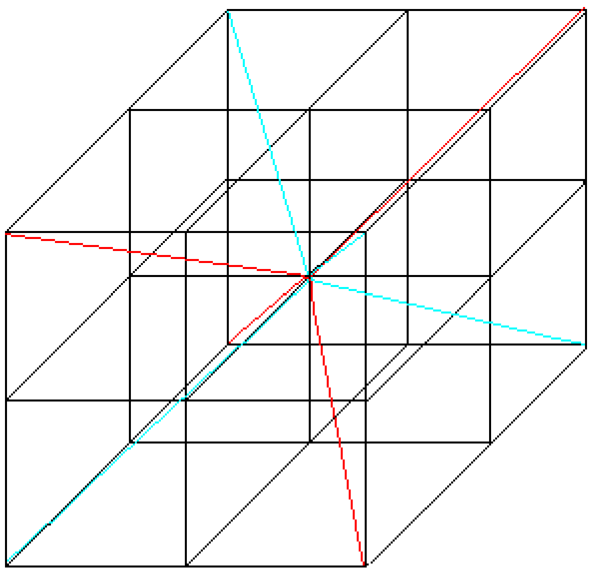
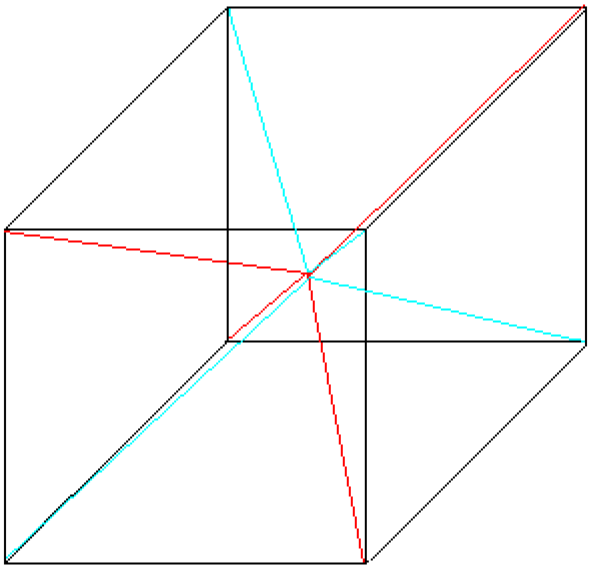


Figure 7

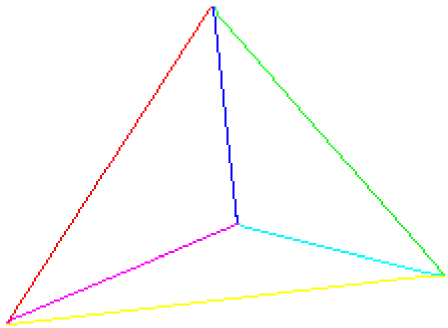


To map between the colour and 3-D representations, we could represent the positive *x*, *y*, *z* axes in Figure 6 by the primary colours, say Red, Green, Blue, against a black background, and the corresponding negative directions by the complementary secondary colours, which, in this case, would be Cyan, Magenta, Yellow. The four Red vectors of Figure 6 would become respectively White, Blue, Red, Green, as determined by the colour coding in Figures 1 and 3, and the four Cyan vectors of Figure 6 would become respectively White, Yellow, Cyan, Magenta, as determined by the colour coding in Figures 2 and 4.

TETRAHEDRAL REPRESENTATION

Yet another 3-D representation (Figure 8) would place the parameters at the vertices of a regular tetrahedron, with the six edges coloured to represent the properties and antiproperties as in Figure 1.

Figure 8



We can consider the faces of the tetrahedron to be the members of the dual group, and, clearly, an alternative representation would reverse primary and secondary colours and / or the roles of faces and vertices. It might be possible to consider the tetrahedron as close-packed with inverted tetrahedra with complementary colour-representation in an all-white solid-space, which can be extended to infinity.

An interesting possibility is that a structure like the one in Figure 8, if flattened out in a 2-dimensional space, could be considered as a 'dart' or 'kite' in a Penrose tiling pattern (the base line being optional to the connections between the vertices representing the parameters). Penrose tiling is, of course, a five-fold symmetry, and, typically, a group of five darts (or kites) will produce a star-shaped pattern with each of the darts joined with all the others at its apex, and with its two nearest neighbours along two of its edges. The star then has five inner and fiver outer vertices, and, surrounding a central star made of darts, we will have ten kites, each joined by two edges to two nearest neighbours, and by another edge to one of the ten outer edges of the star. If we assume that each of the 4 vertices of any dart (3 symmetric and 1 asymmetric) must represent one of space, time, mass and charge, and that joint vertices may only represent one parameter, then putting a 3-dimensional parameter like charge or space, at the centre of the star forces us to choose the inner and outer vertices in such a way that the other 3-dimensional parameter occurs three times in the five inner or outer vertices, while each of the other parameters occurs only once. It is interesting that another five-fold structure (the Dirac algebra) emerging from a combination of two types of 3+1-component units (space-time 4-vectors and mass-charge quaternions) is also forced to 'privilege' one of its two 3-dimensional quantities, space or charge.

THE DIRAC NILPOTENT

In the parameter group, not only are the properties dual, but so is the distribution between the parameters. This is why the minimum representation of the full duality is the Dirac algebra, of order $C_2 \times C_2 \times C_2 \times C_2 \times C_2 \times C_2$, and produced by the 64 possible combinations of the 'double vector', ± **i**, ± **j**, ± **k**, ± ***i***, ± ***j***, ± ***k***. Hidden within this representation, but expressive of the cyclic nature of the operators (and apparent in Figures 6 and 7), are the respective pseudoscalar and scalar terms, ± *i* and ± 1 (and we could, alternatively, use the 'double vector', ± **i**, ± **j**, ± **k**, ± *i**i***, ± *i**j***, ± *i**k***). Leaving out ± signs, the full representations are: space (**i**, **j**, **k**), time (*i*), mass (1), and charge (***i***, ***j***, ***k***). The fact that the Dirac algebra can be derived from a combination of two three-dimensional operators now suggests a further possibility, based on our mathematical and visual representations of the group. This is that one of the two three-dimensional parameters may be mapped on to the other three parameters, represented as the 'dimensions', and, in fact, the smallest set of units from which the full algebra can be derived comes from exactly such a mapping. We take, for example,

time	space	mass	charge
i	**i** **j** **k**	1	***i*** ***j*** ***k***

and, taking, each of the units of charge onto one of the 'dimensions' represented by time, mass and space,

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

create the following combinations:

*i**k***	***i*i** *i***j** *i***k**	***j***

A set of five units of this kind, or pentad, will always generate the entire Dirac algebra of 32 parts (excluding signs). The 32 parts turn out to be 1 and *i*, and six Dirac pentads, three based on the quaternion operators (as here) and three on the vector operators. Any of these sets can be used as the basis for the five gamma matrices in the Dirac equation, but it is most convenient to use the quaternions, as here, because charge is a conserved quantity, and the mathematical structure then has a convenient physical interpretation. (In the case of vector space, the components are not uniquely determined, because the quantity is nonconserved, and can even be arbitrarily reduced to a single one. However, the conservation of charge is directly related to the conservation of angular momentum, and so brings in the spatial rotation simultaneously, as becomes evident in the full explanation of symmetry-breaking.) The combined units take on the physical characteristics of their component quantities. The charge units introduce conservation and discreteness (quantization) to all the quantities. However, the new conserved quantities retain their respective pseudoscalar, vector and real scalar identities, as Dirac energy, Dirac momentum and Dirac rest mass:

$$i\boldsymbol{k} \qquad \mathbf{i}\boldsymbol{i} \;\; \mathbf{i}\boldsymbol{j} \;\; \boldsymbol{i}\mathbf{k} \qquad \boldsymbol{j}$$
$$E \qquad \mathbf{p} \qquad m$$

In a nonconserved form they produce the respective quantum operators:

$$\partial / \partial t \qquad \nabla \qquad m$$

Treating the momentum term as a single quantity, the free-fermion Dirac state vector now becomes a nilpotent $(\pm\, i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\, e^{-i(Et - \mathbf{p.r.})}$, where $(\pm\, i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ expresses the absolute conservation of charge and mass(-energy), and the exponential term, operated upon by $(\pm\, i\boldsymbol{k}\partial / \partial t \pm \boldsymbol{i}\nabla + \boldsymbol{j}m)$, the absolute nonconservation of space and time. The identities of the three 'charge' operators are preserved, even in the combinations of $(\pm\, i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$, for they now become discriminated into ones with timelike (weak), spacelike (strong) and masslike (electric) properties, and the effects can be distinguished physically by the aspects of angular momentum conservation to which they relate. The Dirac algebra, which produces the simplest possible combination of all the dualistic properties required by space, time, mass and charge, generates a broken symmetry in the manifestations of the charges' interactions.

Significantly, the term $(\pm\, i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$, which is expressed most conveniently as a row or column vector with four components (yet another 4-vector mapping, and one which can be accomplished with the four quaternion components, 1, $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$, if required) is a nilpotent or square root of zero, a precise expression of the fundamentally dualistic process of returning 'something' back to 'nothing' through a squaring operation. The classic way of doing this is through the Dirac equation in which the nonconservation operator $(\pm\, i\boldsymbol{k}\partial / \partial t \pm \boldsymbol{i}\nabla + \boldsymbol{j}m)$ is applied to the state vector $(\pm\, i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)\, e^{-i(Et - \mathbf{p.r.})}$, in which $(\pm\, i\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ represents the conserved terms, so producing a zero result. This expresses the fact that our fundamental duality has been represented in terms of conservation and nonconservation, and that the effect of applying both is to maintain the zero totality.

The Dirac nilpotent represents the most concise packaging of the dualistic information contained in the parameter group, the most complete way of parameterizing nature; and, as we have seen, the combination of all the desired physical elements, with all the inherent symmetries, into a single, parameterization of nature, is the same as the process of ‘quantization’ of energy, momentum (or angular momentum) and ‘rest mass’.

The interaction of a fermion with the (infinite) vacuum, or mass-energy continuum, produces an infinite succession of products or superpositions of Dirac nilpotent states. This extends the dualling processes to infinity. Each of the ‘virtual’ states produced also acts in the same way, producing a pattern of the same form as the Conway system of constructed real numbers. The requirement of infinite dualling ensures the entanglement of all states in the universe (although, as with classical interference, decoherence will make this virtually unobservable except in special cases).

It seems that we get something from nothing, not just in a physical way, by perfect symmetry between the parameters denying overall characterization, but also literally, by making the fundamental unit of our characterization a square root of zero, and that this becomes zero in the Dirac equation when we apply to it a differential operator, and generate an exact equal to it as an eigenvalue. The Dirac equation itself expresses the fundamental duality of our view of ‘nature’, for the left-hand term (the differential operator) specifies the nonconserved aspects, and the right-hand term (the wavefunction) the conserved aspects. Any individual nilpotent wavefunction structure ($\boldsymbol{k}E + i\boldsymbol{i}\mathbf{p} + i\boldsymbol{j}m$) must then be unique because a superposition of identical ones would zero the wavefunction of the entire universe, and Pauli exclusion becomes obvious. This specification of uniqueness requires instant correlation, at the same time as the 4-vector nature of the operator requires time-delayed action between discrete sources. It is also a reflection of the uniqueness or local conservation of individual charge components.

THE FACTOR 2

Duality has an astonishingly simple manifestation in physics through the appearance of the factor 2 everywhere where it becomes significant. This is discussed in detail in the paper ‘The physical significance of the factor 2’ [Rowlands, 2002]. The main result of this is that each process that doubles the options available also produces a doubling of the physical effect which can be reduced to simple numerical terms. At the same time this is often balanced by a halving of the options in another direction.

Thus, when we describe a physical process using constant, rather than changing, terms, we are effectively using both sides of the + / – duality at once. This is the case when we use potential, rather than kinetic energy equations, or both action and reaction sides of Newton’s third law, or even relativistic, rather than rest, mass. Relativity itself does not introduce the factor 2, but relativistic equations can often be used as classic examples of changing conditions. The most controversial instance in historical terms is the double bending of light rays in a gravitational field, which can, in fact, be seen as an example of the use of a kinetic, rather than potential, energy equation. Of coursing, halving in one respect may lead to doubling in another. So halving the energy, by using a kinetic term, produces a doubling of the angular deflection; but it is also possible to produce the doubling directly by taking both space- and time-related effects into account. This could be seen as an application of the second route to duality: complexification, or the adding of a complex term to a real one. It is significant here that the group relationship between the physical parameters is so integrated that such apparently alternative explanations emerge without any fundamental contradiction. Both explanations are equally true, and neither has precedence over the other.

Another significant alternative can be seen in the explanation of half-integral fermion spin and its effects. It is possible to derive the resulting magnetic properties using a classical kinetic energy equation, and so using one side of the + / – duality. On the other hand, both the Dirac and Schrödinger equations derive the half-integer value for spin indirectly by using the *doubling* effect produced by the third process of duality: dimensionalization. This comes about in both cases via the anticommuting properties of multivariate momentum vectors, a direct result of 3-dimensionality. What this indicates is that the symmetrical structure applied to physics is organized in such a way that *both these interpretations of the dualling process apply simultaneously*. In effect, this hidden balancing act also operates in yet other, more subtle ways because the virial relation between potential and kinetic energies is specifically one of doubling only when the force laws which apply are those characteristic of 3-dimensional space; and the action and reaction mechanisms which produce the doubled value for potential energy rely on applying vectorial (or dimensional) considerations to the kinetic energy term.

That the doubling mechanism also applies in purely mathematical, as well as in physical, contexts is evident from the topological explanation of the Aharanov-Bohm effect, though the physical and mathematical applications must ultimately have the same origin. Square-rooting and halving have an intimate relationship, which is manifested physically in the relation between vector spin terms of bosons and fermions and their respective uses of double or single nilpotent operators, in addition to the halving approximation used to find the kinetic energy term in the bionomial expansion for relativistic mass. This relationship is determined entirely by the fact that 3-D Pythagorean addition is a dualistic process, with a numerical doubling arising from noncommutativity, and this applies to both the vector operators used for space and momentum, and the quaternion operators used in the Dirac nilpotent.

There are also other possible mathematical connections. It is tempting, for instance, to believe that the uniqueness of the value ½ as the real part of the zero-solutions of the Riemann zeta function has a significance which is physical as well as mathematical, and that, as Hilbert originally conjectured, the solutions represent the eigenvalues and energy levels of an Hermitian operator, which is the Hamiltonian of a quantum mechanical system. It is conceivable that the ½ is related to the zero-point energy term of a series of fermionic harmonic oscillators. It is certainly true that, solving the Dirac nilpotent equation for any spherically-symmetric potential other than a linear or Coulomb one (i.e. under harmonic oscillator conditions) requires a Coulomb or phase term with numerical coefficient ½, which is of the opposite complexity to the rest of the potential, and which can be associated with the zero-point energy or (equivalently) the random directionality of the fermion spin. There may also be some physical significance in the fact that integers, like the fundamental parameters, only add directly to produce other integers or in the form of squares to produce squares of integers, but do so in an infinite progression. Both of these mathematical results suggest the possibility of further fundamental significance in the factor 2.

QUANTUM PHYSICS AND THE CLASSICAL TRANSITION

The definition of the Dirac nilpotent suggests that this is the most efficient way or parameterizing nature while ensuring its total ‘nothingness’. It may be possible to relate this to the aims of *topos* theory, in using a nilpotent Pythagorean structure to create a ‘parameter space’ which contains within itself dynamical and other physical possibilities. The uniqueness of the individual Dirac nilpotents, together with their necessary entanglement with each other and their infinite interaction with the vacuum, suggest that this is a real number space, with the numbers countable in the Robinson or Löwenheim-Skolem sense (see Appendix III). Through the Dirac equation, the nilpotents are then interpreted simultaneously in terms of conservation (the eigenvalue) and nonconservation (the operator). Quantum physics thus becomes a natural consequence of the fundamental meaning of conservation and nonconservation, and its separation from the physics of measurement (classical physics) becomes obvious. It is still necessary, however, to make sense of the classical transition, and also of the relationship between gravity and the other forces. Considerations of such ideas may also suggest the origin of the classical laws of thermodynamics.

Measurement processes are discrete, and involve discrete sources (or charges). They rely on the $SU(3) \times SU(2) \times U(1)$ symmetries which apply to these sources (and whose direct expressions are ‘interactions’, equivalent in principle to the action of classical field terms) producing restrictions on the freedom of the individual wavefunctions to contain infinitely possible variations in space and time coordinates.

A hypothetically isolated system (e.g. a hydrogen atom not interacting with other hydrogen atoms) must be purely quantum. Once we have any classical element or interaction the system is no longer isolated. This is how we make a measurement. We can’t make classical-type observations on an isolated system, otherwise it wouldn’t be isolated. An isolated system conserves E-$\mathbf{p}$-m within the system, linking it with the total $\boldsymbol{k}$, $\boldsymbol{i}$, $\boldsymbol{j}$ charge values, whether 0 or unit, positive or negative. This system must remain coherent – with angular momentum operators aligned, so that addition is effectively scalar, like that of the charge units. If the system interacts with an external system, then it can no longer be defined in an isolated way: the connection between the conservation laws for charge ($\boldsymbol{k}$, $\boldsymbol{i}$, $\boldsymbol{j}$) and angular momentum (E-$\mathbf{p}$-m) is broken. If the system is not isolated, then energy is not conserved *within* it, but some is lost to the ‘rest of the universe’ with which it interacts. Hence, we need the second law of thermodynamics, and, in fact, the first law (where the energy balance is only maintained globally by incorporating the ‘lost’ energy into the equation). The connection between the second law of thermodynamics and the direction of time is now apparent. To make a measurement requires a semi-classical situation with a non-isolated system; as soon as we make a measurement, we lose energy from the system to the ‘rest of the universe’, so increasing the ‘entropy’. The sequence of events behaves as an irreversible sequence because time itself is irreversible, because of its continuity, and a sequence of event ‘measurements’ must follow the same sequence; but this for any known pair of events will always require an increase in entropy.

CONCLUSION

Physics, mathematics and philosophy emerge together out of the basic idea of duality, though the concept applied is more fundamental than this name, with its connotations of a necessary discreteness, would imply. Mathematics is not something ‘applied’ to physics for ‘convenience’. It is, in fact, extremely *inconvenient*, as the mathematical laws of physics are general differential equations, which have to be reinterpreted ('solved', using different boundary conditions) every time a measurement is taken. Observation and theory, in physics, necessarily use incompatible types of mathematics because observation depends only on one member of the parameter group (space), while theory sets up the properties of the other members in opposition. So, mathematics is required only because it is a fundamental component of physics, and the structure of mathematics itself seems to suggest physical boundaries to the type of ideas which can be made mathematically useful (though obviously not in the form of a purely one-to-one correspondence). In fact, physics can become a kind of test of the ultimate value of mathematical structures at the fundamental level.

For example, physics appears to insist on the fact that all discrete quantities must be dimensional. This would not be required of a mathematical theory based, as most are, on the primacy of the integer series. However, if we begin mathematics with the integer series, then we have major problems in accommodating the reals – there is no natural progression – and, if we assume, like most axiomatists, that the most fundamental proposition in mathematics is 1 + 1 = 2, we will come up against the problems that Gödel identified with axiomatic theories or 'rigidly logical systems' which are intrinsically incomplete.

However, in physical terms, we may suppose that the integer series is not primary and that arithmetic, although the most psychologically familiar, is not the most fundamental branch of mathematics; and, further, that, the moment we assume that the number 1 (or even number at all) is the most basic concept in mathematics (or indeed in human thought), we have at the same time brought in a whole package of information that we will never be able to establish from first principles. Physics, in fact, tells us that integers and discrete numbering are not primary – they are associated with dimensionality – and dimensionality only has a meaning in the context of complexity. The integers are really a *codification* of a multiplicity of prior stages in mathematical evolution. To begin with them will necessarily produce an incompleteness in our logical procedures, with key steps appearing merely as assumptions in a circular argument. But, if we begin at the true primary stage, with a zero end product at every stage, we effectively remove the incompleteness in our axiomatization. We also reach a primary stage in which even the word 'dual' loses its meaning, although its convenience for the later stages makes it worth retaining if separated from its numerical associations.

The very applicability of the concept of 'duality' to the process of returning from 'something' to 'nothing' implies that the actual processes of counting and generating numbers are created, along, with ‘addition’, ‘squaring’, and other arithmetical procedures, at the same time as the categories of conjugation, complexification, and dimensionalization are separated from their dualistic counterparts. Defining the integers as an ordinal set within a much more fundamental process allows us to create new mathematical processes in which this ordinal set is applied in other ways, and so we can create types of mathematics where the relation to physical categories is less direct, but the ultimate ‘physical’ or ‘dual’ origin will remain.

From a purely physical point of view, the Dirac nilpotent would appear to be the perfect way of producing something from nothing; its structure also effectively incorporates or generates all the discrete and continuous groups of interest in fundamental physics, from C_2 to E_8 [Rowlands, Cullerne and Koberlein, 2001]; while the infinite imaging of the fermion state in the vacuum and the infinite entanglement of all nilpotent fermion states extends the dualling to infinity, as required. At the other end of the scale, the author, and collaborators, have shown, in many previous papers, how this concept applies to the structure of fundamental particles and the four fundamental physical interactions.

Appendix I Quaternions and multivariate vectors

Quaternions follow the multiplication rules:

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

If the quaternions are complexified we have:

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

Multivariate vectors follow exactly the same multiplication rules:

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

In effect, this means defining a 'full product' for two vectors **a** and **b** of the form

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

The rules for multivariate unit vectors are also exactly identical to those for Pauli matrices, and, through the additional cross term, immediately generate the concept of fermion spin.

Appendix II: Properties and antiproperties of space, time, mass and charge

The inextricable combination of properties and antiproperties means that neither an epistemological conception of 'reality' (we create it by our perception) nor an ontological one (it is 'out there' waiting for us to discover it) is meaningful. The division between epistemology and ontology can be based on the opposition of parameters susceptible or not susceptible to measurement (space versus the rest), or, alternatively, on the opposition nonconserved and conserved parameters (space and time versus mass and charge). In either case, the complete description requires both – 'nothing' is neither epistemological nor ontological. The properties and antiproperties also incorporate all the fundamental types of 'mathematical' number: positive, negative, integer, rational, algebraic, complex, transcendental, denumerable real, nondenumerable real, fixed and variable. They also include a concept of absolute 'uniqueness', which has not yet found its way into conventional mathematics, unless in the properties of the Dirac nilpotent algebra.

(a) Nonconserved / conserved

Physics structures itself by defining systems in which conserved quantities remain fixed while nonconserved quantities vary absolutely. Differential equations show the variation of space and time coordinates while retaining the fixed values of mass and charge, and the quantities which depend upon them: energy. momentum and angular momentum. Both nonconservation and conservation are absolute. The Dirac equation for a free fermion expresses this fact in its most convenient form. Quantum mechanics, in this form, is more explicable than classical mechanics, in that it fully expresses the nonconservation properties of space and time.

Conservation applies to all three types of charge. Lepton and baryon conservation are obvious manifestations, respectively, of pure weak and strong charge conservation, as is the non-decay of the proton. Conservation laws of mass and charge is also *local*, rather than global. Classically, each element of mass or charge has a permanent *identity*. Nonconservation is exactly opposite: space and time elements have no identity whatsoever. Hence, space and time have translation symmetry, with their elements specifically *stated to be indistinguishable* in physical equations. Three-dimensional space also has rotation symmetry; that is no identity for spatial *directions* or unique set of dimensions. The contrasting properties of mass and charge are 'translation asymmetry' (conservation of quantity), and 'rotation asymmetry' (electromagnetic, weak and strong charges independently conserved). Charge, unlike space, is conserved in both quantity and 'direction' (i.e. type). Weak, strong and electric charges are not interconvertible. The axes are fixed, along with the units.

The translation and rotation symmetries, of course, are identified by Noether's theorem with conserved quantities. Time and space translation symmetry are identified respectively with energy (E) and momentum (**p**) conservation, while space rotation symmetry becomes identical to the conservation of angular momentum (**J**). These three conservation laws can be identified further with the conservation laws of mass, value of charge and type of charge, and, in fact, the additional conserved quantities (E, **p**, **J**) can be seen as being 'created' at the same time as the application of the quaternion operators associated with the conserved *w*, *s*, *e* charges to the parameters time, space and mass produces the Dirac state. The conservation of charge type (*w*, *s*, *e*), alternatively rotation asymmetry or charge independence, manifests itself in the mutual independence of the three different aspects of angular momentum conservation (handedness, direction, and magnitude).

Previous work by the author and colleagues has shown that fundamental particles may be defined in terms of their *w*, *s*, *e* charges and rest mass, with the last determined ultimately from the charge structure. The conventional definition of a fundamental particle assumes an irreducible representation of the Poincaré group, or the group of space and time translations and rotations compatible with special relativistic invariance. Here, it can be seen that such translations and rotations are essentially identical to the conservation properties related to charge and rest mass which define a particle in the present theory.

Gauge invariance is a further demonstration of the absolute nonconservation of space and time, and, according to the Yang-Mills principle, is as local as all the principles of conservation. A system which is conservative in relation to mass, charge, energy, momentum or angular momentum, will remain so under arbitrary changes of the coordinates representing the nonconserved quantities, space and time.

(b) Real / imaginary

Pythagorean addition, or addition through squared values, is important to all the fundamental physical quantities. This is a consequence of their origins in quaternion and 4-vector representations. 4-vectors, with three real parts and one imaginary, are a familiar representation of Minkowski space-time. If the vector (or real) part is multivariate, then spin is automatically included. The three components of charge (say, *ie*, *js*, *kw*) can be considered as the 'dimensions' of a single charge parameter, with their squared values used in the calculation of forces added, in the same way as the three parts of space, by Pythagorean addition:

space-time	**i***x*	**j***y*	**k***z*	*it*
mass-charge	***i****e*	***j****s*	***k****w*	*m*

The opposing real and imaginary natures of space and time explain simply why identical masses attract, while identical charges (of any kind) repel, the coupling strength producing respective positive and negative values.

Real numbers can be privileged according to sign; imaginary ones cannot, and there are always simultaneous and equal status + and – solutions to consider. Thus, while real mass can be made unipolar (to ensure that it remains a continuum), imaginary charge always produces solutions of two signs, and 'antifermions' and 'antibosons' (with opposite signs of electromagnetic, strong and weak charge) have the same status as 'fermions' and 'bosons'. For the same reason, imaginary time has two possible signs, of equal status, in physical equations, though, as a continuous quantity, it has only one physical direction, and cannot be reversed.

Charges, as imaginary quantities, are only accessible through their squared values in interactions; mass, as a real quantity, is accessible in terms of its unsquared value (as inertia) as well as through its squared value (as gravitation). Time, as an imaginary quantity, is only accessible physically through its squared value (in acceleration), while space, like mass, is accessible through both its squared and unsquared values. Time 'measurement' always requires acceleration (because uniform velocity is imaginary), while space measurement is always direct. Time also is always the independent variable in physical equations, because we have no direct control over it, while space is the dependent variable.

(c) Countable / noncountable

Noncountable or continuous time and mass-energy have no origin. Continuous energy (vacuum) requires an infinite universe. Continuous time requires one without beginning or end. Continuous quantities cannot be reversed, because they have no origin. Mass becomes unipolar (with a single sign) while time is unidirectional. However, time as an imaginary quantities, has two *mathematical* solutions in equations. Hence, there are two directions of time *symmetry*, while there is only one direction of physical time. The unipolarity of mass is the reason why we have a CPT, rather than an MCPT, theorem, with C standing for charge conjugation, P for space reflection and T for time reversal, each of which has two mathematical sign options.

Only discrete quantities can be multidimensional, because multidimensionality requires origins, even if they aren't fixed (as in the case of space). Also, the discreteness of a quantity like space, with unfixed origins, is only possible through dimensionality. What we call 'measurement' in space requires discontinuity in both quantity and direction, and includes both reversals and changes in orientation. This is why what we call 'measurement' takes place only through space. 'Time'-measuring devices all rely on some kind of repetition of a spatial interval. Special conditions, relying on spatial reversibility and dimensionality, have to be used to set up such measurements, although anything which can be perceived at all can be used to measure space, at any time.

There are two definitions of real numbers in mathematics. In Robinson's non-standard analysis, Skolem's non-standard arithmetic, and non-Archimedean geometry, the reals are denumerable. The Löwenheim-Skolem theorem is significant here, in requiring any consistent finite, formal theory to have a denumerable model, with the elements of its domain in a one-to-one correspondence with the positive integers. In the Cantorian definition, which is related to the standard versions of analysis and geometry, they are a non-denumerable continuum. Essentially, the first definition reflects the properties of space, while the second accords with the properties of time (or mass). Both systems yield identical results, because differentiation is a property linked to nonconservation, and not concerned, in principle, with the difference between absolute continuity and indefinite divisibility. Zeno's paradoxes are solved only by assuming a divisible (or digital) space and a non-divisible (or analogue) time. The solution by limits can only be used if we adopt the time-like, or standard, version of differentiability.

Mathematicians have been unable to decide whether mathematics is ultimately continuous or divisible. This is because both are physical options. The indefinite divisibility of space, though sometimes wrongly termed ‘continuity’, is a different attribute from the absolute continuity of time. Space’s indefinite elasticity, its ‘continual’ recountability, or unending divisibility, are ways of expressing its *nonconserved* nature, the nonfixed nature of its units, but infinite divisibility is, mathematically, the very antithesis of absolute continuity, and the whole process of measurement would be impossible if space, unlike time, were not divisible in this way. Continuity, again, is not related to differentiability – the confusion here arises from the fact that one of the differentiable quantities, time, *is* continuous, and that the continuity of time is a significant contributor to our psychological perception of states of change. Differentiability is, again, a manifestation of nonconservation, a wholly separate physical category, and is equally valid in both discrete and continuous, and in classical and quantum contexts, though it is obviously not valid where the discreteness is fixed (which is where much of our psychologically-based notion of discreteness comes from).

Wave-particle duality and the opposing Schrödinger and Heisenberg versions of quantum mechanics are the result of adopting predominantly continuous or discrete options for physical quantities. Neither is actually physically possible, and each incorporates the alternative in the process of ‘measurement’. The non-quantum-field version of the Dirac theory partially overcomes this by including terms equivalent to those in the dual group, with aspects of the real and imaginary nature of space, time, mass and charge reversed. The nilpotent or quantum field version of the Dirac theory, however, incorporates both discrete and continuous options as duals within its mathematical structure, and no longer requires explicit use of the dual group.

Hamilton was correct, in 1843, in seeing quaternions as being responsible for the three-dimensionality of space, as quaternions introduce the concept of three-dimensionality which vectors subsequently adopt. They also occur one stage earlier than vectors in the evolution of physical concepts via duality. It is significant that the quantized parameters (E, $\mathbf{p}$, m, and, collectively, $\mathbf{J}$) emerge from the application of the charge quaternions to the originally non-quantized space, time and mass. Although quantization may thus appear to be equivalent to converting these three parameters into discrete forms, the discrete versions should be seen rather as *composites*, with which the discrete charge structures are inextricably linked.

(d) The construction of physical laws

The parameter group leads naturally to laws of physics, based on the explicit specification of what is conserved and what is nonconserved. The construction of these laws relies on the fact that every statement about conservation is simultaneously a statement about nonconservation. To relate the conserved and nonconserved aspects, we use the scaling relations between the parameters, and their dual inverses, which are established at the moment that we compress the eight units of the independent parameters into the five of the combined Dirac algebra, and create composite parameters with characteristic aspects of each. The most significant of these are the ones which relate mass and charge, respectively, to time and space, and which can also be shown to be conserved parameters: namely, energy, linear momentum and angular momentum. These are quantities of the same mathematical form as time and space, which are conserved in exactly the same way as those quantities are nonconserved, and form conjugate pairs with them for exchanging statements about conservation into statements about nonconservation and vice versa. Both classical and quantum laws can now be constructed in terms of these conjugate pairs (e.g. via Poisson brackets).

Energy is a pseudoscalar and is conserved in quantity and individual element (i.e. is translation asymmetric) in precisely the same way as the pseudoscalar parameter time is not conserved in quantity and individual element (i.e. is translation symmetric). It may be regarded as the link between time and the real scalar quantity mass (the gravitational source), as the conservation of energy is directly linked to the conservation of mass. Linear and angular momentum are, respectively, vector and pseudovector, and are conserved in quantity and individual element (i.e. are translation and rotation asymmetric) in precisely the same way that the vector parameter space is not conserved in quantity and individual element (i.e. is translation and rotation symmetric). They may be regarded as the respective links between space and the quantitative values and the quaternion operators applied to the different charge types.

Energy and momentum are regarded as a pure 4-vector in special relativity, but, in a quantum system, this is not strictly true, because quantum (Dirac) energy and momentum terms are only fully represented mathematically when each has a (different) quaternion operator applied to its respective pseudoscalar or vector. When taking the invariant scalar product, of course, these operators disappear, but their significance becomes apparent when we introduce a third term, with yet another quaternion operator (rest mass) to convert them into a nilpotent. A pure 4-

vector could not be made into a nilpotent in this way. If we take the Dirac differential operator (∂^μ), we can see that this, also, is not a pure 4-vector, and the same must apply, in general, to relativistic time and space, whether the system is quantum or classical. The nilpotent is completed with a third term, which occupies the same position as rest mass does in completing the differential operator nilpotent. This term is a real scalar, like rest mass, and its squared value must exactly cancel the scalar product of the time and space components. The fact that it can never be negative means that only retarded solutions are possible for the space-time combination. It also means that, numerically, this new term (the 'proper time', τ) is equivalent to the time value with a zero space component. 'Proper time', however, is, strictly, a rest mass-related, rather than time-related concept. Its validity in both classical and quantum contexts is an indication that the link between these two domains is essentially through the scalar additive nature of the rest mass or 'inertia' of the component systems.

Appendix III: Algebraic representations of the parameter group

A simple algebraic representation of the parameter group can be accomplished by representing the properties of space (real, nonconserved, divisible) by, say, x, y, z, with the respective antiproperties (imaginary, conserved, indivisible) represented by $-x$, $-y$, $-z$. The group now becomes:

space	x	y	z
time	$-x$	y	$-z$
mass	x	$-y$	$-z$
charge	$-x$	$-y$	z

With group multiplication rules of the form:

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

and similarly for y and z, we can establish the standard D_2 group multiplication table with space as identity element, and each element its own inverse (the duality of space-time elements and their inverses is, interestingly, a feature of string theory). However, we could just as easily have chosen mass as the identity element by representing the properties and antiproperties:

space	x	$-y$	$-z$
time	$-x$	$-y$	z
mass	x	y	z
charge	$-x$	y	$-z$

and the same applies to time and charge. The ‘multiplication’ rule here only concerns the signs, but the creation of a Dirac nilpotent incorporating both the space-time 4-vector and the mass-charge quaternion suggests that there must also be a direct multiplication rule between the units of the parameters and those of their inverses, which is exactly what we need to require the existence of the fundamental constants *G*, *h* and *c*.

REFERENCES

Atkins, P. [1994], *Creation Revisited*, Harmondsworth, Penguin, p. 23.

Rowlands, P. [1983], The fundamental parameters of physics. *Speculat. Sci. Tech*., 6, 69-80, 1983.

Rowlands, P. [1999], *The Fundamental Parameters of Physics: An Approach Towards a Unified Theory*, PD Publications, Liverpool.

Rowlands, P. [1999], Physics; let’s get down to basics, in K. Bowden (ed.), *Aspects II*, 1999 (*Proceedings of XX ANPA Conference*, Cambridge, September 1998), 123-134.

Rowlands, P. [2001], A foundational approach to physics, arXiv:physics/0106054.

Rowlands, P. [2002], ‘The physical significance of the factor 2’, ANPA Philosophical Papers, presented August 2001.

Rowlands, P. and Diaz, B. [2002], A universal alphabet and rewrite system, arXiv:cs.OH/0209026.

Rowlands, P., Cullerne, J. P., and Koberlein, B. D. [2001], ‘The group structure basis of a foundational approach to physics’, arXiv:physics/0110091.

Young, N. [1988], *An Introduction to Hilbert Space*, Cambridge University Press, p. 59.