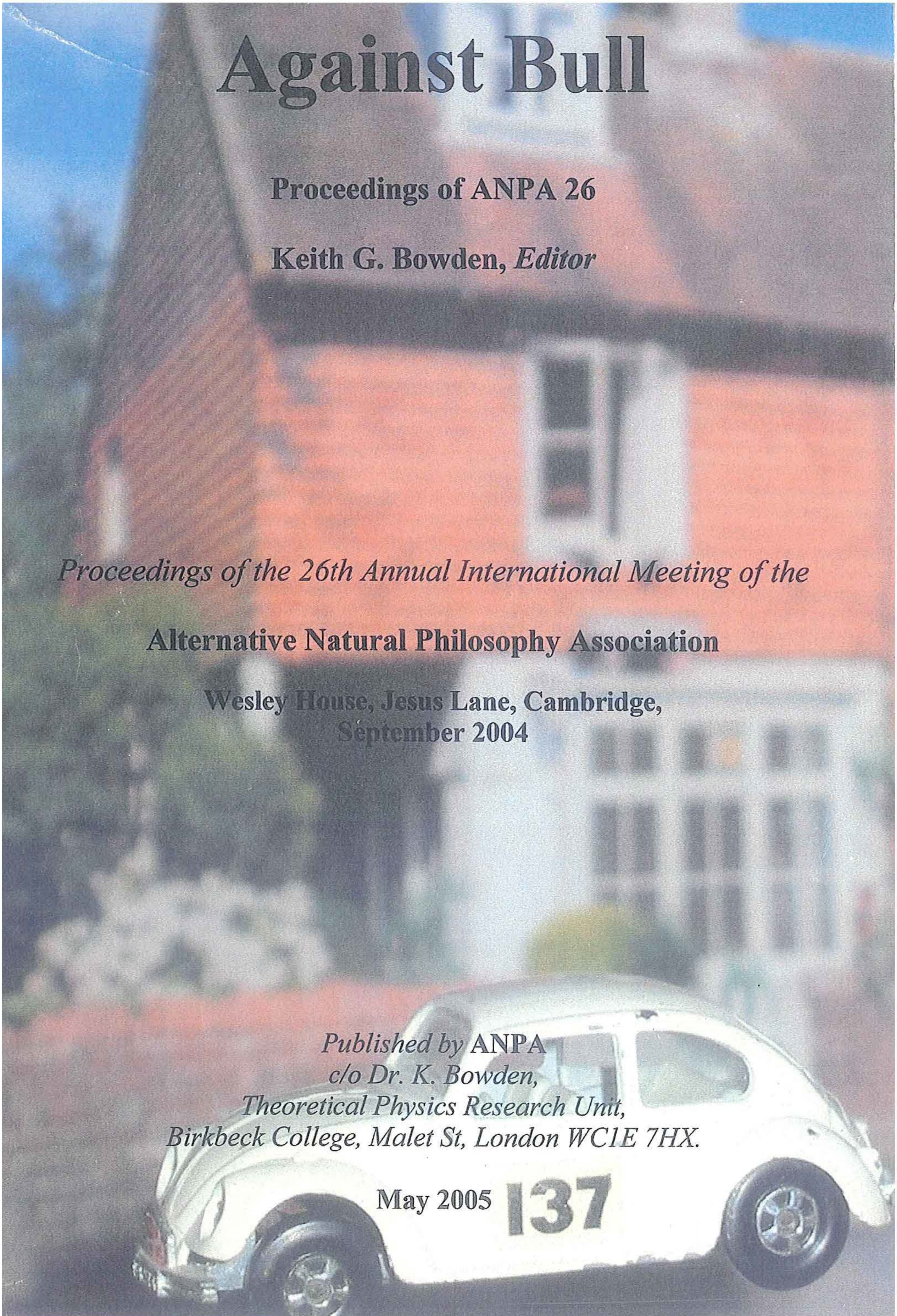


Against Bull

Proceedings of ANPA 26

Keith G. Bowden, *Editor*

Proceedings of the 26th Annual International Meeting of the

Alternative Natural Philosophy Association

Wesley House, Jesus Lane, Cambridge,
September 2004

Published by ANPA
c/o Dr. K. Bowden,
Theoretical Physics Research Unit,
Birkbeck College, Malet St, London WC1E 7HX.

May 2005

Against Bull

Proceedings of ANPA 26

K. G. Bowden, *Editor*

The Alternative Natural Philosophy Association

Against Bull: Proceedings of ANPA 26/Keith G. Bowden, *Editor*

The Proceedings no longer has an ISBN

published by ANPA,
c/o Dr. Keith G. Bowden,
139 Sandringham Rd, Barking,
Essex IG11 9AH, UK.

Contents

HISTORICAL ANNIVERSARY TALKS

In Remembrance of Faruq

The Dalai Lama once said that one's physical body reflects one's preceding incarnation, and one's mind the next. Faruq Abdullah, son of an Imam, surely reflected such a high birth. Besides his non-pareil care of his post as ANPA's secretary for many years, he was also ANPA's champion for investigating alternative medicine and biological connections. He also read Krishnamurti closely, and he often started conversations with me on the subject of suffering. Often. Only gradually did I learn that he had embarked on a years-long campaign to cure his wife of her long-standing schizophrenia. Only then did I finally understand who he was. It was eventually decided that it was best that she return to her family in Pakistan, but, as the very nature of love, and healing, is to take on another's burden, so, unwittingly, had Faruq. It proved to be heavy indeed, and I was shocked to learn, a few years later, that he lay in hospital, his body smashed and the status of his mind uncertain, the result of an unsuccessful leap from height. The tragedy brought us all to stunned silence - that is the nature of tragedy. It is hard to break that silence even now, even in grief. For myself, the news that he was 'active and in good spirits' when he died, after the several years in the shadows, was a great and welcome balm: he went to his Creator, not in the throes of despair, but rather in peace. Thus may he inherit the fruit of his heritage and compassion. Praise be to Allah for this mercy.

-Mike Manthey

Editorial

Against Bull v4.1

Theoretical Physicists consider Physics to be the art (or science) of using models to predict the results of "certain experimental situations". They consider anything else to be Philosophy or Metaphysics. They are not really interested in "reality", except in some specialised technical senses. They (and Philosophers of Physics) consider this to be the remit of the Philosophers of Physics.

If they have a model that works they will not be interested in changing it for a new one unless it gives better predictions, or it is more elegant, accurate or consistent than theirs, in a serious way or, as James Lindesay said in a recent meeting, "If it helps me remember (experiments) better!"

The reasons for this are to do with the difficulty of assigning a meaningful mechanistic, or “realistic” interpretation to the statistical predictions of Quantum Mechanics. A natural conclusion might be that Quantum Mechanics must be wrong. The trouble with this is that Quantum Mechanics is the single most successful theory that we have by a number of orders of magnitude, and has been for a very long time.

For example, consider the phenomenon known as Quantum Correlation. It is a fact, both of Relativity and (seems to be) of our observations of the world, that nothing, including information, can travel faster than light. However it is also a fact, both of Quantum Mechanics and seemingly of our observations of the world, that two spatially (outside the light cone) separated “correlated particles”, made simultaneously under special conditions, are in contact (or “superluminally connected”) instantaneously, in exactly such a way that their joint statistics show instantaneous correlations, but that these statistics cannot be used to send messages faster than light. One can understand intuitively that, although the statistical correlations are instantaneous, the randomness “cancels out” the ability to communicate superluminally.

Consider two (massive) correlated particles a light year apart. At each end an observer takes a regular sequence of appropriate measurements of the particles and communicates them to the other by a light beam. The measurements will be instantaneously correlated (that is, without a time delay). The information arriving in the light beam at either end will always be a year out of date compared to the direct measurement of the local particle. Quantum connections cannot be explained using light mediated ones.

It is an **ASTONISHING** ***mathematical*** fact that it is possible to deform the sensible classical Newtonian-Maxwellian model of the world in such a way that it is possible to create instantaneous statistical correlations between two remote particles, but that we cannot use the correlations to send messages (information). Such correlations are called **noncausal, nonlocal** correlations. A Relativity Theorist might say that the connection has **negative proper time** (or more correctly, proper time *squared,* although both usages are common).

It is the **actual observation**, in the real world, of such Quantum Mechanical predictions (by Alain Aspect in 1982 and many others since), that has given rise to the situation in modern physics today. Once one has realised how astonishing it is that all this is mathematically possible, it is not quite so astonishing to discover that this is actually the world that we live in. But ***it is astonishing***. Anyone who does not realise this does not understand the single main point of Quantum Mechanics today.

It is these noncausal, nonlocal, outside the light cone connections that have given rise to recent technologies such as Quantum Computation, Quantum Teleportation, Quantum Privacy and Quantum Key distribution. To think of Quantum Computation in terms of negative proper time connections may make it easier to understand how Quantum Computation can be faster than Classical Computation, but it is still astonishing. It was just in order to understand this nonlocality that Interpretations of Quantum Mechanics were devised. It is now fifty few years since Bohm's "Hidden Variables" paper was published – it is certain that he later regretted this title – and exactly fifty years since Everett's Relative State Interpretation, but it is remarkable how little has changed in the intervening period, except for the deeper understanding of nonlocality and quantum connections that Quantum Computation has enabled.

I observed recently that there were four stages to the origins of the situation.

Pre Quantum-Mechanics there were Newtonian Mechanics and Maxwellian Field Theory which were generally associated with "reality" although Newton himself was not necessarily of this view regarding, for instance, the reality of the gravitational field. Conversely Maxwell took a fairly mechanical view which was later replaced by field theory as we know it today.

Post Quantum-Mechanics there was a period of confusion until the 1927 Solvay conference and the **Copenhagen Interpretation** (promoted by Neils Bohr) which pronounced Quantum Mechanics (and by extension all of Physics, and in the extreme view all of Science) to be an algorithm for the prediction of the results of certain experimental situations, and ***nothing more.*** Von Neumann even claimed that realistic models of reality were provably not possible. This situation maintained for 24 years.

Such statements are, however, disprovable by example and David Bohm did just this in 1952 with the Quantum Potential model, which became the **Bohm-Hiley Interpretation**, which is a realistic model of the world, albeit non-local. This marked the beginning of the third period in which a return to a Newtonian type position would have been possible.

This period ended in 1957 with H Everett III's **Relative State Interpretation** (similar to the Many Worlds). Now there were TWO realistic models, and one had to choose between them. Neither had the force of being a unique paradigm. Indeed there are now a plethora of them (for instance Consistent Histories), all mathematically equivalent, but with different Philosophical implications. It may well have been this fact, and the very short period (five years) in which a return to a unique mechanistic thinking had been possible, that led to the current situation in which the Copenhagen Interpretation is orthodox thinking.

(Someone commented to me recently that Physicists are programmed and not taught. The problem is that this is what you do with an algorithm. Program it.)

The Copenhagen Interpretation should not be confused with Logical Positivism, it is a much more complex concept! Indeed part of the algorithm of Quantum Mechanics is, ***perversely,*** that we must discuss quantum phenomena using the language of classical mechanics. So that when a Physicist says that a particle undergoes a particular process, he may not mean that the particle in some sense really exists, he is just cranking the algorithm of Quantum Mechanics. Bohr may be indirectly responsible for some of the confusing and confused thinking around today, but he was still a genius!

It is the perversity and positivist aspects of the Copenhagen Interpretation that leads many people to criticise this approach and search for a sensible mechanical picture of the world that they can understand once again. In many cases the search appears to be rather like tilting at windmills in that the Bohm interpretation was the first of many well understood and relatively satisfactory ontological models that already exist, although this fact is perhaps not well enough known.

Finally, let me say that all I have said in the above paper is bog standard Quantum Mechanics as taught to me in Undergraduate courses at the University of Sheffield in the early seventies, and as told to me in meetings with members of the ANPA Advisory Board, the ANPA Executive Council and the Sigma Club over the last twenty years or so. For those people who do not know the Sigma Club, it consists of the Philosophy of Physics groups from the Universities of Oxford, Cambridge and London (Birkbeck, Imperial and University Colleges and LSE). As such the Sigma Club is undoubtedly one of the most influential and knowledgeable Philosophy of Physics groups in the world. Most of the Heads of the Departments involved are or have also been members of the ANPA Advisory Board. For all these things I am very grateful to them, and to their ceaseless fight Against Bull I dedicate this volume.

The title of the next volume of the ANPA Proceedings will be "On Goop". From then on I intend that the titles will better reflect the creativity, expertise and openmindedness of the group.

Keith Bowden, *Birkbeck College, University of London,* January 2005

Reference

J S Bell, Speakable and Unspeakable in Quantum Mechanics, Cambridge, 1987.

ANPA Proceedings Editorial Policy

ANPA has been criticised in the past - in particular by members of its own Advisory Board - for having no formal editorial policy for its Proceedings. This has been balanced by a feeling within ANPA that we should keep ourselves open to all viewpoints. In the last few years as editor I have tried to tighten things up in such a way as I felt would satisfy our critics whilst not compromising our own position. This has been partially successful although for some time I have felt that it is time that there was a formally stated policy. The following has been approved by the Executive Council, although it is open to feedback from all. By "the editor" is meant the Editor or (an) appropriate nominated Referee(s) (note the capital R!)

1. The paper should make a new and original contribution to the fields of ANPA's interest. Survey papers are acceptable.

2. The default use of language for submitted papers in Physics {and Philosophy of Physics}* should be the common language of Physics as usually understood by Physicists {and, in particular, by Philosophers of Physics}*. Any other use of language should be carefully explained at the start of the paper and all appropriate definitions included there.
{* added by KGB}

3. The editor should be satisfied that the paper is ***presented*** in such a way that the majority of the readership will understand the author's intentions. In particular ***it should be clear*** that the author has a correct understanding of the subject matter. Please avoid forward references.

4. "Verbatim" reports will be accepted subject to the above three conditions only, regardless of whether the final draft is an accurate rendition of what was originally said.

5. Theories of any nature are acceptable material, provided they are compatible with the known facts, and provided they are deemed to be of interest to the readership. Theories of alternative, imaginary worlds are also acceptable, provided their nature is made clear.

ANPA Proceedings Notes for Authors

I would like to try to continue conformity of *style* for future issues of the Proceedings. Ideally I would like contributions to be submitted in International Journal of General Systems format (I have some copies of their Notes for Authors) or similar.

Times Roman, 14 point, is compulsory. **10 point is TOO SMALL to be reduced to A5.** If you are sending hard copy please send two copies single sided. Main heading 20 point capitalised and centred, other headings 16 point capitalised to the left. Author's name(s) capitalised and centred. Address italicised and centred. No underlining. At least a one inch bottom margin for footers; page numbers NOT top centre. *Only copy in good English will be considered, and remember, this is a formal Proceedings.* **Remember also to include your name (surprising how many people omit this!), affiliation and full address, email address and the version number (even if it is 1.0) or date of the draft ON THE COPY, centred below the main heading, or in 8 point font (tiny!) after it.** I often get sent more than one version of a paper and invariably mix them up! Send copy **to *KEITH BOWDEN, 139 SANDRINGHAM RD, BARKING, ESSEX IG11 9AH.***

The copy date for the ANPA2005 Proceedings is January 1st 2006. The issue will go to print on April 1st 2006. This will be adhered to rigidly this year.

Keith Bowden,
Theoretical Physics Research Unit,
Birkbeck College,
Malet St,
London WC1E 7HX, UK.

k.bowden@physics.bbk.ac.uk Tel: 0208-594-5064

FOR EDITORIAL ADDRESS SEE THE PENULTIMATE PARAGRAPH OF THE TEXT

THE MATHEMATICS INVOLVED IN CALCULATING THE FINE STRUCTURE CONSTANT

CLIVE KILMISTER

Red Tiles Cottage

High Street

Barcombe

Lewes BN8 5DH

ABSTRACT. I correct my previous paper which calculated the inverse fine structure constant to be 137·036012.. Part 1 (which was not given at the meeting) corrects some algebraic mistakes. These corrections do not change the numerical results. In part 2 I take up the arithmetical details and report on the checking of the calculation in the past year, which now gives a corrected value of 137·036011...

PART 1

In (1) I showed that a correct understanding of Parker-Rhodes' construction which gave rise to an estimate of the inverse fine structure constant as 137 modifies the value to 137·036012... in complete agreement with the latest measured values. The later part of (1) contains algebraic errors. These are worth correcting although they do not affect the numerical value of α . They arise on pp.10-12 over the two alternative techniques used to close an open discrimination system. I have confused them. On p.10 I used Q* to denote the quaternion-like but non-associative loop got by taking (in the notation of (1)) z = 1, y = -1. This is the definition of Q* used to tabulate the next level on p.12 so that table should follow it directly. I went on to "tame the non-associativity" in a different way.

I recognised that the open part of Q* (that is, the set of values of discriminations between elements, not signals) is the open part of one of Arleta's grouplets, being derived from quaternions Q by the rule

$$a\cdot b = b^{-1}a$$

where the left-hand side is the product in Q* and the right-hand that in Q. Once I have chosen this grouplet on p.11 it gives a different table from that on p.12. The reason for this is the behaviour of y, z. Evidently, since in Q z = -1, y = 1 we have

$$a\cdot z = -a, \quad z\cdot a = -a^{-1} = a,$$

$$a\cdot y = a, \quad y\cdot a = -a$$

with $y\cdot z = z\cdot y = y$.

There is a left-unit, z, and a right-unit, y but no two-way unit. The grouplet is not a loop. It is not too hard to re-calculate the table on p.12. I continue to use the k_r notation although it has less pay-off than before when these elements were generators of the centre. It is also convenient to use I = (z,z,z) as well, so that, for example, -D = ID. The table is:

	132' A	321' B	123 C	2'1'3' D	z11 E	2z2 F	33y G
A	I	D	IE	IB	C	G	k_1F
B	ID	I	F	A	k_2G	Ck_2	E
C	E	IF	I	k_1G	Ak_1	B	k_2D
D	B	IA	k_2G	I	k_2F	k_2E	k_1C
E	IC	k_1G	k_1A	Fk_2	I	Dk_1	Bk_3
F	IG	k_2C	IE	Ek_2	k_1D	I	k_2A
G	k_3F	IE	Dk_2	Ck_1	k_3B	Ak_2	I

The structure of the seven 3-element sub-systems is now a little more complicated but does not concern me here.

PART 2.

This second part has four objectives. Firstly, if I am to persuade people outside ANPA that Frederick's construction inevitably leads to $1/\alpha = 137 \cdot 036011..$ (a correction in the sixth decimal place from last year) they will want to check the details. Secondly, in previous presentations through the years I have talked a lot about the ideas - process - the need for a philosophical understanding. As a result the arithmetical details have been glossed over. Thirdly, Keith Bowden generously suggested some form of computing help if I could say clearly what was needed. This paper shows that once one sees clearly what is needed the rest is so easy as not to be worth programming. Lastly, the calculations throw up two purely mathematical problems; I came to the meeting with these as puzzles but Lou Kauffman showed me the solutions, which I include here.

In (1) pp.1-6 I explained the early stages of the calculation. A table is then given on p.7 which comes from (2). There is inadequate explanation of this table in either of those references. I want you to be quite clear about what is happening here. I'll take one line of the table as an example:

	1 a b	ce 2f 3d	1ce 2af 3bd
11. g h i		168 (5)	112 (6)

To explain this more fully: firstly, discrimination (+) has been omitted, so that pq means p + q. The three elements of the next level corresponding to one-element dcss are taken as (1,a,b), (c,2,d), (e,f,3) where a,b,...f are column vectors over $\mathbb{Z}_2$ and are drawn from the set [1,2,12,3,13,23,123]. Concentrate on $11 \cdot 2$ (that is the entry in the second column).

Where did this come from? From the table one can always work backwards. The seven basis elements at the next level are the three mentioned above together with (g,2,3), (1,h,3), (1,2,i) and (1,2,3). To make matters quite clear, the matrix (1,12,23) is $\begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\\0 & 0 & 1\end{bmatrix}$. The entry in the table says that some discrimination or other has finished with a zero. Obviously the first element having ceg = 0 suggests

(c,2,d) + (e,f,3) + (g,2,3) = (ceg,f,d).

But h, i are involved and we note that

(1,h,3) + (1,2,i) = (0,2h,3i)

so evidently the zero is

(c,2,d) + (e,f,3) + (g,2,3) + (1,h,3) + (1,2,i),

the sum of 5 basis elements and that is the reason for the "(5)" in the table.

A priori a,b ... can have 7 values but not all combinations are permitted. For evidently a $\neq$ 1 or the matrix would be singular and a $\neq$ 2 or the matrix would have more eigenvectors than 1. So a,b,c,d,e,f sit in a 5-dimensional space. The five components of a, a_r say, are then

(a_r) = (3,12,13,23,123).

The order of these has been chosen arbitrarily but that is the end of the arbitrariness. The next important idea is that there is a symmetry group (which was not mentioned in (2)). This is generated by (i) interchanging 1 and 2, which I denote by $\binom{1}{2}$ (ii)similarly for $\binom{1}{3}$, and (iii) for $\binom{2}{3}$. Of course the whole group is of order 6 since it is the symmetric group on 3 symbols. Now $\binom{1}{2}$ turns (1,a,b) into (a',2,b') where the dashes denote the new values resulting from the interchange. So I am inclined to say that $\binom{1}{2}$ turns a into c and b into d. But this has to be understood in this sense: if, for instance, a = 13, then it is turned into c = 23. Such qualifications can be put on one side if one first uses the symmetry group to generate the corresponding components of b,c,..f from those of a. Starting with the values of a_r above and applying $\binom{2}{3}$ one finds:

$$b_r = (2,13,12,23,123)$$

and so on. Although a can have any one of 5 values and so for b, (a,b) has, not 25, but 14, because of the conditions of a non-singular matrix with only one eigenvector. It is easy, if a bit tedious, to work out which they are and the results can be expressed in the form of the following symmetric matrix A (for "allowed"):

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix} .$$

We can write this in abbreviated notation used before as

(345,45,145,123,123)

(stating which rows are filled with a 1 in each column). But there is a difference from the earlier use of this notation for the 1's in this matrix will be treated not as bits but a natural numbers. The way in which this matrix is to be interpreted is: for a_1 (i.e. a = 3) the possible b_r have r = 3, 4, 5 (i.e. b = 12, 23, 123)

Now for the details of how to calculate the table in (1). If a choice of the 7 matrices at the next level is not a linearly independent one, then some sum of them must come to zero. It is explained in (2) why the triad ones cannot do this by themselves, so it must be the sum of the unary ones and a possible triad and possibly the identity. The case of 11·2 set out above is an example. What is needed to calculate α is how many of each of these zeros there are. Most of the numbers in the printed table were originally found by me by just counting up the possibilities, a crude process which could, and did, lead to error. Now what I am doing here is to count up the number of solutions of certain sets of equations, over $\mathbb{Z}_2$. In the case of 11·2, the three equations are:

$$ce = g, \quad f = 2h, \quad d = 3i.$$

Since a, b are notinvolved, the 14 cases of (a,b) are free and so we can expect to find, not the 168 zeros mentioned in the table, but 168/14 = 12. In this case, then and in many others but not unfortunately all, the labour is much reduced.

To tackle this counting, I consider graphs associated with each set of equations. These correspond to a way in which one could go about solving the equations. For example, for 11·2 one could start with any d, which (from the matrix A) would give some c's; these c's determine some e's (from the first equation) and each of those e's will give f's (using A again) which must then be consistent with the second equation. The graph would be:

d —┬— c —┬— e —┬— f —— (h)
    A   (g)   A

One will need more matrices for the other relations. For example, for ce = g it is easy to find (in abbreviated notation)

R = (45,3,24,13,1)

Similarly one can find a matrix Q which encapsulates d = 3i or equally f = 2h. A better version of the graph is then:

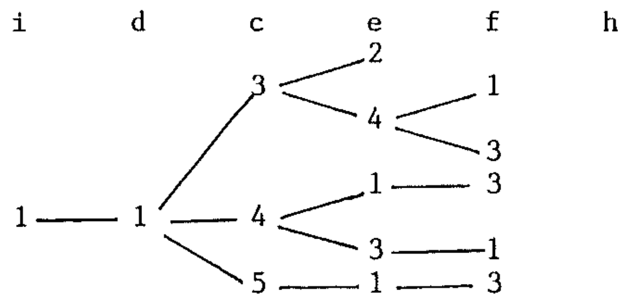
$$i \xrightarrow{Q} d \xrightarrow{A} c \xrightarrow{R} e \xrightarrow{A} f \xrightarrow{Q} h$$

This shows i being turned into h by QARAQ, remembering that in this product the 1's are natural numbers, not bits. Then

$$QARAQ = \begin{bmatrix} 2 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

I would like you to bear this in mind while I show you another way of getting the result.

I first set out the general idea, and then how it may be made more systematic. The various possibilities are set out like this:



Here I have shortened things a little; I have entered nothing under h but have noted that f has to be 1 or 3 because of the second equation. Also this is only half the story; another tree grows from taking i = 123, that is d = 12 or, in the suffix notation used here d is 3. But notice that the five end-products are 2 1's and 3 3's; and you will find that the other tree gives 3 1's and 4 3's. So you can see that not only does the sum of the entries in QARAQ give the required total, 12, but the entries in each column give the number of the appropriate element.

After several false starts I have devised a tableau method of setting this out systematically.

DEFINITIONS. A tableau is a rectangular array of up to six columns of sites, and an indefinite number of rows. The sites in the left-hand column are called initial and those in the right-hand column are called final. A site may be filled (by one of the suffixes 1,2,..5) or empty. Every non-final site has an immediate successor (the next site to its right) and other successors (the sites in the next column to its right). If t is a successor of s, the s is a predecessor of t. Between each pair of columns is a square 5x5 matrix of 0's and 1's which relates the suffix filling a site with those allowed suffixes that can fill its successor sites. Any suffix can fill initial sites. A site may be active or passive, defined in this way:

(a) A filled site is passive;

(B) a site is active if and only if it is not the case that its immediate successor is passive.

RULES FOR FILLING SITES.

Initially all sites are empty (and so active).

1. An active initial site is to be filled by the next unused suffix.

2. When any non-final site has been filled, the next step is to fill its highest active successor with the next allowed suffix. If all such allowed suffixes have already been used, then the original site is called saturated (in writing out the tableau, I usually underline the suffix in a saturated site).

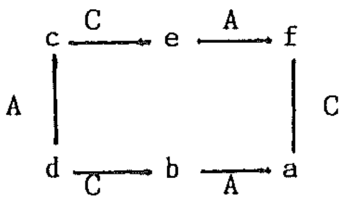
3. When a final site has been filled, the next step is to return to its predecessor and repeat 2. If the predecessor is saturated, then return to its predecessor and repeat 2, and so on. The tableau for 11·2 is as follows:

d	A	c	R	e	A	f
1		3		2		–
				4	1	
						3
		4		1		3
				3		1
		5		1		3
3		1		4		1
						3
				5		1
						3
		4		1		3
				3		1
		5		1		3

You will observe the 2 + 3 1's and 3 + 4 3's as before. So much for the calculation when the graph is a simple linear one. I shall skip the details for the other case when the graph closes on itself. For example, 0·3 in the table has the three equations:

$$ce = 1, \quad af = 2 \quad bd = 3$$

and the graph will then have the form:



where C is the matrix (2,1,5,0,3). Now starting with c say the matrix $(AC)^3$ must return one to c again

This means that we are concerned with the trace; in fact,

$$(AC)^3 = \begin{bmatrix} 1 & 2 & 4 & 0 & 2 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 2 & 4 & 4 & 0 & 3 \\ 2 & 4 & 4 & 0 & 3 \end{bmatrix}$$

and a tableau does indeed give a 1, a 2, 3 3's and 3 5's, 8 in all. So the total number is the trace 8. As I said before, I came to the meeting puzzled about how to prove these methods worked, although I was sure they did. But Lou Kauffman showed me that what is being used here is just the incidence matrix of a graph. If the vertices of a graph are numbered off 1,2,.. the incidence matrix A is defined by $A_{ij} = 1$ if there is an edge joining i and j and 0 otherwise. Then it is well known and easy to prove that $(A^n)_{ij}$ is the number of paths of length n from i to j.

Now you may be wondering how all this pays off because (as was already said in (2)) you just improve 137·033.. or David

McGoveran's 137·037.. to 137·0351.. This is still way out from the experimental value. Is the game worth the candle? An affirmative answer is provided in (1), section 6 where the difference produced by realising that discrimination need not be commutative is exploited to alter the probabilities. As I showed there the answer is changed to 137·036012..(as is stated in (1)). But in fact the careful use of the techniques of this paper has allowed me to correct the arithmetic so as to give a definitive value of 137·036011...

REFERENCES.

(1) C.Kilmister. The fine structure constant in Spin, Proceedings of ANPA 25 edited K.G.Bowden, 2004.

(2) T.Bastin, C.Kilmister. Combinatorial Physics. Singapore, World Scienfic, 1995.

PROCESS

TED BASTIN

Maesllwyn, Tan y Groes, Cardigan SA43 2JF

edward@bastinsooo1.fsnet.co.uk

ABSTRACT

It is shown that the hierarchy theory, and in particular the experimental successes with the scheme of coupling constants, depends on understanding the construction in *process* terms. By this we mean that the scheme of levels is constructed iteratively; the point of reference moves continually through the levels and back. This idea originated in an early hierarchical model giving the coupling constants known as "Program Universe". The observer is not separate from this process, however the way in which the discoveries made through the iteration are recorded so that a publicly available space comes into being, is not dealt with in this paper beyond assuming a background which presents new entities to the basic construction. This background may plays a similar role to the vacuum of the standard model of particles. A parallel development of the process philosophy due to Cahill is briefly discussed.
Conclusion: if one accepts the evidence of the constants' calculations then one is forced to an iterative process view.

1. PROCESS AS A NECESSARY PRINCIPLE

I start from Parker-Rhodes' construction. He based it upon the need to express algebraically the construction of successive new sets of entities out of the operations upon the elements of a previously existing set. As I repeatedly point out it was not at all in his mind to fiddle about mathematically to procure a particular sequence of special numbers. At that time the process idea was already implied since one had to acquire these sets in some order. With a view to physical application two courses were open. One could regard the construction process as a once-for-all thing, or one could see it as able to happen over and over again. The latter is the view I take as the proper implementation of the process idea.

II. CMULATIVE SUMS

From the outset it was clear that we needed to relate the numbers of physically effective things in the stages of the construction with important physical numbers. Parker-Rhodes thought of these things by analogy with the eigenvalues of quantum theory. I believe he never took the process idea seriously. The 4 elements at level 1 had a straightforward interpretation as the dimensional structure of space-time. We soon found it necessary to see the constructed elements as *discriminately closed subsets*. To get the numbers right for experimental identification it was necessary to add those of the different stages together. This empirical necessity meant that the second of the meanings for the process idea just given was unambiguously the right one. If the successive levels of construction were just elaborations of the one before then it was illogical to add them up.

III. ITERATION AND ALGEBRAIC STRUCTURE

Iteration entails the idea that there is something stable and capable of being run time and time again with internal changes which tell us about the world. (I avoid saying 'give information' because that is become a bit of a cliché.) So the hierarchy algebra appears as a set of rules which constrain the development but do not prescribe it. There is then a very big question to answer about the degree of permanence of the world that is constructed. However we answer that question, we do see why the constants that are derived from just those rules can be identified with experimental numbers –what we call the scale-constants.

IV. PROGRAM UNIVERSE

It almost looks as though we are recapitulating the work that was done years ago on Program Universe, where one had to find a way to store information. Process: the history. Hints came from the addition +3+7 etc. Then the inevitability of taking the construction seriously. Was it an algebraic device merely or did it have a counterpart in the world? Sometimes it was said to be just a model, but if so what was it a model of? If the overpowering suggestion of iteration were rejected then it seemed inevitable that all the interesting construction had to be done once for all in an instant. Was this a model of the Big Bang, or what?

VI. PERCEPTION

The iterative character of the hierarchy algebra comes as a surprising step already because physical space is usually thought of as a platonic receptacle into which new experimental information is continually fitted. That it should itself be subject to construction is something new. Further thought about the part played by the mind of the experimenter in the iteration makes the step more than surprising – completely strange. The step may be rejected out-of-hand as being altogether outrageous. The passive observer has to go. If we are to speak of an observer at all, then he has to go along with the construction and deconstruction in the iteration. There may be ways in which sense can be given to a community of experience and so some sort of public world, but all that is a lot further on.

I expect that it is at once clear that we are proposing a world of pan-psychism since a wide variety of sorts of spatial and perhaps temporal configurations may come to seem the best presentations of what is going on in what we should usually call objective space and time. Hence paranormal phenomena are to be expected as normal rather than outside rationality.

VI. ITERATION, THE STATISTICAL BACKGROUND

We have to get any dynamics we want for the purposes of comparison with the standard model from the bare dimensionless numbers of the hierarchy and from such pretty basic ideas as come from the construction of those numbers. This theoretical policy follows because we see these numbers as the paradigm case of measurement and from which the ordinary kinds of measureable numbers –being more complicated- have to be constructed. In this way we get as near to the continuum as we find the need, and the continuum is a word meaning that that road is open to go further.

The meaning of 'statistical' and 'random' need to be discussed. Our position is that these terms always refer to the state of our knowledge. If we use a random variable, that is because we do not know the causation, and the randomization is necessary to avoid giving misleading information about what is known to exist. This view is contradictory to the current observer/state philosophy which has to assume that the choice of randomness is just the right one out of a series of mathematical options. They (the state/observer theorists) have to say that because otherwise there would be possible alternatives to their way of explaining discreteness. Hidden variable theorizing would become possible.

It is usually said that coupling constants specify the strengths of the fundamental fields or fundamental interactions. This could mean that they are *ratios* of some experimental measures of those fields. This interpretation works approximately for the comparison of the electromagnetic and the gravitational fields since the hierarchy factor 2^{137} or approximately 10^{39} is of the correct order of magnitude. However there is no obvious meaning for the pure numbers found from the hierarchy standing alone. One is tempted to use the two simplest levels (giving 3, 10) to try to form ratios with the strengths of the strong field, but the experimental interpretation for these coupling constants are at best vague.

The historical route to the understanding of the fine-structure constant from the quantum theory of around 1920 seems to have given rise to the simple view that of course there is an adequate experimental meaning for that quantity upon which more refined theories can build. The problem is to see how from that beginning one can interpret the constant as in itself (and without the help of any ratio formed with other such constants like the gravitational one) one can regard it as a specification of an interaction strength or collision cross-section.

Questions like this call into play a general belief in a background of particles collectively called the vacuum. Since the vacuum has no ordinary spatial concommittants the particles are said to be 'virtual'. Something of this sort, however, seems to be a logical necessity. Last year I quoted statements of Peter Rowlands, and, still needing guidance through this problem, I make no apology for quoting him again. He puts the cards on the table in a way that you do not find much in most treatments :-

"The idea of an alpha is that if you have an isolated charge it still interacts with the vacuum to produce virtual bosons (photons or whatever) by emission or absorption, and, where another charge is present, the bosons can be emitted by one and absorbed by the other in a two way process.... Virtual interaction with the vacuum is similar to real interaction with a real field (which, of course, is simply a distribution of real charge of some kind) -the vacuum acts as a virtual field." And:- "The measurement of the coupling between charge and field is the coupling constant alpha. Now, in the case of a coupling involving one charge emitting (or radiating) a photon and another charge absorbing it (in a mutual process that goes both ways simultaneously), the rate of the interaction, or probability per unit time that the process will happen, is proportional to alpha." He quotes Mike Houlden as saying that 'a particle with charge e randomly radiates photons at a rate that is a constant of nature'. Scattering and emission /absorption are essentially the same process on a Feynman diagram."

VIII. A PARALLEL DEVELOPMENT -CAHILL

A parallel development to ours has become evident recently. Reg. Cahill and two or three associates have written extensively on a theory which they base unambiguously on process.

The windowless monad. Cahill appeals to monads but does not say how the windows come. It is THE background problem for us too. In conventional physics (what Cahill calls it the geometrical modelling of space and time. He also speaks of the historical or 'being' model of reality. Also the Eleatic model. (In contrast to the becoming or processing model). I think he also speaks of 'object based physics' but I can't find it now. He observes that Heraclitus argued that it was the appearance of stability, not change, that needed to be explained.

Cahill has an engaging quote from a person much in the mainstream of analytic thought - Charles Peirce. "The one intelligible theory of the universe is that of objective idealism, that matter is effete mind, inveterate habits becoming physical laws. But before this can be accepted it must show itself capable of explaining the tridimensionality of space, the laws of motion, and the general characteristics of the universe, with mathematical clearness and precision, for no less should be demanded of every philosophy."

Our statistical background is to be compared with Cahill's 'quantum foam', and both views see a place for a pre-spatial 'vacuum'. Cahill appears to side with the observer/state philosophy by taking the randomness as a matter of principle, though I think he may go too far on this point in view of the fact that he attributes all process to the intervention of the statistical background which he calls the 'quantum foam'. Since he firmly believes in the reality of that, I think it would be right for him merely to insist that we can only discover the effects that it leads to through the very process that he is describing. He does indeed need to insist that you can't avoid that construction by any back-door route, but that position is something else again. I think that Cahill's view of randomness is essentially different from that embraced in the observer/state philosophy because there randomness is necessary to preserve the construction of quantum discreteness, and must not be mitigated. My view is that whenever we attribute randomness to a process we are stating out ignorance of what goes on, and proposing to use the mathematics of randomness to avoid committing ourselves to any specific interpretation.

It is easiest to quote Cahill about how he gets off the ground:-
"Here we describe a model for a self-referentially limited neural network and in the following sections we see how such a network results in emergent geometry and quantum behaviour, and which, increasingly, appears to be a unification of space and quantum phenomena. Process

physics is a semantic informational system and is devoid of *a priori* objects and their laws and so it requires a subtle bootstrap mechanism to set it up. We use a stochastic neural network, … having the structure of real-number valued connections or relational information strengths B_{ij} (considered as forming a square matrix) between pairs of nodes or pseudo-objects i and j." . I think it is best to start with his matrix linking of two symbols a, b.

There is a big jump in Cahill's account where I think a lot of work is needed. Numbers are attached to the elements of the quantum foam in virtue of their appearing in the process. As I see it, and certainly assume in the hierarchy theory, something else is needed before these numbers can define spatial position –even to build his 'gebits' from neural nets. In particular Cahill bases inertia on a sort of drift through the foam as though the elements of the foam had positions out there as well as numerical values. However the positions would need another unrelated set of values. It is this notion of inertia (which indeed we too need) that he uses to kick off his very classical looking gravitation theory. Even merely to label the elements uses up the measure.

It is not simple to relate Cahill's construction with ours. They are different. Cahill's 'quantum foam', which is like our statistical background and is related to the vacuum fluctuations, is inaccessible except insofar as it influences the development of what he calls neural nets. These -called gebits or (pre)geometrical bits- enable the construction of more complex things using continual iteration which means keeping on starting all over again. Some structure that is constructed in this way has a degree of permanence, though none is irrevocable, and much is quite transient. Again he reminds us of our use of the statistical background. This construction is meant to be supported by a division between semantics and syntax. He appeals as well to Godel incompleteness and Turing – evidently to explain the origin of finiteness

Then there is an astonishing development. Cahill seems to feel the need for a constant which plays the part of the fine-structure constant (by contrast with our having it forced upon us). He finds it in some small gravitational effects which come to light in measurements of G in boreholes, and he identifies the anomaly with that constant. He is able to give an approximate numerical for it correct to within 3 or 4 percent. From a conventional point of view he is riding roughshod over the barrier between mechanical and electromagnetic forces, since from that point of view you cannot correct gravity using electromagnetics.

Cahill writes "As for alpha turning up in gravity that is a real surprise but then, as shown in the attached paper CahillDM.pdf, that then explains a great deal. However I don't necessarily interpret this as evidence for EM effects in gravity, rather alpha is most likely a measure of a deep randomness (an aspect of deep processes - what I call Self-Referential Noise), and so it shows up in gravity, where it gives the strength of spatial self-interaction, and also in QED where it is the measure of the probability of charged-particle photon emission/absorption. Clearly we are seeing evidence of a much deeper and unified physics."

There is one other effect to which Cahill gives foundational status. This is the now much discussed absoluteness of position in space. He has written and even experimented on the Michelson-Morley experiment. It turns out that the non-detection of motion holds only in empty space.

On Nothing, or an Exploration of the Vacuum

Peter Rowlands

Department of Physics, University of Liverpool, Oliver Lodge Laboratory, Oxford St, Liverpool, L69 7ZE, UK. email p.rowlands@liverpool.ac.uk

Abstract. Vacuum is an essential component of a physics which includes the second law of thermodynamics. Its origins can be seen in the Dirac nilpotent structure which arises from the concept of zero totality, and its many manifestations can be explained using the mathematical structure provided by the nilpotent formalism.

Physics and observables

There is a long-standing belief among some physicists that physics is only about observables, but there is at least one law which denies this. One way of stating the second law of thermodynamics is to say that physics cannot be structured purely on observables. Observable 'time', used in classical mechanics, electromagnetic theory and relativity, is reversible, but 'thermodynamic' time is not. We know from the second law that the absolute order of events cannot be changed. In principle, this involves an infinite number of possible occurrences, though only a finite number will be observable. Every time we observe a finite sequence, it appears to correspond to what could be constructed from an infinite one, but we can only assert this as a supposition; it cannot be observed.

The unchangeable and absolute order of physical events is what Newton meant by 'absolute time', although his words have been distorted ever since to apply to the very different concept of time *measurement* which he specifically distinguished as 'relative time'. However, Einstein used exactly the same concept in defining a concept of *causality*, strictly outside the kinematic definition of special relativity, to determine that the absolute order of events could not be changed by changing one's frame of reference. For Einstein, this was manifested by the inability of transporting 'information' (more specifically, energy) at greater than the 'speed of light' (c). In more formal terms, this means defining proper time (τ) as an independent parameter, alongside space and time in the equation

$$r^2 - c^2t^2 = c^2\tau^2 .$$

In the terms I have used on previous occasions (with $c = 1$, $\eta = 1$), this is equivalent to defining ($\pm$ ***k****t* $\pm$ ***i*i**r + ***i****i*τ) or its conjugate expression, ($\pm$ ***k****E* $\pm$ ***i*i**p + ***i*j***m*), as a *nilpotent*, or square root of zero.

The Newtonian concept of absolute time was based on a universe structured on an instantaneous and universal gravitational action at a distance. The Einsteinian concept can now be seen to depend on the instantaneous correlation required in quantum mechanics, as the relationship between t and **r** enshrined in the definition of c, emerges only with the definition of the nilpotent structure also incorporating τ. The instantaneous correlation itself is also a necessary consequence of nilpotent structure as each 'fermionic' term of the form ($\pm$ ***k****E* $\pm$ ***ii*p** + ***ij****m*) must be defined as instantly distinguishable from all others to preserve a nonzero total wavefunction. Significantly, 'observable' time ($\pm$ ***k****t*) is not pure time (t), while ($\pm$ ***k****E*) is not pure energy (E). In quantum mechanical terms, ($\pm$ ***k****t* $\pm$ ***ii*r** + ***ii***τ) is only a classical approximation, with the quantum ($\pm$ ***k****E* $\pm$ ***ii*p** + ***ij****m*) really representing ($\pm$ ***k***$\partial/\partial t$ $\pm$ ***ii***∇ + ***ij****m*), where t is not an observable. Significantly, *CPT* invariance, which is defined to simultaneously preserve relativity *and* causality, is an obvious consequence of the structure ($\pm$ ***k****E* $\pm$ ***ii*p** + ***ij****m*) or ($\pm$ ***k****t* $\pm$ ***ii*r** + ***ii***τ), precisely because each of the E, **p**, m or t, **p**, τ terms is preceded by one of the quaternion operators ***k***, ***i***, ***j***, which make it nilpotent.

Another way of accommodating instantaneous correlation, either in the Newtonian or quantum mechanical sense, is to describe time and energy in terms of 'continuity', as suggested by the Klein-4 symmetry between space, time, mass and charge, put forward in previous papers [Rowlands, 1983, 1999, 2001]; and another way of describing continuity of either time or energy is to use the concept of 'filled vacuum', at least in a virtual sense. (Interestingly, in particle physics, it is the filled vacuum that leads to the creation of rest mass in discrete particle states, in the same way as the m term, or 'proper energy', is the 'causality' term in the fermionic state vector.) And so, ultimately, the second law of thermodynamics is an expression of the filled vacuum concept; and, while 'pure' relativity, defined in a kinematic sense, is about space and time only, and so does not need a vacuum, as Einstein showed, the extended concept, including proper time and causality, definitely requires one. Nevertheless, this filled vacuum idea is one of the most peculiar aspects of physics, and its acceptance has always been problematical. The reason is that the 'continuity' it evokes is necessarily an indefinite thing. It talks about what we don't know, not what we know. In a sense, when we set out an idea of physics, necessarily based on observables, we unavoidably map out only a part of the total picture. This is what we mean by discreteness, finiteness, or even a defined 'system'. To make this work, however, we have to find some way of specifying 'the rest', the part that cannot be defined. This is what we mean by 'vacuum'. We have to take the undefinable 'rest' (of the universe) into account. But though we cannot define it directly, we can define it indirectly in terms of what it *is not*; and, in many cases, this reflects the structure of the part that we can define.

The nilpotent state vector

The origins of the Dirac or fermionic state have been seen in the mathematical structure which emerges from trying to define a concept of zero totality (this is detailed in the Appendix). Out of this structure emerge the four fundamental parameters space, time, mass and charge, and their relationship as a Klein-4 group structure.

mass	conserved	real	continuous (nondimensional)
time	nonconserved	imaginary	continuous (nondimensional)
charge	conserved	imaginary	discrete (dimensional)
space	nonconserved	real	discrete (dimensional)

The Dirac state then becomes the most efficient packaging of the respective scalar, pseudoscalar, quaternion and multivariate vector units of these quantities to create a single physical unit. In the packaging process, the respective pseudoscalar, vector and scalar units of the first three quantities are compactified by applying one of the three quaternion charge units separately onto each one. So

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

are reconstructed as new pseudoscalar, vector and scalar quantities, energy, momentum and rest mass (E, **p**, m):

ik	***i*i *j*i *k*i**	1***j***
E	**p**	m

At the same time, the symmetry between the quaternion charge units is broken to create weak, strong and electric charges (w, s, e), with respective pseudoscalar, vector and scalar characteristics.

w	s	e

This implies that the combined Dirac state (± ***k****E* ± ***i*i**p + ***ij****m*) (where the terms are multiplied throughout by i by convention) must be an expression of the charge state of a fermion as well as of its energy state.

The combined fermionic or Dirac state (± ***k****E* ± ***i*i**p + ***ij****m*) is a nilpotent, or square root of zero, since (± ***k****E* ± ***i*i**p + ***ij****m*) (± ***k****E* ± ***i*i**p + ***ij****m*) = $E^2 - p^2 - m^2 = 0$ is simply the standard relativistic relation between energy, momentum and rest mass. Here the momentum vector **p** is taken to be multivariate, incorporating the concept of spin. In this case, **pp** = **p.p** + i**p** × **p** = (σ.**p**) (σ.**p**) = $(-p)(-p) = p^2$. Creating the nilpotent state as a single package requires simultaneous generation of the fundamental constants η, c and G to ensure that the terms have the appropriate scale relationships. In effect, the nilpotent structure demands that both special relativity and quantization are required simultaneously, and that each requires the other – it is, in fact, the unique way of achieving this. The E and $\mathbf{p}$ terms in the state may represent either eigenvalues or operators, depending on whether we choose a conserved or nonconserved state representation. Using the operator version of the state ($\pm\, \boldsymbol{k}\partial/\partial t \pm \boldsymbol{i}i\nabla + \boldsymbol{i}\boldsymbol{j}m$), we may write down an equivalent *conjugate metric* ($\pm\, \boldsymbol{k}t \pm \boldsymbol{i}i\mathbf{r} + \boldsymbol{i}\boldsymbol{j}\tau$), where τ is proper time. In the classical limit, this becomes the definition of special relativity in nilpotent form.

Interestingly, the 8 fundamental units – 4 'real' and 4 'imaginary' – allow a twistor or, alternatively, an octonion-type representation, as does the combination of the Klein-4 group with its imaginary dual, which reverses the real / imaginary status of the group parameters. Such structures are intrinsically symmetrical, but the realisation that the 5 composite units are a more efficient compactification of the 8 fundamental units creates, for the first time, a fundamental asymmetry in the structure, which is especially manifested in the pseudoscalar term i. This term is not only necessary to nilpotency, it also seems to be responsible for the asymmetry manifested in nature through time, energy and the weak interaction. In addition, the 5-fold nature of the structure within an overall 3-dimensionality may be related to the breaking of symmetry in 3-D space manifested by Penrose tiling, where the patterning is as unique and unperiodic as fermionic composition [Rowlands, 2002].

The nilpotent in operator form defines the entire state, both amplitude and phase, because the phase term is uniquely determined as that required to make the eigenvalue (or amplitude) nilpotent. This extra constraint is unique to the nilpotent formulation, and ensures that neither the Dirac equation nor any other is needed to describe the state. This is especially significant where the E and $\mathbf{p}$ operators are replaced by *covariant* derivatives, or ones including field terms, because the same principle then applies. Solution of the Dirac equation in these instances is replaced by the process of finding the phase term which makes the amplitude of the Dirac state a nilpotent. In the particular case of a point charge of any kind, where spherical symmetry is a minimum requirement, the nilpotency requirement leads to a minimum condition equivalent to an inverse linear (Coulomb) potential [Rowlands, 2004a].

By incorporating the extra symmetries due to nilpotency, the fermionic state is automatically second quantized, with an in-built supersymmetry. Amplitude and phase are determined uniquely by the same operator, and all are quantized in the same way. The ± signs in front of $\boldsymbol{k}E$ and $\boldsymbol{i}i\mathbf{p}$ represent the four simultaneous 'solutions' for the Dirac fermionic state – fermion / antifermion ($\pm\, \boldsymbol{k}E$), spin up / down ($\pm\, \boldsymbol{i}i\mathbf{p}$) – and the full representation of the nilpotent operator is as a 4-term row or column vector, in the same way as the conventional Dirac spinor. From this, we may define an antifermionic state as one with the form (μ *kE* ± *ii***p** + *ijm*). A spin reversed fermion becomes (± *kE* μ *ii* **p** + *ijm*); a spin 1 boson (± *kE* ± *ii***p** + *ijm*) (μ *kE* ± *ii***p** + *ijm*); and a spin 0 boson (± *kE* ± *ii***p** + *ijm*) (μ *kE* μ *ii***p** + *ijm*). A Bose-Einstein condensate state, Cooper pair, or other 'bosonic' two-fermion (Berry phase) equivalent, will then have the form (± *kE* ± *ii***p** + *ijm*) (± *kE* μ *ii***p** + *ijm*).

It is significant that spin 0 bosons, unlike those of spin 1, cannot be massless, as their state vectors would automatically multiply to zero. Also the fermion and antifermion components of spin 0 bosons must have the *same* helicity (which is determined from the combination the signs of *E* and **p**), which can only be true for a particle with non-zero mass. All weak, strong and electric interactions between fundamental particles can, in one way or another, be represented by vertices in which a bosonic state is created by the combination of a fermionic with an antifermionic state vector (or, in the case of Berry phase, fermionic with fermionic), and the required modification of their *E*, **p** and *m* terms.

In the case of baryons, we make specific use of the vector nature of the **p** operator, and construct an entangled states of which the first row in the spinor has the form ($kE \pm ii\mathrm{p}_x + ijm$) ($kE \pm ii\mathrm{p}_y + ijm$) ($kE \pm ii\mathrm{p}_z + ijm$). The baryonic state thus has three components (conventionally described as 'quarks') representing three six possible 'phases' of a recognizably *SU*(3) symmetry in which **p** is respectively $\pm ii\mathrm{p}_x$, $\pm ii\mathrm{p}_y$, $\pm ii\mathrm{p}_z$. The three components of the composite baryon will then have the usual properties attributable to the components of vectors, and can be no more separated than the dimensions of space or momentum. A gauge invariant nonlocal 'transfer' of **p** between the phases will then take place at a constant rate, irrespective of any concept of physical separation. Such a constant rate of change of momentum is equivalent to a constant force or a potential which is linear with distance. In addition, the symmetry requirement which makes both + and – **p** terms necessary, and hence both left- and right-handed helicities, is equivalent to requiring the baryon to have non-zero mass.

A nilpotent state is naturally Pauli exclusive because (± *kE* ± *ii***p** + *ijm*) (± *kE* ± *ii* **p** + *ijm*) = 0, and nonlocality is another automatic consequence. QED, QCD, QFD are all derivable in a straightforward way, while propagators defined through nilpotent states remove the infrared divergence [Rowlands and Cullerne, 2001a, b]. Each nilpotent state is necessarily unique, because it would instantly zero if multiplied by one exactly like itself. So, nonlocality is immediately implied by a nilpotent structure. The formation of any new state, which is determined by the nature of all other nilpotent states, is a creation event within a unique birthordering. Each 'creation' event (which includes any change in parameters, as well as entirely new fermionic creations) also necessarily changes all existing states to some degree. In this sense, a nilpotent structure uniquely allows us to conceive of the infinite while only observing the finite.

Specific field quantization becomes superfluous because the nilpotent terms are already second quantized quantum field operators. They are also exactly supersymmetric, with the operators $Q = (\pm kE \pm ii\mathbf{p} + ijm)$ and $Q^{\dagger} = (\mu\ kE \pm ii\mathbf{p} + ijm)$ respectively converting bosons to fermions, and fermions to bosons or bosons to antifermions. An exact supersymmetry suggests that the particles are their own supersymmetric partners, which implies a vacuum connection. The respective *C*, *P*, *T* transformations can be represented by: $-\boldsymbol{j}\ (\pm kE \pm ii\mathbf{p} + ij)\ \boldsymbol{j} = (\mu\ kE\ \mu\ ii\mathbf{p} + ijm)$; $\boldsymbol{i}\ (\pm kE \pm ii\mathbf{p} + ij)\ \boldsymbol{i} = (\pm kE \pm ii\mathbf{p} + ij)$; $\boldsymbol{k}\ (\pm kE\ \mu\ ii\mathbf{p} + ij)\ \boldsymbol{k} = (\mu\ kE \pm ii\mathbf{p} + ijm)$, with *CPT* invariance an easily-derived consequence. The half-integral spin of fermions can be derived in a standard formal way, for example:

$$[\hat{\sigma}, \mathcal{H}] = [-\mathbf{1}, -j\,(\mathrm{i}p_1 + \mathrm{j}p_2 + \mathrm{k}p_3) + ikm] = 2ij\,\mathbf{1} \times \mathbf{p}$$

$$[\mathbf{L}, \mathcal{H}] = -ki\,[\mathbf{r}, \mathbf{1}.\mathbf{p}] \times \mathbf{p} = -j\,[\mathbf{r}, \mathbf{1}.\mathbf{p}] \times \mathbf{p} = -ij\,\mathbf{1} \times \mathbf{p},$$

since $[\mathbf{r}, \mathbf{1}.\mathbf{p}]\ \psi = i\mathbf{1}\ \psi$. Then $[\mathbf{L} + \hat{\sigma}/2, \mathcal{H}] = 0$, which makes $\mathbf{L} + \hat{\sigma}/2$ a constant of the motion. However, the significance of half-integral spin in a physical sense seems to be the same as is implied by the natural exact supersymmetry: the pure fermionic state may be considered in some way as incomplete without its vacuum partner.

Dirac vacuum operators

The 4-component spinor representing the nilpotent Dirac state incorporates four creation / annihilation operators:

fermion creation spin up	$(kE + ii\mathbf{p} + ijm)$
fermion creation spin down	$(kE - ii\mathbf{p} + ijm)$
antifermion creation spin up	$(-kE - ii\mathbf{p} + ijm)$
antifermion creation spin down	$(-kE + ii\mathbf{p} + ijm)$

Taking any of these operators, we can also specify vacuum operators which (assuming an appropriate normalization) leave the state unchanged. For example, $(kE + ii\mathbf{p} + ijm)$ is unchanged if postmultiplied by $\boldsymbol{k}\ (kE + ii\mathbf{p} + ijm)$ any number of times:

$$(kE + ii\mathbf{p} + ijm)\ \boldsymbol{k}\ (kE + ii\mathbf{p} + ijm)\ \boldsymbol{k}\ (kE + ii\mathbf{p} + ijm)\ \boldsymbol{k}\ (kE + ii\mathbf{p} + ijm) \ldots$$

However, since $\boldsymbol{k}\ (kE + ii\mathbf{p} + ijm)\ \boldsymbol{k}$ is identical to the antifermion creation operator $(-kE + ii\mathbf{p} + ijm)$, so we can also write this expression with alternate terms representing fermion / antifermion creation:

$$(kE + ii\mathbf{p} + ijm)\ (-kE + ii\mathbf{p} + ijm)\ (kE + ii\mathbf{p} + ijm)\ (-kE + ii\mathbf{p} + ijm) \ldots$$

or as a process of alternate fermion / boson creation through the supersymmetry operators $QQ^{\dagger}QQ^{\dagger}$ .... The antifermionic creation state, here, acts as a vacuum 'reflection' of the fermionic creation state, and vice versa, while the real fermion and its virtual reflection combine to create a supersymmetric bosonic partner, which is identical to the original fermion creation. And we can extend the argument to say that a real bosonic creation state, such as $(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)(-\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$, would produce simultaneous supersymmetric virtual antifermion and fermion states as the respective vacuum reflections of the component creation operators $(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$ and $(-\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$.

The expression $\mathbf{k}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$, however, is not unique in defining a 'vacuum' state: $\mathbf{i}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$ and $\mathbf{j}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$ have the same properties, and must also be considered vacuum operators. In the case of $\mathbf{i}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$, the vacuum 'reflection' requires a change in spin orientation. In the case of $\mathbf{j}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$, the fermion reflects as an antifermion with an additional change of spin, and preservation of helicity. Each of the three cases produce supersymmetric bosonic-type states, which are respectively spin 1, spin 0, and Bose-Einstein condensate, for the coefficients $\mathbf{k}$, $\mathbf{j}$ and $\mathbf{i}$. There is, however, no *discrete* vacuum equivalent to $1(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$ because this is Pauli excluded: $(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)\,1\,(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m) = 0$. At the same time, two different vacuum states can only combine to produce the third through the mediation of the physical part of the coefficient, that is E, $\mathbf{p}$ or m.

Defining vacuum states through the coefficients $\mathbf{k}$, $\mathbf{j}$ and $\mathbf{i}$ suggests a new meaning for the four 'solutions' characterizing the Dirac state. The first row of the spinor represents the fermionic / antifermionic state, while the other three rows are the three discrete vacuum reflections. The three coefficients can also be seen as resulting in (or being responsible for) the concept of discrete (point-like) charge.

$\mathbf{k}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$	or	$i\mathbf{k}E(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$	weak vacuum
$\mathbf{i}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$	or	$\mathbf{i}\mathbf{p}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$	strong vacuum
$\mathbf{j}(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$	or	$\mathbf{j}m(\mathbf{k}E + \mathbf{i}i\mathbf{p} + \mathbf{i}\mathbf{j}m)$	electric vacuum

Charge, in this interpretation, becomes a *manifestation* of the vacuum; and, as in the case of the charges, the three vacua are entirely independent of each other, not recognizing each other's existence.

The nilpotent state vector can now be seen as incorporating both real and virtual components, in the same way as it incorporates both mass and charge; and *zitterbewegung* may be interpreted as a switching between them. This is why the state vectors are supersymmetric. The real fermion and its set of dual vacuum images combine to produce a single-valued bosonic spin state, analogous to a conserved physical system, simultaneously incorporating the action and reaction sides of Newton's third law of motion or a virial doubling of the kinetic energy in a potential energy term. This is why fermion and antifermion state vectors have identical components, with only the *order* privileging either $+E$ or $-E$ states as the 'real' ones.

The 'real' fermion creation is then distinguished by its real coefficient (1) from the vacuum 'image' states, which are induced by the weak, electric and strong elements, and are specified by quaternion coefficients. The first term in the Dirac 4-spinor thus has a different status to all the others, just as the time coordinate has a different status to the three space coordinates in the conventional Minkowski 4-vector. In the case of a *free* fermion (or boson), this status is especially significant, for the vacuum terms then make no contribution to the particle's energy, meaning that renormalization is not required – as the nilpotent version of QED demonstrates. In renormalization, only the effects of the 'image' terms cancel, leaving those of the 'real' term unchanged.

Zero-point energy

One interpretation of vacuum is as 'the rest of the universe', the 'reaction' half of Newton's third law. This is how we can define it by reference to the 'image' charge or 'reflection' of a discrete source. For the discrete weak, strong and electric vacua, it means that part of the rest of the universe recognized by the appropriate charge, and it is an effective negation of that component. The total vacuum, however, is the *continuous* vacuum produced by the real (gravitational) component, and, for any given fermion, produces a state vector equivalent to $-1(\boldsymbol{k}E + i\boldsymbol{i}\mathbf{p} + i\boldsymbol{j}m)$, with negative energy. The combination of fermion plus total vacuum then produces a zero totality and zero state vector. 'Continuity', in this context, can only mean the absence of discrete energy levels, and it is this property which gives rise to the infinite virtual energy density and virtual energy of $\frac{1}{2}\eta\omega$ for every possible mode of vibration, the so-called zero-point energy. The continuous vacuum is thus constituted out of the mirror image states of *all possible* fermion states, and it is this continuous vacuum which makes possible the nonlocal connection required by Pauli exclusion. Each possible state provides a virtual vacuum energy of $\frac{1}{2}\eta\omega$, like the ground state of a harmonic oscillator – which, of course, is precisely what it is. To create a real fermion state, we excite a virtual vacuum state of $-\frac{1}{2}\eta\omega$ up to the level $\frac{1}{2}\eta\omega$, using a total energy quantum of $\eta\omega$ The continuous vacuum, however, can never be observed directly precisely because it *is* continuous, and so the concept of continuity will necessarily remain a 'potential' or virtual one.

The weak vacuum

We can consider the function of charge as 'partitioning' the continuous vacuum, that we never observe directly, in a way which, being discrete, can be observed. Charge thus becomes a kind of vacuum state, associated with the quantum field nature of the state vector, and the different charges are associated with qualitatively different vacuum states through their association with respective pseudoscalar, vector and scalar coefficients. The three discrete vacua describe only that part of the vacuum which the associated type of charge sees.

The total vacuum which the charges partition originates as an expression of the continuous or noncountable nature of mass-energy. Continuity necessarily makes mass-energy unidimensional and unipolar, and, because it is also real, restricts it to a single mathematical sign, which is usually taken as positive. We can interpret this as implying a non-symmetric ground state or a filled vacuum, which is that of negative energy or antifermions. Physically, it manifests itself in the spin 0 Higgs field, which breaks charge conjugation symmetry for the weak interaction, and gives rest masses to the fermions and weak gauge bosons.

The derivation of the Higgs field usually begins with the Goldstone theorem, according to which the breaking of a continuous symmetry of a physical system necessarily leads to the appearance of a massless scalar or spin 0 boson state. This theorem is, in fact, a natural consequence of nilpotent structure. Bosonic states can be considered as resulting from the transformation of one fermion into another or, sometimes, into itself. Where a symmetry is *complete*, however, with no degeneracy permitted, there is no mechanism for a fermion to transform into itself. An example occurs in the colour singlet baryon, which cannot transform into itself via the colour force, meaning that no gluon can be formed out of a combination of $R\bar{R}$, $G\bar{G}$ and $B\bar{B}$, with the result that the number of independent *SU*(3) generators is reduced from 9 to 8. In the nilpotent formalism, this principle ensures that a *free* fermion state (as its own supersymmetric partner) remains unchanged as a result of self-interaction via the vacuum, and so requires no charge or energy renormalization.

An ideal vacuum would have the most complete symmetry possible, with exact and absolute *C*, *P* and *T* symmetries. The maintenance of these symmetries would not, therefore, require the creation of a scalar boson out of a vacuum fermion (with $\pm\, kE \pm ii\mathbf{p}$) and its charge-conjugated antifermionic partner (with $\mu\; kE\; \mu\; ii\mathbf{p}$). However, the breaking of charge conjugation symmetry (*C*) and / or either parity (*P* = *CT*) or time-reversal symmetry (*T* = *CP*), which would be equivalent (with the alternative option of maintaining *C* by the simultaneous violation of *P* and *T* available only to bosons), would make the existence of such a scalar boson necessary, thus justifying the Goldstone theorem. Of course, in the nilpotent formalism, the spin 0 scalar must necessarily be massive, though, being, again, its own supersymmetric partner, there would be no mechanism for an observable self-interaction, and so no renormalization problem concerning the mass.

We can also see how the Higgs boson gives mass to fermion and boson states through the necessary vacuum connection, if we imagine it as being 'produced' at the vertex of an interaction between a massless antifermion and a massive fermion, or between a massless antifermionic boson component and a massive fermionic boson component. If we suppose a pure (massless) antifermion at the vertex with only right-handed helicity, then the production of a boson whose fermion component also has right-handed helicity (which is forbidden in the pure state) must necessarily require a fermion at the vertex with an element of right-handed helicity. So, the boson 'gives' mass to the fermion. (The reverse, of course, is always true, and we could equally see the existence of a positive mass term in fermion states as breaking charge conjugation symmetry, because no mechanism existe to reverse the sign of m.)

Because, it is the $\boldsymbol{k}$ operator which changes fermion to antifermion, the weak vacuum is the one associated with fermion / antifermion annihilation / creation. The pseudoscalar aspect implies that the vacuum or charge state, or potential, may be complex, which is the requirement for *CP* violation. The pseudoscalar representation also naturally implies dipolarity because of the fundamental mathematical duality of $\pm\, i$, and the indistinguishability of the two signs under weak charge conjugation violation. This special property of the weak interaction appears to be the ultimate source of different phases of matter and phase transitions, when the indistinguishability of sign is allowed to effectively eliminate the weak component in fermion-fermion combinations, and so overcome aspects of Pauli exclusion. It is certainly the origin of the Berry phase where the spin 0 'bosonic' state is such as would be required in a pure weak transition from $-i\boldsymbol{k}E$ to $+\, i\boldsymbol{k}E$, or its inverse. Because the spin 0 state is necessarily massive, time reversal symmetry (the one applicable to the transition) must be broken in the weak formation or decay of states involving the Berry phase.

The use of the $\boldsymbol{k}E$ term for the weak vacuum ensures that it has to be through the weak vacuum that we express the continuity of mass-energy and the conjugate irreversibility of time. No *physical* state can be defined corresponding to $-E$, although charge-conjugated $-i\boldsymbol{k}E$ states can be defined by reversing the sign of the $i\boldsymbol{k}$ operator. In principle, this leads to weak charge conjugation violation, which means that the weak interaction is indifferent to the *sign* of the weak charge, and can only distinguish between fermion and antifermion. To preserve *CPT* symmetry, either parity or time-reversal symmetry must also be violated.

Assuming that the requirement for continuous vacuum energy (or thermodynamic time) ensures that a physical bias exists in favour of matter over antimatter, means that the vacuum should have a weak dipole moment, exhibited as a single-handedness of rotation, and represented by the $\frac{1}{2}\eta\omega$ modes of vibration or zero-point energy. In principle this may be seen as the origin of left-handed fermion spin, the fermion being created simultaneously with its vacuum reflection. If the weak vacuum is in a continual state of proclaiming its filled state, by creating weak dipoles which have a dipole moment or specific handedness, we may also expect that 'fluctuations' in this vacuum will be the same thing as the production or annihilation of a weak dipolar fermion-antifermion pair, each of spin ½, via a harmonic oscillator creation-annihilation mechanism. Fluctuations of this kind are responsible for the Casimir or Van der Waals force from the zero point energy, corresponding to the potential for a fluctuating dipole-dipole interaction.

The strong vacuum

The strong interaction, as we know it, is manifested, through a nonlocal gluon sea, with switching of momentum components in terms of both sign and direction to incorporate the six phases. This is exactly what is provided by the ± *i***p** term in the state vector. It is notable that baryon structure is essentially affine, dissolving into component gluons and combinations of virtual baryons *ad infinitum*. This is exactly what we would expect from the affine nature of the **p** operator, whose components can be no more separated or fixed than the dimensions of space. The vector nature of the strong operator also means that the strong vacuum is the only one with explicit relative phases. In strongly interacting systems, the phases are associated with the presence or absence of electric and weak charge components. Where the phases associated with these components coincide, there is no way of distinguishing the phases, and consequently no strong interaction.

The electric vacuum

Fermionic states are ones in which weak charges are present. There are, however, two types of fundamental fermionic state: quark and lepton. For quarks, the **p** phases are explicit, and *s* charges are present; for leptons they are nonexplicit, and *s* charges are absent. As different types of charge and vacuum exist entirely independently of each other, weak charge must be indifferent to the presence or absence of the strong charge. So weak and electric charge allocations for quarks and fermions must follow the same pattern; hence, the fractional electric charges allocated to quarks are merely an expression of the perfect gauge invariance of the strong interaction – analogous to the process of fractional charge creation in the quantum Hall effect – and are not an intrinsic aspect of quark structure. At the same time, the weak charge must be indifferent to the presence or absence of *e*.

The fermionic states with and without electric charge are conventionally described as the $SU(2)_L$ states (up / down, neutrino, electron, etc.); these must be made *explicitly* indistinguishable under the weak interaction. Conventionally, we use the third component of weak isospin (t_3), by analogy with the $SU(2)$ of spin, as the quantum number for distinguishing these states. For the two isospin states, $t_3 = \pm \frac{1}{2}$, but only in half the total number of states (the left-handed ones). For free fermions, the quantum number for the electric force takes the value $Q = -1$, where electric charge ($-e$, taken as negative by convention) is present, again in half the number of states (though a different half). If the weak and electric interactions are described by any grand unifying gauge group, then orthogonality and normalization conditions require the mixing ratio,[5] defined as $sin^2 \theta_W$, to be determined by Tr (t_3^2) / Tr (Q^2), which in this case must be 0.25.

The ratio, however, cannot apply only to free fermions if the weak interaction is indifferent also to the presence or absence of the strong charge. So exactly the same mixing proportion, with $sin^2 \theta_W = 0.25$, should exist also for quark states, and separately for each 'colour' phase, or momentum direction, so that the weak interaction cannot be detected through 'colour'. Interpreting 'colour' through momentum phases or directions, allows the instantaneous existence of only one quark phase in three. So we find that the charge variation 0 0 $-e$ must be taken against either an empty background or 'electric vacuum' (0 0 0) or a full background (*e e e*), so that the two states of weak isospin in the three colours become:

$$\begin{matrix} e & e & 0 \\ 0 & 0 & -e \end{matrix}$$

The most obvious manifestation of an electric vacuum would, therefore, appear to be in the $SU(2)_L$ for the weak interaction. The weak vacuum, which is full and cannot be reversed, effectively contrasts with the electric vacuum, which can be filled or emptied, or reversed in the case of antifermions. However, while the $SU(3)$ and $U(1)$ structures come directly from the vector **p** and scalar *m* terms in the Dirac state, the $SU(2)$ structure for the weak interaction is only related to the $SU(2)$ spin structure, associated with the pseudoscalar E, in an indirect way. This is because the E term in the equation doesn't fully express the asymmetry of the physical E. The $SU(2)$ for E is the $SU(2)$ for helicity, and is related to the $SU(2)_L$ for weak isospin only via a matrix (like the CKM matrix) involving rest mass. It is the mass dependence, relating to the filled nature of the vacuum through the Higgs mechanism, which makes $SU(2) \rightarrow SU(2)_L$. It arises because the gravitational / inertial component is not specifically included through a specific term in the compactified Dirac state. We can express the $SU(2)$ / $SU(2)_L$ relationship as follows:

changing	neutrino / up quark	to	electron / down quark
requires	zero $-e$	becoming	nonzero $- e$
	≡ isospin up		≡ isospin down
	100 % weak		partially weak
	left-handed		partially right-handed
	100 % fermionic		partially antifermionic
	say 100 % spin up		partially spin down
	no added m		added m
	unchanged E / **p** sign		partially changed E / **p** sign

The mixing of E and **p** terms, or right-handed and left-handed components, in the nilpotent state vector, is also equivalent to the mixing of e and w charges, or electric and weak vacua, but this mixing cannot affect the weak interaction as such, which has no right-handed component for fermions. So, the weak interaction must be simultaneously left-handed for fermion states and indifferent to the presence or absence of the electric charge, which introduces the right-handed element.

Particle states

Particle states are determined essentially by charge structure. The presence or absence of units of charge is not determined by the *magnitudes* of E, **p**, m, but by the relative *phases* of the three components of **p**, with respect to those of E and m. By this, we mean the phases associated with the conservation of the handedness and magnitude of angular momentum, mapped onto the phases for direction. If the handedness and magnitude phases are aligned, then the particle is a lepton (L), with no strong charge; if the handedness and magnitude phases are not aligned, then the particle becomes a quark in one of the three 'colour' phases A, B, C.

A

		B	G	R
u	$+e$	$1j$	$1j$	$0i$
	$+s$	$1i$	$0k$	$0j$
	$+w$	$1k$	$0i$	$0k$
d	$-e$	$0j$	$0k$	$1j$
	$+s$	$1i$	$0i$	$0k$
	$+w$	$1k$	$0j$	$0i$

B

		B	G	R
u	$+e$	$1j$	$1j$	$0k$
	$+s$	$0i$	$0k$	$1i$
	$+w$	$1k$	$0i$	$0j$
d	$-e$	$0i$	$0k$	$1j$
	$+s$	$0j$	$0i$	$1i$
	$+w$	$1k$	$0j$	$0k$

C

		B	G	R
u	$+e$	$1j$	$1j$	$0k$
	$+s$	$0i$	$1i$	$0j$
	$+w$	$1k$	$0k$	$0i$
d	$-e$	$0j$	$0k$	$1j$
	$+s$	$0i$	$1i$	$0k$
	$+w$	$1k$	$0j$	$0i$

L

		$\bar{e}$	$\bar{e}$	ν_e
	$+e$	$1j$	$1j$	$0j$
	$+s$	$0k$	$0i$	$0i$
	$+w$	$0i$	$0k$	$1k$
		$\bar{\nu}_e$	$\bar{\nu}_e$	e
	$-e$	$0i$	$0k$	$1j$
	$+s$	$0j$	$0i$	$0i$
	$+w$	$0k$	$0j$	$1k$

It is possible to incorporate all the information in these tables, which principally derive from those for the *SU*(2) charge allocations, into a single unified representation for the entire set of charge structures for quarks and leptons, and their antistates [Rowlands, 2003]:

$$\sigma_z.\left(\boldsymbol{i}\hat{\mathbf{p}}_a(\delta_{bc}-1)+\boldsymbol{j}(\hat{\mathbf{p}}_b-\mathbf{1}\delta_{0m})+\boldsymbol{k}\hat{\mathbf{p}}_c(-1)^{\delta_{1g}}g\right)$$

Here, the quaternion operators ***i***, ***j***, ***k*** represent the respective strong, electric and weak charge units; σ_z is the spin pseudovector component defined in the *z* direction; $\hat{\mathbf{p}}_a, \hat{\mathbf{p}}_b, \hat{\mathbf{p}}_c$ are units of quantized angular momentum, selected randomly and independently from the three orthogonal components $\hat{\mathbf{p}}_x, \hat{\mathbf{p}}_y, \hat{\mathbf{p}}_z$.

The other terms are logical operators, representing the $SU(3) \times SU(2)_L \times U(1)$ structure of the 3 nongravitational interactions. In principle, the term attached to the ***i*** operator represents quark confinement; the one attached to ***j*** represents weak isospin, and the one attached to ***k*** represents weak charge conjugation violation. If we assume that σ_z is naturally left-handed for fermions, and right-handed for antifermions, then anything in the expression which tends towards reducing the degree of this one-handedness, by zeroing any of the terms or changing their signs, can be considered a mass generator, and so we find 3 separate processes for generating mass by the Higgs mechanism. These are: the missing (zero) charges in composite (baryon and boson states) (***i***, ***j*** and ***k***); the empty electric vacuum state ($-\mathbf{1}\delta_{0m} = 0$) (***j*** only); the reversal of sign of the weak charge in the second and third generations (***k*** only).

The gravitational vacuum

Three terms in the Dirac 4-spinor for a fermion represent its three discrete vacuum 'reflections'; the fourth (conventionally placed in the first row) represents the particle creation itself. Because the three vacuum reflections are generated by terms which are also charge operators, it is natural to conclude that charge is fundamentally a vacuum generator. At the same time, the mass of the fermion and the related vacuum energy may be assumed to be 'generated' by the 'mass' operator (1). So we can consider gravity, the force generated by mass, as being in some way represented by a vacuum operator of the form $1(\pm\, \boldsymbol{i}kE \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$. However, it is more probable that the gravitational vacuum has the form, $-1(\boldsymbol{k}E + i\boldsymbol{i}\mathbf{p} + i\boldsymbol{j}m)$, the term which zeros the Dirac state.

Many people have assumed that gravity is a discrete force. It comes, however, from a continuous vacuum and is the only serious candidate for the carrier of nonlocality, for the instantaneous quantum correlation of Dirac states, and for the source of the infinite zero-point energy spectrum. The use of the coefficient 1 may be taken as equivalent to the statement that the gravitational vacuum cannot be quantized directly. One way of representing this would be to define gravitational energy as negative (because of the attractive force) and to refer to the filled vacuum for negative energy states, as proposed in the original positron theory of Dirac, and as appears to be required to explain the absence of antimatter from the universe's ground state. There is, nevertheless, a discrete vacuum representation related to mass. This is the inertial component, related to the discrete rest mass, which itself originates in the fermionic or bosonic charge structure. In the Higgs mechanism, it is signalled by a nonlocal finite energy level for the weak vacuum. The inertial component may be seen as a discrete local reaction specified by $1(\pm\, \boldsymbol{i}kE \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$ to the continuous nonlocal gravitational energy specified by $-1(\pm\, \boldsymbol{i}kE \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$. The total zero energy of the 'universe' could then be said to come from the combination of a positive nilpotent (inertia, sum of charges) with a negative one (gravity).

Gravity and inertia

The instantaneous nonlocal connection between fermionic states requires a mechanism which has the attributes of classical gravity. This is what we mean by saying that the combined gravitational vacuum must be continuous. (It is interesting, in this connection, that Newtonian gravity can be described as a quantum theory; because it has a vacuum and nonlocal connection, while 'quantum gravity', as conventionally conceived, without continuity, cannot.) However, the discrete structure of the nilpotent state requires a space-time relation which is equivalent to a finite transmission of information at a fixed ratio (traditionally symbolized by *c*, the 'speed of light'). Another expression which is equivalent to 'information' is *measurement*. Such concepts only have meaning in terms of discrete sources (such as charges). However, they can also be applied to manifestations of the fundamentally continuous energy state, such as mass and radiant energy, *as long the truly continuous nature of the energy, that is the instantaneity of the interaction between energy elements, is made explicit*. This requires a combination of instantaneous interaction and time-delayed measurement, and is the origin of the concept of inertia.

Inertia arises from the fact that the Lorentz-invariance which derives from nilpotency is already *assumed* within our system of measurement or information; but is not constructed to describe instantaneous interactions. That is, Newtonian equations which are inertial for instantaneous interaction will not be inertial in the Lorentzian space-time required for measurement between discrete sources. Lorentzian space-time is, in effect, a *noninertial frame* for an instantaneous interaction, and using it in this context will produce effects, which may be considered analogous to those of the classical inertial (centrifugal and Coriolis) forces. To find equations appropriate to instantaneous interaction, expressed in a Lorentzian space-time (and, more specifically, in a *gravitationally-affected* Lorentzian space-time), we need to find the inertial force terms, or equivalent, which define the straight-line position. In fact, equations of this kind already exist within the formal structure of general relativity.

In a paper published in 1992 I wrote: 'The result of making the equations for inertial force linear is that the quantum of interaction for the inertial force is not the spin 2 graviton but the spin 1 photon, or virtual photon; of course, as quantum field theory predicts, spin 2 intermediaries would not be possible for a repulsive force between like particles. As soon as we drop the necessity of the spin 2 intermediary, we ensure that our quantum theory becomes renormalisable in the same way as QED.' [Rowlands, 1992] In principle, if we can demonstrate that the equations derivable from the inertial process parallel those of classical electromagnetic theory, then the quantization of the inertial field becomes identical to that for the electromagnetic field. The inertial field, of course, like the photon field, may be gravitationally affected, but, in the case of point-particles interacting with other point particles (which is the only true definition of a quantum system), the Schwarzschild solution ensures that this effectively results in one doubling due to length contraction and another due to time dilation, increasing the total inertial effect by the numerical value 4, but leaving it unchanged in structure.

It has already been demonstrated in previous publications that we can treat the purely linear effects predicted by the Schwarzschild solution for point sources as inertial effects due to the use of the gravitationally-affected metric of measurement for a system which does not require it, and the effect has been described as the 'aberration of space' [Rowlands, 1992, 1994a, b]. The fact that measurement between discrete sources takes place at the velocity of light, or some gravitationally-affected version of it, does not prove that gravity requires a Lorentzian metric for its transmission. It is completely consistent with all known facts and mathematical theories to describe the general relativistic 'curvature' equations as a convenient way of describing the inertial effects produced by choosing a metric determined by measurement, in the most general form, rather than as an intrinsic description of the relationship between space and time required by gravity itself.

The result of this will be a shift in our fundamental perception of the relationship between gravity and inertia, and General Relativity will only be meaningful in terms of the combined theory. Gravity will be, in principle, unobservable directly except through the fictitious inertial reaction it produces between discrete sources. This inertial reaction will, of course, be repulsive, like other inertial forces. (It is fictitious in the sense that it is produced by a noninertial frame of reference, not in the sense that it is not directly felt.) Because inertia is repulsive, there will be no nonlinearity and no spin 2 gravity. Gravity will not be quantized, as such, but the gravitational inertial reaction will be, and will be a direct analogue of the electromagnetic force. The inertial interaction between identical discrete sources will be repulsive, and the quantum of the interaction will be spin 1, like the photon, the twenty-fifth generator in a $U(5)$ grand unification [Rowlands and Cullerne, 2001c, d]. The static component of the inertial reaction between discrete sources will be identical in magnitude to the gravitational interaction, but will be transmitted at the speed of light (the static component, of course, carries no information about speed of transmission).

Formal development of the theory produces a fictitious 'gravomagnetic' field which is purely inertial, and has no gravitational component, though it is gravitationally *affected*, from the localised component of gravity from discrete sources. There will be a corresponding set of Maxwell equations and a corresponding quantum dynamics, but, as a result of the action of gravity, producing both length contraction and time dilation, the 'magnetic' terms, though qualitatively similar to those in electromagnetic theory, differ numerically, for the 'quantum' case of interactions between point particles with spherical symmetry, by the factor 4. [Rowlands, 1992, 1994a, b] The equations, however, will duplicate in form those of QED, and the quantum theory will be renormalizable. The presence of semi-classical gravitational fields involving the interactions of many particles may be accounted for using the curved space-time of general relativity. It is significant here that effects attributable to the 'curvature of space-time' in the Schwarzschild (point-source) solution, such as the double bending of photons by a point gravitational source, can be considered as incorporating vacuum effects, in the same way as fermionic spin requires a double rotation, and the Aharonov-Bohm effect requires a double circuit in the multiply connected space surrounding a discrete component. Both gravitational light deflection and electron spin can be derived relatively simply using Newtonian kinetic energy equations, demonstrating that they represent the 'action' half of a process, whose 'reaction' half comes from the vacuum.

Mach's principle

The linear gravitational theory, which requires a universe that is flat, Euclidean, and infinite, incorporates a version of Mach's principle, in which the inertia of a body will *appear* to be due to the action of forces produced by all other bodies in the universe. (Mach's principle then becomes equivalent to the equivalence principle in asserting the identity of gravitational and inertial mass.) This immediately produces the Hubble redshift, with an apparent acceleration, with deceleration parameter $q_0 = -1$, and no continuous creation of matter. This acceleration was predicted in earlier work, and as is now apparently observed. Sciama, as early as 1953, had investigated the possibility that Mach's Principle could be explained in terms of a gravitational analogue of the acceleration-independent inductive force of electromagnetic theory,

$$F = \frac{Gm_1 m_2 \sin\theta}{c^2 r} \frac{dv}{dt}.$$

To attribute the inertia of a body of mass m to the action of the remaining matter within its event horizon (mass m_u, radius r_u) requires the total sum of all such force terms to be identical with the Newtonian inertial force, $m dv/dt$. Assuming isotropy (to remove the angular dependence, $\sin\theta$), and, making the 'static' component of this force additionally equal in magnitude to the static gravitational attraction due to the same distribution of matter, we find that an object of mass m at a distance r from an observer at the centre of a sphere defined by r_u will be seen to experience an inertial force

$$\frac{Gmm_u}{c^2 r}\frac{dv}{dt} = \frac{Gmm_u}{r_u^{\,2}},$$

with acceleration $$a = \frac{dv}{dt} = \frac{c^2 r}{r_u^{\,2}}.$$

Writing this in the form $$v\frac{dv}{dr} = \frac{c^2 r}{r_u^{\,2}},$$

and integrating with respect to r, we obtain

$$v = \frac{cr}{r_u}, \qquad a = \frac{v^2}{r}.$$

The result, which is independent of any specific cosmological model, and was predicted long before the experimental discovery of the redshift acceleration, also implies that $ar/v^2 = 1$, which is analagous to a value of – 1 for the deceleration parameter, as usually defined, in terms of the scale factor R,

$$q_0 = \frac{R\ddot{R}}{\dot{R}^2}.$$

There is no appeal here to a cosmological constant or a mysterious 'dark energy'. The repulsive force is entirely explicable in terms of similar fictitious forces found throughout physics. It is, however, a vacuum effect, resulting from the connection between the discrete and continuous aspects.

The effect, here, might be described as 'centrifugal'. However, it is likely that there will also be a Coriolis effect (of the form $2m\omega \times v$), for large-scale systems, such as galaxies and galactic clusters, under gravitational rotation. It is widely assumed, for example, that the rotation curves of galaxies and galactic clusters can only be explained by haloes of a mysterious 'dark matter' surrounding them, because the graphs of v against r flatten out of the after an initial linear increase. However, this effect seems to occur, in the case of galaxies, at the point where v^2 / r becomes of order c^2 / r_u, as though it might be reaching a limit, or observational event horizon, at the value of some universal inertial acceleration for rotating gravitating systems. The dramatic increase of the ratio of dark / visible matter ratio with the size of object, from stars, to galaxies, to galactic clusters and superclusters, means that it is impossible to explain the dark matter component of the larger objects with the dark matter contribution of the component smaller ones; however, such an effect would be totally expected on an explanation based on an inertial event horizon determining the structures that are gravitationally possible. (A constant maximum value of v^2 / r might imply that the masses of large gravitating systems, calculated as v^2r / G, would show a proportionality to r^2, leading to the 'fractal' structuring of galaxies and clusters claimed by Roscoe [1999].)

The calculations in this section suggest that inertial mass can be considered as equivalent to the value which *would be* produced by Mach's principle, and, is, in this sense, a manifestation of the vacuum. If gravity and inertia are linked by the Machian theory described here, then the Planck mass will represent a specifically *inertial* quantum, and an event horizon for the time-delayed inertial force, through which inertial mass must *appear* to be generated. So, the event horizon provides a natural cut-off in deriving mass and energy values by renormalization. If the Casimir effect is fundamental in the way described below, then charges must also be a manifestation of Mach's principle in the same way as inertial mass, and the values of the coupling (or fine-structure) constants, like the rest mass values, must be vacuum-determined. Renormalization is precisely this calculation in the case of both mass and charge, and, though it is not usually done this way, it is possible, to calculate self-energies through nonperturbative calculations from zero bare values.

Quantized gravitational inertia

If we assume that 3-dimensionality is the sole source for discreteness in physics [Rowlands 2004b], the mathematical object called a 4-vector will have no physical realisation at the quantum level. Instead, we use the *3-dimensional* nilpotent structure, represented by the terms $\pm ikt \pm i\mathbf{r} + j\tau$ and $\pm ikE \pm i\mathbf{p} + jm$. The first term may be described as the 'quantum metric; the second is the realisation of the Dirac state, and may be regarded as the phase space version of the first. No other fundamental structure is both fully quantum and fully relativistic, and the 3-dimensionality of the structure is essential to both of these conditions.

The interpretation of inertia as the result of the interaction between discrete matter and the continuous gravitational vacuum suggests that it is gravitational inertia rather than gravity which is subject to quantization. On the basis that 3-dimensionality is the sole source for discreteness in physics, we can develop a more formal mathematical theory of quantum gravitational inertia, in which the key structure becomes the 3-dimensional nilpotent structure, represented by $\pm ikt \pm i\mathbf{r} + j\tau$, or its phase-space equivalent, $\pm ikE \pm i\mathbf{p} + jm$. The 3-dimensionality comes from the fact that, for gravitational inertial interactions involving individual fermionic states at the quantum level, there is, of course, an effective reduction or compactification of the spatial dimensions, to a single well-defined parameter (**r**). In this case, we can construct a 2 + 1 theory of gravitational inertia, based on a 3 × 3 'quantum metric' (with **r** and τ representing the 'real' parts and *it* the imaginary), which would be both quantizable and renormalizable.

One way of quantizing gravitational inertia in this way is via the discrete gravity theory based on the idea of *extended causality*, which has been presented by Koberlein [2001, 2002]. Koberlein's presentation (to which I am significantly indebted in this section) can be translated (or even transcribed) almost immediately into nilpotent terms and, in principle, quantized, without changing any of the significant details, because it uses explicit proper time and a single direction for **r**. In this formulation, a single object (particle or field) at two points in Minkowski space-time (represented by the 4-vector x) must satisfy the causality constraint

$$\Delta\tau^2 + \Delta x^2 = \Delta(ikt)^2 + \Delta(i\mathbf{r})^2 + \Delta(j\tau)^2 = 0 \quad , \qquad (1)$$

which defines a hypercone for the object. Extended causality then applies when we shift τ and x by infinitesimal steps $d\tau$ and dx. We then obtain

$$(\Delta\tau + d\tau)^2 + (\Delta x + dx)^2 = (ik\Delta t + ikdt)^2 + (i\Delta\mathbf{r} + id\mathbf{r})^2 + (j\Delta\tau + jd\tau)^2 = 0 \ (2)$$

Defining f as a fibre in the 'spacetime' $x \equiv (i\mathbf{r}, ikt)$, where $f_\mu = dx_\mu \,/\, d\tau = id\mathbf{r} \,/\, jd\tau + ikdt \,/\, jd\tau$ and $f^\mu = -\, d\tau \,/\, dx^\mu = -\, kd\tau \,/\, d\mathbf{r} + iid\tau \,/\, dt$, we may combine (1) and (2) to obtain:

$$\Delta\tau + f\,.\,\Delta x = j\Delta\tau + (id\mathbf{r} \,/\, jd\tau + ikdt \,/\, jd\tau)(i\Delta\mathbf{r} + ik\Delta t) = 0 \; ,$$

while extended causality now requires

$$(\tau - \tau_0) + f_\mu\,(x - x_0)^\mu = j(\tau - \tau_0) + (id\mathbf{r} \,/\, jd\tau + ikdt \,/\, jd\tau)(i\mathbf{r} + ikt - i\mathbf{r}_0 - ikt_0) = 0 \; . \; (3)$$

Suppose now that we introduce a massless scalar field $\phi_f\,(x, \tau) \equiv \phi_f(i\mathbf{r}, ikt, j\tau)$, the extended causality in (3) will constrain it to the hypercone generator. Equation (3) will also induce a direction to the field derivatives, so that

$$\partial_\mu\phi_f = (\partial_\mu - f_\mu\partial_\tau) \equiv \nabla_\mu\, \phi_f \; .$$

With this expression we can now derive a discrete field equation. If χ is the coupling constant and $\rho(x, \tau) \equiv \rho(i\mathbf{r}, ikt, j\tau)$ the discrete field source, then, using standard methods, the action of the field is given by

$$S_f = \int i\, d\mathbf{r}\, dt\, d\tau\, \{ \tfrac{1}{2}\, \eta^{\mu\nu}\, \nabla_\mu\, \phi_f \nabla_\nu\, \phi_f - \chi\, \phi_f\, \rho(i\mathbf{r}, ikt, j\tau)\} \; .$$

The field equation then becomes

$$\eta^{\mu\nu}\, \nabla_\mu \nabla_\nu\, \phi_f\,(i\mathbf{r}, ikt, j\tau) = \rho(i\mathbf{r}, ikt, j\tau) \;\; ,$$

with energy tensor

$$T^{\mu\nu}_f = \nabla^\mu\phi_f \nabla^\nu\phi_f - \tfrac{1}{2}\, \eta^{\mu\nu}\, \nabla^\alpha\phi_f\, \nabla_\nu\phi_f \; .$$

The solution can be conveniently expressed in terms of a Green's function. Here we write:

$$\phi_f(i\mathbf{r}, ikt, j\tau) = \int id\mathbf{r}\, dt\, d\tau\, (i\mathbf{r}' + ikt')\, G_f(i\mathbf{r} + ikt + j\tau - i\mathbf{r}' - ikt' - j\tau')\, \rho(i\mathbf{r}', ikt', j\tau')$$

and $$\eta^{\mu\nu} \nabla^{\mu} \phi_f \nabla^{\nu} \phi_f G_f(\mathbf{i}\mathbf{r} + ikt + j\tau) = i\delta\mathbf{r}\,\delta t\,\delta\tau \ .$$

Using the Heaviside step function, Θ, with $b = \pm 1$, the Green's function is then

$$G_f(x, \tau) = \tfrac{1}{2}\, \Theta(bf^4 t)\, \Theta(b\tau)\, \delta(\tau + f \, . \, x)$$

or

$$G_f(\mathbf{i}\mathbf{r} + ikt + j\tau) = \tfrac{1}{2}\, \Theta(b(\mathbf{i}d\mathbf{r} \,/\, jd\tau + ikdt \,/\, jd\tau)ikt)\, \Theta(bj\tau)$$
$$\delta(j\tau + (\mathbf{i}d\mathbf{r} \,/\, jd\tau + ikdt \,/\, jd\tau)(\mathbf{i}\mathbf{r} + ikt)) \ . \qquad (4)$$

Even in classical, discrete form, this equation is independent of the transverse components of the field, paralleling the quantum reduction to a single well-defined direction of spin.

If we now take a single scalar charge $q(\tau)$, with world line $z(\tau)$, as a field source, then:

$$\rho(x, t_x = t_z) = q(\tau_z)\, \delta^{(3)}(x - z(\tau_z))\, \delta(\tau_x - \tau_z) \ ,$$

and the solution for the emitted field becomes:

$$\phi_f(x, t_x) = \chi \int d\tau_y\, \Theta(t_x - t_y)\, \Theta(\tau_x - \tau_y)\, \delta[t_x - t_y + f \, . \, (x - y)]\, q(\tau_z) \ ,$$

which reduces to

$$\phi_f(\mathbf{i}\mathbf{r}, ikt, j\tau) = \chi\, q(\tau)\, \Theta(t)\, \Theta(\tau) \big|_f \ ,$$

or

$$\phi_f(\mathbf{i}\mathbf{r}, ikt, j\tau) = \chi\, q(\tau) \big|_f \ ,$$

when $\tau \geq 0$ and $t > 0$. In the discrete model, the emission or absorption of a field causes a discrete change in q, and we can apply the standard techniques appropriate to quantum field theory, and, in particular, QED, to develop the formalism for interactions at higher orders, the 2 + 1 nature of the theory ensuring its renormalizability.

Zero-point energy and acceleration

If the inertial component is itself gravitationally affected, the quantum metric can be seen as generating off-diagonal 'curvature' terms, which will be entirely equivalent to the effects of potentials on the phase space or Dirac state equivalent, and these will, in turn, determine the nature of the functional term required to make the Fourier transformation between the quantum metric and the Dirac state. Since the Dirac state directly determines the nature of the vacuum which responds to it, we can also consider the process in relation to the zero-point energy, where it becomes equivalent to generating an accelerated frame, with proper acceleration a with respect to a Lorentz frame. In these circumstances, the spectrum is no longer Lorentz-invariant, the plane waves become distorted, and the spectral density becomes

$$\rho(\omega) = \frac{dU}{d\omega} = \frac{\omega^2}{\pi^2 c^3}\left(1 + \left(\frac{a}{\omega c}\right)^2\right)\left(\frac{\eta\omega}{2} + \frac{\eta\omega}{exp(2\pi c\omega / a) - 1}\right),$$

an expression determined from the space-time properties of an accelerating reference frame. The accelerating observer now sees a black body spectrum of the form

$$\left(\frac{\eta\omega}{2} + \frac{\eta\omega}{exp(2\pi c\omega / a) - 1}\right) = \left(\frac{\eta\omega}{2} + \frac{\eta\omega}{exp(\eta\omega / kT) - 1}\right),$$

with temperature T defined by $\eta a / 2\pi k$ (the Davies-Unruh effect). Hawking radiation is a special consequence of the general phenomenon, in which the acceleration is provided by the local gravitational field $g = - GM / r_s^2$, at the edge of a black hole. (Because of the curvature, vacuum, or 'kinetic energy' effect, only one half of the fermion-antifermion pair produced by the temperature increase is released.) The Davies-Unruh effect provides an opportunity for detecting an absolute inertial acceleration against a uniform temperature background, a way of defining a universal reference frame, in the same way as the almost equally isotropic cosmic microwave background radiation. It is significant, however, that it cannot detect a gravitational one, for, in general relativity, only gravitational acceleration is relative, inertial acceleration requiring an absolute space-time.

In addition to the black-body correction to $\frac{1}{2}\eta\omega$, acceleration also causes a change to the density of states, with an additional acceleration-dependent term:

$$\frac{\omega^2}{\pi^2 c^3} \to \frac{\omega^2}{\pi^2 c^3}\left(1 + \left(\frac{a}{\omega c}\right)^2\right).$$

Puthoff [1989] has proposed that this correction may be used, together with a cut-off in the zero-point field (presumably at the Planck mass), to generate a long-range 'Van der Waals'-type force with the inverse-square characteristics of gravity. We can, in fact, see this force as the static inertial repulsion for discrete sources, which is numerically equal to the gravitational interaction, but is generated by a cut-off in the zero-point field, as gravity itself is not.

Following on from this, Haisch and Rueda [1994] have shown that the Davies-Unruh anisotropy in the distribution of vacuum fluctuations is equivalent to a non-vanishing Poynting vector, which leads to a resistance to acceleration which may be interpreted as inertia or inertial mass, according to Newton's second law, $F = ma$. Haisch and Rueda also use a zero-point cut-off, and their employment of the Poynting vector may be considered as equivalent in principle to the use of a 'gravomagnetic' inertial field to generate the inertial mass within the discrete event horizon defined by the limiting velocity c. [Rueda and Haisch, 1998 a, b] This argument depends on the linearity of the gravitational field being determined by the fact that the zero-point cut-off applies only to the discrete inertial repulsion, and not to the continuous gravitational attraction.

In fact, the general conception of continuity for energy suggests that the real energy density (and gravitational energy) is infinite in infinite space, while the parallel continuity of time suggests that it also exists in infinite time. The infinite gravitational energy is, of course, never observed because observation requires discreteness (and discreteness introduces the cancelling energy of inertia), but infinite energy density is required even where we observe a finite energy value, such as the Higgs field, for the discrete weak 'partition', because this has a constant value at every point in space. A consequence of the continuity which makes energy infinite is that this energy is never actually 'transferred'. Virtual vacuum energy does not move. It is identical in all places at all times. What we actually observe being transferred or localized is the discrete partitioning due to charge or inertia (measured through angular momentum), and this transfer is absolute against the background.

Charge occupancy

Charge 'occupancy' (1 / 0) is determined by phase [Rowlands, 2003]. We can write baryon nilpotents in the general form:

$$\begin{array}{c} \textit{inertial} \\ \textit{strong} \\ \textit{weak} \\ \textit{electric} \end{array} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ -ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \end{pmatrix} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ -ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \end{pmatrix} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ -ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \end{pmatrix}$$

or

$$\begin{array}{c} \textit{inertial} \\ \textit{strong} \\ \textit{weak} \\ \textit{electric} \end{array} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \\ -ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \end{pmatrix} \begin{pmatrix} ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ -ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \end{pmatrix} \begin{pmatrix} ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_3 + jm \\ -ikE \pm i\boldsymbol{\sigma}.\mathbf{p}_2 + jm \\ -ikE \,\mu\, i\boldsymbol{\sigma}.\mathbf{p}_1 + jm \end{pmatrix}$$

However, for any given direction of σ (in principle, defining the inertial phase) only certain $\mathbf{p}$ phases will be 'active'. These representations are similar to those for charges in the particle tables. The strong charge goes through all possible phases, while the weak and electric remain relatively (although not absolutely) fixed on single phases. In a baryon, the weak and electric phases must be on different quarks. The inertial element again goes through all possible phases in fixing the direction of spin σ. Mesons have the same structure as baryons, except that they are single fermions combined with the corresponding antifermions and the three phases ('colours') should be considered in a purely temporal sequence. The weak and electric charges 'switch on / off' as the phase changes through the components 1, 2, 3.

For free fermions (leptons), the phases are purely the inertial phases. Only the direction of vector properties of **p**, of course, define a strong phase – the magnitude is determined by the combination of *E* and *m*. For free fermions, there is no strong charge because no information is carried about direction, and there is no *SU*(3) symmetry. Leptons have weak and electric occupancy on the same phase, with a temporal cycle, 1-2-3, as the structure rotates through the three directions involved in **p**.

We can consider *iE*, **σ.p** and *m* as the respective coefficients for the weak, strong and electric vacuum terms. There are two pseudoscalar terms ± *iE*; six vector terms $\pm \boldsymbol{\sigma}.\mathbf{p}_1$, $\pm \boldsymbol{\sigma}.\mathbf{p}_2$, $\pm \boldsymbol{\sigma}.\mathbf{p}_3$; and one scalar term *m*. The weak component switches in such a way as to make *iE* into –*iE*. The strong switches in such a way as to make $\boldsymbol{\sigma}.\mathbf{p}_1$ convert to $-\boldsymbol{\sigma}.\mathbf{p}_1$, and also to $\boldsymbol{\sigma}.\mathbf{p}_2$; $-\boldsymbol{\sigma}.\mathbf{p}_2$; $\boldsymbol{\sigma}.\mathbf{p}_3$; and $-\boldsymbol{\sigma}.\mathbf{p}_3$. The weak transition involves dipolarity. The strong transition requires a constant rate of change of **p**, which is equivalent to a linear potential. The electric component preserves *m*. The respective group structures are *SU*(2), *SU*(3) and *U*(1).

Only *iE* and **σ.p** vary, and only **σ.p** varies in magnitude with phase. The *E* terms are, therefore, always global. There are two ways of setting up these transitions, both using covariant derivatives for the operators *iE* and **σ.p**. We can either set up a combination of (pseudo)scalar and vector group generators; or, using the idea that these groups represent the spherical symmetry of a point source, and are covariant, we can replace the scalar and vector parts by scalar potential functions of *r*, associated with *E*. In the first case, the scalar parts are scalar phase (Coulomb) terms; the vector parts are the generators that make the individual interactions have the *SU*(3), *SU*(2) or *U*(1) symmetries associated with the **p**, *E* and *m* operators in the nilpotent, or with the direction, handedness or radial magnitude of the angular momentum. In the second case, we can group all terms related to a single particle under a single representation of *E* as a scalar function of *r*, which applies globally to the entire state.

The two *SU*(2) states – filled electric and empty electric background – being global, are automatically set with respect to *E*. That is, the background is incorporated as the potential producing a scalar or *U*(1) phase in the *E* term. In the case of spherical symmetry, this becomes a Coulomb potential. It is possible to combine all the information into a single expression by using the fact that Lorentz invariance, in the case of a purely point source with spherical symmetry, allows us to transfer all the information contained in **σ.p** to the E term by adding a potential function of r which reproduces the specific aspect of spherical symmetry ($SU(3)$, $SU(2)$ or $U(1)$) incorporated in the covariant part of **σ.p**, that is, the part responsible for the interaction. When the frame is chosen such that all this information is transferred to the E term, then all specific phase information is lost. The rotation of vector **p** terms ensures that the strong term is a linear function ($\propto r$). The scalar nature of m ensures that the electric term is a scalar phase (Coulombic) ($\propto 1 / r$). The dipolarity of $\pm iE$ ensures that the weak term is a dipolar equivalent of the scalar phase ($\propto 1 / r^3$). These options are also evident as a result of directly applying the condition of spherical symmetry to the fermionic state.

A string theory without strings

Superstring and membrane models have been claimed as offering the best candidate for a unified theory of physics, though it is also believed that the unifying theory will probably not be any of the five classes of string theory currently known, but a more fundamental, unifying theory of which these are model-dependent approximations under specific assumptions. An ideal string or membrane theory would, then, appear to be one which removes the model-dependent assumptions, in effect a string theory without strings.

10 space-time dimensions are apparently required to construct a quantum field theory of superstrings in which all anomalies cancel out, while an eleventh dimension is required to extend to supermembranes embedding all the different classes of string theory. The nilpotent Dirac theory provides exactly these requirements. Each nilpotent represents 10 conserved quantities and so could be constructed in a 10-dimensional phase space:

energy	3 components of momentum	rest mass
weak charge	3 components of strong charge	electric charge

This set of 10 'dimensions' incorporates a fundamental duality involving vacuum. All particles are dual with the vacuum, and only exist in relation to it (*zitterbewegung* being the dynamic manifestation of this), and so we require ten pieces of information at the same time for a full specification of a particle state. The energy and the charge components appear as mutually exclusive occupiers of the vacuum and material aspects. In principle, one set of five components represents the particle and the other set the dual vacuum state, or one set represents amplitude and the other phase. All ten, however, are needed to specify the state, and to convert from phase-space to 'real' space, we would simply use the conjugate metric ($\pm kt \pm ii\mathbf{r} + ij\tau$). It is significant that six 'dimensions' (all except E and $\mathbf{p}$) are fixed or compactified, exactly as required in string theory, and also that they constrained to symmetries that are spherical in origin, like the $U(1)$ symmetry in Kaluza-Klein theory, which corresponds here to the special case of electric charge.

The eleventh or 'membrane' dimension is the commutative Hilbert space linking all the nilpotents (see Appendix), which, significantly, is connected to gravity and instantaneous correlation. We should, nevertheless, recognize that both 10- and 11-dimensional models are, really, manifestations of a more fundamental 3-dimensionality, in being determined by the three quaternion operators $\boldsymbol{k}$, $\boldsymbol{i}$, $\boldsymbol{j}$. A quantum nilpotent structure can always be given a 3-dimensional representation, through the affine nature of the $\mathbf{p}$ or s operator. Only one direction for spin is well defined and only one state for the colour charge. This understanding, as we have seen, allows us to develop a renormalizable theory of quantum gravity or quantum gravitational inertia.

String theories are, by definition, also supersymmetric. It is, of course, the unbroken supersymmetry of the Dirac nilpotent which allows us to define energy and charge states simultaneously. In principle, an unbroken supersymmetry requires the vacuum to have zero total energy, which is what we expect if the total energy due to the matter is exactly cancelled by negative gravitational energy. Spontaneous symmetry-breaking, in this interpretation, is not, then, due to the overall state of the vacuum, but to the discrete weak vacuum, which, via the Dirac / Higgs mechanism, privileges $+E$ states over $-E$ for discrete matter. Significantly, the superspace needed for supersymmetry postulates four antisymmetric coordinates as superpartners of space-time; here they become mass and the three charges. The eight coordinates together constitute the superspace, which, in this formalism, takes on the character of the nilpotent Dirac algebra.

The Casimir effect

The manifestation of the continuous vacuum that we *do* observe, is the well-known Casimir force of attraction between uncharged metal plates of area A and small separation d:

$$F = \frac{\pi hc\ A}{480\,d^4}\,.$$

Because of the dependence on $1\,/\,d^4$, this force manifests itself over the range 1 μm as a dipole-dipole interaction, of exactly the same kind as the Van der Waals force of cohesion between molecules. This interpretation assumes zero-point fluctuations of virtual photons in the space between the plates or molecules, but it is equally possible to obtain the same result using zero-point fluctuations of the electrons in the metal surfaces. In this case it becomes the London dispersion interaction. Yet another picture (Hellmann-Feynman) sees the quantum charge clouds in the two plates, molecules or other objects becoming deformed as they approach, corresponding to a change in the expectation values of their charge distributions. In this case, the force is identical to that of chemical bonding due to the classical electrostatic force.

The Casimir force is thus not a distinct phenomenon, but an aspect of the classical electromagnetic interaction. While Peterson and Metzger [2004] use this as a means of removing such unobservables as quantum fluctuations from the argument, we can turn the argument round so that the ordinary electromagnetic force becomes a vacuum projection. An inverse fourth power Casimir force between objects, which are electrically neutral globally but composed locally of electrostatic dipoles, would then imply an inverse square force between the individual charged particles of which they are composed. And related effects of aggregated matter, such as nuclear forces, could be described in similar terms as Casimir-type manifestations of vacuum fluctuations, either fermionic or bosonic, as much as interactions between discrete charges specified by expectation values. It may well be that the left-handed chirality which is observed in aggregated matter (particularly organic) is due to the weak (Van der Waals force) ultimately producing the states of matter.

If we describe the forces due to discrete charges (electric, strong, weak) as Casimir-type manifestations of the vacuum, then we have a direct physical interpretation for the respective use of the quaternion operators $\boldsymbol{j}$, $\boldsymbol{i}$, $\boldsymbol{k}$ both for these three charges, and for the operation of the respective electric, strong, weak vacua via $\boldsymbol{j}(\pm\, \boldsymbol{i}\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$, $\boldsymbol{i}(\pm\, \boldsymbol{i}\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$, $\boldsymbol{k}(\pm\, \boldsymbol{i}\boldsymbol{k}E \pm \boldsymbol{i}\mathbf{p} + \boldsymbol{j}m)$. Because the operators are attached respectively to pseudoscalar E, vector $\mathbf{p}$ and scalar m, in the state vector, then their vacua will have different effects, and so the forces will behave differently. However, the key driving mechanism in all Casimir calculations is that they are the result of separating out *discrete* objects from a *continuous* background, and that they only have meaning in the context of object pairs. Creating a discrete object pair at some finite separation generates a force because it creates a discrete space which is shielded from some of the modes of vacuum vibration outside this space. In principle, therefore, all interactions between discrete charged objects, and even the values of the charged coupling constants, can be seen as resulting from the *existence* of the rest of the universe as a vacuum state, exactly in line with renormalization and Mach's principle for the parallel case of inertial mass.

In this interpretation, the Casimir and related effects become the way in which the discrete charged vacua manifest themselves in relation to the continuous total vacuum background; they represent the partitioning of the vacuum through the three types of charge state. The occupation status of the charge states (that is, whether the charges have unit or zero values) is established on the basis of relative phases between the components of the state vector. This then determines particle type and possible interactions. The vacuum, however, becomes the mechanism by which this process becomes manifested; the creation of discrete units with non-zero occupation status creates the 'distortions' of vacuum, which we call interactions, in the same way as the presence of discrete sources creates the vacuum response or distortions of simply-connected space which we call the Aharonov-Bohm effect and the Berry phase. The charges thus act through the Casimir effect in the vacua created by the quaternion operators, and these vacua have the characteristics necessary to make them components of the total gravitational vacuum.

Masses of particles

Charge and inertial mass are alternate localizations of the vacuum. So, we might expect the inertial masses generated by the Higgs mechanism to be determined by the fine structure constants which fix the vacuum energy scales attributable to the charges. The grand unification process described in earlier work [Rowlands and Cullerne, 2001c, d] allows the calculation of the four fine structure constants, α, α_2, α_3, α_G, to be made absolutely at any given value of energy, the four unknowns being determined uniquely from four basic equations, or their extensions at higher orders of perturbation theory. In principle, we can define, also, an energy (mass) scale, at which α_3, the only constant representing an unbroken symmetry, has unit value, so defining a strong charge / mass, or any idealized perfect charge / mass equivalent, with perfect phase rotation. At this value of α_3, however, perturbation theory breaks down and the calculation can only be done approximately. The evidence indicates, however, that there is a fundamental value of mass appropriate to the zero charges in composite states possibly equal to about 70 MeV. Let us describe this as m_f.

The substitution of units of m_f for missing charges seems to be successful in providing the main mass values for the baryon and meson states composed of quarks of the first two generations, as structured in the tables. Each zero charge is replaced by inertial mass m_f, within the $SU(3)_f$ and $SU(4)_f$ representations. The masses for the meson octet and the baryon octet and decuplet have been derived and discussed extensively in earlier publications. The principle, however, can also be extended to the weakly interacting bosons. If we take the Higgs boson as the complete vacuum bosonic state for all zero charges and all representations (A, B, C, L), we have 6 flavours, 6 anti-flavours, 3 colours (or equivalent phase states), 3 charge types for each quark / antiquark, and 2 for each quark-antiquark pairing, over 4 representations (*A*, *B*, *C*, *L*), we obtain a total of 2592 zeros and a mass of 2592 m_f. This suggests $m_H = 182$ GeV, which is in the range of values required for supersymmetry to remain unbroken (approximately 170-180 GeV), as the nilpotent representations suggests it should. If we assume that the electroweak interaction cannot distinguish between *A*, *B* and *C*, and so perform the same calculation over 2 representations, we obtain $m_Z = 91$ GeV, and can then use $M_W = M_Z \cos\theta_W$ (at the appropriate energy) to find M_W. If m_H really is 182 GeV then we should expect the most favoured decay mode for the Higgs boson to be via the two Z^0 or four lepton route.

In the standard treatment of the Higgs mechanism, fermion masses *m* are generated by (harmonic oscillator) couplings g_f to the Higgs field of the form

$$g_f = \frac{e}{\sqrt{2}\, sin\theta_W} \frac{m}{m_W}$$

Maximal weak coupling may be expected to occur in the third generation, where the total mass appears to be $\Sigma m = M_H = 2\, M_Z$. Taking $M_W = M_Z \cos\theta_w$, $\sin\theta_w$ at $M_Z = 0.5$, and the weak coupling constant, $g = e / \sin\theta_w$, we obtain, for this coupling

$$\sum g_f = \frac{g}{\sqrt{2}} \frac{2}{cos\,\theta_W} = \sqrt{\frac{8}{3}} g\,,$$

so making the total coupling producing the fermion masses in this generation *directly determined* by the weak coupling. The masses can then be related to M_H in the ratio of the Higgs coupling to the weak coupling:

$$\frac{m}{m_H} = \frac{g_f}{g} \sqrt{\frac{3}{8}}\,.$$

Assuming that the weak coupling remains fixed, irrespective of generation, it would seem that the electroweak masses of the two lighter fermion generations are separated from this one by successive factors of the order of the electromagnetic fine structure constant, α.

Since the vacuum charge state determines the inertial mass value, we can imagine that the mass deriving from the filled electric vacuum is related to the fundamental mass m_f by the factor α, the electric coupling, and so is given by αm_f. Since the electron appears to be the purest example of mass being acquired by this mechanism, we may equate αm_f with m_e, and so $m_f = m_e / \alpha$. According to this argument, the mass of the electron comes from the electric vacuum state being empty, whereas it is filled in the case of the neutrino. Standard theory also determines the vacuum expectation value for the Higgs field at $f = M_W / 2g$, where g is the weak coupling constant expressed in charge units. Since $g = e / \sin \theta_W$, with e the electrical coupling, then $g = 2e$ if $\sin^2 \theta_W = 0.25$. Taking account of the fact that unit electric and weak charges are divided between three phases, we obtain f of order 3 M_W or 246 GeV.

The definition of m_f, and hence of m_e / α, as a fundamental unit of mass, also gives fundamental status to the unit of length known as the 'classical radius' of the electron, which is defined as $r_c = e^2 / 4\pi\varepsilon_0 m_e c^2 = \eta / m_f c$, and becomes the Compton wavelength for m_f. The ratio of the Planck mass to m_f, which is also the ratio of the 'classical radius to the Planck length, is a significant dimensionless number, of order 1.74×10^{20}. The 'classical radius', of course, has no physical meaning in relation to the point-like electron, but a uniform spherical distribution of charge e would produce a mass equivalent m_e over the radius $r_e = 3\, r_c / 2 = 4.2289 \times 10^{-15}$ m. If we consider the (electric) vacuum as the virtual equivalent of a uniform distribution of this charge, then it is interesting to note that the gravitational density of m_e over this radius, $3Gm_e^2 / 4\pi r^4 = 4.132 \times 10^{-14}$ Jm^{-3}, and this is *exactly* the energy density that would be produced by a uniform distribution of black body radiation at a temperature of 2.72 K, and determined by aT^4, where $a = 7.56 \times 10^{-16}$ $\text{Jm}^{-3}\text{K}^{-4}$.

Appendix: The origins of the nilpotent representation and vacuum

From the nilpotent representation, it is possible to show that there must be such things as vacuum states. But what is the real origin and meaning of vacuum? To answer this question, we may invoke a process for generating the parameter group and the nilpotent state using a principle which is analogous to the concept of rewrite mechanism, as used in computing [Rowlands and Diaz, 2002; Diaz and Rowlands, 2004]. Making no prior mathematical assumptions, in particular any conventional concept of mathematics / numbering, the mathematical structure underlying physics can be shown to rewrite itself from zero totality. We assume 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. The system rewrites its own rules on the basis that each state is a zero totality state but is not a unique zero.

The rewriting proceeds by defining a zero totality system to contain nothing new within it (symbolized by $\rightarrow$) and a new zero totality outside it (symbolized by $\Rightarrow$). If we describe the totality, as so far understood, as an 'alphabet', then an examination (here represented by 'concatenation') 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) $\rightarrow$ (alphabet) *i.e. there is nothing new*
(alphabet) (alphabet) $\Rightarrow$ (new alphabet) *i.e. the zero totality is not unique*

An absolutely key fact is that the maintenance of a total zero state requires 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. 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:

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

Applying $\rightarrow$ to this alphabet,

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

and so

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

However, (R, R*) (R, R*) will, of course, require a new conjugated alphabet, whose additional terms must chosen in such a way that the subalphabets yield nothing new. So

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

Applying $\rightarrow$ to this new alphabet means that

$(R) \quad (R, R^*, A, A^*) \rightarrow (R, R^*, A, A^*) \equiv (R, R^*, A, A^*)$
$(R^*) \quad (R, R^*, A, A^*) \rightarrow (R^*, R, A^*, A) \equiv (R, R^*, A, A^*)$
$(A) \quad (R, R^*, A, A^*) \rightarrow (A, A^*, R, R^*) \equiv (R, R^*, A, A^*)$
$(A^*) \quad (R, R^*, A, A^*) \rightarrow (A^*, A, R^*, R) \equiv (R, R^*, A, A^*)$

That is, to maintain an unchanged alphabet, we arrange that each term cycles into another. NB No concept of 'ordering' is required by concatenation. So, applying $\rightarrow$ to the new alphabet means that

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

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, we can only ensure $\rightarrow$ applies if concatenated *terms*, such as AB, AB*, are in the alphabet:

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

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*) → (R, R*, A, A*, B, B*, AB, AB*)
(R*) (R, R*, A, A*, B, B*, AB, AB*) → (R*, R, A*, A, B*, B, AB*, AB)
(A) (R, R*, A, A*, B, B*, AB, AB*) → (A, A*, R, R*, AB, AB*, B, B*)
(A*) (R, R*, A, A*, B, B*, AB, AB*) → (A*, A, R*, R, AB*, AB, B*, B)
(B) (R, R*, A, A*, B, B*, AB, AB*) → (B, B*, AB, AB*, R, R*, A, A*)
(B*) (R, R*, A, A*, B, B*, AB, AB*) → (B*, B, AB*, AB, R*, R, A*, A)
(AB)(R, R*, A, A*, B, B*, AB, AB*) → (AB, AB*, B, B*, A, A*, R, R*)
(AB*)(R, R*, A, A*, B, B*, AB, AB*) → (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*) → (R, R*, A, A*, B, B*, AB, AB*)
(R*) (R, R*, A, A*, B, B*, AB, AB*) → (R, R*, A, A*, B, B*, AB, AB*)
(A) (R, R*, A, A*, B, B*, AB, AB*) → (R, R*, A, A*, B, B*, AB, AB*)
(A*) (R, R*, A, A*, B, B*, AB, AB*) → (R, R*, A, A*, B, B*, AB, AB*)
(B) (R, R*, A, A*, B, B*, AB, AB*) → (R, R*, A, A*, B, B*, AB, AB*)
(B*) (R, R*, A, A*, B, B*, AB, AB*) → (R, R*, A, A*, B, B*, AB, AB*)
(AB) (R, R*, A, A*, B, B*, AB, AB*) →(R, R*, A, A*, B, B*, AB, AB*)
(AB*)(R, R*, A, A*, B, B*, AB, AB*) → (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) → (R) (*commutative*)
(AB) (AB) → (R*) (*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 (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 A, B, C, D …-type term of anticommuting with them. All other such terms will commute.

And 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, 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, 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 $\Rightarrow$ 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 key connection between 3-D and discreteness is now revealed.

Dimensionality, i.e. anticommutativity, is the source of discreteness in a zero totality universe. An anticommutative system has the property of *closure*, whereas a commutative system remains open to infinity. Geometry is logically prior to algebra.

Defining the integer series through the anticommutative default option additionally allows us to *reinterpret* the undefined (R) in terms of real numbers, and the process of concatenation as multiplication, and that of conjugation as algebraic negation. We can also set up a system of integral units for the entire algebraic structure, in which the units of A, B, C ... would become quaternionic square roots of -1.

order 2 ± 1

order 4 $\pm 1, \pm i_1$

order 8 $\pm 1, \pm i_1, \pm j_1, \pm i_1j_1$

order 16 $\pm 1, \pm i_1, \pm j_1, \pm i_1j_1, \pm i_2, \pm i_2i_1, \pm i_2j_1, \pm i_2i_1j_1$

order 32 $\pm 1, \pm i_1, \pm j_1, \pm i_1j_1, \pm i_2, \pm i_2i_1, \pm i_2j_1, \pm i_2i_1j_1, \pm j_2, \pm j_2i_1,$
$\pm j_2j_1, \pm j_2i_1j_1, \pm j_2i_2, \pm j_2i_2i_1, \pm j_2i_2j_1, \pm j_2i_2i_1j_1$

order 64 $\pm 1, \pm i_1, \pm j_1, \pm i_1j_1, \pm i_2i_1, \pm i_2i_1, \pm i_2j_1, \pm i_2i_1j_1, \pm j_2,$
$\pm j_2i_1, \pm j_2j_1, \pm j_2i_1j_1, \pm j_2i_2, \pm j_2i_2i_1, \pm j_2i_2j_1, \pm j_2i_2i_1j_1,$
$\pm i_3, \pm i_3i_1, \pm i_3j_1, \pm i_3i_1j_1, \pm i_3i_2, \pm i_3i_2i_1, \pm i_3i_2j_1,$
$\pm i_3i_2i_1j_1, \pm i_3j_2, \pm i_3j_2i_1, \pm i_3j_2j_1, \pm i_3j_2i_1j_1, \pm i_3j_2i_2,$
$\pm i_3j_2i_2i_1, \pm i_3j_2i_2j_1, \pm i_3j_2i_2i_1j_1$

The infinite series of alphabets can be reinterpreted in these terms as an infinite series of quaternion systems produced by successive processes of conjugation (A), complexification (B) and dimensionalization (C):

	(A)	(B)	(C)	(B)	(C)	(B)
order 2	$(1, -1)$					
order 4	$(1, -1)$	$\times (1, i_1)$				
order 8	$(1, -1)$	$\times (1, i_1)$	$\times (1, j_1)$			
order 16	$(1, -1)$	$\times (1, i_1)$	$\times (1, j_1)$	$\times (1, i_2)$		
order 32	$(1, -1)$	$\times (1, i_1)$	$\times (1, j_1)$	$\times (1, i_2)$	$\times (1, j_2)$	
order 64	$(1, -1)$	$\times (1, i_1)$	$\times (1, j_1)$	$\times (1, i_2)$	$\times (1, j_2)$	$\times (1, i_3)$

Here, *conjugation* introduces opposite algebraic signs; *complexification* multiplies throughout by a single imaginary term; dimensionalization multiplies again by a new imaginary term completing the quaternion set

order 2	conjugation	×	$(1, -1)$
order 4	complexification	×	$(1, i_1)$
order 8	dimensionalization	×	$(1, j_1)$
order 16	complexification	×	$(1, i_2)$
order 32	dimensionalization	×	$(1, j_2)$
order 64	complexification	×	$(1, i_3)$

At the fourth stage, significantly, the cycle begins to repeat. So we have:

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

To incorporate the first four as independent units of a single universal system we need order 64 (the Dirac algebra), and we can recognize the four as *introducing* the units of the physical quantities which make up the nilpotent state vectors:

order 2	mass	real scalar
order 4	time	pseudoscalar
order 8	charge	quaternion
order 16	space	multivariate vector

The mathematical structure we have generated using the rewrite system is related closely to Clifford algebra, and it is possible to see that this can be taken as an origin for the conventional mathematics based on defining a number system. The conversion to a specifically physical application occurs when we apply the whole structure (i.e. all possible zero totality alphabets) at once, while using the fact that a repeating structure can be recognised after the first four alphabets. This requires the constraint that the information contained within the alphabets allows an immediate return to zero within a universal perspective. It is significant that the Dirac state uses a compactified form of the Dirac algebra, with the components projected onto a 3-dimensional operator. This enables the terms of the first two alphabets to be *extrinsically* structured according to the discrete numbering system which is introduced only with the third, while retaining their *intrinsically* continuous (unstructured) character, and so preserves the interpretation of higher order alphabets as extensions of lower order ones.

In thus applying a unit or numbering structure (derived from 3-dimensionality) to the original alphabetic categories, we specify that they must be algebraically related. Thus, we force an algebraic connection between the units ***i***, ***j***, ***k***, 1 and *i*, and the only structure which can accommodate this connection is itself 3-dimensional. So we end with an interlocking of 3 × 3-D systems, one of which remains incomplete, leading to two of the four parameters remaining unconjugated (nonconserved). The Dirac state, therefore, is not complete in itself. It is only completely specified in partnership with the 'rest of the universe'. In principle, this must create zero totality. In the special case where ($\pm$ ***ik****E* $\pm$ ***i*****p** + ***j****m*) is a nilpotent, the conjugate or dual state, – ($\pm$ ***ik****E* $\pm$ ***i*****p** + ***j****m*), becomes the 'rest of the universe' because both the superposition, ($\pm$ ***ik****E* $\pm$ ***i*****p** + ***j****m*) – ($\pm$ ***ik****E* $\pm$ ***i*****p** + ***j****m*), and the combination – ($\pm$ ***ik****E* $\pm$ ***i*****p** + ***j****m*) ($\pm$ ***ik****E* $\pm$ ***i*****p** + ***j****m*) = 0 for a nilpotent state.

Summarising, we could say that the requirement of simultaneous validity for the first four alphabetic structures creates problems in representing discrete quantities with complete quaternionic structures and continuous quantities with incomplete ones. In effect, we can only create zero totality for all the first four alphabetic structures simultaneously if they are packaged in such a way that we obtain nilpotent solutions for the combined state ($\pm$ ***ik****E* $\pm$ ***i*****p** + ***j****m*), with *E*, **p**, *m* real. The packaging of the four physical components at order 64 in the Dirac nilpotent, however, allows us to introduce a new level of closure, again determined by anticommutativity and 3-dimensionality, with an immediate (potential) return to zero.

With E, **p** and m in the Dirac state represented by real numbers (from the parent quantities, time, space and mass), we can define solutions for which the state is unique (0 if squared). However, the infinite variation within the universal rewrite system determines that there are infinitely many quaternionic systems beyond the three used in creating the Dirac state. How do we incorporate these into the physical picture? The answer is that these are commutative to those used in the Dirac state, so there will be infinitely many Dirac states, each with different commutative coefficients. So the infinite variation within the universal rewrite system determines that each new state introduces a coefficient which is necessarily commutative with those of the nilpotent system. The recipe here is for a quantum mechanical universe based on nilpotent states within an infinite-dimensional Hilbert space (which is conveniently defined by a Grassmann algebra). Only in such a system can the uniqueness of zero, as an infinitely degenerate concept, be maintained. In mathematical form, individual antisymmetric nilpotents ψ_1, ψ_2, ψ_3, …, have coefficients which are unrepeated but arbitrary units or strings of commutative units.

This generates an infinite-dimensional Grassmann algebra, with successive outer products defined by the Slater determinant, with

$$\psi_1 \wedge \psi_1 = 0 \qquad \text{and} \qquad \psi_1 \wedge \psi_2 = -\, \psi_2 \wedge \psi_1 \qquad \text{etc.}$$

The nilpotent units ψ_n must be both nilpotent and antisymmetric, and each must be unique to avoid the trivial case of R = 0 (an immediate return to zero). This infinite-dimensional algebra is equivalent to the complex Hilbert space of conventional quantum theory, with algebraic and nonlocal superposition of fermionic or Dirac states throughout the entire universe. (A possible consequence of an infinity of Dirac nilpotent states with commutative coefficients, constructing a 'universe', is that, from within such a universe, we would have no idea of the internal structure available within the external commutative coefficients. A populist construction of the idea might suggest they could even contain 'parallel universes', though the idea is probably meaningless.)

REFERENCES

Diaz, B. and Rowlands, P. [2004], *Journal of Computing Anticipatory Systems*, in press.

Haisch, B., Rueda, A. and Puthoff, H. E. [1994], *Phys. Rev.* A, **49**, 678.

Koberlein, B. D. [2001], 'Extended causality as a model of discrete gravity', in K. G. Bowden (ed.), *Implications* (*Proceedings of ANPA 22*), 56-63.

Koberlein, B. D. [2002], 'Extended causality as a model of discrete gravity', in K. G. Bowden (ed.), *Correlations* (*Proceedings of ANPA 23*), 18-23.

Peterson, R. and Metzger, R. M. [2004], *Int. J. Chem.*, 7, 1-4.

Puthoff, H. E. [1989], *Phys. Rev.* A, **39**, 2333.

Roscoe, D. E. [1999], *Astron. Astrophys.*, **343**, 696-704, 788-800, 1999.

Rowlands, P. [1983], *Speculat. Sci. Tech.*, 6, 69-80.

Rowlands, P. [1992], 'Is the gravitational field really nonlinear?', in *Proceedings of III Conference on Physical Interpretations of Relativity Theory*, British Society for Philosophy of Science, 311-318.

Rowlands, P. [1994a], *A Revolution Too Far*, PD Publications, Liverpool.

Rowlands, P. [1994b], 'Experimental testing and the physical interpretation of general relativity', *Proceedings of IV Conference on Physical Interpretations of Relativity Theory*, British Society for Philosophy of Science, 298-312.

Rowlands, P. [1999], 'Physics; let's get down to basics', in K. Bowden (ed.), *Aspects II* (*Proceedings of ANPA 20*), 123-134.

Rowlands, P. [2001], arXiv:physics/0106054.

Rowlands, P. [2003], arXiv:quant-ph/0301071.

Rowlands, P. [2004a], *AIP Conference Proceedings*, 78, 102-115.

Rowlands, P. [2004b], 'The origin and meaning of 3-dimensionality', in K. G. Bowden (ed.), *Spin* (*Proceedings of ANPA 25*), 292-323.

Rowlands, P. and Cullerne, J. P. [1999], 'A derivation of particle structures and the Dirac equation from fundamental symmetries', K. G. Bowden (ed.), *Aspects II* (*Proceedings of ANPA 20*), 155-91.

Rowlands, P. and Cullerne, J. P. [2000a], 'The Dirac algebra and charge accommodation', in K. G. Bowden (ed.), *Participations* (*Proceedings of ANPA 21*), 89-111.

Rowlands, P. and Cullerne, J. P. [2000b], '*SU*(5) and Grand Unification', in K. G. Bowden (ed.), *Participations* (*Proceedings of ANPA 21*), 124-146.

Rowlands, P. and Cullerne, J. P. [2001a], arXiv:quant-ph/0103036

Rowlands, P. and Cullerne, J. P. [2001b], arXiv:quant-ph/0109069

Rowlands, P. and Cullerne, J. P. [2001c], arXiv:quant-ph/0106111

Rowlands, P. and Cullerne, J. P. [2001d], *Nuclear Physics* A 684, 713-5.

Rowlands, P. and Diaz, B. [2002], arXiv:cs.OH/0209026.

Rueda, A. and Haisch, B. [1998 a, b], *Found. Phys.*, **28**, 1057; *Phys. Lett.* A, **240**, 115.

Sciama, D. W. [1953], *Monthly Notices RAS*, **113**, 34.

SLAC-PUB-10571
26 July 2004

SCIENTIFIC COSMOLOGY: A Calculation of Cosmological Scale from Quantum Coherence* †

James V. Lindesay‡, jlslac@slac.stanford.edu
H. Pierre Noyes, noyes@slac.stanford.edu
Stanford Linear Accelerator Center MS 81, Stanford University
2575 Sand Hill Road, Menlo Park CA 94025

Abstract

We use general arguments to examine the energy scales for which a quantum coherent description of gravitating quantum energy units is necessary. The cosmological dark energy density is expected to decouple from the Friedman-Lemaitre energy density when the Friedman-Robertson-Walker scale expansion becomes sub-luminal at $\dot{R} = c$, at which time the usual microscopic interactions of relativistic quantum mechanics (QED, QCD, etc) open new degrees of freedom. We assume that these microscopic interactions cannot signal with superluminal exchanges, only superluminal quantum correlations. The expected gravitational vacuum energy density at that scale would be expected to freeze out due to the loss of gravitational coherence. We define the vacuum energy which generates this cosmological constant to be that of a zero temperature Bose condensate at this gravitational de-coherence scale. We presume a universality throughout the universe in the available degrees of freedom determined by fundamental constants during its evolution. Examining the reverse evolution of the universe from the present, long before reaching Planck scale dynamics one expects major modifications from the de-coherent thermal equations of state, suggesting that the

*Work supported by Department of Energy contract DE-AC02-76SF00515.

†Presented at The Twenty-Sixth Annual Meeting of the ALTERNATIVE NATURAL PHILOSOPHY ASSOCIATION Cambridge, England, 31 July - 5 August, 2004

‡Permanent address, Physics Department, Howard University, Washington, D.C. 20059

pre-coherent phase has global coherence properties. Since the arguments presented involve primarily counting of degrees of freedom, we expect the statistical equilibrium states of causally disconnected regions of space to be independently identical. Thus, there is no luminal "horizon" problem associated with the lack of causal influences between spatially separated regions in this approach. The scale of the amplitude of fluctuations produced during de-coherence of cosmological vacuum energy are found to evolve to values consistent with those observed in cosmic microwave background radiation and galactic clustering.

1 Introduction

There is general (although not universal) agreement among physical cosmologists that the current expansion phase in the evolution of our universe can be extrapolated back toward an initial state of compression so extreme that we can neither have direct laboratory nor indirect (astronomical) observational evidence for the laws of physics needed to continue that extrapolation. Under these circumstances, lacking a consensus "theory of everything", and in particular a theory of "quantum gravity", we believe that the prudent course is to rely as much as possible on general principles rather than specific models. This approach is adopted in this paper.

We believe that the experimental evidence for currently accepted theories of particle physics is relevant up to about 5 TeV—the maximum energy or temperature we need consider in this paper. We further assume that our current understanding of general relativity as a gravitational theory is adequate over the same range, and consequently that the cosmological Friedman-Lemaitre (FL) (Hubble) dynamical equations are reliable guides once we have reached the observational regime where the homogeneity and isotropy assumptions on which those equations are based become consistent with astronomical data to requisite accuracy. Although the elementary particle theories usually employed in relativistic quantum field theories have well defined transformation properties in the flat Minkowski space of special relativity, we hold that their fundamental principles still apply on coordinate backgrounds with cosmological curvature. In fact there is direct experimental evidence that quantum mechanics does apply in the background space provided by the Schwarzschild metric of the earth thanks to the beautiful experiments by Overhauser and collaborators[1, 2]. These experiments show that the interference of a single neutron with itself changes as expected when the plane of the two interfering paths is rotated from being parallel to being perpendicular to the "gravitational field" of the earth.

Since quantum objects have been shown to gravitate, we expect that during some period in the past, quantum coherence of gravitating systems will qualitatively alter the thermodynamics of the cosmology. Often, the onset of the importance of quantum effects in gravitation is taken to be at the Planck scale. However, as is the case with Fermi degenerate stars, this need not be true of the cosmology as a whole. By quantum coherence, we refer to the entangled nature of quantum states for space-like separations. This is made evident by superluminal correlations (without the exchange of signals) in the observable behavior of such quantum states. Note that the exhibition of quantum coherent behavior for gravitating systems does not require the quantization of the gravitation field.

The (luminal) horizon problem for present day cosmology examines the reason for the large scale homogeneity and isotropy of the observed universe. The present age of the universe can be estimated from the Hubble scale to be $H_o t_o \cong 0.96 \Rightarrow t_o \cong 13.2 \times 10^9$ years$= 4.16 \times 10^{17}$ seconds. If the size of the observable universe today is taken to be of the order of the Hubble scale $\frac{c}{H_o} \approx 10^{28} cm$, then if the universe expanded from the Planck scale, its size at that scale would have been of the order $\sim 10^{-4} cm$ at that time. Since the Planck length is of the order $L_P \sim 10^{-33} cm$, then there would be expected to be $(10^{29})^3 \sim 10^{87}$ causally disconnected (for luminal signals) regions in the sky. Further, examining the ratio of the present conformal time η_o (which multiplied by c is the distance a photon can travel in a given time) with that during recombination $\eta_*, \frac{\eta_o}{\eta_*} \sim 100$, the subsequent expansion is expected to imply that light from the cosmic microwave background would come from $100^3 = 10^6$ disconnected regions. Yet, angular correlations of the fluctuations have been accurately measured by several experiments[3].

Our approach is to start from well understood macrophysics and end at the onset of microcosmology. We refer to this period as gravitational de-coherence. The FRW scale factor is used to compare cosmological scales with those microscopic quantum scales we are familiar with, which define the lengths of rulers, ticks of clocks, mass of particles, and temperatures of thermodynamic systems. We will insist that our calculations not depend on the present particle horizon scale, which is an accident of history. It will be argued that the equilibration of microscopic interactions can only occur post-decoherence. Global quantum coherence prior to this period solves the horizon problem, since quantum correlations are in this sense supraluminal.

Present data examining the luminosities of distant Type Ia supernovae, which have an understood time and frequency dependency, indicate clearly that the rate of expansion of the universe has been accelerating for several billion years[4]. This conclusion is independently confirmed by analysis of the Cosmic Microwave Background radiation[5]. Both results are in quantitative agreement with a (positive) cosmological constant fit to the data. Our interpretation of this cosmological "dark energy" will be due to the vacuum energy of a quantum coherent cosmology.

One physical system in which vacuum energy density directly manifests is the Casimir effect[6]. Casimir considered the change in the vacuum energy due to the placement of two parallel plates separated by a distance a. He calculated an energy per unit area of the form

$$\frac{\frac{1}{2}\left(\Sigma_{modes}\,\hbar c k_{plates} - \Sigma_{modes}\,\hbar c k_{vacuum}\right)}{A} = -\frac{\pi^2}{720}\frac{\hbar c}{a^3} \tag{1.1}$$

resulting in an attractive force of given by

$$\frac{F}{A} \cong \frac{-0.013 dynes}{(a/micron)^4} cm^{-2}, \tag{1.2}$$

independent of the charges of the sources. Lifshitz and his collaborators[7] demonstrated that the Casimir force can be thought of as the superposition of the van der Waals attractions between individual molecules that make up the attracting media. This allows the Casimir effect to be interpreted in terms of the zero-point motions of the sources as an alternative to vacuum energy. Boyer[8] and others subsequently demonstrated a repulsive force for a spherical geometry of the form

$$\frac{1}{2}\left(\sum_{modes}\hbar c k_{sphere} - \sum_{modes}\hbar c k_{vacuum}\right) = \frac{0.92353\hbar c}{a}. \tag{1.3}$$

This means that the change in electromagnetic vacuum energy *is* dependent upon the geometry of the boundary conditions. Both predictions have been confirmed experimentally.

The introduction of energy density ρ into Einstein's equation introduces a preferred rest frame with respect to normal energy density. However, as can be seen in the Casimir effect, the vacuum need not exhibit velocity dependent effects which would break Lorentz invariance. Although a single moving mirror does experience dissipative effects from the vacuum due to its motion, these effects can be seen to be of 4th order in time derivatives[9].

Another system which manifests physically measurable effects due to zero-point energy is liquid ^{4}He. One sees that this is the case by noting that atomic radii are related to atomic volume V_a (which can be measured) by $R_a \sim V_a^{1/3}$. The uncertainty relation gives momenta of the order $\Delta p \sim \hbar/V_a^{1/3}$. Since the system is non-relativistic, we can estimate the zero-point kinetic energy to be of the order $E_o \sim \frac{(\Delta p)^2}{2m_{He}} \sim \frac{\hbar^2}{2m_{He}V_a^{2/3}}$. The minimum in the potential energy is located around R_a, and because of the low mass of ^{4}He, the value of the small attractive potential is comparable to the zero-point kinetic energy. Therefore, this bosonic system forms a low density liquid. The lattice spacing for solid helium would be expected to be even smaller than the average spacing for the liquid. This means that a large external pressure is necessary to overcome the zero-point energy in order to form solid helium.

Applying this reasoning to relativistic gravitating mass units with quantum coherence within the volume generated by a Compton wavelength λ_m^3, the zero point momentum is expected to be of order $p \sim \frac{\hbar}{V^{1/3}} \sim \frac{\hbar}{\lambda_m}$. This gives a zero point energy of order $E_0 \approx \sqrt{2}mc^2$. If we estimate a mean field potential from the Newtonian form $V \sim -\frac{G_N m^2}{\lambda_m} = -\frac{m^2}{M_P^2}mc^2 << E_0$, it is evident that the zero point energy will dominate the energy of such a system.

For the reverse time extrapolation from the present, we adopt the currently acccepted values[5][10] for the cosmological parameters involving dark energy and matter:

$$h_0 \cong 0.73; \quad \Omega_\Lambda \cong 0.73; \quad \Omega_M \cong 0.27. \tag{1.4}$$

Here h_0 is the normalized Hubble parameter. Note that this value implies that the universe currently has the critical energy density $\rho_c = 5.6 \times 10^{-4} GeV\ cm^{-3}$. We can make the backward time extrapolation with confidence using known physics in the customary way back to the electro-weak unification scale $\sim 100\ GeV$, with somewhat less confidence into the quark-gluon plasma then encountered and beyond the top quark regime, and expect that — unless unexpected new particles and/or new physics are encountered —we can continue up to an order of magnitude higher energies with at most modest additions to the particle spectrum. In this radiation-dominated universe this backward extrapolation (which taken literally *must* terminate when the Friedman-Robertson-Walker scale factor $R(t)$ goes to zero and its time rate of change $\dot{R}(t)$ goes to infinity) is guaranteed to reach the velocity of light $\dot{R}(t_c) = c$ at some finite time t_c when the scale factor $R(t_c)$) still has a small, but finite, value.

As we discuss more carefully below, using our extrapolation beyond the limit just established (i.e. $\dot{R}(t_c) = c$) would seem to conflict with our basic methodological assumption that we invoke no unknown physics. It is true that *as a metric theory of space-time* the curved space-times of general relativity used in the homogeneous and isotropic cosmological models we employ are not restricted in this way. However, if we wish to drive these models by mass-energy tensors derived from either particulate or thermodynamic models relying on some equation of state, and hence the hydrodynamics of some form of matter, we must not use them in such a way as to allow causal signaling at speeds greater than c by non-gravitational interactions. The exception to this stricture which *is* allowed by known physics is that *coherent quantum systems* have supraluminal *correlations* which cannot be used for supraluminal *signaling*. Consequently, we are allowed — as we assert in this paper — to start our examination of the universe at the $\dot{R} = c$ boundary if it is a fully coherent quantum system. Note that the beautiful experiments by Overhauser and collaborators already cited[1, 2] justify our invocation of such systems when they are primarily (or even exclusively) dependent on gravitational interactions. Thus we claim that it is consistent to start our cosmology with the cosmological *decoherence* of a quantum system at the $\dot{R} = c$ boundary. This quantum decoherence process is discussed in detail in Section 3:*Dark Energy De-coherence.*

In Section 2:*Background* we examine an earlier paper[11] which gave cosmological reasons why $\sim 5Tev$ might be the threshold for new physics. This paper was based on much earlier work by E.D.Jones[12, 13, 14] and a more recent collaboration with L.H.Kaufmann and W.R. Lamb[15]. We find that, in contrast to Jones' *Microcosmology* which starts with an expansion from the Planck scale, we can identify the transition from speculative physics to a regime which can be reached with some confidence by the backward extrapolation from the present already assumed above. This section shows that the identification of the critical transition with the de-coherence of a cosmological gravitationally quantum coherent system allows us to recover the semi-quantitative results of the Jones theory *without* having to introduce speculative physics. However, in the development of these results, we were motivated to re-examine the problem from a fresh perspective, as is done in the following section.

In Section 3:*Dark Energy De-coherence* we motivate our explanation of the dark energy driving the observed acceleration of the cosmology as the gravitational vacuum energy density of a zero temperature Bose condensate. Here we are able to more quantitatively reproduce the results of the prior section in an independent manner. Prior to a sub-luminal rate of expansion of the FRW scale factor, we assert that only gravitational and quantum coherence properties are relevant to the dynamics of the expanding cosmology. We will develop a single parameter model in terms of the cosmological constant, and use this to predict a mass and temperature scale for decoherence. We give an argument to support our assumption that the condensate remains at zero temperature during the pre-coherent phase of the cosmology. We will end by examining the expected amplitudes of density fluctuations if such fluctuations are the result of dark energy de-coherence.

Finally, in Section 4: *Discussion and Conclusions*, we discuss the nature of cosmological dark energy, especially with regards to the distant future. It is especially interesting to question the constancy of vacuum energy density after a future gravitational re-coherence event $\dot{R} \geq c$. Some thoughts on our present and future efforts will be given.

2 Motivation

2.1 Jones' Microcosmology

Our present work originated in the re-examination of a paper by one of us[11] emphasizing the likelihood of some threshold for new physics at $\sim 5\ TeV$. This in turn was based on a discussion with E.D. Jones in 2002[14]; our understanding of this discussion and his earlier ideas[12, 13] has been published by us in collaboration with L.H.Kaufmann and W.R.Lamb[15]. Briefly, Jones envisages an extremely rapid ("inflationary") expansion from the Planck scale (i.e.

from the Planck length $L_P = \frac{\hbar}{M_P c} \cong 1.6 \times 10^{-33} cm$, where $M_{Pk} = [\frac{\hbar c}{G_N}]^{\frac{1}{2}} \cong 2.1 \times 10^{-8} kg \cong 1.221 \times 10^{19} GeV/c^2$, and $G_N = \frac{\hbar c}{M_P^2}$ is Newton's gravitational constant) to a length scale $R_\epsilon \sim \frac{1}{\epsilon}$. For the reader's convenience, we will also display the Planck time $T_P = \frac{L_P}{c} \cong 5.4 \times 10^{-44} sec$ and the Planck temperature $\theta_P = \frac{M_P c^2}{k_B} \cong 1.4 \times 10^{32}\,^o K$. Unless necessary for clarity, we will generally choose units such that $\hbar = 1, c = 1, k_B = 1$. This expansion, whose details are not examined in this paper, is characterized by the dimensionless ratio

$$\mathcal{Z}_\epsilon \sim \frac{R_\epsilon}{L_P} \sim \frac{M_P}{\epsilon}. \tag{2.5}$$

When this expansion has occurred, the virtual energy which drives it makes a thermodynamic equilibrium transition to normal matter (i.e. dark, baryonic and leptonic, electromagnetic,...) at a mass-energy scale characterized by the mass parameter m_θ and length scale $\frac{1}{m_\theta}$. Jones uses the energy parameter ϵ as a unit of energy which he calls one *Planckton* defined as one Planck mass's worth of energy distributed over the volume $V_\epsilon \sim \frac{1}{\epsilon^3}$ measured by the scale parameter $R_\epsilon \sim \frac{1}{\epsilon}$. The virtual energy is assumed to consist of N_{Pk} Plancktons of energy ϵ giving total mass $N_{Pk}\epsilon$, corresponding to an energy density $\sim \frac{N_{Pk}\epsilon}{V_\epsilon}$. This virtual energy makes an (energy-density) equilibrium transition to normal matter, so that

$$N_{Pk}\epsilon^4 \sim m_\theta^4. \tag{2.6}$$

By utilizing the Dyson-Noyes argument which follows in the next section, Jones then argued that the number of Plancktonic energy units in the region was related to the dimensionless ratio by $N_{Pk} = Z_\epsilon^2$.

It is assumed that a virtual energy unit ϵ spread throughout the region R_ϵ is left unthermalized (not transitioned to normal matter), and hence can be interpreted as the cosmological constant density ρ_Λ at this and succeeding scale factors. It then can be approximately evaluated at the present day using $\Omega_\Lambda \sim 0.73$ as we show in the Sec. 2.4. In this sense both Jones' theory and ours can be thought of as *phenomenological* theories which depend only on a *single* parameter (outside of the constants from conventional physics and astronomy). We will review in the concluding section some additional *observed or potentially observable* facts that might be *predicted.*

2.2 Dyson-Noyes-Jones Anomaly

We note that our sketch of the Jones theory as expressed by Equation 2.6 introduces two new parameters (N_{Pk},m_θ) which must be expressed in terms of ϵ if we are to justify our claim that this is a single parameter theory. These parameters refer to a very dense state of the universe. According to our methodology, we must be able extrapolate back to this state using only current knowledge and known physics. Jones assumes that this dense state can be specified by using an extension to gravitation[17] of an argument first made by Dyson[19, 20]. Dyson pointed out that if one goes to more that 137 terms in e^2 in the perturbative expansion of renormalized QED, and assumes that this series also applies to a theory in which e^2 is replaced by $-e^2$ (corresponding to a theory in which like charges attract rather than repel), clusters of like charges will be unstable against collapse to negatively infinite energies. Schweber[20] notes that this argument convinced Dyson that renormalized QED can never be a *fundamental theory.* Noyes[17] noted that *any* particulate gravitational system consisting of masses m must be subject to a similar instability and could be expected to collapse to a black hole.

We identify $Z_{e^2} = 1/\alpha_{e^2} \cong 137$ as the number of electromagnetic interactions which occur within the Compton wave-length of an electron-positron pair ($r_{2m_e} = \hbar/2m_e c$) when the Dyson bound is reached. If we apply the same reasoning to gravitating particles of mass m (and if we are able to use a classical gravitational form), the parameter α_{e^2} is replaced by $\alpha_m = G_N m^2 = \frac{m^2}{M_P^2}$; the parameter fixing the Dyson-Noyes (DN) bound becomes the number of *gravitational* interactions within $\hbar/mc$ which will produce another particle of mass m and is given by

$$\mathcal{Z}_m \sim \frac{M_P^2}{m^2} \tag{2.7}$$

If this dense state with Compton wavelength $\lambda_m \sim 1/m$, contains $\mathcal{Z}_m$ interactions within λ_m, then the Dyson-Noyes-Jones (DNJ) bound is due to the expected transition transition $\mathcal{Z}_m m \to (\mathcal{Z}_m + 1)m$, indicating instability against gravitational collapse due to relativistic particle creation.

It is particularly interesting to examine the production channel for the masses m due to $\mathcal{Z}_m$ interactions within λ_m. This is diagrammatically represented in Figure 1. If there are $\mathcal{Z}_m$ scalar gravitating particles of mass m within the Compton wavelength of that mass, a

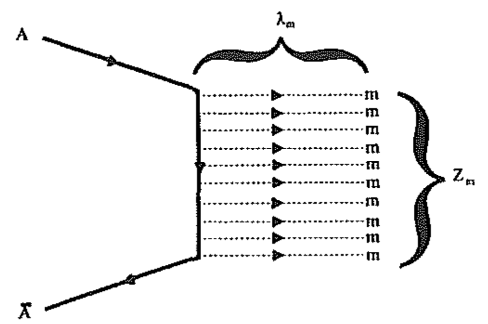


Figure 1: Noyes-Jones collapse of gravitating quanta

particle falling into that system from an appreciable distance will gain energy equal to mc^2, which could produce yet another gravitating mass m. Clearly, the interaction becomes anomalous.

Generally, when the perturbative form of a weak interaction becomes divergent, it is a sign of a phase transition, or a non-perturbative state of the system (eg bound states). One expects large quantum correlations between systems of mass m interacting on scales smaller than or comparable to the Compton wavelength of those masses. We will assume that if the (intensive) number of gravitational interactions (with no more than a Planck mass worth of interaction energy per Planckton, as defined in Section 2.1) of mass units m which can occur within a region of quantum coherence is greater than the DNJ limit, a phase transition into systems with quantum coherence scales of the Compton wavelength of those mass units will occur. We then expect de-coherence of subsystems of vastly differing quantum coherence scales when the Jones transition occurs.

We could be concerned that such a concentration of mass might form a black hole. Although we will be primarily assuming an FRW global geometry, if the Schwarzschild radius of a concentration of mass is considerably less than the proper radial size of the cosmology, one can sensibly discuss smaller regions which approximate a Schwarzschild geometry. For a system of a large number $\mathcal{Z}_m$ of gravitating particles (given by the onset of the DNJ anomaly), the Schwarzschild radius of these masses is given by

$$R_S = \frac{2G_N(\mathcal{Z}_m m)}{c^2} = 2\mathcal{Z}_m \frac{m^2}{M_P^2}\frac{\hbar}{mc} = 2\lambda_m \tag{2.8}$$

Therefore, such a gravitating system of masses would be expected to be unstable under gravitational collapse. We determine the maximum number $\mathcal{Z}_m$ of coherent interaction energy units m beyond which the system will become unstable under gravitational collapse as

$$R_S = \lambda_m \Rightarrow \mathcal{Z}_m = \frac{M_P^2}{2m^2} \tag{2.9}$$

This argument does not assume Newtonian gravitation.

A general comparison of the dependence of the Compton wavelength and Schwarzschild radius on the mass of the system as shown in Figure 2 gives some insight into regions of quantum coherence. To the left of the point of intersection of the two curves, the Compton wavelength is larger than the Schwarzschild radius, and for such localized mass distributions the quantum coherence properties are important in any gravitational considerations. On the other hand, for masses much larger than the Planck mass, the gravitational distance scales are well outside of the quantum coherence scales for isolated masses. For the present

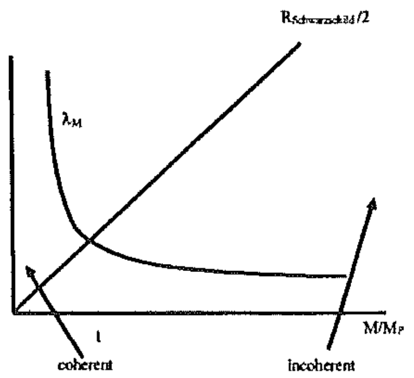


Figure 2: Functional dependence of Compton wavelength and Schwarzschild radius on system mass

discussion, it is the transition region that is of interest. An elementary particle is not expected to have a mass greater than a Planck mass. If the mass were greater than M_P, then the Compton wavelength would be less than the Schwarzschild radius of the particle, thereby dis-allowing coherence (for local experiments) of the particle due to Hawking radiation, as will be discussed.

In our previous discussions[15], we considered N_m particles of mass m within $\sim \lambda_m$ and hence $\mathcal{Z}_m = \frac{N_m(N_m-1)}{2}$ interacting pairs. The Schwarzschild radius in this case was considerably smaller than λ_m

$$R_S \cong 2G_N\sqrt{2\mathcal{Z}_m}m \sim \frac{\lambda_m}{\sqrt{\mathcal{Z}_m}} << \lambda_m \qquad (2.10)$$

However, in the present discussion, $\mathcal{Z}_m$ counts the number of interactions carried by a quantum of mass m. Such interactions are expected to have a coherence length of the order found in the propagator of a Yukawa-like particle, $\lambda_m = \frac{\hbar}{mc}$ The DN argument applied to gravitation does allow us to partition 1 Planck mass worth of energy into $\mathcal{Z}_m$

interactions within the Compton wavelength λ_m. One way of examining Dyson's argument is to note that if one has $\mathcal{Z}_{e^2} \cong 137$ photons of approriate energy incident on an electron, all within its Compton wavelength λ_{m_e}, we expect a high likelihood of pair creation. By analogy, if there are $\mathcal{Z}_m$ coherent masses m within λ_m, there is high likelihood of the production of a scalar mass m.

2.3 Coherent Gravitating Matter

As noted in the Introduction, our approach is to examine the physical principles that we feel most comfortable using, and then extrapolate those principles back to the earliest period in the evolution of the universe for which this comfort level persists. Those conclusions that can be deduced from these principles will in this sense be model independent. We will refrain from engaging in constructing micro-cosmological models during earlier stages.

Following Jones we have associated a Planckton (cf. Sec.2.1) with a region that has quantum coherent energy of one Planck mass M_P. There are expected to be many Planck units of energy within a scale radius of the universe. Planckta will be considered to be internally coherent units that become incoherent with each other during the period of de-coherence.

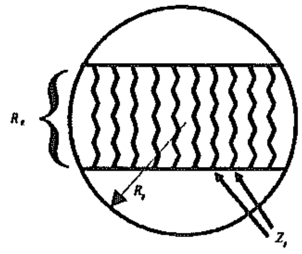
In general, we define ϵ as a gravitational energy scale associated with the scale factor R of the FRW metric. On a per Planckton basis, the average number $\mathcal{Z}_\epsilon$ of (virtual quantum) energy units $\epsilon < M_P$ that are localizable within a region of quantum coherence $R_\epsilon \sim 1/\epsilon$ is given by

$$\mathcal{Z}_\epsilon \epsilon = M_P \Rightarrow \epsilon \sim \frac{R_\epsilon}{L_P} = R_\epsilon M_P \qquad (2.11)$$

Note that for us this understanding of the meaning of $\mathcal{Z}_\epsilon$ *replaces* the Jones approximate identification of Z_ϵ motivated by quantum gravitational energy and scale relationships. This allows us to *start* our "cosmological clock" at a *finite* time calculated by backward extrapolation to the transition. In this paper we need not consider "earlier" times or specify a specific $t = 0$ achievable by backward extrapolation.

The localization of interactions has to be of the order $\hbar/mc$ in order to be able to use the DNJ argument. This allows us to obtain the number of gravitational interactions that can occur within a Compton wavelength of the mass m which provide sufficient energy to create a new mass or cause gravitational collapse, namely $\mathcal{Z}_m \sim \frac{M_P^2}{m^2}$. The process of de-coherence occurs when there are a sufficient number of available degrees of freedom such that gravitational interactions of quantum coherent states of Friedman-Lemaitre (FL) matter-energy could have a DNJ anomaly. We assume that a Planck mass M_P represents the largest meaningful scale for energy transfer at this boundary. A single Planckton of coherent energy in a scale of R_ϵ will have $\mathcal{Z}_\epsilon$ partitions of an available Planck mass of energy that can constitute interactions of this type. This is shown schematically in Fig. 3. Since there is global gravitational de-coherence at later times, the quantity $\mathcal{Z}_\epsilon$ can only be calculated prior to and during de-coherence.

When the number of partitions of a given Planckton energy unit



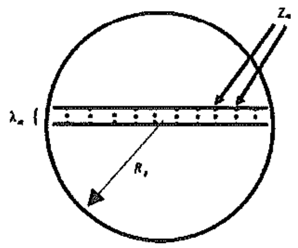


Figure 3: Counting of gravitating quanta during de-coherence

equals the DNJ limit, in principle there could occur a transition of the DNJ type involving interaction energies equal to a Planck mass (or less than a Planck mass at later times), thus allowing us to conclude that de-coherence gives a mass scale from the relationship

$$\mathcal{Z}_\epsilon = \mathcal{Z}_m \equiv \mathcal{Z}. \tag{2.12}$$

Expressed in terms of the energy scales, this gives the fundamental equation connecting (via the current value of Ω_Λ) the observable parameter ϵ to the mass scale at decoherence m

$$m^2 \approx \epsilon M_P. \tag{2.13}$$

This connection is a succinct summary of Jones' theory; henceforth we will refer to it as the *Jones equation.* Note that, viewed in this way, we no longer need the (Jones) thermodynamic Eqn. 2.6 to derive Eqn.

2.13. Therefore the temperature scale we called m_θ need no longer be *directly* identified with the particulate mass m which is associated with our *quantum decoherence* transition.

2.4 Correspondence with the Measured Cosmological Constant

We use the present day measurement of the cosmological density ρ_Λ to determine the energy scale m of cosmological de-coherence. The present cosmological constant energy density ρ_Λ is usually given in terms of the critical density ρ_c and the reduced Hubble parameter h

$$\rho_c \equiv \frac{3H_o^2}{8\pi G_N} \approx 1.0537 \times 10^{-5}\, h^2\, GeV/cm^3. \tag{2.14}$$

As already noted, we will take the value of the reduced Hubble parameter to be given by $h = 0.73$; current estimates of the reduced cosmological constant energy density $\Omega_\Lambda \equiv \frac{\rho_\Lambda}{\rho_c} \approx 0.73$, which means that the vacuum energy sets the de-coherence scale as

$$\rho_\Lambda \approx 4.10 \times 10^{-6} \frac{GeV}{cm^3} \approx 3.13 \times 10^{-47} GeV^4 \sim \epsilon^4 \tag{2.15}$$

Using the value $\hbar c \cong 1.97 \times 10^{-14} GeV\ cm$ we can immediately calculate the de-coherence scale energy and FRW scale radius

$$\epsilon \sim 10^{-12} GeV \quad , \quad R_\epsilon \sim 10^{-2} cm. \tag{2.16}$$

The Planck energy scale and DNJ limit at this scale is given by

$$\mathcal{Z}_\epsilon \equiv \frac{M_P}{\epsilon} \sim 10^{30} \sim \mathcal{Z}_m \tag{2.17}$$

The equality of the interaction factors $\mathcal{Z}$ gives the Jones equation $m^2 = \epsilon M_P$ from which we calculate a value for the mass scale for quantum de-coherence

$$m \sim 5\, TeV/c^2 \tag{2.18}$$

The number of Planck energy units per scale volume during de-coherence is given by

$$N_{Pk} \sim Z^2 \sim 10^{60}. \tag{2.19}$$

At this point we will examine the Hubble rate equation during this transition. If we substitute the expected energy density into the Freedman-Lemaitre equation, we obtain a rate of expansion given by

$$\dot{R}_\epsilon \sim R_\epsilon \sqrt{\frac{8\pi G_N}{3}(\rho_m + \rho_\Lambda)} \sim \frac{1}{\epsilon}\sqrt{\frac{1}{M_P^2}\mathcal{Z}_\epsilon^2 \epsilon^4} \sim c. \tag{2.20}$$

This is a very interesting result, which implies that the transition occurs near the time that the expansion rate is the same as the speed of light. In Section 3 on dark energy de-coherence, we will develop this argument as the primary characteristic of this transition.

2.5 Gravitating massive scalar particle

If the mass m represents a universally gravitating scalar particle, we expect the coherence length of interactions involving the particle to be of the order of its Compton wavelength with regards to Yukawa-like couplings with other particulate degrees of freedom present. If the density is greater than

$$\rho_m \sim \frac{m}{\lambda_m^3} \sim m^4 \qquad (2.21)$$

we expect that regions of quantum coherence of interaction energies of the order of m and scale λ_m will overlap sufficiently over the scale R_ϵ such that we will have a macroscopic quantum system on a universal scale. As long as the region of gravitational coherence is of FRW scale

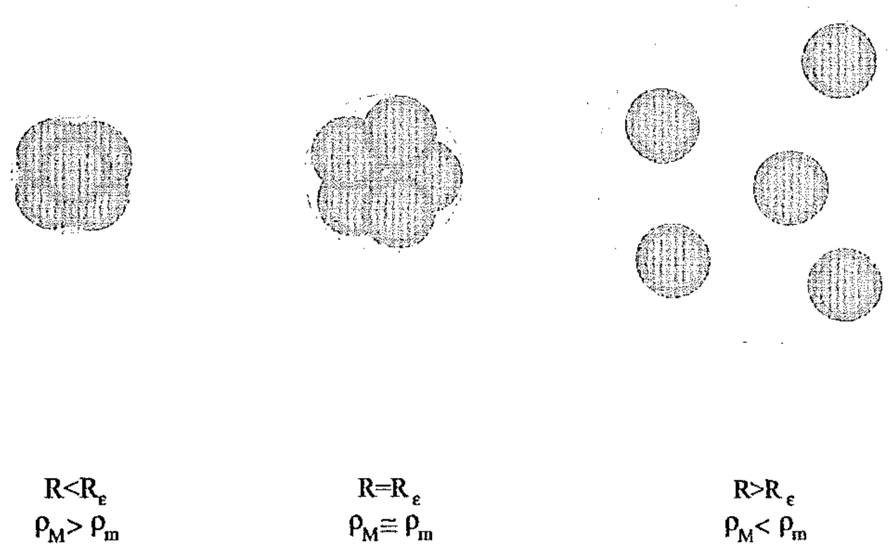


Figure 4: Overlapping regions of coherence during expansion

$R < R_\epsilon$, cosmological (dark) vacuum energy is determined by this scale. However, when the density of FL energy becomes less than ρ_m, we expect that since the coherence length of the mass m given by its Compton wavelength is insufficient to cover the cosmological scale, the FL energy density will break into domains of cluster decomposed (AKLN de-coherent[21]) regions of local quantum coherence. This phase transition will decouple quantum coherence of gravitational interactions on the cosmological scale R_ϵ. At this stage (de-coherence), the cosmological (dark) vacuum energy density ρ_Λ is frozen at the scale determined by R_ϵ. The cosmological dark energy contribution to the expansion rate is so small, and its coupling to de-coherent FL energy so insignificant, that its value is frozen at the value just prior to de-coherence given by

$$\rho_\Lambda \sim \frac{\epsilon}{R_\epsilon^3} = \epsilon^4 \tag{2.22}$$

We should note that using the DNJ argument, if the mass m were engaged in active cosmological energy exchanges involving non-gravitational microscopic interactions prior to decoherence, then, since those interactions have considerably larger coupling constants, it can be concluded that $\mathcal{Z}_g = \frac{1}{g^2} < \mathcal{Z}_m$. This would mean that this interaction would have broken coherence prior to our expected gravitational de-coherence event. We expect the mass m to be dark during this period (as must be all other particles).

During de-coherence, we assume that the FL energy contained in R_ϵ is given by N_{Pk} Planck mass units appropriately red-shifted to the de-coherence epoch. This gives an intensive FL energy density (for a spatially flat universe) of the form

$$\rho_{FL} \sim N_{Pk}\epsilon^4 \tag{2.23}$$

Since de-coherence is expected to occur when this density scale is given by the quantum coherence density scale for the mass by ρ_m, we obtain the following relationship between the de-coherence energy scale ϵ and the scalar mass m:

$$N_{Pk}\epsilon^4 \cong m^4 \tag{2.24}$$

This allows us to consistently relate the number of Plancktonic energy units in the region of coherence R_ϵ to the DN counting parameters:

$$N_{Pk} \cong \frac{m^4}{\epsilon^4} = \frac{m^4}{M_P^4}\frac{M_P^4}{\epsilon^4} = \frac{\mathcal{Z}_\epsilon^4}{\mathcal{Z}_m^2} = \mathcal{Z}^2, \tag{2.25}$$

which insures that all quantities relevant to our theory can be reduced to a single parameter in the Jones equation $m^2 \cong \epsilon M_P$.

It has already been suggested[11] that if we identify the Jones mass parameter m with a massive, scalar gravitating particle, this could be a candidate for particulate dark matter. Unfortunately if we assume that the mass m interacts *only* gravitationally, such a particle would be difficult to discover in accelerator experiments due to the extremely small coupling of gravitational scale forces.

We hope to be able to estimate the expected dark matter to photon number ratio from available phenomenological data if the mass is known. The FL equations satisfy energy conservation $T^{\mu\nu}_{;\nu} = 0$, which implies $\dot{\rho} = -3H(\rho + P)$. The first law of thermodynamics relates the pressure to the entropy density $\dot{P} = \frac{S}{V}\dot{T}$, which then implies an adiabaticity condition on the expansion given by

$$\frac{d}{dt}\left(\frac{S}{V}R^3\right) = 0. \tag{2.26}$$

Assuming adiabatic expansion, we expect $g(T)\,(RT)^3$ to be constant far from particle thesholds. Here, $g(T)$ counts the number of low mass particles contributing to the cosmological entropy density at temperature T. This gives a red shift in terms of the photon temperature during a given epoch

$$\frac{R_o}{R} \equiv 1 + z = (1 + z_{dust}) \left(\frac{g(T)}{g_{dust}}\right)^{1/3} \frac{T}{T_{dust}}, \tag{2.27}$$

where z_{dust} is defined as the redshift at equality of radiation and pressureless matter energy densities. In terms of photon temperature, we can count the average number of photons using standard results from black body radiation

$$\frac{N_\gamma}{N_{\gamma o}} = \frac{(RT)^3}{(R_o T_o)^3} \cong \frac{g_o}{g(T)}. \tag{2.28}$$

This allows us to write a formula for the dark matter - photon ratio at the temperature of dark matter number conservation (T_{freeze}), in terms of its mass and the measured baryon-photon ratio:

$$\frac{N_{dm}}{N_\gamma} \cong \frac{\Omega_{dm}}{\Omega_{baryon}} \frac{N_{b o}}{N_{\gamma o}} \frac{m_N}{m} \frac{g(T)}{g_o}, \tag{2.29}$$

which gives $\frac{N_{dm}}{N_\gamma} \cong 2.9 \times 10^{-9} \frac{g(T_{freeze})}{g_o} \frac{m_N}{m}$.

If there is DNJ collapse, one can estimate the lifetime of the resulting black holes. These collapsed objects would be expected to emit essentially thermal low mass quanta at a rate determined by the barrier height near the horizon $\sim MG_N$ and the wavelength of the quanta $\sim (MG_N)$ giving a luminosity of order $dM/dt \sim -1/M^2 G_N^2$. This can be integrated to give a lifetime of the order

$$t_{evaporation} \sim M^3 G_N^2. \tag{2.30}$$

This means that a collapsed DNJ object has an approximate lifetime of

$$\tau_{BH} \sim \frac{(\mathcal{Z}_m m)^3}{M_P^4} \sim \mathcal{Z}_m \frac{\hbar}{mc^2} \tag{2.31}$$

Substituting the expected mass, the lifetime is expected to be $\tau_{BH} \sim 400sec$, which is long compared to the inverse Hubble rate $H_\epsilon^{-1} \sim 10^{-13}sec$ during decoherence.

We can also estimate the number of low mass quanta that would result from the evaporation. We will examine the quantum mechanics of massive scalar particles $g_{\mu\nu}p^\mu p^\nu = -m^2$ in a Schwarzschild metric.

$$ds^2 = -(1 - \frac{2G_N M}{r})dt^2 + \frac{dr^2}{1 - \frac{2G_N M}{r}} + r^2 d\theta^2 + r^2 sin^2\theta d\phi^2. \quad (2.32)$$

Using tortoise coordinates r^*, where $r^* \equiv r + 2G_N M log\left(\frac{r}{2G_N M} - 1\right)$, the action is taken to be

$$W = \frac{1}{2}\int dtdr^* d\theta d\phi r^2 \left(1 - \frac{2G_N M}{r}\right) sin\theta \left[\frac{-\chi_t^2 + \chi_{r^*}^2}{1 - \frac{2G_N M}{r}} + m^2\chi^2 + \frac{\chi_\theta^2}{r^2} + \frac{\chi_\phi^2}{r^2 sin^2\theta}\right]. \quad (2.33)$$

The equation of motion generated for the reduced radial function $\psi \equiv r\chi$ for stationary states is given by

$$\psi_{r^*r^*} - \left(1 - \frac{2G_N M}{r}\right)\left(m^2 + \frac{\ell(\ell+1)}{r^2} + \frac{2G_N M}{r^3}\right)\psi = \psi_{tt} = -(m^2 + k_\infty^2)\psi. \quad (2.34)$$

The effective potential barrier height is seen to be of the order of the inverse Schwarzschild radius. The asymptotic solution satisfies $\psi_{rr}(r \to \infty) \cong -k_\infty^2\psi$, whereas the solution near the Schwarzschild radius (for s-waves) is given by $\psi_{r^*r^*}(r \to 2G_N M) \cong -(m^2 + k_\infty^2)\psi$. We expect that when the temperature is above mass threshold, the particle can be radiated, and that for temperatures above the barrier height $V_{max} \sim \frac{0.3}{2G_N M}$ the transmission rate of particles is of order $\frac{k_\infty}{\sqrt{m^2 + k_\infty^2}}$.

Writing the luminosity and number rate as

$$\begin{aligned} \frac{dM}{dt} &= \frac{\eta(M)}{M^2G_N^2} \\ \Rightarrow t &= \frac{1}{3}\frac{M^3G_N^2}{\bar{\eta}} \\ \frac{dN}{dt} &= \frac{\eta(M)}{MG_N} \cong \frac{\bar{\eta}^{2/3}}{(3G_N t)^{1/3}} \end{aligned} \tag{2.35}$$

Here $\eta(M)$ is expected to be a slowly varying function of the temperature that counts the number of low mass thermal states at the temperature of the black hole. This factor is expected to be essentially constant between particle thresholds. The solution then takes the form

$$N(M) \cong \frac{1}{2}\frac{M^2}{M_P^2}. \tag{2.36}$$

More generally, the total number of low mass quanta resulting from evaporation from mass M to mass M' is expected to satisfy

$$N(M \to M') \cong \frac{M^2 - M'^2}{2M_P^2}. \tag{2.37}$$

If the black hole has formed due to DNJ collapse, substituting $M = \mathcal{Z}_m m$ gives

$$N(\mathcal{Z}_m m) \cong \frac{1}{2}\mathcal{Z}_m^2\frac{m^2}{M_P^2} \cong \mathcal{Z}_m. \tag{2.38}$$

Therefore, the intermediate quanta in the collapse are expected to produce an essentially equal number of low mass quanta during evaporation.

We can estimate the relative number of quanta of mass m evaporated by a black hole formed by DNJ collapse. If the mass cannot be radiated prior to temperature T_m, this ratio is given by

$$\frac{N_m}{N_{Total}} \cong \frac{1}{g(T_m)}\left(\frac{M(T > T_m)}{M}\right)^2, \tag{2.39}$$

where $g(T_m)$ is the number of low mass states available for radiation at temperature T_m. Substituting $M = Z_m m$, $T_m = \frac{1}{8\pi G_N M(T_m)} \cong m$, and using the Particle Data Group expectation[5] for the number of low mass states at this temperature $g(T_m) = \frac{427}{4}$ gives an estimate of $\frac{N_m}{N_{Total}} \approx 10^{-5}$ from each black hole thermalization. This is all that can be concluded at present relevant to the dark matter-photon ratio during thermalization.

To summarize, our re-examination of the Jones theory has led us to the conclude that the $\mathcal{Z}_\epsilon = \mathcal{Z}_m$ relation is best interpreted in that context as the equality of the (intensive) number of gravitational quanta of mass m exchanged between all gravitating systems between the cosmological scale R_ϵ and the particulate scale λ_m, when the DNJ bound $\mathcal{Z}m \rightarrow (\mathcal{Z}+1)m$ is reached. One way of examining Dyson's argument is to note that if one has $\mathcal{Z}_{e^2} \cong 137$ photons of approriate energy incident on an electron, all within its Compton wavelength λ_{m_e}, we expect a high likelihood of pair creation. By analogy, if there are $\mathcal{Z}_m$ coherent masses m within λ_m, there is high likelihood of the production of a scalar mass m. This interpretation *requires* us to be talking about *quantum coherent systems* when the Jones transition from microcosmology to a universe where we can use conventional physics and cosmology takes place. This line of reasoning suggested that this transition itself *must* in some sense correspond to *quantum decoherence* and to the title of this paper. The consequences of pursuing this line of thought constitute the rest of this paper.

3 Dark Energy De-coherence

We will now make quantitative arguments to develop the general ideas motivated by the previous sections. Although the arguments are independent of those in the previous section, we will derive very similar results. In most of what follows we will assume flat spatial curvature $k = 0$. Prior to the scale condition $\dot{R}_\epsilon = c$, which we will henceforth refer to as the time of dark energy de-coherence, gravitational influences are propagating (at least) at the rate of the gravitational scale expansion, and microscopic interactions (which can propagate no faster than c) are incapable of contributing to cosmological scale equilibration. Since the definition of a temperature requires an equilibration of interacting "microstates", there must be some mechanism for the redistribution of those microstates on time scales more rapid than the cosmological expansion rate, which can only be gravitational.

3.1 Dark Energy

As we have discussed in the motivation section, we expect that dark energy de-coherence occurs when the FRW scale is $R_\epsilon \sim 1/\epsilon$. The gravitational dark energy scale associated with de-coherence is given by ϵ, independent of the actual number of energy units N_ϵ in the scale region. Since the dark energy density must be represented by an intensive parameter which should be the same for the universe as a whole, we will express this density as the coherent vacuum state energy density of this macroscopic quantum system. In the usual vacuum state, the equal time correlation function

$$< vacuum|\Psi(x,y,z,t)\,\Psi(x',y',z',t)|vacuum >$$

does not vanish for space-like separations. (For example, for massless scalar fields, this correlation function falls off with the inverse square of the distance between the points). Since we assume no physical distinction between spatially separated points, our correlation functions would be expected to be continuous, with periodic boundary conditions defined by the cosmological scale factor. Given a cosmological scale factor R_ϵ, periodic boundary conditions on long range (massless or low mass) quanta define momentum quantization in terms of this maximum wavelength. The energy levels associated with these quanta would satisfy the usual condition

$$E_{N_\epsilon} = \left(N_\epsilon + \frac{1}{2}\right)\hbar\omega = (2N_\epsilon + 1)\,\epsilon. \tag{3.40}$$

This is associated with quanta of wavelength of the order of the cosmological scale factor with vacuum energy density, given by

$$\rho_\Lambda \equiv \frac{\epsilon}{(2R_\epsilon)^3} = \left(\frac{k_\epsilon}{2\pi}\right)^3 \frac{\sqrt{m^2_{condensate} + k^2_\epsilon}}{2} \tag{3.41}$$

for a translationally invariant universe with periodicity scale $2R_\epsilon$. In effect, this provides the infrared cutoff for cosmological quantum coherent processes,

$$k_\epsilon = \frac{2\pi}{\lambda_\epsilon} = \frac{\pi}{R_\epsilon}. \tag{3.42}$$

We will begin by examining a condensate of massless or very low mass particles.

The vacuum energy scale associated with a condensate of gravitationally coherent massless quanta is given by

$$\epsilon = \frac{1}{2}\hbar\omega_\epsilon = \frac{1}{2}\hbar k_\epsilon c = \frac{\pi}{2}\frac{\hbar c}{R_\epsilon}, \tag{3.43}$$

and the vacuum energy density for such massless quanta is

$$\rho_\Lambda = \frac{\epsilon^4}{(\pi)^3}. \tag{3.44}$$

We might inquire into the nature of the dark energy, in the sense as to whether it is geometric or quantum mechanical in origin. From the form of Einstein's equation

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = \frac{8\pi G_N}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu} \quad , \quad (\Lambda g^{\mu\nu})_{;\nu} = 0 \tag{3.45}$$

if the term involving the cosmological constant should most naturally appear on the left hand side of the equation, we would consider it to be geometric in origin. If the cosmological term is geometric in origin, we would expect it to be a fundamental constant of the cosmology which scales with the FRW/FL cosmology consistently with the vanishing divergence of the Einstein tensor. However, if the previous arguments are interpreted literally, the dark energy density freezes out to a constant determined by the period of last quantum coherence with the FL energy density and the onset of the equilibration of states involving microscopic non-gravitational interactions, supporting its interpretation as a gravitational quantum vacuum energy density. This means that it is fixed by a physical condition being met, and thus would not be a purely geometric constant.

3.2 Rate of Expansion during De-Coherence–massless condensate

We will next examine the rate of the expansion during the period of de-coherence. We will make use of the Friedmann-LeMaitre (FL)/Hubble equations, which relates the expansion rate and acceleration to the densities

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G_N}{3}(\rho + \rho_\Lambda) - \frac{k}{R^2}, \tag{3.46}$$

$$\frac{\ddot{R}}{R} = -\frac{4\pi G_N}{3}(\rho + 3P - 2\rho_\Lambda), \tag{3.47}$$

where we have written

$$\rho_\Lambda = \frac{\Lambda}{8\pi G_N}, \tag{3.48}$$

and ρ represents the FL energy density. The only scale dependent term in this equation involves the spatial curvature k. If k is non-vanishing, we have no reason to assume that any scale other than R_ϵ at de-coherence determines the cosmological scale. In our discussion, the dark energy density will have negligible contribution to the FL expansion during de-coherence, but will become significant as the FL energy density ρ decreases due to the expansion of the universe.

It is unclear whether one can speak of causal horizons and causal communications in the usual ways prior to the period of de-coherence, since the scale expansion rate is larger than c. Assuming that local inertial physics satisfies the principle of equivalence with a limiting velocity of c, it would be difficult to extrapolate the type of physics we do presently into a domain with local expansion rates greater than c. Only the FRW gravitation interacts with rates which can equilibrate states defining a thermal system in this domain, since the other interactions cannot have super-luminal exchanges, only super-luminal quantum correlations. If the expansion rate is super-luminal $\dot{R} > c$, scattering states cannot form decomposed (de-coherent) clusters of the type described in reference [21]. We see from the above discussion that, assuming the validity of an FL universe back to the stage of de-coherence, our usual ideas of microscopic causality become obscure beyond this period.

Since we find the expansion rate equation $\dot{R}_\epsilon = c$ a compelling argument for the quantitative description of gravitational de-coherence, it is this relationship that we will use to determine the form for the energy density during dark energy de-coherence ρ_{FL}, which counts the number of gravitating quanta above vacuum energy in the condensed state. The Hubble equation takes the form

$$H_\epsilon^2 = \left(\frac{c}{R_\epsilon}\right)^2 = \frac{8\pi G_N}{3}\left(\rho_{FL} + \rho_\Lambda\right) - \frac{kc^2}{R_\epsilon^2} = \frac{8\pi G_N}{3}\left(2N_\epsilon + 1\right)\rho_\Lambda - \frac{kc^2}{R_\epsilon^2}. \tag{3.49}$$

We see that $2N_\epsilon$ counts the number of Jones-Planck energy units per scale factor in the pre-coherent universe (referred to by Jones as $N_{Planckton}$), and it defines the ratio of normal to dark energy density during decoherence.

If this condition is to describe the onset of dark energy de-coherence, we can see that a so called "open" universe ($k = -1$) is excluded from undergoing this transition. In this case, the cosmological constant term in equation 3.46 already excludes a solution with $\dot{R}_\epsilon \leq c$.

Likewise, for a "closed" universe that is initially radiation dominated, we can compare the scale factors corresponding to $\dot{R}_\epsilon = c$ and $\dot{R}_{max} = 0$. From the Hubble equation

$$\frac{c^2}{R_{max}^2} = \frac{8\pi G_N}{3}(\rho + \rho_\Lambda) \cong \frac{8\pi G_N}{3}\rho_\epsilon \frac{R_\epsilon^4}{R_{max}^4} \Rightarrow R_{max}^2 \cong 2R_\epsilon^2. \quad (3.50)$$

Clearly, this closed system never expands much beyond the transition scale. For this reason, henceforth we will only consider flat spaces.

We will assert that de-coherence cannot occur prior to $\dot{R} = c$ since incoherent decomposed clusters [22] cannot be cosmologically formulated. Using the equation

$$\left(\frac{c}{R_\epsilon}\right)^2 = \frac{8\pi G_N}{3}(2N_\epsilon + 1)\,\rho_\Lambda \quad (3.51)$$

and the form of ρ_Λ from equation 3.41 we can directly determine the number of quanta in the condensed state

$$N_\epsilon = \frac{1}{2}\left(\frac{3}{2}\frac{M_P^2}{\epsilon^2} - 1\right) \cong \frac{3}{4}\mathcal{Z}_\epsilon^2 \quad (3.52)$$

where, as before $\mathcal{Z}_\epsilon \equiv \frac{M_P}{\epsilon}$. The energy density during dark energy decoherence is therefore given by

$$\rho_{FL} = 2N_\epsilon \rho_\Lambda \cong \frac{3}{2\pi^3} M_P^2 \epsilon^2. \quad (3.53)$$

3.3 Phenomenological correspondence–massless condensate

To make correspondence with observed cosmological values, we will utilize parameters obtained from the Particle Data Group[5]. The value for the critical density is given by

$$\rho_c \equiv \frac{3H^2}{8\pi G_N} \cong 5.615 \times 10^{-6}\, GeV/cm^3 \cong 4.293 \times 10^{-47} GeV^4. \quad (3.54)$$

The cosmological dark energy density parameter $\Omega_\Lambda = \frac{\rho_\Lambda}{\rho_c}$ is taken to have the value $\Omega_\Lambda \cong 0.73$. We will therefore use the value

$$\rho_\Lambda \cong 4.099 \times 10^{-6} GeV/cm^3 \cong 3.134 \times 10^{-47} GeV^4. \quad (3.55)$$

This gives a dark energy scale and FRW scale given by

$$\epsilon \cong 5.58 \times 10^{-12}\, GeV \quad (3.56)$$

$$R_\epsilon \cong 5.54 \times 10^{-3}\, cm. \quad (3.57)$$

The Friedman-Lemaitre energy density from equation 3.53 is then given by

$$\rho_{FL} \cong 2.93 \times 10^{55}\, GeV/cm^3 \cong 2.24 \times 10^{14}\, GeV^4, \quad (3.58)$$

with the Planck energy partition $\mathcal{Z}_\epsilon$ and number of "gravons" N_ϵ at dark energy de-coherence given by

$$\mathcal{Z}_\epsilon \cong 2.19 \times 10^{30} \quad (3.59)$$

$$N_\epsilon \cong 3.59 \times 10^{60}. \quad (3.60)$$

By *gravons* we will mean gravitationally coherent Bose states. The coherent mass density scale $\frac{m}{\lambda_m^3} = m^4$ corresponding to the FL density

ρ_{FL} is given by $m \sim 3800\, GeV/c^2$. However, we will obtain a precise determination of the coherent mass scale in terms of the UV cutoff scale of the gravitational dynamics when we discuss the thermal ground state shortly.

We will next estimate the time of dark energy de-coherence assuming a radiation dominated expansion prior to this period. Although we are dubious about using a standard radiation dominated equation of state prior to dark energy de-coherence, we can get some feeling for the time scale of this transition. The period after de-coherence is radiation dominated until the dust driven epoch, with the scale factor satisfying

$$R(t) \;=\; R_\epsilon \left(\frac{t}{t_\epsilon}\right)^{1/2}. \tag{3.61}$$

This means that $\frac{\dot{R}(t)}{R(t)} = \frac{1}{2t}$, resulting in an estimate for the time

$$t_\epsilon \;\cong\; \frac{R_\epsilon}{2c} \;\cong\; 9.30 \times 10^{-14}\, sec. \tag{3.62}$$

This also gives a Hubble rate of

$$H_\epsilon \;=\; \frac{c}{R_\epsilon} \;\cong\; 5.38 \times 10^{12}/sec \;\cong\; (1.86 \times 10^{-13} sec)^{-1}. \tag{3.63}$$

This Hubble rate gives a minimal lifetime for any gravitating mass scale m that can equilibrate during de-coherence. If the mass is to have meaningful coherence during the period of de-coherence, its lifetime in the thermal bath must be of an order greater than the inverse Hubble rate. This means that

$$\tau_m \;>\; \frac{1}{H_\epsilon} \;\sim\; 10^{-13} sec. \tag{3.64}$$

If the mass scale is associated with the Higgs scalar of the symmetry breaking, this mass could ONLY couple to electro-weak bosons to generate mass, since the Yukawa coupling to masses comparable to the top quark mass would give a width well in excess of this scale.

The estimates are only slightly modified for vector and tensor gravons. Substituting spin degeneracy corresponding to the particle type, the scale factor of de-coherence becomes $R_V \cong 7.3 \times 10^{-3} cm$ for vector gravons, and $R_T \cong 8.3 \times 10^{-3} cm$ for tensor gravons, with the other calculated quantities varying accordingly. We will assume scalar quanta for our further calculations.

3.4 Thermal Ground State

For a hot, thermal system, the ground state is not that state which satisfies $\hat{N}|0>=0$ for all modes (zero occupation), but instead is constructed of a thermal product of occupation number states, weighted by a density matrix. Unlike the zero occupation number state, this ground state need not generally be translationally invariant in time. Examining the low energy modes at high temperatures, the thermally averaged occupation of those modes $<\hat{N}_n>\cong \frac{k_B T}{E_n}$ demonstrates large numbers of low energy massless quanta, giving these modes a large number of degrees of freedom. For our system, there are natural infrared and ultraviolet cutoffs provided by the macroscopic scale k_ϵ and microscopic scale m. We expect macroscopic gravitational physics involving gravitating masses m to be cutoff for momenta $k_{UV} \sim m$.

It is of interest to calculate the energy of the zero-occupation number state using these cutoffs,

$$\hat{H}|0,0,...,0> = \sum_{\vec{k}=\vec{k}_\epsilon}^{\vec{k}_{UV}} \frac{1}{2}\hbar c\hat{k}\,|0,0,...,0> . \qquad (3.65)$$

Inserting the density of states to approximate the sum gives an energy density of the form

$$\frac{E_{|0,...,0>}}{V} \sim \frac{1}{(2\pi)^3}\int_{k_\epsilon}^{m}\frac{1}{2}\hbar ck\;4\pi k^2 dk \;\sim\; m^4-\epsilon^4 \;\sim\; \rho_{FL}, \qquad (3.66)$$

which is essentially the Jones equilibrium condition equation 2.6. This means that the vacuum energy density corresponding to zero occupancy of the gravitational modes corresponds to the energy density of the normal gravitating matter just after decoherence if the ultraviolet cutoff of the long range modes in the superfluid is chosen to be the mass scale m. We will therefore proceed recognizing that the mass scale provides an ultraviolet gravitational cutoff for the decoherent cosmology.

As has been previously discussed, the vacuum energy associated with the condensate is given by $\rho_\Lambda = \epsilon/(2R_\epsilon)^3$, which for a massless condensate gives $\rho_\Lambda = \epsilon^4/\pi^3$. However, once the expansion rate is sub-luminal, global gravitational coherence is expected to be broken due to interactions that propagate at the speed of light. This means that all available modes must thereafter be included in calculations of the vacuum energy. As is the case with superfluids, we will assume that there is an ultraviolet cutoff associated with the (scalar) mass scale m with coherence length $\lambda_m = \hbar/mc$. The vacuum energy density associated with this cosmology transition during dark energy de-coherence from ρ_Λ (at pre-coherence) to ρ_{vac} (at de-coherence) is given by

$$\begin{aligned} \rho_{vac} &= \frac{E_{vacuum}}{V_\epsilon} = \int \frac{g_m}{2}\hbar c|\vec{k}|\frac{d^3k}{(2\pi)^3} \\ \frac{\hbar c}{(2\pi)^2}\int_{k_\epsilon}^{k_m} k^3 dk &= \frac{1}{4}\frac{\hbar c}{(2\pi)^2}(k_m^4 - k_\epsilon^4), \end{aligned} \tag{3.67}$$

where $k_m = \frac{2\pi}{\lambda_m} = \frac{2\pi mc}{\hbar}$, $k_\epsilon = \frac{\pi}{R_\epsilon} = \frac{2\epsilon}{\hbar c}$, and the spin degeneracy g_m will be taken to be unity. Therefore, the vacuum energy at de-coherence is taken to be

$$\rho_{vac} = \pi^2\left(m^4 - \left(\frac{\epsilon}{\pi}\right)^4\right) \cong \pi^2 m^4. \tag{3.68}$$

If we presume minimal parametric input to this model, then this vacuum energy thermalizes as the FL energy density in equation 3.49 for the cosmology $\rho_{FL} = \rho_{vac}$, giving a relationship for the mass scale of dark energy de-coherence

$$m^4 = \frac{3}{2\pi^5}M_P^2\epsilon^2. \tag{3.69}$$

This gives an expected mass scale given by

$$m \cong 2183 GeV/c^2. \tag{3.70}$$

If g_m is the spin degeneracy associated with m, then the left hand side of equation 3.70 is modified by a factor of $g_m^{1/4}$.

The existence of a gravitational mass associated with de-coherence introduces the possibility that the vacuum energies should be calculated in terms of the vacuum states of this mass rather than in terms of the long range excitations (gravons) treated previously. More generally, the mass scale associated with the condensate need not be the same as that of the ultraviolet cutoff, which introduces yet another mass scale. For instance, the cutoff mass m could be associated with a dark matter mass, while the condensate mass could be associated with the symmetry breaking scale. In the present context, we will associate these two scales as identical. Thus far, there is nothing in our discussion preventing the use of vacuum energy as

$$\epsilon_m \equiv \frac{1}{2}\sqrt{m^2 + k_\epsilon^2} = \frac{1}{2}\sqrt{m^2 + \left(\frac{\pi}{R_\epsilon}\right)^2}. \qquad (3.71)$$

The post-decoherence vacuum energy then is given in general by

$$\rho_{vac} = \frac{E_{vacuum}}{V_\epsilon} = \int_{k_\epsilon}^{k_m} g_m \frac{\sqrt{m^2 + |\vec{k}|^2}}{2} \frac{4\pi|\vec{k}|^2 dk}{(2\pi)^3}, \qquad (3.72)$$

(where V_ϵ is the scale volume of the region under consideration) which can be used to solve for the mass m self-consistently by setting $\rho_{vac} = \rho_{FL}$.

3.5 Phenomenological correspondence–massive condensate

We will recalculate the phenomenological parameters for a pre-coherent condensate of massive particles of mass m. The self-consistent mass that satisfies the condition $\rho_{vac} = \rho_{FL}$ is given by $m \cong 19.74 GeV$. The dark energy scale and FRW scale is given by

$$\epsilon \cong 9.87\, GeV \qquad (3.73)$$

$$R_\epsilon \cong 67.0\, cm. \qquad (3.74)$$

The Friedman-Lemaitre energy density from equation 3.53 is then given by

$$\rho_{FL} \cong 2.01 \times 10^{47}\, GeV/cm^3 \cong 1.54 \times 10^6\, GeV^4, \qquad (3.75)$$

with the Planck energy partition $\mathcal{Z}_\epsilon$ and number of condensate particles N_ϵ at dark energy de-coherence given by

$$\mathcal{Z}_\epsilon \cong 1.24 \times 10^{18} \tag{3.76}$$

$$N_\epsilon \cong 2.45 \times 10^{52}. \tag{3.77}$$

The time estimate for a radiation dominated cosmology is given by

$$t_\epsilon \cong \frac{R_\epsilon}{2c} \cong 1.13 \times 10^{-9}\, sec. \tag{3.78}$$

This gives a Hubble rate of

$$H_\epsilon = \frac{c}{R_\epsilon} \cong 4.45 \times 10^8/sec \cong (2.25 \times 10^{-9} sec)^{-1}. \tag{3.79}$$

Again, the estimates are only slightly modified for vector and tensor masses. Substituting spin degeneracy corresponding to the particle type, the scale factor corresponding to de-coherence becomes $R_V \cong 62.0cm$ for vector masses $m \cong 15.6$ GeV, and $R_T \cong 59.8cm$ for tensor masses $m \cong 14.0$ GeV, with the other calculated quantities varying accordingly. We will assume scalar masses for our calculations.

3.6 Thermalization

We will next examine the thermalization of the coherent gravitating cosmology into the familiar particulate states. De-coherence is presumed to occur adiabatically into a radiation dominated cosmology. For each low mass particle state, the standard black body relationships are satisfied:

$$\frac{U}{V} = g\frac{\pi^2}{30}\left(\frac{k_B T}{\hbar c}\right)^3 k_B T = \rho c^2 = 3P \tag{3.80}$$

$$\frac{S}{V} = g\frac{2\pi^2}{45}k_B\left(\frac{k_BT}{\hbar c}\right)^3 \tag{3.81}$$

$$\frac{N}{V} = g^*\frac{\zeta(3)}{\pi^2}\left(\frac{k_BT}{\hbar c}\right)^3 \tag{3.82}$$

where the statistical factors are given by

$$g/\#\,spin\,states = \begin{cases} 1 & \text{bosons} \\ \frac{7}{8} & \text{fermions} \end{cases}, \quad g^*/\#\,spin\,states = \begin{cases} 1 & \text{bosons} \\ \frac{3}{4} & \text{fermions} \end{cases} \tag{3.83}$$

The temperature of radiation with density ρ_{FL} is given by

$$\rho_{FL} = g(T_\epsilon)\frac{\pi^2}{30}\frac{(k_BT_\epsilon)^4}{(\hbar c)^3} \Rightarrow T_\epsilon \cong \frac{T_{crit}}{g^{1/4}(T_\epsilon)} \tag{3.84}$$

For a massless condensate, $T_{crit} \cong 5111$ GeV, and for temperatures above top quark mass, the degeneracy factor $g(m_t) = \frac{429}{4}$ is relatively weakly dependent upon any new degrees of freedom. To a few percent, the temperature of de-coherence is determined to be

$$k_BT_\epsilon \cong 1592\,GeV. \tag{3.85}$$

For a cold massive condensate, we will see in the subsection on Bose condensation that the critical temperature is $T_{crit} \cong 104$ GeV, and (assuming a degeneracy factor of $g(m_b) = \frac{345}{4}$) the temperature of de-coherence is

$$k_BT_\epsilon \cong 15.18\,GeV. \tag{3.86}$$

We will next estimate the present day scale corresponding to the dark energy de-coherence scale R_ϵ. We will assume a relatively rapid transition from a radiation dominated expansion to a matter dominated expansion at the dust transition red-shift, corresponding to equal energy densities of the (present day) relativistic and non-relativistic particles. We will use a value calculated from standard references[5] for $z_{eq} \equiv z_{dust} \cong 3629$, where the red shift satisfies the usual formula

$$\frac{\nu(z)}{\nu_o} \equiv 1 + z = \frac{R_o}{R(z)}. \tag{3.87}$$

After relativistic radiation falls out of equilibrium, its temperature satisfies $T \sim 1/R$. Photons fell out of equilibrium at last scattering $z \sim 1100$, whereas neutrinos fell out of equilibrium much sooner at a temperature $T \sim 1MeV$. The present cosmic background photon temperature is $2.725K \cong 2.35 \times 10^{-13} GeV$, and that of neutrinos is about 1.9K. We will calculate the red shift from CMB photon temperature to de-coherence temperature using equation 2.27.

For a massless condensate, the red shift at decoherence is found to be $z_\epsilon \approx 10^{16}$, whereas for a massive condensate $z_\epsilon \approx 10^{14}$. We can use these redshifts to determine the present scale associated with the de-coherence scale R_ϵ. For a massless condensate $R_o \approx 10^{14}$ cm, which is about the distance of Satern from Earth. For a massive condensate, this scale is given by 10^{16} cm, two orders of magnitude larger.

We next examine the entropy of the system during the de-coherence period. The Fleisher-Susskind[25] entropy limit considers a black hole as the most dense cosmological object of a given size, limiting the entropy according to

$$S \leq S_{\substack{black \\ hole}} = \frac{k_B c^3}{\hbar} \frac{A}{4G_N} \tag{3.88}$$

For a radiation dominated cosmology at de-coherence, the entropy is proportional to the number of quanta, and is related to the energy density $\left(\frac{c}{R_\epsilon}\right)^2 \cong \frac{8\pi G_N}{3}\rho_{FL}$ by

$$\frac{S}{V} = \frac{4}{3}\frac{\rho_{FL}}{T_\epsilon} \Rightarrow S = \frac{4}{\pi G_N}\frac{R_\epsilon}{T_\epsilon}. \tag{3.89}$$

Examining this for the space-like area given by the box $A = 6(2R_\epsilon)^2$ the ratio of the entropy in a thermal environment to the limiting entropy during thermalization is given by

$$\frac{S}{A/4G_N} \cong \frac{2}{3\pi}\left(\frac{1}{R_\epsilon T_\epsilon}\right) \sim 10^{-16}. \tag{3.90}$$

Clearly this result satisfies the FS entropy bound regardless of the mass of the condensate.

3.7 Bose condensation

We next calculate the critical temperature for condensation of a non-interacting gas of massless Bose quanta just prior to dark energy decoherence. At temperature T, such a gas has energy density satisfying the relation

$$\rho = \rho_{GS} + \frac{\pi^2}{30}\left(\frac{k_B T}{\hbar c}\right)^3 k_B T. \tag{3.91}$$

Here ρ_{GS} is the density of the condensate (the density of bosons in the ground state). Critical temperature is defined when the second term is insufficient to contain all particles. For the pre-coherent state, this is given by

$$\rho_{FL} = \frac{\pi^2}{30}\frac{(k_B T_{crit})^4}{(\hbar c)^3}, \tag{3.92}$$

where $\rho_{FL} = \frac{3}{8\pi}\left(\frac{M_P}{R_\epsilon}\right)^2 - \rho_\Lambda$ as before. The critical temperature therefore satisfies $T_{crit} = (g(T_\epsilon))^{\frac{1}{4}} T_\epsilon$. This corresponds to a temperature of around $T_{crit} \approx 5109$ GeV for a pre-thermalized system consisting of only gravons. We expect the system to remain in a zero temperature state prior to de-coherence, defining a vacuum energy density ρ_Λ just prior to de-coherence. The thermodynamics after the availability of sub-luminal degrees of freedom will define the temperature of thermalization using

$$\rho_{FL} = (g(T_\epsilon) + 1)\frac{\pi^2}{30}\frac{(k_B T_\epsilon)^4}{(\hbar c)^3} + \rho_{GS}, \tag{3.93}$$

where $g(T_\epsilon)$ counts the degrees of freedom available to luminal and sub-luminal interactions. Because of the availability of the new degrees of freedom, one expects a solution without condensate, i.e. $\rho_{GS} = 0$, to be consistent at these temperatures.

Just as de-coherence begins, we expect the fraction of condensate to thermal gravons to satisfy

$$\frac{\rho_{condensate}}{\rho_{FL}} = 1 - \left(\frac{T_\epsilon}{T_{crit}}\right)^4 \tag{3.94}$$

$$\frac{N_{condensate}}{N_{thermal}} = 1 - \left(\frac{T_\epsilon}{T_{crit}}\right)^3, \tag{3.95}$$

where the total number of thermal gravons satisfies

$$\frac{N_{thermal}}{V_\epsilon} = \frac{\zeta(3)}{\pi^2}\left(\frac{k_B T_{crit}}{\hbar c}\right)^3. \tag{3.96}$$

The de-coherent temperature is considerably lower than this temperature due to the new degrees of freedom $g(T_\epsilon)$, giving $\frac{\rho_{condensate}}{\rho_{FL}} \cong 0.99$ and $\frac{N_{condensate}}{N_{thermal}} \cong 0.97$ starting thermalization. Expressing ρ_{FL} in terms of the pre-coherent condensate, we obtain a relationship between the pre-coherent and thermal gravons, most of which initially remain in condensate form:

$$\rho_{FL} = \frac{\pi^4}{30\zeta(3)} \frac{N_{thermal}}{V_\epsilon} k_B T_{crit} = 2 N_\epsilon \rho_\Lambda. \tag{3.97}$$

This gives a large ratio of pre-coherent to thermal gravons given by

$$\frac{N_\epsilon}{N_{thermal}} = \frac{\pi^4}{60\zeta(3)} \frac{k_B T_{crit}}{\rho_\Lambda V_\epsilon} \sim 10^{15}. \tag{3.98}$$

Therefore, pre-coherent gravons would be required to rapidly thermalize a large number of states.

We next examine the properties of a (non-interacting) Bose gas of particle of mass m for a system with temperature $T \lesssim m$. At temperature T, such a fluid has energy density satisfying the relation

$$\rho_m(T) = \rho_{GS} + \frac{\zeta(3/2)\Gamma(3/2)mc^2 + \zeta(5/2)\Gamma(5/2)k_B T}{(2\pi)^2 \hbar^3} (2mk_B T)^{3/2} \tag{3.99}$$

where $\rho_{GS} = \frac{N_{condensate}}{V_\epsilon}\sqrt{m^2 + k_\epsilon^2} \cong m\frac{N_{condensate}}{V_\epsilon}$. Critical temperature for the pre-coherent state is again determined when $\rho_{GS} = 0$. This corresponds to a temperature of around $T_{crit} \approx 103.8$ GeV for a pre-thermalized system consisting of only scalar particles $m \cong 19.74$GeV. The thermodynamics after the availability of sub-luminal degrees of freedom will define the temperature of thermalization using

$$\rho_{FL} = g(T_\epsilon)\frac{\pi^2}{30}\frac{(k_B T_\epsilon)^4}{(\hbar c)^3} + \rho_m(T_\epsilon). \tag{3.100}$$

Again, a solution without condensate ρ_{GS} is consistent after thermalization, at a temperature of de-coherence given by $T_\epsilon \cong 15.18$ GeV.

Just as de-coherence begins, we expect the fraction of condensate to thermal scalar masses to satisfy

$$\frac{N_{condensate}}{N_{thermal}} = 1 - \left(\frac{T_\epsilon}{T_{crit}}\right)^{3/2}, \qquad (3.101)$$

where the total number of thermal scalars satisfies

$$\frac{N_{thermal}}{V_\epsilon} = \frac{\zeta(3/2)\Gamma(3/2)}{(2\pi)^2\hbar^3}\left(2mk_B T_{crit}\right)^{3/2}. \qquad (3.102)$$

As decoherence begins, the condensate density fraction is given by $\frac{\rho_{condensate}}{\rho_{FL}} \cong 0.98$ and $\frac{N_{condensate}}{N_{thermal}} \cong 0.94$ starting thermalization. After de-coherence, $\frac{\rho_m}{\rho_{FL}} \cong 0.018$, which means that less than 2 % of the thermalized matter is made up of masses m. The relationship between the pre-coherent and thermal scalars is then given by

$$\rho_{FL} = \left(m + \frac{\zeta(5/2)\Gamma(5/2)}{\zeta(3/2)\Gamma(3/2)}k_B T_{crit}\right)\frac{N_{thermal}}{V_\epsilon} = 2N_\epsilon \rho_\Lambda. \qquad (3.103)$$

This gives a ratio of pre-coherent to thermal masses m given by

$$\frac{N_\epsilon}{N_{thermal}} \sim 5. \qquad (3.104)$$

This means that the inbalance in the ratio of the number of pre-coherent to de-coherent states in the thermalization process for a massive condensate is not as severe as that for a massless condensate.

3.8 Pre-coherence

Although up to this point we have avoided examining the cosmology prior to dark energy de-coherence, it is useful to conjecture on the continuity of the physics of this period. Since the cosmological scale expansion is supraluminal, only gravitational interactions are available for cosmological equilibrations. We will assume that the cosmological scale excitations will have energies that satisfy the usual Planck relation, only with propagation speed determined by the expansion rate:

$$E_\epsilon = h\nu = \frac{h\dot{R}}{\lambda} = \frac{h\dot{R}}{2R} = \pi\hbar H \tag{3.105}$$

For the scalar long range gravitating quanta (collective modes) discussed previously, the density of states is expected to be of the form

$$\Delta^3 n = \frac{V}{(2\pi)^3} d^3k = \frac{4\pi}{(\pi\hbar)^3}\frac{E^2 dE}{H^3} \tag{3.106}$$

If there is thermal equilibration, we therefore expect the usual forms for a scalar boson, with the substitution $\hbar c \rightarrow \hbar\dot{R}$. In particular, the energy density takes the form

$$\rho = \frac{\pi^2}{30}\frac{(k_B T)^4}{(\hbar\dot{R})^3} = \frac{\pi^2}{30}\left(\frac{1}{\hbar R}\right)^3 \frac{(k_B T)^4}{H^3}. \tag{3.107}$$

We assume that the FL equation continues to drive the dynamics, which allows substitution of the Hubble rate in terms of density

$$\rho = \frac{\pi^2}{30}\left(\frac{1}{\hbar R}\right)^3 \frac{(k_B T)^4}{\left[\frac{8\pi G_N}{3}(\rho + \rho_\Lambda)\right]^{\frac{3}{2}}}. \tag{3.108}$$

Since these gravons are expected to behave like radiation $\frac{\rho}{\rho_{FL}} = \left(\frac{R_\epsilon}{R}\right)^4$ (as any condensate is likewise expected to consistently scale), we determine the scaling of temperature with cosmological scale

$$\left(\frac{R_\epsilon}{R}\right)^7 = \left(\frac{T}{T_\epsilon}\right)^4 \tag{3.109}$$

Thus, we see that the scaling of temperature with inverse FRW scale factor no longer holds. As suspected, the equation of state is considerably altered prior to dark energy de-coherence.

If we consistently continue this conjecture to determine the critical temperature for Bose condensation of the gravons, the number of quanta in a scale volume is given by

$$N = N_{condensate} + \frac{\zeta(3)}{\pi^2}\left(\frac{T}{\hbar H}\right)^3, \tag{3.110}$$

where $\zeta(3) \approx 1.202$ is the given Riemann (Euler) zeta function. As usual, the ratio of condensate to "normal" state satisfies

$$\frac{N_{condensate}}{N} = 1 - \left(\frac{T}{T_c}\right)^3. \tag{3.111}$$

The critical temperature is given by

$$k_B T_c = \left(\frac{\pi^2}{8\zeta(3)}\right)^{1/3} \hbar\sqrt{\frac{8\pi G_N}{3}\rho}. \tag{3.112}$$

Therefore, since the energy density is expected to scale like R^{-4}, we can conclude that the critical temperature scales as

$$\frac{T_c}{T_{c\epsilon}} = \left(\frac{R_\epsilon}{R}\right)^2 = \left(\frac{T}{T_\epsilon}\right)^{\frac{8}{7}}. \tag{3.113}$$

Since T_c increases more rapidly than T at higher temperatures, such a system would remain condensed at early times. This means that a system obeying this behavior would have a suppressed vacuum energy due to the condensation into the lowest momentum mode until thermalization during de-coherence. For such a system, the coherence of a supraluminal horizon need not be driven by the rapid expansion rates, but rather is a direct consequence of the global quantum coherence of the macroscopic quantum system.

Since the ratio of the temperature to the critical temperature becomes vanishingly small for the earliest times

$$\frac{T}{T_c} \sim 0.31 \left(\frac{T_\epsilon}{T}\right)^{1/7} \Rightarrow 0 \tag{3.114}$$

we feel justified in asserting that the pre-coherent cosmology which starts completely condensed will remain a zero temperature condensate until de-coherence. This condition is required to justify the use of the lowest momentum mode only in the evaluation of cosmological vacuum energy density at de-coherence.

3.9 Fluctuations

Adiabatic perturbations are those that fractionally perturb the number densities of photons and matter equally. For adiabatic perturbations, the energy density fluctuations grow according to[5]

$$\delta = \begin{cases} \delta_\epsilon \left(\frac{R(t)}{R_\epsilon}\right)^2 & radiation - dominated \\ \delta_{dust} \left(\frac{R(t)}{R_{dust}}\right) & matter - dominated. \end{cases} \tag{3.115}$$

Temperature fluctuations are expected to be related to density fluctuations using $\frac{\delta T}{T} \cong \frac{1}{3}\frac{\delta\rho}{\rho} = \frac{1}{3}\delta$. This allows us to write an accurate estimation for the scale of fluctuations during de-coherence in terms of those at last scattering

$$\delta_\epsilon = \left(\frac{R_{dust}}{R_{LS}}\right)\left(\frac{R_\epsilon}{R_{dust}}\right)^2 \delta_{LS} \cong \frac{z_{dust} z_{LS}}{z_\epsilon^2}\delta_{LS} \tag{3.116}$$

if the fluctuations are "fixed" at dark energy de-coherence. Assuming the values for z_{dust} and z_ϵ calculated previously, along with the red shift at last scattering $z_{LS} \approx 1100$, this requires fluctuations fixed at dark energy de-coherence to have a value

$$\delta_\epsilon \approx 8.46 \times 10^{-27} \delta_{LS} \sim 10^{-31} \quad massless \tag{3.117}$$

$$\delta_\epsilon \approx 1.63 \times 10^{-22} \delta_{LS} \sim 10^{-27} \quad massive. \tag{3.118}$$

If quantum coherence persists such that the fluctuations are fixed at a later scale R_F, this relation gets modified to take the form

$$\delta_F \cong \frac{z_{dust} z_{LS}}{z_F^2} \delta_{LS} \approx \delta_\epsilon \left(\frac{R_F}{R_\epsilon} \right)^2 . \tag{3.119}$$

We expect the energy available for fluctuations to be of the order of the vacuum energy. This energy drives the two-point correlation function for the squared deviations from the average density, which means that we should expect the amplitude of the fluctuations to be of the order

$$\delta_{DC} \sim \left(\frac{\rho_\Lambda}{\rho_{FL} + \rho_\Lambda} \right)^{1/2} \cong \left(\frac{1}{2N_\epsilon} \right)^{1/2} , \tag{3.120}$$

regardless of the specifics of the condensate. This form also appears in the literature on fluctuations[26]. Indeed, we obtain the correct order of magnitude for fluctuations at de-coherence for either massless or massive condensates

$$\delta_{DC} \cong \begin{cases} 3.7 \times 10^{-31} & massless \\ 4.5 \times 10^{-27} & massive. \end{cases} \tag{3.121}$$

At last scattering this gives

$$\delta_{LS} \cong \begin{cases} 3.0 \times 10^{-4} & massless \\ 2.8 \times 10^{-5} & massive, \end{cases} \tag{3.122}$$

whereas for present day observations, this fluctuation is given by

$$\delta_o \cong \begin{cases} 0.3 & massless \\ 0.03 & massive, \end{cases} \tag{3.123}$$

if the fluctuation grows only linearly (which is not the case for late times). The amplitude of galaxy fluctuations is expected to be $\sigma_8 \cong 0.84$, which is the linear prediction theoretical prediction for the amplitude of fluctuations within 8 Mpc/h spheres[27]. We see that the massive condensate best matches fluctuations at last scattering (i.e. $\sim 10^{-5}$), but exploration of the agreement with present day fluctuations requires more than our simple extrapolation from last scattering.

4 Discussion and Conclusions

We feel that we have given a strong argument for the interpretation of cosmological dark energy as the vacuum energy of a zero temperature condensate of bosons. Prior to de-coherence, the scale of gravitational quantum vacuum energy is given by the Friedman-Robertson-Walker (FRW) scale $R(t)$. We have asserted that dark energy de-coherence occurs when $\dot{R} = c$, which is only consistent with a spatially flat cosmology. During de-coherence, the gravitational coherence scale of the Friedman-Lemaitre (FL) density changes considerably (most likely to be the Compton wavelength of the mass m associated with the Bose condensate, which is much less than the coherence scale of the dark energy), resulting in a gravitational phase transition, and the onset of new thermal degrees of freedom. This means that microscopic thermal interactions between components of the FL energy will break gravitational coherence, freezing the value of the gravitational dark energy. We have assumed that the quantum vacuum state for gravitation is an intrinsic state, with an energy density scale given by the vacuum energy density of the zero temperature Bose condensate during the period of last quantum coherence (given by $\epsilon = \frac{1}{2}\sqrt{m^2 + k_\epsilon^2}$, which is determined by the current value of the cosmological constant). When the cosmology has global coherence, the gravitational vacuum state is expected to evolve with the contents of the universe. When global coherence is lost, there remains only local coherence within independent clusters, and the prior vacuum state loses scale coherence with the clusters as the new degrees of freedom become available. This dark energy scale will be frozen out as a cosmological constant of positive energy density satisfying $\rho_\Lambda = \frac{\Lambda}{8\pi G_N} = \frac{\epsilon k_\epsilon^3}{(2\pi)^3}$ in terms of the present day cosmological constant.

To determine the ultimate fate of the universe, one needs an understanding of the fundamental nature of the quantum vacuum. The Wheeler-Feynman interpretation of the propagation of quanta irreducibly binds those quanta to their sources and sinks. A previous paper by these authors directly demonstrates the equivalence of the usual Compton scattering process calculated using photons as the asymptotic states in standard QED with a description that explicitly includes the source and sink of the scattered photons in a relativistic three-particle formalism[23]. According to Lifshitz and others[7], the zero temperature electromagnetic field in the Casimir effect can be derived in terms of the zero-point motions of the sources and sinks upon which the forces act. In the absence of a causal connection between those sources and sinks, one has a difficult time giving physical meaning to a vacuum energy or Casimir effect. Since the zero-point motions produce classical electromagnetic fields in Landau's treatment, these fields propagate through the"vacuum" at c. This would mean that one expects the Casimir effect to be absent between comoving mirrors in a cosmology with $\dot{R} > c$. If the regions in a future cosmology whose expansion are driven by vacuum energy are indeed causally disjoint, then there could be no driving of that expansion due to the local cosmological constant. Such an expansion requires that gravitational interactions propagate in a manner that causally affects regions requiring super-luminal correlations. The expected change in the equation of state for the cosmology as a whole should modify the behavior of the FRW expansion in a manner that would require reinterpretation of the vacuum energy term, as seems to be necessary during the pre-coherence epoch.

This means that we do not view the cosmological constant as the *same* as vacuum energy density. In our interpretation, cosmological vacuum energy changes during pre-coherence and post-coherence. The cosmological constant is frozen at de-coherence due to the availability of luminal degrees of freedom. Since this vacuum energy density is associated with regions of global gravitational coherence, it is interesting to consider whether subsequent expansion in space-time will re-establish coherence on a cosmological scale.

We are in the process of examining the power spectrum of fluctuations expected to be generated by de-coherence as developed. It is our hope that further explorations of the specifics of density and temperature fluctuations will allow us to better differentiate between massless vs various massive condensates. In addition, we have begun to examine the scale of the symmetry breaking involved in coherence, especially with regards to the mass scales involved. It is our belief that this approach will reduce the parameter set needed to describe a consistent cosmology.

Acknowledgment

We are most grateful, once again, to E.D.Jones for permitting us to discuss and extend his theory before his own paper is available. We also extend our gratitude to Walter Lamb, whose careful reading and criticism has enhanced the clarity of our presentation.

References

[1] A.W.Overhauser and R.Colella, Phys.Rev.Lett. **33**, 1237 (1974).

[2] R.Colella, A.W.Overhauser and S.A.Werner, Phys.Rcv.Lctt. **34**, 1472 (1975).

[3] C.L. Bennett, et.al., Astrophys.J.Supp. 148, 1 (2003)

[4] A.G.Riess et al, Astron.J. 116, 1009 (1998); P. Garnavich et al, Astrophys.J. 509, 74 (1998); S. Perlmutter et al, Astrophys.J. 517, 565 (1999)

[5] Particle Data Group, **Astrophysics and cosmology**, as posted (2004).

[6] H.B.G.Casimir, Proc.K.Ned.Akad.Wet. 51, 793 (1948)

[7] E.M. Lifshitz, Soviet Phys. JETP 2, 73 (1956). I.D. Dzyaloshinskii, E.M. Lifshitz, and I.P. Pitaevskii, Soviet Phys. Usp. 4, 153 (1961). I.D. Landau and E.M. Lifshitz, *Electrodynamics of Continuous Media*, pp368-376 (Pergamon, Oxford, 1960)

[8] T.H. Boyer, Phys. Rev. 174, 1764 (1968)

[9] C. Genet, A. Lambrecht, and S. Reynaud, "Casimir effect and vacuum energy," quant-ph/0210173 (2002) 10 pages

[10] Particle Data Group, **Astrophysical Constants**, as posted (2004).

[11] H.P.Noyes, "A Cosmological Calculation Suggesting a Threshold for New Physica at 5 Tev", hep-th/0307250v1 24 Jul 2003,

[12] E.D.Jones, conversations with HPN starting in the 1990's when there was no enthusiasm in cosmological circles fo a "cosmological constant" other than zero, let alone a positive one.

[13] E.D.Jones, private communication to HPN c. 1997. The result of Jones' calculation which gave $\Omega_\Lambda = 0.6 \pm 0.1$ is quoted in *Aspects II (Proc. ANPA 20)*, K.G.Bowden, ed. (1999), p. 207, and cited again in[18], p. 561.

[14] Private presentation by E.D. Jones at SLAC to L.H.Kauffman, W.R.Lamb and HPN on 29 April 2002.

[15] H.P.Noyes, L.H.Kauffman, J.V.Lindesay and W.R.Lamb, "On E.D.Jones' MICROCOSMOLOGY"[16], pp 129-148, and astro-ph/0301176 v1 10 Jan 2003,

[16] *Boundaries: Scientific Aspects of ANPA 24*, K.Bowden, ed.; published by ANPA, c/o Dr. K. Bowden, Theoretical Physics Research Unit, Birkbeck College, Malet St. London WC1E 7HX, July 2003.

[17] H.P.Noyes, "Non-Locality in Particle Physics" prepared for an unpublished volume entitled *Revisionary Philosophy and Science*, R.Sheldrake and D.Emmet, eds; available as SLAC-PUB-1405 (Revised Nov. 1975); reprinted as Ch. 1 in [18].

[18] H.P.Noyes, *Bit-String Physics*, World Scientific, Singapore 2001, Vol. 27 in *Series on Knots and Everything.*

[19] F.J.Dyson, *Phys.Rev.*, **85**, 631 (1952), and a seminar at about that date.

[20] S.S.Schweber, *QED and the Men Who Made It*, Princeton, 1994, Sec. 9.17, pp. 564-65.

[21] M.Alfred, P.Kwizera, J.V.Lindesay and H.P.Noyes, "A Non-Perturbative, Finite Particle Number Approach to Relativistic Scattering Theory, hep-th/0105241, Found.Phys.34:581-616(2004). This general theory of a non perturbative relativistic, finite par-

ticle number, quantum mechanics, has the appropriate particle-antiparticle symmetries[24], and gives the usual, manifestly covariant amplitude for the Compton scattering relevant here in the weak coupling limit[23].

[22] J.V.Lindesay, A.Markevich, H.P.Noyes, and G.Pastrana, "Self-Consistent Poincare-Invariant and Unitary 3-Particle Theory", Phys.Rev.D.33,2339-2349(1986).

[23] J.V.Lindesay and H.P.Noyes, "Non-Perturbative, Unitary Quantum-Particle Scattering Amplitudes from Three-Particle Equations", hep-th/0203262.

[24] J.V.Lindesay and H.P.Noyes, "Construction of Non-Perturbative, Unitary Particle-Antiparticle Amplitudes for Finite Particle Number Scattering Formalisms", nucl-th/0203042.

[25] W. Fischler and L. Susskind, Holography and Cosmology, hep-th/9806039, 7 pages. L. Susskind, The World as a Lologram, hep-th/9409089, J.Math.Phys. 36, 6377-6396 (1995). D.Bigatti and L.Susskind, *TASI lectures on the Holographic Principle*, hep-th/0002044.

[26] E.R. Harrison, "Fluctuations at the Threshold of Classical Cosmology," Phys.Rev. D1, 2726-2730 (1970)

[27] D.N. Spergel, et.al., "First Year WMAP Observations: Determination of Cosmological Parameters", astro-ph/0302209 (2003) 51 pages

SLAC-PUB-11063
18 March 2005

SCIENTIFIC ESCHATOLOGY*

H. Pierre Noyes and James V. Lindesay†
Stanford Linear Accelerator Center MS 81,
Stanford University, 2575 Sand Hill Road, Menlo Park CA 94025,
noyes@slac.stanford.edu

Abstract

The future evolution of the universe suggested by the cosmological model proposed earlier at this meeting by the authors is explored. The fundamental role played by the positive "cosmological constant" is emphasized. Dyson's 1979 paper entitled *Time Without End* is briefly reviewed. His most optimistic scenario requires that the universe be *geometrically* open and that biology is structural in the sense that the current complexity of human society can be reproduced by scaling up its (quantum mechanical) structure to arbitrary size. If the recently measured "cosmological constant" is indeed a fundamental constant of nature, then Dyson's scenario is, for various reasons, ruled out by the finite (De Sitter) horizon due to exponential expansion of the resulting space. However, the finite temperature of that horizon does open other interesting options. If, as is suggested by the cosmology under consideration, the current exponential expansion of the universe is due to a phase transition which fixes a *physical* boundary condition during the early radiation dominated era, the behavior of the universe after the relevant scale factor crosses the De Sitter radius opens up still other possibilities. The relevance of Martin Rees' apocalyptic eschatology recently presented in his book *Our Final Hour* is mentioned. It is concluded that even for

*Work supported in part by Department of Energy contract DE–AC03–76SF00515.

†Permanent address: Department of Physics and Computational Physics Lab, Howard University, Washington, D.C. 20059, jlslac@slac.stanford.edu

the far future, whether or not cultural and scientific descendants of the current epoch will play a role in it, an understanding (sadly, currently lacking) of community and political evolution and control is essential for a preliminary treatment of what could be even vaguely called *scientific* eschatology.

A preliminary version of this paper was presented at
The Twenty-Sixth Annual Meeting of the
ALTERNATIVE NATURAL PHILOSOPHY ASSOCIATION
Cambridge, England, 31 July - 5 August, 2004

1 Introduction

On the Silver Jubilee of the founding of the Alternative Natural Philosophy Association it was possible to present a new piece of natural philosophy[1, 2, 3] which the authors believe has profound implications for the new field of particle-astrophysics/cosmology; SLAC has recently made this discipline an important part of its mission. What follows is a discussion of a few of the important questions raised by the future implications of this cosmology if it, and other cosmologies which share with it a positive "cosmological constant", survive the rigorous scientific scrutiny they are now receiving. The authors believe that the way some of these implications are absorbed by our science and the broader cultures in which our science is embedded will have relevance to the enormous problems the children and grandchildren of those present here will have to face in the coming decades.

During the last fifty years, scientific cosmology has moved from being a speculative field which could be viewed askance by many scientists not intimately associated with it to being one of the "cutting edge" disciplines — both theoretically and observationally — of 21st century science. But most of the technical work in cosmology is confined, necessarily, to the study of the past. The study of the future of the universe rarely gets such careful attention. One goal of this paper is to try to motivate more working scientists to bring about a similar change of attitude and practice in what we call here *scientific eschatology.*

Coincidentally with the founding of ANPA, Dyson[4] published a paper with the same objective. At that time he could take a geometric view of the problem, defined by whether the universal curvature parameter is positive (closed universes) or negative (open universes). One of his objectives in the paper was, clearly, to make a case for the possibility that rational, scientific, ecological communities of living organisms can look forward to a "time without end". He easily shows that if the members of such communities depend on the biochemistry with which we are familiar, or, in more colloquial terms, contain "flesh and blood" organisms as an essential constituent, there is no hope that his objective can be reached within either class of cosmologies he considers. He therefore extends the definition of "biology" to include "organisms" (and by implication, ecosystems composed of them) that have the same *structure* as those we have become familiar with on our planet. This was done by postulating a scaling law for quantum mechanical interactions which gave precision to his meaning of *structure*. Even in this context, he was unable to meet his objective of preserving rationality forever within closed ("big crunch") cosmologies. But for open cosmologies he found a way to show that appropriate "biological" strategies could provide a (subjective) *time without end* for organisms and ecosystems of ever increasing complexity. We will summarize his arguments below.

At the time Dyson wrote, observational evidence for a "cosmological constant" did not exist. Einstein's motivation for introducing this constant (i.e. to preserve, in the large, a static and infinite universe satisfying his *cosmological principle*) had evaporated with discovery of the red shift for distant galaxies and Hubble's law. So there was no reason for Dyson to include in his discussion the De Sitter cosmologies produced by a positive cosmological constant. We now know that until about 5 Giga-years ago the rate of expansion of the universe was *decreasing.* We also know that more recently the rate has been *increasing.* From very early times until the present ($\sim 13.6\ Gyr$) the observational data can be fitted by a positive cosmological constant Λ together with an evolving matter density which can be checked in other ways. The data are consistent with Λ being strictly constant throughout this period. If Λ is indeed a "constant of nature", this fact would render Dyson's analysis moot.

A review of Dyson's paper presented here reaches the conclusion that, under his assumptions (which did not include the possibility of a cosmological constant), a "time without end" for the type of "biology" he defines was then a live possibility. Next, the consequences of having a cosmological constant fixed for all time will be examined. At first sight, these consequences are indeed dismal. They could be met by the tragic remarks of Bertrand Russell[5] in *A Free Man's Worship* or the defiant challenge with which Dyson closes his own article:

> As Haldane (1924) [6] the biologist wrote some fifty years ago, "The human intellect is feeble, and there are times when it does not assert the infinity of its claims. But even then:
>
> Though in black jest it bows and nods
> I know it is roaring at the gods
> Waiting the last eclipse,"

Fortunately, we believe there are possible routes to "time without end" which might be reached from where we are in an accelerating expansion scenario. The physics in the class of models presented earlier in this meeting[1] need not *require* the "cosmological constant" to be fixed forever, — only over a well defined and finite period of time. This opens up a new set of possibilities for biological strategies, which might allow our intellectual and cultural heirs to win through to an indefinitely extendable future. The authors of this paper have only scratched the surface of this fascinating subject, but this analysis suggests that heroic action may well be required, not just now, but for many Giga-years into the future. That prompt action in this century is clearly required has been recently argued by the Astronomer Royal[7]. Older members of ANPA will recall, and more recent members should be made aware, that Martin Rees played a highly significant role in supporting the first tottering steps of ANPA toward the robust success we now enjoy. The authors of this article trust he will not mind too much if we take his measured analysis of the current desperate situation as the basis for a plea to the scientific community at large to recognize the tremendous need for a rigorous *political* science; this argument clearly goes beyond the case Dyson has made for including an extended scientific *biology* in order to establish an adequate basis for *scientific eschatology.* But to make this case we must now complete the steps we have already indicated that lead us to that conclusion.

2 Dyson's "Time without end"

We now take a closer look at Dyson's pioneering paper. As already noted he rapidly concludes that a "big crunch" universe docs not provide enough scope for an optimistic view of the far future. Even in an open (or flat) universe, Dyson's argument closes the door on the continuation of "flesh and blood" biology of a complexity comparable to that currently experienced on our planet for a *physical* time without end. This focuses his interest on the biology of "...sentient black clouds[8], or sentient computers[9]...",— structures which we know some (most?) ANPA members dismiss as impossible. To make his argument plausible and quantitative, he proposes a "biological scaling hypothesis" that has as its first consequence the conclusion that the "...appropriate measure of time as experienced subjectively by a living creature is not physical time t but ..." an integral from zero to t of the temperature function defined by his scaling law ([4], Eq.56). This is called *subjective* time. "The second consequence of the scaling law is that any creature is characterized by a quantity Q which measures its rate of entropy production per unit of subjective time." For a human being living long enough to pronounce "Cogito, Ergo Sum" this works out to be $\sim 10^{23}$ bits, and for our species $\sim 10^{33}$ bits. This sets a fixed lower bound for the temperature θ, which is ([4], Eq. 73) $\theta > (Q/N)\ 10^{-12} deg$, where N is the number of electrons available to the society of complexity Q. For our current biosphere $N = 10^{42}$, so our current social complexity cannot be maintained at temperatures lower that $10^{-23} deq$.

Here a few comments are in order. Dyson's biological scaling hypothesis ([4], p. 454) is

> *Biological Scaling Hypothesis. If we copy a living creature, quantum state by quantum state, so that the Hamiltonian of the copy is*
>
> $$H_c = \lambda U H U^{-1}\,,$$
>
> *where H is the Hamiltonian of the creature, U is a unitary operator, and λ is a positive scaling factor, and if the environment is similarly copied so that the temperatures of the environments of the creature and the copy are respectively T and λT, then the copy is alive, subjectively identical to the original creature, with all its vital functions reduced in speed*

by the same factor λ.

This hypothesis is made in a context which should be spelled out further. Dyson starts his section on biology by posing three deep questions concerning the nature of life and consciousness:

(i) Is the basis of consciousness matter or structure?

(ii) Are sentient black clouds, or sentient computers, possible?

(iii) Can we apply scaling laws in biology?

As is clear in his article, if the answer to (i) is "matter", Dyson takes this as, in more colloquial language, tying consciousness inexorably to "flesh and blood" and in the cosmological context to certain death. There are traditional ways (eg religious rather than scientific) to escape this dismal conclusion, but the authors of this article follow, instead, Dyson's optimistic attitude and assume that the basis of consciousness is structural rather than material. The reader should consult his article (and other writings) to see how he justifies his own attitude. He notes that we (i.e. our culture) do not yet know how to answer these three questions "But they are not in principle unanswerable. It is possible that they will be answered fairly soon as a result of progress in experimental biology."

As just explained, Dyson's tentative answer to question (i) is "structure". The fact that "quantum computers" are now claimed to have greatly extended powers compared to "classical computers" lends weight to his tentative conclusion. His tentative "yes" answer to (ii) puts him in the "strong AI" camp. The authors of this article are not sure this is the right (even tentative) answer today, but HPN feels it likely that *communities* of computers starting in environments with appropriate resources and with inheritance mechanisms and survival pressures sufficiently similar to those encountered in biology would evolve into recognizably conscious beings. The technical point here is that a community that has evolved to the point where its members have independent choices of action available (i.e. have "free will") is obviously not an *algorithmic computer*.

Support for Dyson's tentative "yes" answer to (iii) is provided by recent work by Fred Young[10], a former president of ANPA. Young makes a strong case that the basic structures of biology can be represented by scaling laws using small-protein concentration ratios for physiological modelling at the cellular level, tissue and system organizational modelling at the organism level, biological species organization at the ecological level, etc. The authors of this article find Young's results to date quite compelling and extremely promising for the future of his approach to biology.

We now return to Dyson's paper at the point where he has established the finite temperature bound $\theta > (Q/N)\ 10^{-12} deg$ below which any society of complexity Q bits controlling N electrons must not fall. This bound is arrived at by asserting that the rate of energy dissipation (i.e. use of energy by the society) must not exceed the power that can be radiated away into space. At that point it can still dissipate the energy which it must expend to keep on operating. Below that point it must decrease its complexity or increase its controlled number of electrons. Since the supply of energy available to the society is assumed finite, it must reach this point at a finite time, and Dyson remarks ([4], p. 456):

> We have reached the sad conclusion that the slowing down of metabolism described by my biological scaling

hypothesis is insufficient to allow a society to continue indefinitely.

Here Dyson examines a possible way to insure "time without end" for entities who "live" in this cold and forbidding future. This is, quite simply, the biological strategy of *hibernation.* Life can metabolize at a higher temperature and then hibernate at a much lower temperature to stretch out its *subjective* time. To quote Dyson again ([4], p 4550])

> Suppose then that a society spends a fraction $g(t)$ of its time in its active phase and a fraction $[1 - g(t)]$ hibernating. The cycles of activity and hibernation should be short enough so that $g(t)$ and $\theta(t)$ [Here $\theta(t)$ is a function Dyson has already assumed technologically available subject to explicit thermodynamic constraints] do not vary appreciably during any one cycle. Then (56) and (59) [previous constraints] no longer hold. Instead subjective time $[u(t)]$ is given by
>
> $$u(t) = f \int_0^t g(t')\theta(t')dt',$$
>
> [rather than $u(t) = f \int_0^t \theta(t')dt'$]

and is no longer bounded. The parameter f was chosen by Dyson to have a value of $(300\, deg\, sec)^{-1}$ as a scale comparable to that of human society to make the subjective time rate $u(t)$ dimensionless. Hence the constraints which led to his dismal conclusion no longer applied. In this way he was able to achieve a system with an infinite subjective time in an expanding universe, while expending only a finite amount of energy.

One matter where the finite energy limitation is serious is in the storage of memory. As Dyson remarks ([4], p. 456)

> I would like our descendants to be endowed not only with an infinitely long subjective lifetime but also with a memory of endlessly growing capacity. To be immortal with a finite memory is highly unsatisfactory; it seems hardly worth while to be immortal if one must erase all trace of one's origins in order to make room for new experience.

Digital memory with the finite energy resources available to a periodically hybernating society is obviously not an option which meets this requirement when *digital* memory storage in condensed matter is employed. Dyson turns to analog memory and claims that the angles between a finite number of structures (e.g. stars) in an expanding universe can be used for an *analog* memory storage of ever increasing capacity. So far as we can see here, the point to focus on is that for indefinitely expanding storage capacity for *information* to be available, the system must have no finite upper bound on the *entropy*. In Dyson's expanding universe scenario, even though the energy available to the society is finite, the unbounded expansion of the volume over which this energy is distributed means that there is, indeed, no *a priori* upper limit on the entropy. Consequently analog storage might meet the requirement he imposes. As with much of this discussion, meeting *technological* challenges this requirement poses starting from any particular configuration could prove to be daunting.

3 Asymptotically De Sitter Universes

Recent observational results have tentatively convinced most of the experts that the energy density of our universe is currently partitioned into approximately 73% dark energy, 23% dark matter and 4% ordinary matter and radiation, in a space that is flat rather than either "closed" or "open". The simplest way to fit the data is to assume that the various interlocking pieces of evidence which lead to this picture constitute an actual *discovery* of Einstein's cosmological constant Λ as a new universal constant *and* a measurement of that constant to a couple of percent. In the approach taken in the paper presented earlier in this meeting[1], these authors prefer to think of it as a *phenomenological* constant specified as constant only over a finite interval in universal time. That point of view is explored eschatologically in the next section. In this section we adopt the more naive approach.

The specific consequences of interest here which follow from this assumption are that:

a) The matter energy density that drives the cosmological expansion in the dynamical Friedman-Lemaitre (FL) equation, which we call ρ_{FL}, will eventually become insignificant compared to the cosmological constant density $\rho_\Lambda = \frac{\Lambda c^4}{8\pi G_N}$ (here G_N is Newton's gravitational constant).

(b) Consequently, the FL (Hubble) equation is replaced asymptotically by $[\frac{\dot{R}}{R}]^2 \simeq \frac{8\pi G_N}{3c^2}\rho_\Lambda = \Lambda c^2/3$, which implies that $R(t) =$

$R_E e^{\sqrt{\frac{\Lambda}{3}}ct} \equiv R_E e^{\frac{ct}{R_\Lambda}}$. Here R_Λ is sometimes called the De Sitter radius or horizon, and R_E is set at a time when ρ_{FL} is negligible compared to ρ_Λ. Although objects of the scale of the gravitationally bound super-cluster have dynamics which are determined by local matter densities, the late time exponential expansion defines a cosmological De Sitter horizon $R_\Lambda = \sqrt{\frac{3}{\Lambda}} \simeq 16.6\ Glyr$ which serves as a causal boundary, i.e. anything which crosses this boundary can never re-establish luminal contact with our region of the universe. Our galaxy appears to be close to the edge of, and probably bound to a super-cluster with a radius of about 50 mega-parsecs $\approx 0.16\, Gly$. If we therefore take the current FL scale parameter ($R_0 = R(t_0)$ at time t_0) in our (gravitationally bound) locality to be about 100 mega-parsecs, we find that this scale parameter will cross the De Sitter horizon (i.e. $R(t) = R_\Lambda$) when $t \simeq t_0 + 65Gyr$. Here the current time t_0 is usually taken to be about 13.6 Gyr. This means that all other galaxies not in our local super-cluster will vanish within about 65 Gyr. This consequence already precludes useful discussion of Dyson's analog memory storage and far-ranging communications strategies if Λ is indeed a universal constant on a par with $\hbar, c, G_N$ and k_B. Both of these strategies rely on scaling laws that assume causal (i.e. luminal) contact can always exist if the society waits long enough. For further discussion see (c), (i) and Section 4 below.

(c) Since the De Sitter horizon has a finite area, the causal region of the De Sitter space will have[11] a finite entropy $S_\Lambda = k_B \frac{\pi R_\Lambda^2}{L_P^2}$, where $L_P = (\hbar G_N)^{\frac{1}{2}} c^{-\frac{3}{2}}$ is the Planck length and k_B is Boltzmann's constant.

(i) Because finite entropy implies finite information storage capacity, this fact precludes, simply using counting arguments, any way of realizing Dyson's analog storage method for constructing an indefinitely extendable memory. As a counting argument, this holds for quantum-coherent systems (eg quantum computers) as well as for digital computers.

(ii) Systems with finite entropy undergo Poincaré recurrences[12]. Such recurrences are due to the finite number of configurations (microstates) available to a system with finite entropy. Because there are only a finite number of configurations that the system fluctuates among, the system will eventually return to any given initial configuration. Since these recurrences have only to do with counting of states, this fact applies to both classical and quantum statistical systems. Such recurrences are maximally destructive of information.

(iii) One strategy of despair that has sometimes been suggested is to find a way to pass clues about our experience in this universe through its fiery destruction to provide information that could prove useful to new societies evolving in the cycle that might emerge after we are consumed. Clearly, the destruction of information precludes placing any hope in this possibility.

(iv) This list of unpleasant facts about a De Sitter universe could easily be extended. This might explain why Dyson has not, to our knowledge, extended his analysis to such universes.

There may still be a finite strategy arising from the fact that in such universes the De Sitter horizon maintains a finite temperature. The existence of a horizon creates an information deficit within the space bounded by that horizon. The subtle quantum correlations in states near the horizon must be described statistically in terms of the degrees of freedom and parameters accessible in the causal region. Whenever there are such statistical degeneracies in ways to describe a particular physical state, the concept of temperature becomes a meaningful tool to describe macroscopic states. There are many states across any horizon which can describe any given measured state within the causal region, thus associating an entropy and temperature with that horizon. Consequently, this could serve as an inexhaustible source of energy to run a *steady state* (constant rate of energy throughput) forever. For instance a sufficiently clever technological society could focus the thermal energy coming from some finite area of the horizon on a "boiler" which could be maintained at a higher temperature than the horizon. That society could then employ a thermodynamically viable engine to extract useful work (e.g. turn heat into low temperature matter or energy density) by a cycle between the "boiler" and a "condenser" at some temperature intermediate between the boiler temperature and the horizon temperature. To be viable, the time scale of the work cycle must be much less than the equilibration time of the heat engine[13]. This engine will, of course, be a non-equilibrium element of a much larger thermodynamic system which includes the De Sitter horizon. The condenser would require an efficient radiator (essentially already assumed possible by Dyson) to get rid of the thermodynamically required waste heat for the cycle. So far as we can see, such a device is not in conflict with the second law of thermodynamics. This society would have the inescapable informational problems arising from finite memory, and political problems arising from having to decide how the finite resources are budgeted between current and future needs. However, we will have to find solutions to the political problems if we are to get through the next century, as we discuss later.

A more ambitious, and more speculative possibility is that the "steady state" societies we have considered so far are not energy limited. Consequently, they might even sequester more energy than they need to keep going, and become energy *accumulating* societies containing expanding resources of condensed matter, part of which could be devoted to expanding local resources. Perhaps it might be possible to use accumulated resources to modify the cosmology itself.

To explore this possibility, examine Einstein's equation, which describes the connection of the local geometry to the local energy densities:

$$G_{\mu\nu} = \frac{8\pi G_N}{c^4} T_{\mu\nu} + \Lambda g_{\mu\nu}. \tag{3.1}$$

We assume that the local, gravitationally bound, energy densities which remain have clustered in a spherically symmetric manner. The space-time metric for a system with spherical symmetry has the form $ds^2 = g_{tt}c^2dt^2 + g_{rr}dr^2 + g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2$. Upon inserting this form into Einstein's equation, the radial component of the metric can be shown to be given by

$$(g_{rr})^{-1} = 1 - \frac{8\pi G_N}{c^4\, r} \int_0^r (\rho(\tilde{r}) + \rho_\Lambda)\tilde{r}^2 d\tilde{r}. \tag{3.2}$$

The time-time component g_{tt} of the metric and the inverse of the space component g_{rr}^{-1} will have the same zero, which determines the location of the horizon, and fixes the center as the same as the center of the mass distribution.

For example, if there is a mass M associated with the galactic super cluster more or less localized near the center of the region, the horizon scale r_H associated with the local cosmology satisfies

$$0 = 1 - \frac{2G_N M}{c^2 r_H} - \frac{r_H^2}{R_\Lambda^2} = 1 - \frac{R_M}{r_H} - \frac{r_H^2}{R_\Lambda^2}, \qquad (3.3)$$

where R_M is the Schwarzschild radius $R_M \equiv \frac{2G_N M}{c^2}$. Note that if $M = 0$, the horizon is located at R_Λ as expected. It is of interest to note that if $R_M > \frac{2R_\Lambda}{3\sqrt{3}}$, there is no horizon in the causal region. One might hope to be able to eliminate the De Sitter horizon by increasing the mass M, however the society would be trapped between two horizons which are drawn together as the mass M is increased. This would indeed be a hostile environment for finite beings.

Alternatively, if the mass density ρ_M is uniformly distributed throughout the region, the horizon scale satisfies

$$r_H = \sqrt{\left(\frac{3c^4}{8\pi G_N}\right)\frac{1}{\rho_m + \rho_\Lambda}}. \qquad (3.4)$$

The entropy associated with this horizon is given by $S = k_B \frac{\pi r_H^2}{L_P^2}$, whereas the temperature is given by $T = \frac{\hbar c}{2\pi k_B r_H}$. Any societal activity which is capable of increasing the energy density within the causal region would be expected to utilize processes that preserve entropy, converting entropy from the larger horizon (at a cooler temperature) into the lower horizon entropy (of smaller area and higher temperature) added to the local entropy densities associated with the increased local energy densities. Such adiabatic processes will unfortunately not change the overall finite entropy, and therefore not prevent recurrences or solve information limits. However, if such processes can occur, the society has access to increasing energy supplies, increasing temperatures, and the associated increasing rates of subjective time as defined by Dyson.

3.1 A possible societal strategy

We see that if a society is able to modify the overall energy contained within the causal region, then the regional cosmology can be changed. It remains to demonstrate whether such activities can be fruitful even in principle. Suppose the society disperses heat engines uniformly throughout the region, and uses the cold materials produced by those engines to construct other engines. The associated energy densities would then satisfy

$$\frac{d\rho_M}{dt} = \alpha\rho_M \Rightarrow \rho_M = \rho_{Mo}e^{\alpha t} , \qquad (3.5)$$

where α is determined by the efficiency of the engines in converting horizon heat into cold materials. The initial available density of engines ρ_{Mo} is expected to be a fraction of the galactic super cluster density relative to the De Sitter horizon. Therefore, the society would make significant modifications to the De Sitter cosmology in the region in a time of the order

$$t_{modify\ DeSitter} = \frac{1}{\alpha} log \left(\frac{\rho_\Lambda}{\rho_{Mo}} \right), \tag{3.6}$$

as long as this time is significantly less than the Poincaré recurrence time.

We will estimate the scale of this time by assuming that the thermal distribution of particles associated with the De Sitter horizon is the same as that associated with a black body cavity at that temperature. The efficiency of a Carnot engine running between the optimal temperatures is given by $e = \frac{W_{out}}{Q_{in}} = 1 - \frac{T_H}{T_{boiler}} < 1$. The thermal distribution of low mass particles in the cavity satisfies

$$n(\epsilon)d\epsilon = \frac{g}{2\pi^2(\hbar c)^3} \frac{\epsilon^2 d\epsilon}{e^{\epsilon/k_B T} - 1}, \tag{3.7}$$

where $n(\epsilon)$ is the number of quanta per unit energy per unit volume, and g is the degeneracy of the thermal quanta. We will design the collectors on the heat engines so that they are efficient at collecting radiations of wavelengths ($\lambda_{collected} \leq \lambda$

) comparable and smaller than the dimension of the collector of area $\lambda \times \lambda$. We will assume that the collector scale is considerably smaller than the horizon scale, which means that the exponential in the Planck formula is much larger than 1. Only a fraction f_Q of those quanta in a region of space will have directions toward the collector, and the rate at which usable quanta strike the collector can be estimated to be

$$\frac{\Delta \bar{N}_{collected}}{\Delta t} \approx f_Q \left(\int_{hc/\lambda}^{\infty} n(\epsilon) d\epsilon \right) \times c \times Area_{collector}. \qquad (3.8)$$

Since the area of the collector defines the wavelength λ^2, the rate of the collection of quanta is estimated to be

$$\frac{\Delta \bar{N}_{collected}}{\Delta t} \approx f_Q \frac{g}{\pi} \left(\frac{c}{R_\Lambda} \right) e^{-(2\pi)^2 \frac{R_\Lambda}{\lambda}}. \qquad (3.9)$$

This rate is clearly exponentially small in the horizon scale compared to the scale size of the heat collector. We expect that the rate of heat conversion α should be related to the rate of quantum collection multiplied by the average energy per quantum collected relative to the rest energy mc^2 of each heat engine, in order to satisfy Eq. 3.5. Thus, an estimate of the rate of exponential growth in energy density in the society due to the heat engines is given by

$$\begin{aligned} \alpha &\sim f_Q \tfrac{g}{\pi} \left(\tfrac{c}{R_\Lambda}\right) \tfrac{hc/\lambda}{mc^2} e^{-(2\pi)^2 \frac{R_\Lambda}{\lambda}} \\ &\approx f_Q 2g \left(\tfrac{c}{R_\Lambda}\right) \tfrac{\lambda_m}{\lambda}\, e^{-(2\pi)^2 \frac{R_\Lambda}{\lambda}}, \end{aligned} \tag{3.10}$$

where λ_m is the Compton wavelength of the mass of each heat engine.

We now can compare the time scale required for such heat engines to have (local) cosmological significance with the Poincaré recurrence time scale of that cosmology. Recurrences are expected to occur stochastically on time scales given by[12]

$$t_{recurrence} \cong t_{reshuffle}\, e^{\frac{S_\Lambda}{k_B}}, \tag{3.11}$$

where $t_{reshuffle}$ is the typical time scale associated with the microscopic reshuffling of those configurations that give rise to the finite entropy S_Λ. The configurations are counted by the thermodynamic weight Ω in the Boltzmann identification $S = k_B log\Omega$. This reshuffling time can be estimated to be a fraction f_R of the causal transit time $\frac{R_\Lambda}{c}$ across the De Sitter patch (causal region), giving recurrence times of the order

$$t_{recurrence} \cong f_R \left(\frac{R_\Lambda}{c}\right) e^{\pi\left(\frac{R_\Lambda}{L_P}\right)^2}. \tag{3.12}$$

Comparison of the recurrence time with the energy accumulation rate of the heat engines given by

$$t_{modify\,DeSitter} \approx \frac{\pi}{f_Q\, g\, log\left(\frac{\rho_\Lambda}{\rho_{Mo}}\right)} \left(\frac{R_\Lambda}{c}\right) \frac{\lambda}{\lambda_m}\, e^{(2\pi)^2 \frac{R_\Lambda}{\lambda}} \tag{3.13}$$

gives design constraints on the size of the heat engines:

$$\frac{\lambda}{\lambda_m}\, e^{(2\pi)^2 \frac{R_\Lambda}{\lambda}} < \frac{f_R\, f_Q\, g\, log\left(\frac{\rho_\Lambda}{\rho_{Mo}}\right)}{\pi}\, e^{\pi\left(\frac{R_\Lambda}{L_P}\right)^2} . \tag{3.14}$$

Although on the left hand side of the equation, the macroscopic size of the collector λ is expected to be decades of orders of magnitude larger than the Compton wavelength associated with the mass of the heat engine, the exponentiation of the square of the horizon scale relative to the Planck length on the right hand side of the equation should give considerable flexibility in the design of a workable engine.

Since there seem to be no inconsistencies in building engines which can utilize the heat of the horizon to store energy density, we can speculate on the consequences of such actions. As the horizon scale shrinks, the temperature associated with the horizon increases. Using Dyson's association of the subjective time rate with the temperature of the organism/society, any evolving "biological" organisms will have increasing rates of subjective time. However, if the process of energy density storage is indeed adiabatic, the overall entropy of the local cosmology will remain fixed, leaving the Poincaré recurrence time unchanged. Such a strategy could indeed increase *relative* subjective time without bound. In such an environment, Dyson's strategy of hybernation would not be useful, since it would only serve to slow down relative subjective time. No mention is here made of the formidable engineering tasks involved in constructing heat engines which function efficiently in increasingly hot spaces. We have been unable to come up with any *a priori* reasons against such societal intervention on a cosmological scale. Note that in contrast to Dyson's scenario which accommodates biology to a changing cosmological environment, this society alters both the cosmology and itself to manipulate its subjective time.

4 Is the "Cosmological Constant" Really Constant?

The scenario presented at this meeting only infers that Λ be constant during the finite time period for which the rate of expansion of the relevant FL scale factor is sub-luminal. This interval is illustrated for a specific choice of scale factor[1] in Fig.1, but our remarks here apply to the whole class of models which

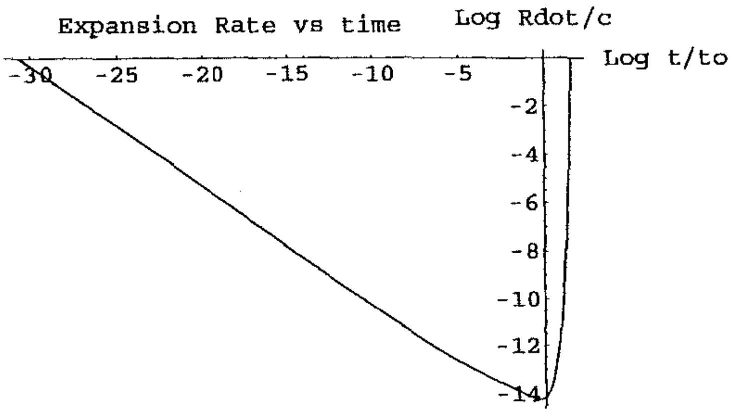


Figure 1: Finite time period of sub-luminal expansion rate

give a reasonable fit to current data and meet this requirement. In the model discussed in[1], the earlier time corresponds to a phase transition characterized by quantum de-coherence such as occurs when the ground state of a Bose-Einstein condensate starts to "evaporate" due to the confining density falling below the critical density at that temperature. Note that in such a model, density is falling more rapidly than temperature, so that it is consistent to assume that all of the energy is locked up in the (quantum-correlated) ground state of the condensate (lowest spacial frequency mode, using spacially periodic boundary conditions) *prior* to de-coherence. What happens at the future time t_E when this scale radius crosses the putative De Sitter horizon, or even whether there *is* a De Sitter horizon at that time, is the focus of the discussion in this section.

To focus our thinking, we will initially assume that the "dark energy" which is driving the exponential expansion at late times (but prior to the second time that $\dot{R} = c$) is "vacuum energy" due to zero point motions of sources[14]. The vacuum energy in the Casimir effect depends only on $\hbar c$ and the boundary conditions, independent of the *coupling constant* to the electromagnetic field. Nevertheless, it is possible to calculate the measured force effect using the charges and currents in the conducting bounding surfaces due to the fluctuations arising from the uncertainty principle. This gives the same result because, physically, the boundary conditions necessarily require that the boundaries themselves be made of *material* objects which act as electromagnetic conductors or dielectrics. Thus, from a *physical* point of view, we might expect "dark energy" effects to disappear once the scale radius of the gravitationally bound portion of the universe we are considering has crossed the (now putative) De Sitter horizon and is, *ipso facto*, out of luminal contact. For instance, Casimir plates separated by a

De Sitter horizon are *not* expected to exhibit the Casimir effect. Using the interpretation of dark energy espoused here, we question whether it will continue to manifest as a cosmological constant at late times. The De Sitter cosmology requires that the cosmological constant be in fact a constant. If this is not the case, then expectations and predictions of a horizon, with its associated properties, are premature.

At even sooner times, we expect the scale associated with the cosmological inhomogeneities responsible for galactic clustering and the fluctuations in the cosmic microwave background radiation to become comparable to or cross the De Sitter scale radius. As stated in the previous section, if the expansion is primarily due to dark energy during the intervening period, this is expected to occur in about 65 Gyr. Generally, we expect local geometry to be determined by local energy densities as described using Einstein's equation $G_{\mu\nu}(x) = 8\pi G_N T_{\mu\nu}(x) + \Lambda g_{\mu\nu}(x)$. For dynamically significant periods of time prior to this crossing, it is clear that the homogeneity and isotropy assumptions inherent in a Friedman- Lemaitre cosmology do not hold on the scale of galactic clustering. This means that the local geometry generated, $G_{\mu\nu}(x)$, is neither pure (cosmological) FRW-Lemaitre nor the Schwarzschild-like region of an isolated galactic cluster in Minkowski space (which would have no space-time expansion from dark energy). For instance, our local gravity is primarily the Schwarzschild space-time generated by Earth, with negligible influence from the overall cosmological acceleration due to the dark energy (or else we would be leaving the surface of the Earth!). This means that our local space-time is not undergoing the exponential expansion associated with a cosmological constant, despite our presence in an accelerating cosmology. We expect the evolution of our local scales to be determined by our local energy (and dark energy) densities, appropriately matching asymptotic boundary conditions. Likewise, on scales for which the cosmological matter inhomogeneities are important, the local densities are expected to have significant influence on the behavior of the geometry relative to cosmological dynamics. As the scale of relevance to galactic clustering crosses the De Sitter scale radius, one must take care in describing the De Sitter scale as a horizon. It is not unreasonable to suggest that the association of a given scale distance with supra-luminal rates of expansion could be only a temporary phase in the evolution of a cosmology that contains radiation, matter, and dark energy.

Once all cosmologically "co-moving" matter associated with the current exponential expansion has lost (luminal) causal contact with the finite, gravitationally bound system we believe will be left behind, we expect the regional situation to change. One conjecture is that from then on there is no reason to believe that the De Sitter scale radius, having no physical system to support it, should remain a "horizon" (meaning that regions beyond this "horizon" which were receding supra-luminally would again move sub-luminally). More precisely, the horizon (which is a global concept) never really existed, but there would only be a temporary loss of luminal contact as objects cross into a region which will recess at supra-luminal rates. After that time, objects which achieve (necessarily sub-luminal) escape velocity from the finite system (which now will hardly be describable as having "uniform density" on scales comparable to the De Sitter radius) would presumably continue to spread out into the *flat space* which is then the appropriate boundary condition for describing escape velocity. They could well be out of reach in terms of intact recovery as objects. However they would still remain in (eventual) luminal contact. This conjecture raises eschatologically important opportunities and issues which we now explore.

Fortunately, there would be no information horizon limiting the potential complexity of memory. This means that the entropy of the system need no longer be finite, and hence there would be no "big crunch" due to a Poincaré recurrence.

Unfortunately there would be no cosmological heat source to draw upon for energy, so that the finite energy crisis would be exacerbated. This implies that in the earliest stages of this scenario all efforts should be made to collect energy resources to be utilized during the cold far future. Hibernation would be a very bad idea until AFTER causal contact begins to be re-established with the remaining accessible parts of the universe.

Our position on the fringes of the bound galactic super cluster is advantageous for the transitional stage of the eschatology. Regions in the bound cluster nearest the outer orbits are well positioned with regards to communications, access to external information, and energy required for transportation. It looks as though we are destined to be on the dynamic frontier.

5 Apocalyptic Eschatology

Up to now the assumption has been implicitly made that the eschatological problem of primary concern has been whether "biology" in Dyson's sense can continue indefinitely (at least in subjective time) at the level of complexity our civilization has already achieved and with an ever growing memory. As has been seen, the technological challenges are formidable, but nothing in the laws of physics as now known precludes this possibility with anything like certainty. Implicit also in this analysis is the assumption that strict causality, or in theological terms predestination, does not hold. In other words, such societies are assumed to have, in some effective sense, *free will*, that is to make choices which have meaningful consequences relevant to the survival of themselves and/or their heirs. For better or worse, we so far only have knowledge of *one* such society that has passed the technological threshold needed to even envisage the far stretches of time involved. Hence we have no *scientific* way to estimate the probability of success. But we do have available reasonably well understood examples of societies at somewhat lower levels of technological development which have not had the foresight to avoid collapse, or even extinction[15, 16].

For a thoughtful analysis of the current situation we turn to the Astronomer Royal, Martin Rees[7], who will be remembered by older ANPA members for his invaluable assistance in getting ANPA going. He comes to the shocking conclusion that our global society has only a 50-50 chance of surviving the challenges we will meet in the current century. What strikes the authors of this paper as most depressing in the picture he paints is not just the individual problems — which are threatening enough — but the fact that he, in common with Dyson's earlier treatment of the far future, fails to discuss the fact that even if a technological means of meeting the problems is conceivable, there is no *global* decision making process in prospect, let alone available, that can bring together the planetary resources needed and direct them into the search for and implementation of the action needed on the time scale available. A promising start on the analysis of the problem of environmental collapse has been made by Jared Diamond, using a broad enough sample of examples to be meaningful. He finds, somewhat to his surprise, that there are no cases of collapse due solely to environmental change ([16], p. 11). Four of the five different sets of factors needed to make the analysis (environmental damage, climate change, hostile neighbors, friendly trading partners) can be more or less important or even absent, but understanding the society's responses to its environmental problems *invariably* is needed to understand the result. In other words, the basic problem falls squarely in the political arena, as our own analysis of the situation had concluded prior to encountering this recent development in his work. In short, the *political science* needed for the task has yet to be created. This is not the place to suggest how that might be achieved, other than the stale remark that without such a guide to global mobilization, the future is bleak.

To end on a more cheerful note, once our species succeeds in meeting the political problem, and avoiding the threatening apocalypse, the future for our intellectual and cultural heirs could well continue as long as the political will to do so persists.

References

[1] J.V.Lindesay and H.P.Noyes, Scientific Cosmology, *Proc. ANPA 26*, 2004 (in press); arXiv:astro-ph/0407535v1 26 July 2004.).

[2] J.V.Lindesay and H.P.Noyes, "Cosmic Microwave Background Fluctuation Amplitude from Dark Energy DeCoherence", astro-ph/0410217v2 25 Oct 2004.

[3] J.V.Lindesay, H.P.Noyes and E.D.Jones, "CMB Fluctuation Amplitude from Dark Energy Partitions", astro-ph/0412477v2 11 Feb 2005.

[4] F.J.Dyson, "Time without end: Physics and biology in an open universe", *Rev.Mod.Phys* **51**, 447-460 (1979).

[5] B.Russell, "A Free Man's Worship" (c.1905), in *Mysticism and Logic.*

[6] J.B.S.Haldane, *Daedalus, or Science and the Future* (1924).

[7] Martin Rees, *Our Final Hour: A Scientist's Warning: How terror, error, and environmental disaster threaten*

humankind's future in this century — on Earth and beyond, Basic Books, New York (2003).

[8] F.Hoyle, *The Black Cloud*, Harper, New York (1957).

[9] K.Capek, *R.U.R.*, translated by Paul Selver, Doubleday, New York (1923).

[10] F.S.Young, "Theoretical Reduction of Biological Complexity: A Method for the Modeling and Simulation of Complex Systems" (work in progress).

[11] L.Suskind and J.Lindesay, *AN INTRODUCTION TO BLACK HOLES, INFORMATION and the STRING THEORY REVOLUTION: The Holographic Universe*, World Scientific, Singapore, 2005, p. 121, Eq. 11.5.26.

[12] L.Dyson, J.Lindesay, and L.Suskind, "Is There Really a de Sitter/CFT Duality", JHEP 0208 (2002) 045; hep-th/0202163.

[13] The authors are indebted to Tom Etter for pointing this out.

[14] E.M. Lifshitz, Soviet Phys. JETP 2, 73 (1956). I.D. Dzyaloshinskii, E.M. Lifshitz, and I.P. Pitaevskii, Soviet Phys. Usp. 4, 153 (1961). I.D. Landau and E.M. Lifshitz, *Electrodynamics of Continuous Media*, pp368-376 (Pergamon, Oxford, 1960)

[15] J.Tainter, et. al, *The Collapse of Complex Societies*, Cambridge University Press (1990).

[16] J.Diamond, *COLLAPSE: How Societies Choose to Fail or Succeed*, Viking (2005).

25 years of the Combinatorial Hierarchy Anniversary Talks

Pierre Noyes

I thought this would be a good occasion to tell you something of my history and the history of ANPA that probably are not in your consciousness. A lot of it is not in the public record. I think I want to share this with you. The first thing that became clear to me in preparing this talk, particularly because of Ted's talk, is that I have a very different intellectual background than most of the members of ANPA, and that this should be taken into account.

When I was a freshman in high school my first interest in science was in evolution, geology and paleontology. But when I was a sophomore the Spanish Republic was fighting our battle against fascism and in my view the people of Spain were betrayed both by the neutrality of my own country and more actively by the capitalist powers in the West. So I identified with the cause of the Spanish Republic when I was 14. I wasn't quite old enough to go to Spain; I hope I would have if I had been a couple of years older.

Another thing that is relevant is my family background. My father, born in 1857 in Iowa from puritan New England stock, was one of the first international chemists in the USA. My eldest half-brother served in the Signal Corps at the front in the A.E.F., was discharged in France and got his PhD in Chemistry from the Sorbonne without ever having a bachelor's degree. He was not satisfied until he had won all the honors in chemistry my father had; he was one up on the old man when he got a post retirement job in 1965 and then he began to relax. He was quite a guy.

My second half brother, from a different mother than the first and different from my own mother, was forced to take a lot of chemistry in college. He was really interested in a literary career and he graduated at 18. My parents had supported my eldest brother in France, so my second half brother demanded that they support him in Paris for a couple of years. He wrote (but could not sell) three novels at one point; he worked on LaGuardia's first mayoral campaign in New York. Clearly he was the maverick of the family. I first got to know him when I was in high school. My politics at that time were not as far left as his, but knowing him contributed to my decision to go into Political Science when I applied for a scholarship at Harvard. I had had a very good history teacher, who arranged for me to read the *Communist Manifesto.* My grandfather on my mother's side was an abolitionist quaker; his mother ran a station on the Underground Railway in southern Indiana, and he served with Sherman's army on the March to the Sea through Georgia, as I learned while I was in high school. He was also one of the founders of the study of Political Science in the United States. So I have this very peculiar background to put it mildly and it is not a background that was leading me into science. My own brother I grew up with who was only 4-1/2 years older than me had graduated *summa cum laude* from Harvard in Chemistry. So I had a father and two brothers as Chemists, and one brother who was a maverick; he ended up being quite successful materially, but that is another story.

In April 1940 when I had my scholarship interview for Harvard, I made it quite clear to the interviewer that I was a Socialist and was going into Political Science. I though that I had blown it. I did get the scholarship. I still thought of the war as a rerun of World War I and a capitalist plot, and did not understand the military consequences of the Nazi-Soviet pact. A few months later came the fall of France and the Battle of Britain. I realized that for all their faults the French and British had been fighting our battle as well and that now Britain stood alone. So I became an interventionist. I had found out that my earlier political analysis was lousy. I was only 16 remember. So what was I doing going into political science as a career? I decided that I would probably make fewer mistakes if I went into Chemistry. So, when I entered Harvard in the fall of 1940 I registered as a major in Chemistry.

So how did I become a physicist? At the start of my sophomore year, on the advice of my brother who had gone to Harvard in Chemistry, I took the first half of the course for Physics majors rather than the watered down physics course most chemists took. This was beginning to convince me that physics rather than chemistry is the fundamental science, and more to my taste. The week of Pearl Harbor they told everybody in the professional physicists' course I was taking that was not registered in physics: "Switch your major to physics; you will be more useful in this war." I took that as an order and did so. I got in fact no upper division training in physics. I had five graduate courses in electronics and finished my training in three years (May 1943). I was not subject to the draft until the last six months and got a deferment by agreeing to work at a particular secret laboratory when I graduated.

So I ended up at the MIT rad lab in the antenna group designing radar antennas That was 1943 - I was 19. I very soon found out that the major work of that lab had already been been done. As we now all can know, it was the 3 centimeter microwave radar developed at MIT that in 1943 broke the German U-boat campaign in the North Atlantic and saved the western allies' war effort. Planes equipped with this radar were catching U-boats on the surface at night without the Germans realizing they were being detected. Actually bombers with similar radar sets were being shot down in Germany, but the German air force had not communicated this fact to the head of the Germany Navy for a year. As soon as the German navy knew about our 3 cm radar they took countermeasures and we never got another U-boat - fortunately most had been sunk by then.

With this going on I was doing sort of empirical antenna work at the lab. I didn't feel right not being in the military on active duty at my age. Most of my classmates were in the service in one way or another. In fact they had the heaviest losses of any class at Harvard during World War II. Most of those lost were flyers. So when I had a chance I went into the Navy. I was an aircraft radio tech for two years. I was scheduled for the invasion of Japan as an enlisted man. I had applied for a commission and was turned down [probably for political reasons, as I was able to surmise from an FBI interview in 1956]. My duty in the invasion of Japan. might have been in an unarmed plane over the coast of Kyushu at the start of action on November 1, 1945. But that's another story.

I was discharged in the summer of 1946 after six months Navy shore duty in the Philipines. I had learned because of the switch from Chemistry to Physics that Physics was what I wanted to do. I was admitted to both Chicago and Berkeley but I went to Berkeley because I had heard of Oppenheimer — in fact I had a copy of the Smythe Report on military use of atomic weapons in my sea bag when I went to the Pacific — but I knew what I wanted to do - I wanted to go into Physics - that was the fundamental science, I was convinced. I came to it very late compared to most physicists and with a different background. I did my graduate work in Berkeley. This becomes relevant to some of the intellectual history - first that summer before I started graduate work I read Russell's "History of Western Philosophy" - that was my introduction to philosophy per se. I had also at an earlier period read Bridgman's "Logic of Modern Physics" thanks not to any of my scientist brothers but to the one who had been forced to take science and had read this as a layman. He thought it was important and so I read it on his recommendation. So Bridgman and Russell formed my most significant intellectual background in the philosophy of science. It was also important that I learned my quantum mechanics from several different professors with differing points of view; all of them had been trained close enough to the start of the quantum revolution to have a lively understanding of the fundamental difficulties it raises. When it comes to Eddington, when I was a graduate student I had looked at looked at his *Fundamental Theory.* I got through chapter 1 or 2 in which he concludes that the functional behavior of the nuclear potential has to be "Gaussian", i.e. the harmonic oscillator potential. Since my own research at that time showed that such a potential could not fit low energy proton-proton scattering data, I lost interest. In contrast to both Ted and Clive my initial impression of Eddington was *not* favorable.

Skipping lightly over the intervening years, I note first that I had the good fortune to be the physicist Tom Kuhn trusted to read and comment on a draft of his *Structure of Scientific Revolutions.* I tried to fit my work with Geoff Chew into that framework, and already had the *paradigm shift* model in mind when I encountered the Combinatorial Hierarchy. As least as important in opening me up to radical ideas in science was my re-radicalization by the Vietnam war; representative of my public actions was an attempt to get Nixon tried for war crimes in the U.S.courts using the U.S.Constitution and the Nuremberg, Tokyo, and Manilla precedents.

My first contact with Ted was accidental. Pat Suppes, as some of you know, was head of the Philosophy Department at Stanford, also a courtesy professor in three other departments, and is still interested in foundations of quantum mechanics. He had Ted over to give a seminar at Stanford 1n 1972. He introduced Ted to me when I was at lunch at the Faculty Club. Ted talked about his *pie in the sky* (from my point of view) way of getting the 137 and I thought to myself *"What is the best way to get rid of this nonsense?"*. I didn't have a chance then but the next time Ted came back to Stanford — also on Pat Suppes initiative — I went to Ted's seminar planing to ask "How can a physically measured number number be generated by an algebriac algorithm constructed with no emprical input?". I went in to the seminar room with a certain mind set shall we say - this is important - because the first thing that impressed me about Ted when I heard him actually talking on his subject was that he sounded rational. He was talking sense. Then I knew he was talking sense when he said "The first quantization is the quantization of mass." Ted does not remember saying that but I remember it very distinctly. He actually said it and it got across to me that that was a very interesting thought - I had never heard it put that way before. And I agree with Ted - I do believe it myself. The quantization of mass is at least a very fundamental fact that we still, even to this day, don't really understand. Then he went through the combinatorial hierarchy construction. I was listening quite carefully because I knew he might be saying something worthwhile. How can a counting number have anything do to with an empirical number that isn't even an integer? And that was when I got the answer. Dyson had published a paper in 1952 which impressed me very much. In that paper he makes a calculation which can be interpreted as showing that 137 is indeed the *number* of charged particle pairs within a Compton wavelength which necessarily will produce another pair. This explains the fact that the perturbation theory of quantum electrodynamics is not uniformly convergent. So you have identified 137 as a counting number of actual things you can count. Then I realized at essentially the same time - okay, the big number (i.e $2^{127} \sim 10^{38}$) then is the number of things that you can count gravitationally inside a Compton wavelength that would form a black hole under Newtonian gravity. So then I was hooked and I could hardly restrain myself before I got up in the question period and explained all this to the people who had been listening to Ted.

From then on we kept in touch and I made it a point of seeing Ted whenever I came to England. In 1973 or 1974 I was in Europe to see my son, the wine-maker, when he was working in a chateau in the Gironde. I got a flight from there to England and got a train up to King's Lynn as instructed by Ted. I was told his group were having a meeting up there so I went to this famous windmill where the Epiphany Philosophers were meeting and got blisters wearing hip boots out into the tide water. Ted didn't realize what he was doing to me until he had to practically carry me back. They were planning the volume *Revisionary Philosophy and Science* - that never came off. Then there was also a session back at Milton Close in Cambridge before I took off and I think it was at that meeting that Parker-Rhodes first presented his idea of indistinguishables. That was my introduction to Fredrick as well. And I agreed to write a chapter on the idea (Dyson extended to gravitation) I had had when I first grasped how *I* could make sense out of what Ted was saying. I wrote a first draft of *Non-Locality in Particle Physics*, my contribution to the volume, on the plane back to California. The volume never appeared, but my chapter was put out as a SLAC publication in 1975.

So much for that piece of history. I kept being more and more involved with Ted and Clive and Fredrick. I went on seeing them every time I came to England during the era prior to the founding of ANPA — the event whose silver jubilee we are now celebrating. I was coming to Europe two or three times a year for scientific reasons and I would make space and come to Cambridge and talk to them. It turns out that Ted and John had gone to the meeting organized by that (self-confessed) German nuclear weaponeer von Weizsacher in 1976 and were going to go back again in 1978. Ted asked if I would go and present a paper with him. It turned out I had some time to work on it because I was going to spend a few days in Finland prior to von Weizsacher' 1978 Tutzing conference. My initial contact in Finland was because of my quantum mechanical three body problem work, but the head of theoretical physics (K.V.Laurikainen) was also a friend of Fredrick's and willing to arrange this quiet time for me to work on the combinatorial hierarchy.

I found out when I was talking to Ted about this paper we were going to give in 1978 - (a) that was when Parker-Rhodes came out with his calculation of the electron-proton mass ratio - of course we were going to present, but (b) I found out that nobody had even proved the existence of the hierarchy. This had fallen between stools. So I spent my time in Finland working out by hand a construction that would actually map the hierarchy and thus show that it existed. Later John Amson and Clive managed to get a really rigorous mathematical existence proof. So I realized things were not in as great shape about the foundations of this work as I had assumed. I had taken the rigor of the construction for granted because these people are professionals in the field - I'm not - I am essentially a Chemist upgraded to phenomenological data analysis in physics. Still am.

Now we are up to 1979 and the foundation of ANPA. I had given a course on Galileo in connection with the Social Thought and Institutions program at Stanford, and also a freshman seminar on Galileo comparing the way he was condemned with the Oppenheimer trial. I had ventured to do this because of a very good book by Desantillana called "The Crime of Galileo". The preface to that book points out the extreme similarity in terms of institutions of the structure of the Galileo trial and the Oppenheimer trial. It turned out that a friend of mine who was working in the freshman western culture course asked me to give a talk on Galileo and he had me and a few other friends who had helped him out in that course to dinner at the Faculty Club.

The husband of one of the English Department people who was in that course was an investment banker - not exactly the company I normally keep. We hit it off — and he let it drop during the evening that he was giving away about 78% of the money for people who were making charitable donations to non-profit organizations. At just that time I had received a desperate letter from Ted because of certain problems he was having. I thought: "Well, okay, here may be a way to raise money for Ted's research." Dugal Thomas, the investment banker, continued to be congenial when I had him to lunch later that week, just the two of us. I told him about Ted's problem, and asked "Why don't we do something about this?" He said: "Well, you can't just give the guy [i.e. Ted] money. That's not legal. But if you make an organization ...."

So that is when I wrote out the first ANPA budget and draft proposal for about $ 300.000 a year. I gave this proposal to Dugal Thomas as a plan for an organization to raise money which could then hire Ted as coordinator. Dugal was between jobs at that point and promised to work on it for a week. That was how it happened. Ted and Clive, Fredrick and John were interested. Dugal Thomas made a good faith effort. Unfortunately the wife of the prospect he was really counting on came down with cancer and wasn't thinking about charity at that point. So that scheme fell through. And Dugal only had a week to work on it, while he was moving from the West coast to the East coast. He kept trying to turn up something for us after he went East, but he wasn't really concentrating that much. I have never been able to raise funds myself.

Then I got the Humboldt award to spend a year in Germany. Because of various complications with my history at Stanford I didn't have to take half pay - I had a full year's pay coming. I could take the Humboldt stipend at that time along with my SLAC pay. So for the first time I was ahead of the game financially. At this point I said I could put up 5000 marks German money if Fredrick would put up at 1000 pounds and that would give us enough to get going. So it was on that basis that the first ANPA meeting was called. That is the straight story of how it all began. I thought people should be aware of it. What has kept it going? that's another question. But I thought you should have the beginning part of it.

[Question: "What did we do with that money?"]

We used it to pay Ted a very small stipend for a few months, and after that found that we could keep the organization going on annual dues and meeting fees. Paying back the original loan turned out to be simple. Fredrick was holding the mortgage on Ted's farm and Ted was paying 1000 pounds a year. Fredrick thought of it as forgetting the payment for one year and extending the payment period for one year. It wasn't really something he was worried about. Eventually I got my money back over the years from dues. It was all settled very naturally. We have never been very good at raising money until the Epiphany Philosophers came to our aid.

At some point in this history I realized that I didn't really understand what Ted and Clive meant by a *process* theory, and how formulating that idea in logical terms led to the combinatorial hierarchy. Also, I didn't understand how this answered the objection that — at least from a conventional point of view — one can't obtain an empirical number like the fine structure constant from a mathematical algorithm. I guessed that I just didn't have the background to understand what they were saying. So I thought these people [Ted, Clive, John, and Fredrick] know something that I don't know. I counted on there being a really axiomatic system in the mathematical sense which justified their reasoning. Well I think it's no secret there is no such thing. In fact, when I finally asked Clive why they had *not* provided a *rigorous* background for the hierarchy he admitted that there is no way to prove that this happens in a logical sense, in a mathematical, or in a philosophical sense. Further, he said tat the reason they had not worked harder to provide one was that at bottom they were afraid that the whole scheme would dissolve if they pushed it too hard. And Clive confirms - he did say that, too!

I think I have pulled out enough skeletons by now. I just want to thank all of you in ANPA past and present and future for giving me this marvelous opportunity to work with ideas I would never have encountered in any other way - to work seriously and honestly - with people who think very differently than I do and with whom I would never think alike, yet all of whom I can learn something from. I think I have had my reward at this meeting because certainly from my point of view what James and I reported to you is the best work I have done in my life and it never would have been possible without ANPA. So thank you all!

HISTORICAL TALKS 2

EARLY PHILOSOPHICAL IDEAS

Ted Bastin,

Maesllwyn, Tan y Groes, Cardigan SA43 2JF
edward@bastinsooo1.fsnet.co.uk

I am going to talk about some philosophical ideas that we used and that heavily influenced our mathematical work during the twenty years or so before the impact of Frederick Parker-Rhodes and John Amson. The 'we' is Clive Kilmister and I.

To start with, our problem was set by Eddington's attempt to find reasons for fundamental dimensionless constants and why they had the values they had. What Eddington had done was to start with the numbers of dimensions in well-known structures: in particular the space-time structure and the energy-momentum tensor of squared order. To this we said 'dimensions!'; can you do this with dimensions - abstracting numbers from their physical meanings and using them in an unrelated context; what does it mean?' This led to the further question of what the word 'dimension' itself means. Eddington's use of dimension number did not make sense

We thought that when we used the term '3-dimensional' we were drawing attention to a particular symmetry. We found that symmetry in the *quadratic* group of three elements together with the unit element – a, b, c, e., which is the outer product of two cyclic groups of order 2. The multiplication table is completely specified by a x b = c, and the other relationships that can be formed using those symbols in any way. And that is the point. The equation remains true with the elements in any order. The structure of the group does not permit us to define an order – the very symmetry characteristic that we picked on as the essential for dimensionality. Clive showed that this property of the a, b, c in the quadratic group was unique, and developed a formalism to set it in its place.

The quadratic group is closely related to the quaternion group whose physical applications are better known, but for our purposes we do not want the extra bits in the quaternion group. Indeed we really want to ignore the unit element in the quaternion group though it has to be there to provide transformations. It took a long time to make it clear that it didn't presage the representation of space-time. We used the term *theory language* to describe a structure of this kind. In the effort to explain how a theory –language fitted the world we got support from the "logic of facts" arguments of Wittgenstein –particularly as in the Tractatus. Thus; Tractatus 3.1431; "The essential nature of the propositional sign becomes very clear when we imagine it made up of spatial objects (such as tables, chairs, books) instead of written signs." Here there seemed to be support for our view that basic elements of theory were conveyed by the use we made of bits of matter. Another correspondence with Wittgenstein appears in the very first sentence of the Tractatus; "Die Welt ist alles, was der Fall ist." . "The standard Ogden translation; "The world is everything that is the case" is totally static and it is no accident that 'Fall' has dynamic overtones, and it seems that the all-pervading process view that emerged much later, fits. Dan Kurth says that 'Fall' appears in German as 'case' in legal process but with a sense of 'happenning'. So perhaps we should translate as '...everything that happens'.

The theory-language arose quite separately from the attaching of numbers to measurements, and that seemed to make it possible to think of it as a description which could be –as it were- 'lifted off' the normal metrical background, which no longer has to be necessary for its very definition, so that one could discuss its appropriateness to the use being proposed for it. Consider the Big Bang. It is universally acknowledged that physical concepts cease to have their classical meanings in the neighbourhood of the Big Bang, yet it seems impossible to avoid their use. One seemingly *had* to ask what happened before the Big Bang, or perhaps when it ceased to be reasonable to use that language. One got into the most terrible muddles. For all its strangeness the theory-language gives us a way out. The Tractatus comes in handy again. "Death is not an event of life; death is not lived through." Suppressing the impulse to say 'How do you know it is not?' one might apply the sentiment to time at the Big Bang: "The Big Bang is not an event in time. The Big Bang is not lived through."

It was very difficult to get this separation of the underlying structure from the numerical representation of measurements even contemplated, back then, for reasons I never understood, and I doubt if thing are very different now. I thought that the steady state theory should be regarded as providing a *regulative principle* (to use a phrase then current) simply because it afforded a comprehensible way to talk amid the slaughter of the usual concepts by the Big Bang. It was a different kind of alternative. This didn't go down at all well with Bondi (my supervisor) who had been reading Popper and thought that the most important thing about the steady state theory was that it should be falsifiable. Nevertheless the theory-language seemed to me to be a proper tool for discussing the limits to extrapolation, which would include those on the cosmological scale as well, of course.

Where were we to go from here so as to use the theory-language to represent different points in a space? The obvious thing was to set up a multiplicity of theory languages, one for each point or something like that. However this was obviously facile. We had somehow to make the theory-language itself more complex. At that time we were thinking about test-particles; that being a concept much in use then for specifying fields. The test-particle of mechanics is different from that for electromagnetism. The mechanical field accelerates the particle only in the line in which it is traveling whereas if we want to speak of an acceleration in a dimension at right-angles to the direction of the field then we enter electromagnetism, and there is a basic enrichment of our descriptive power and we have entered a new theory-language. That is where the idea of a hierarchy started, and it was integral to it that the simpler level should be incorporated into the next. We thought that the new should be constructed out of the operations existing at the first. That was Frederick Parker-Rhodes' starting point. I remember his saying "You have the wrong tool for what you are trying to do: binary algebra is the right one."

It was tempting at that time to think we had the essentials of electromagnetism and second order tensor structures to hand, but of course that was a long way ahead. All we should say was that without the level step we could not formulate the electromagnetic or the general relativistic concepts. To get there we had to wait for a long development starting from our regarding discrimination as the entrance of new things from a statistical background and making an impact on the development of structure. That would be a new chapter in our story.

An outlook that impinges on the theory-language idea is called 'structuralism'. Structuralism is the view that all we know is of the structure of things rather than of their intrinsic nature, and that the group is the clear way to express structure. This view was embraced by Eddington and led to a lively exchange between him and R.B.Braithwaite which was analysed by Steven French in the recent Eddington memorial conference (CRASSH). French argues the case put by Braithwaite and also M.H.A. Newman, saying that the group structure is only given once the relevant transformations have been specified (i.e. whether we are talking about rotations or permutations for example), but to do this is to supply *content* and so we no longer have pure structure. This did controvert Eddington's argument that the use of group theory allows us to abstract away the "pattern" or structure of relations between them. It might seem at first that this criticism would also invalidate our levels. In fact it affords support for there is one clear difference. The theory-language depends essentially on hierarchical construction, and hierarchical construction is the way toward the expression of things in their particularity..

Newman argues that all that could significantly be attributed to the world from the group structuralism would be its cardinality. We should be happy to go along with that given the wider scope of a hierarchy of groups with a principle of upwards continuation, for we can get from there, as far as we need, in the provision of the traditional physical variable. Indeed, looking forward to our present position, our policy for understanding the high energy particle is precisely to insist that there has to be a first stage when we have only the developing set of these cardinal numbers, and that their most important physical properties are entailed in that. Moreover, cutting ourselves adrift from the conventional background of conventionally defined dynamical concepts is something that needs to be done.

MY PERSISTENCE WITH ANPA.

CLIVE KILMISTER

Red Tiles Cottage

High Street

Barcombe

Lewes BN8 5DH

At this pleasant Silver Jubilee occasion I thought I would give some explanation of why I have stuck with ANPA so long. Looking round at so many familiar faces suggests this is otiose but I have learnt a lesson from Viv Pope and will give a potted autobiography. I shall choose as a starting point my interest in Eddington. After some of his popular books I toiled through THE MATHEMATICAL THEORY OF RELATIVITY as a schoolboy, early in the War. Around 1941 I found out about, and was able to borrow from the library, RELATIVITY THEORY OF PROTONS AND ELECTRONS. This was much more of a shock. Remember I was young: I believed everything that great men said. I couldn't make much of his arguments but I was fascinated by the way that, somehow, numbers came out. More recently I have concentrated on his guess that $\hbar c/e^2$ is about 137 but at the time I was much more impressed by his "calculation" of the proton-electron mass ratio as the ratio of the roots of the equation $x^2 - 136x + 10 = 0$. Of course this ratio is about 1848 and m_p/m_e is about 1836; this did not discourage me at the time.

Now to the personal details. I was a Ph.D student at QMC from 1948-50 and wanted to borrow a copy of FUNDAMENTAL THEORY which was in the college library. It had to be got back from someone who had had it out for a long time and that was how I came into contact with Ted Bastin. We came together because we were both trying to understand Eddington. Already by the early fifties we were convinced that what we slowly isolated as Eddington's programme of constructing a rationale for placing these numbers prior to measurement (and so casting their character as it were onto subsequent measurements)was the way to go. We came to reject his other (and inconsistent) view that one should try to bolt on the algebraic structure (Clifford algebras) to the rigmarole of differential equations in orthodox quantum mechanics, so as to be able to say "You see, it comes in here, so this must be $\hbar c/e^2$...".

In 1954 and later we were engaged in constructing (abelian) groups with significant and unusual symmetry properties which we hoped would justify Eddington's kind of arguments. We were beginning to run into the sand. We could see the need for a hierarchy of such groups but we couldn't see the precise relationship between the levels. Then one day Ted and Frederick Parker-Rhodes came to lunch with me at college to explain this odd construction which gave, unforcedly, the numbers 3, 10, 137 10^{39} and then stopped. The stop was the knockout argument for me. It addressed one of Eddington's problems; he had a scheme for getting numbers but it was a bit like the Sorcerer's Apprentice. You could keep on turning the handle and producing more numbers of no obvious physical significance.

I spent a very cold Lent term's leave in Cambridge with Ted in 1963 but for my part I was still experimenting with the algebraic details. For example, a lot of time was spent on investigating cycles, - sequences like u, Au, A^2u,...u. Much computing but no progress. It took me quite a long time to face up to the real problems: (a) What is the relation of the construction to physics? (b) Why is 137 different from the measured value (137·0360..). Our paper, often referred to as PITCH, was an attempt (with Pierre Noyes and John Amson) to answer (a) but at first neither question was answered satisfactorily because both needed the notion of process. In fact it was not until ANPA was formed in 1978 that progress began.

I think the story of how the first meeting came to be held has been often told. As I remember, plans for a small conference in Cambridge had fallen through and my wife, Peggy, suggested using our weekend cottage, as it was then but it is now where I live. Once the four of us (John Amson couldn't come at the last minute) had constructed the heady atmosphere that ANPA has had ever since, I was hooked. At first Ted was appointed as coordinator. The early meetings in Cambridge did not need to have a time-table; we just took it in turn to talk. As things grew more elaborate we owed a debt to Faruk Abdullah for undertaking the organisation (and finding Wesley House for us). I took over from him for a couple of years and we had great help from Mary Pope and now from Arleta and Keith.

So what next in this personal story? My original quest for numbers has been answered in part by my calculating (somewhat to my surprise) that the inverse fine structure constant is 137·036011.... But what of m_p/m_e? I have made no progress at all, perhaps because I can't get a hold on the notion of mass. Evidently there is still plenty to keep me busy. Long life to ANPA!

AVOIDING RELATIVITY'S PARADOXES

Cynthia K. Whitney
Editor, Galilean Electrodynamics
141 Rhinecliff Street, Arlington, MA 02476, USA

ABSTRACT

A hundred years of scientific and philosophical literature about Einstein's special (and general) theories of relativity leaves us today with a variety of puzzles and paradoxes to ponder. The paradoxes are all too often discussed in the literature without any universally satisfactory resolution. This paper develops a novel analysis of the paradoxes, one based on concepts borrowed from the literature of control engineering. It takes the typically engineering viewpoint that physical reality is one thing, while our conceptual model for it is quite another. The two do not necessarily match, and if they do not, then we will make wrong inferences from data. Wrong inferences can very well be inconsistent, and that can lead to apparent paradoxes.

The possibility for a wrong physical model to be embedded in Einstein's relativity theories is traced to the historical sequence of developments: in his time, Einstein knew Maxwell's electromagnetic theory, but not the full roster of empirical facts that were discovered later, and which motivated the development of modern quantum theory. The paper discusses a post-Maxwell model for light that harmonizes with those facts. This new model forms the basis for a new postulate about light, which can be used as an alternative to what Einstein's actually did use, his 'second postulate', specifying constant light speed c. The new alternative postulate reproduces the main kinematic features of special relativity, but without any paradoxes.

INTRODUCTION

In the literature there recur many questions about Einstein's Special Relativity Theory (SRT), and about his General Relativity Theory (GRT) too. A partial list of includes the following:

- Do moving clocks really run slow? Is there really a 'Twins Paradox'? It may seem so, but is it right to interpret that situation in terms of 'time dilation'?

- Do moving lengths really contract? Did Ehrenfest really find a Paradox with the rotating disc? What about the barn, the train and all the rest of the related paradoxes?
- The observed speed of an object is limited to light speed *c*, but is this really a physical limit? Does 'mass increase' occur to enforce the limit?
- In SRT, rotation is sometimes considered within scope, sometimes out of scope. Which way is it?
- GRT seems to be an aether theory. Can SRT's rejection of the aether really stand?
- How does GRT fit with Quantum Mechanics (QM)?

A variety of attitudes toward these questions is found in the physics and philosophical literature. Among Einstein's defenders, the most common may be characterized as 'denial': Many authors argue that there really are no paradoxes. Another common response may be characterized as 'evasion': the problems occur in the very macro realm of cosmology or the very micro realm of QM, and it is argued that human intuition applies only to human-scale phenomena. This line of thought extends naturally to high speed, intense gravity, and associated relativistic phenomena. It leads naturally to what may be characterized as 'tolerance': in modern mathematics it is a known fact that rational analysis is fundamentally limited in how far it can go, so a similar condition is attributed to QM. Why not also include relativistic domains? Tolerance leads naturally to 'embrace': the Copenhagen school of thought about QM holds that Nature is fundamentally capricious, and many think the same of relativity.

HOW CAN PARADOX HAPPEN?

The purpose of the present paper is to articulate an attitude different from any of the currently common ones. The viewpoint expressed traces to the author's experience in the world of control engineering. There, it is recognized that physical reality is one thing, while our conceptual model for it is quite another, and the two may not necessarily match. That is, the engineer is always dealing with a 'plant' that has some 'state' that is not known directly, but rather has to be inferred from 'data', using a mathematical model for how the plant state relates to the data. If the actual plant and the mathematical model for the plant do not match, then inferences from the data to the plant state will be wrong. Wrong inferences can very well be inconsistent, and that can very well lead to paradoxes.

The following Sections illustrate how this mechanism could apply in relativity theories. The development argues that, because it was conceived so early, Einstein's second postulate may constitute an incorrect 'plant model' for the physical process of light propagation. It suggests a different model incorporating the new facts learned in the time since the advent of relativity theories. It shows that if that if the new model would in fact be right, then use of the old model would produce incorrect inferences from all observations of moving objects. It shows how those incorrect inferences could produce the well-known oddities and paradoxes of SRT.

THIS IS *NOT* EINSTEIN'S POSTULATE

For all scientists, the fundamental working situation is that there is *something that we do not know*. To even begin thinking about this unknown, we must make one or more guesses about it. This is necessary in order to define the point of departure from which to pursue a logical analysis. That is why we make a postulate: it is basically one reasonable guess.

Here I want to use language a little more precisely than is commonly done. To me, a postulate is different from a hypothesis; a hypothesis can be tested experimentally, and maybe rejected, but a postulate is un-testable, at least within the technology that exists in its own time. And a postulate is different from an axiom, which is about a mental mathematical world, and can be tested only in the logical sense, for consistency or inconsistency with other axioms. A postulate can be, and in this case is, about a material, physical world – and probably can, eventually, be tested physically.

Einstein's Postulate Was...

Like all humans, Einstein could see objects and changes of position or state, but he could *not* see space or know time. His postulate about light propagation [1] filled those conceptual gaps for him. His particular prescription, *i.e.* constant light speed c for all inertial observers, made sense in the historical context within which Einstein worked. His historical age was dominated by the successes of Maxwell's equations, manifest in the developing electrical technology of the day. Maxwell's equations do imply a constant light speed in vacuum.

But now, a century beyond Einstein, we know many more things beyond Maxwell. Most particularly, we know all those mysterious experimental facts that motivated the development of QM. For example, light comes with characteristic frequencies, atoms do not collapse due to continuous radiation loss, and light sometimes shows inexplicable distant correlation. In short, Maxwell's equations do not cover all the apparently electromagnetic phenomena that exist in the natural world.

A change of status from 'postulate' to 'hypothesis', followed by testing against actual reality, is exactly what can happen now in regard to Einstein's postulate. There exist modern experiments suggesting that separated, but correlated-at-birth, photons can affect each other; that is, *within* light itself, there exists some sort of non-local, instantaneous action at a distance. So the idea of universal c for all communication is apparently factually incorrect. We still need a postulate today, and light propagation is indeed a suitable topic for it, but we need something different than what Einstein used.

Alternative Postulates Have Been...

Einstein's formulation of his 'second postulate' evolved over time [1], and that fact suggests that further evolution after Einstein is not out of order. I am not the first person to think so. For many years now, Prof. Domina Eberle Spencer, her family, and her students [2] have been offering, testing, and refining postulates concerning light. I have a much smaller but similar history myself [3].

It should be noted that Einstein put his emphasis on the observer. For him, that meant the receiver of light, because observation requires reception. His c was with respect to the receiver (constrained only to behave inertially). By contrast, the original Moon and Spencer paper [1a] established for the subsequent evolution of postulates about light speed a focus of attention on the source of the light. They did not mean that in the way that Ritz had meant it: c relative to the source at the moment of emission; they meant on-going attachment to the source, however it might move.

My Own Alternative Postulate:

I think the truth contains elements from both Einstein and from Moon-Spencer: part of the propagation scenario functions with respect to Moon-Spencer's source, and part of it functions with respect to Einstein's receiver. And I believe wholly in Moon-Spencer's on-going attachment. Maxwell's theory is about disembodied light, but real light is never disembodied; it is always attached to something – a material system that gives it, or receives it, or relays it. The characteristic speed is not a property of Maxwell's aether, it is a property of the material system attached to the light. My arbitrary observer does not own the propagation parameter; first the source owns it, but not forever; eventually, the receiver owns it.

My earliest postulate, from [3c], articulated the following post-Maxwellian observations about light:

1) Light always originates from some sort of a source. Let the process be called 'emission'. The word evokes the image of something expanding, probably spherically, from a point-like source. If we wanted to say something about the 'something', we would probably turn to Maxwell's equations, and select a wave solution that expands spherically from a point source. But to test any solution, we would have to make the scenario more complicated; we would have to add something, because:

2) Light always terminates at some sort of a receiver. Let the process be called 'absorption'. The word evokes the image of something converging, probably spherically, on a point-like receiver. The absorption process appears to be just like the emission process, only time-reversed.

3) The symmetry between emission and absorption processes suggests that propagation overall could consist of a perfectly balanced combination of the two. That is, the symmetry suggests that propagation from a point A to a different point B could consist of expansion from A until contact at B, followed by collapse to B.

These three observations invited the following preliminary analysis. Note that if the two points A and B are fixed at separation L, then the overall propagation process takes time L/c. If that process consists of two steps that are exactly the same except for time reversal, then each step takes half the time, $L/2c$. This deduction implies that the expansion step occurs as if at speed $2c$, and the collapse step also occurs as if at speed $2c$. (I said 'as if' because the individual steps, separated from the process as a whole, are unobservable. We could never know if some identifiable 'thing' actually moves at $2c$.)

My newer postulate refines the earlier one by taking account of the following additional observations:

4) There exists one particularly long-standing question: 'With respect to *what reference* is the characteristic light speed specified?' With Einstein's postulate, the question pertained to the c, and his answer was: the anthropocentric observer. With the new postulate, the question should be asked with regard to the $2c$, and answered in a non-anthropocentric way.

5) The emission process involves only the source, not the receiver, so the expansion speed $2c$ must be relative to the source. Conversely, the recep-

tion process involves only the receiver, not the source, so the collapse speed $2c$ must be relative to the receiver.

6) Modern experiments suggest that *within* light itself there exists some form of non-local action, to enforce distant correlations. The facts provide a precedent for proposing further that the switch of reference from source to receiver must be instantaneous.

To summarize then, my postulate about light is that:

> **Light propagation consists of one, or more, sequences of two steps each: expansion at speed $2c$ relative to a source, followed immediately by collapse at speed $2c$ relative to a receiver.**

CONSEQUENCES

This new postulate makes light neither the familiar extended **wave**, nor the familiar point **photon**. Instead, its key characteristic is its finite **envelope**. This envelope generally has some **spatial extent** to it, and implies some built-in **non-locality**; it is quite different from any kind of object that is well-treated in SRT. This light envelope also has **temporal evolution** to it. And at each moment of evolution, the envelope is attached to some bit of matter, either a source or a receiver, but always an **anchor particle**. That is, I think about light as if it were **an evolving, extended body, always attached to an anchor particle.**

The observed phenomenology clearly drives us in this direction; only the paucity of mathematical tools dissuades us. Neither Maxwell's equations nor any other linear wave equation will describe two-step propagation without unbelievable boundary conditions. Neither Newton's laws nor any other description of point-particle mechanics will model two-step propagation. We just have to build a mathematical model from scratch.

Now let us see where this refined postulate takes us. Consider first the simplest case, where there is no relative motion between light emitter and light absorber. The source is at point A at $x = 0$, and the receiver is at point B at $x = L$. To account for the total time known to be required for the whole propagation process, $T = L / c$, allow half the total time for each step: $T_1 = L / 2c$ for expansion, and $T_2 = L / 2c$ for collapse.

Now let the previously fixed points A and B become a possibly moving source and receiver, and let them have arbitrary Galilean velocities $\mathbf{V}_A$ and $\mathbf{V}_B$ relative to an arbitrary observer describing the scenario. The source-to-receiver separation L must change from a constant to a variable, dependent on time T. For the sake of definiteness, suppose $\mathbf{V}_A$, $\mathbf{V}_B$, and the line $\overline{AB}$ are all aligned. Suppose A and B are initially separated by distance L_0, and that the separation increases as the scenario progresses (*i.e.* $V_A < V_B$). Let the increase in separation be computed, first assuming the usual one-step light moving at universal speed c relative to the observer, and second assuming two-step light, each step at speed $2c$ relative to its particular reference (source or receiver).

With the one-step light model, the propagation scenario would have just one step.

1) In the special case that the observer is tied to the source, the source-to-receiver separation would increase during propagation outgoing to intercept the escaping receiver to

$$L_1 = L_0 / (1 - V_B / c) \quad (1a)$$

2) In the special case that the observer is tied to the receiver, V_A would be negative, and the separation would increase during the propagation to

$$L_1 = L_0 \times (1 - V_A / c) \quad (1b)$$

3) In general, the separation would increase to

$$L_1 = L_0 \times (1 - V_A / c) / (1 - V_B / c) \quad (1c)$$

This end result (1c) is dismaying because it is certainly observer dependent. Observer-dependent results constitute the fundamental **PROBLEM** with the Einsteinian postulate about light propagation: it leads to disagreements among different observers in any scenario involving propagation of light or any other electromagnetic effects. All reality becomes only relative. We have mysteries like 'length contraction', 'time dilation, 'relativity of simultaneity', 'the twins paradox', *etc.* Are those things 'real' or 'apparent'? Are they 'physical' or 'kinematic'? 'Ether effects' or 'observer effects'?

With two-step light, the propagation scenario involves two steps. They constitute:

1) Expansion from the source to intercept the retreating receiver. During this step, the operative 'light speed' is $2c + V_A$. The source-to-receiver distance increases because the receiver retreats, but it may also decrease because the source pursues. Overall, the changed separation is modeled by

$$L_1 = L_0 \times \frac{1 - V_A/(2c+V_A)}{1 - V_B/(2c+V_A)} \to L_0 \times \frac{1}{1 - V/2c} \qquad (2a)$$

where $V = V_B - V_A$. Note that the choice of observer coordinate frame has dropped out. Note too the formal similarity but actual difference between Eqs. (2a) and (1a).

2) Collapse to the contacted receiver. Now the operative 'light speed' is $2c + V_B$. The source-to-receiver separation changes to

$$L_2 = L_1 \times \frac{1 - V_A/(2c+V_B)}{1 - V_B/(2c+V_B)} \to L_1 \times \left[1 + V/2c\right] \qquad (2b)$$

Observe the similarity/difference between (2b) and (1b).

Altogether then,

$$L_2 = L_0 \times (1 + V/2c)/(1 - V/2c) \qquad (2c)$$

This end result (2c) for two-step light is surely not at all the same as the result (1c) for one-step light: it does not involve anything but *relative* speed V, and it admits only one possible analysis outcome, regardless of what observer chronicles the scenario. This is a highly desirable attribute. Producing observer-independent results is the fundamental **PROMISE** of the new two-step postulate about light propagation.

PARADOXES CAN BE ELIMINATED

If light propagation is two steps in physical reality, rather than one step as is assumed in present mathematical models, then any measurements based on using light signals to infer parameters of physical entities will generally be in error. For example, attributions of time coordinates for distant events can be in error because signal propagation delay can be incorrectly figured. In general, one has only two choices: either **1)** work out propagation-retarded time coordinates using the two-step light model, or else **2)** modify the one-step light model in a way that compensates for the error. It turns out [4] that the required modification is to replace the generic c by a modified c' defined by

$$c' = c / (1 + V / 4c) \quad (3)$$

where V is the component of relative motion driving the receiver away from the source, and so opposing the propagation.

So while the two-step $2c$ is independent of relative motion between source and receiver, the one-step c' is not. This is the key factor that makes the two-step process very different from the one-step model that SRT assumes. The discrepancy between c and c' changes the assignment of time coordinates to events located away from a spatial coordinate origin where 'the observer' resides. The time coordinates of distant events are generally not measured by a clock actually located at the event; they are **inferred** by assuming that the image of the event propagated at speed c to the observer at the origin. But if the 'event' is the passage of a source moving through a fixed spatial point, then correct inference of time requires c', not c. So for a **moving source**, time coordinates inferred assuming c are **biased**; *i.e.* **wrong**. This fact alone can account for a lot of strange results in SRT.

Example: The Twins Paradox

According to SRT, a moving clock looks slow. But in choosing *a priori* between two moving clocks, one can well wonder which of them is **really** moving. This question leads to the well-known Twins Paradox: twins are imagined to separate, one going on a journey while the other stays at home. SRT seems to predict that the traveling twin ages less than the sedentary one. All of the possible attitudes listed in the Introduction of this paper are to be found in the extensive literature about this paradox.

A simple resolution to the whole situation emerges from consideration of two-step light propagation. The image of a clock passing through x at T is observed at $t_o = T + x / c'$, and wrongly inferred to have passed through x at $t = t_o - x / c = T + Vx / 4c^2$. With $x = VT$, we have $t = (1 + V^2 / 4c^2)T > T$. So T progresses slower than t, but only because inferred time t is simply wrong.

Example: The Ehrenfest Paradox

According to SRT, a moving rod contracts. This is usually understood to be an observational effect, but if it is instead considered to be a real effect, then there is a paradox. According to Ehrenfest, the perimeter of a rotating disk is like a sequence of rods. So does the rotating disk shatter at the rim?

A resolution to the paradox emerges as follows. A rod has two ends, $x_i, i = 1,2$. From $t_i = T_i + Vx_i / 4c^2$, equal inferred t_i's has to mean unequal true T_i's. If we pair the ends for unequal T_i's, we get a wrong length L', Indeed if $x_1 = VT$ and $x_2 = VT + L$, then equal t_i's mean $T_2 = T_1 - VL' / 4c^2$, and $L' = L - V^2 L' / 4c^2$ so $L' = L / (1 + V^2 / 4c^2) < L$. So a moving rod *looks* short, but only because inferred times are wrong.

Example: Limited Particle Speed

It is asserted by SRT that Galilean particle speed is limited to light speed c. The Galilean speed V is defined as $\Delta x / \Delta T$. But without infinitely many clocks, ΔT is not directly observable. The next best thing is speed expressed in terms of coordinate time, $v = \Delta x / \Delta t$. Small observable v is not the same as large Galilean V; the two speeds are related through $v = V / (1 + V^2 / 4c^2)$. This function has a maximum value of $v = c$ located at $V = 2c$. So while Galilean V can be completely unlimited, observable v is limited to the maximum value of c. Thus particle speed *looks* limited to light speed c, but only because inferred times are wrong.

Note that this analysis has distinguished between observable v and Galilean V, a distinction not recognized in SRT. Needless to say, having only one name for two different things can lead to paradoxical situations in SRT.

Note too that the distinction between V and v is of great technological importance.

Example: Rotation Effects

In SRT, light speed is supposed to be c. But the well-known Sagnac effect belies that supposition. The Sagnac effect constitutes rotation sensing by means of the rotation-induced shift in the interference pattern produced by two light beams traveling in opposite directions around an optical loop. The implication is that light has speed c only in a non-rotating coordinate frame.

The troublesome Sagnac effect can be excluded from the domain of SRT by excluding the rotational motion that produces it. The argument goes: rotation involves acceleration, and hence motion that is not inertial. Also true although not usually mentioned, rotation invites rotating coordinate frames, and, for large enough radii, rotating coordinate frames imply relative motions at speeds in excess of c, precluded by the speed limitation accepted by SRT.

But rotation exclusion is not really feasible. The SRT Lorentz group of coordinate transformation naturally includes rotations, and non-collinear Lorentz transformations *generate* rotations. Indeed, those details are involved in explaining *another* experimental effect: the so-called anomalous magnetic moment of an electron in an atom.

All these things can become consistent only within the context of two-step light propagation. Although it goes beyond the scope of a single paper to discuss all possible rotation related effects, let us consider one of them here: the above mentioned anomalous magnetic moment of an electron in a hydrogen atom. The situation is that from the vantage point of the electron, the proton nucleus is seen to circulate around and create a magnetic field. The problem is that the energy associated with the electron coupling to this magnetic field is only about half what would be expected on the basis of the energy associated with the electron coupling to some exogenous magnetic field.

The resolution emerges as follows. Two-step light propagation implies coordinate transformation matrices that are not the same as Lorentz transformations. [3b] Instead of the $\gamma, \gamma\beta$ coefficients characteristic of Lorentz transformations, there are coefficients $1 + V^2 / 4c^2$, V / c, and $(V / 2c) \times (1 + V / 8c)$. This latter coefficient determines the magnetic field seen by the electron due to the nucleus in the atom 'circulating', and it is about half the $\gamma\beta$ coefficient SRT uses. Thus the magnetic field communicated by two-step light propagation in this situation is only about half what SRT predicts. So the electron doesn't really have an anomalous magnetic moment, but rather it sees an 'anomalous' *magnetic field.* But this field is 'anomalous' only as judged by SRT.

BUT HOW STRANGE IS TWO-STEP LIGHT?

Two-step light is decidedly novel, and some readers of earlier presentations have thought it is pretty strange. To help explore the conceptual strangeness, Fig. 1 illustrates a much more arbitrary scenario. The source A is doing a complicated dance, and the receiver B is doing an even more complicated dance. Universal time is denoted by T, and the two trajectories are $\mathbf{R}_A(T)$ and $\mathbf{R}_B(T)$, and the two Galilean velocities are $\mathbf{V}_A(T)$ and $\mathbf{V}_B(T)$. The growing and then shrinking arrows illustrate the light propagation from A to B. The light ray leaves A at time T_0, expands until it makes contact with B at time T_1, and then collapses to B, completing the propagation at time T_2.

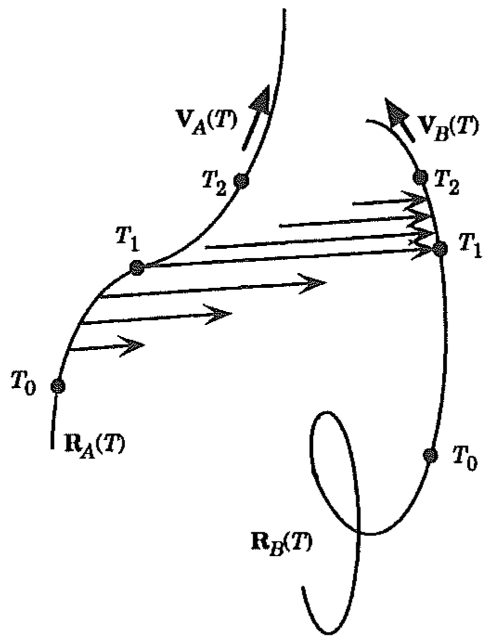


Figure 1. Two-step light propagation in a very arbitrary scenario involving two participants, A and B.

The two-step light postulate clearly has some similarities with Einstein's postulate [1] and with Moon-Spencer's postulate [2]. The more arbitrary scenario can now help expose the differences. Einstein began in 1905 [1a] with "Each ray of light moves in the coordinate system 'at rest' with the definite velocity V [our speed c] independent of whether this ray of light is emitted by a body at rest or a body in motion." After encountering difficulties with acceleration, he refined that in 1907 [1b] to "...clocks can be adjusted in such a way that the propagation velocity [speed] of every light ray in vacuum - measured by means of these clocks – becomes everywhere equal to a universal constant c, provided that the coordinate system is not accelerated." So if B is accelerating, Einstein's postulate does not say at what speed light approaches B. Obviously, the two-step light postulate has no such acceleration exclusion, and that fact may seem rather strange.

Moon and Spencer began developing their postulate in 1956 [2a], and they and their collaborators have developed their idea over nearly the last fifty years [2b]. They considered "all possible motions of A and B." That means they had no acceleration exclusion with respect to A or B – a major step in the right direction, in my opinion.

Moon and Spencer have questioned the usual idea that "once a pulse of radiation has left the source, this pulse has independent existence and is unaffected by subsequent motion of the source," and considered the possibility for "the radiation to remain in some way coupled to its source even after it has been emitted." This novel idea I have adopted wholeheartedly.

They have also refined the definition of the acceleration issue. Their requirement is just that the coordinate frame in which light speed is stated be not rotating. This is much less restrictive than prohibiting all acceleration. Einstein's broader exclusion means that in principle SRT is a theory that in the real world cannot ever be tested! So I also adopt the Moon-Spencer inclusion of acceleration.

But alas, two-step light is still rather strange in the Moon-Spencer context. They have only one step where I have two; they have no second step in which the radiation is already in some way coupled to its receiver even before it has been fully absorbed. And their rotation exclusion applies to a coordinate frame attached to the source, whereas mine applies to the coordinate frame in which the scenario-chronicling observer resides.

The two-step light postulate is not so much a postulate about the *speed of light*, but rather more about the *process of light propagation*. It may seem very strange to thinkers who have previously focused exclusively on the issue of light speed. Could people get used to this new postulate? Well, not so far. Indeed, the new postulate provokes a lot of passion. Critics are apt to plead "How *could* Nature possibly *work* like that?????"

That kind of question has a venerable history. It goes back at least to Newton, and recurs with Maxwell, Einstein, Planck, and others. These were all better men than I am! So I think I would be well-advised to study all of them and select one of them to follow.

Newton's critics did not like his purely mathematical model because of its apparent instantaneous action at a distance; they wanted some sort of a mechanical aether in between to do the action. Newton's attitude seems to have been to recognize the question as a request for a speculation, one not involved in the mathematical model itself, but necessary for his audience. So he didn't confront the question in his *Principia*; he dealt with it elsewhere: in private letters, and somewhat in *Opticks*. That segregation was, I think, a good approach.

The 'how' question came up again with Maxwell. He began with Faraday's experiments and his own predilection to explain them physically, in terms of an aether. His mathematical ideas were all entangled with his aether ideas. His followers expunged the resulting complexity, and probably expunged more than they should have done. So I think that wasn't such a good approach for dealing with the 'how' question.

The 'how' question came up yet again with Einstein; for example, in regard to the constant c. He pretty much ignored the question. He focused on operationalism, dismissed Newton's absolute space and time as metaphysical, and dismissed Maxwell's aether as superfluous. The trouble was that he later needed something like an aether for his general relativity theory (GRT). So the name got changed – 'four-dimensional space-time' or 'metric tensor' instead of 'aether'. This is embarrassing. So I don't like the dismissive approach to the 'how' question.

And again the 'how' question came up with the whole quantum revolution. Planck introduced his quantum hypothesis in the context of blackbody radiation, and it plainly described something true throughout the domain of atomic-scale phenomena. But to this day people wonder how things could work that way. The prevailing attitude, that of the so-called 'Copenhagen school', is that certain questions are simply disallowed. This attitude is one step harsher even than Einstein's: it is absolutely crazy-making. I do not like the Copenhagen attitude at all.

So the approach I choose is Newton's. You will have to write to me personally if you want to hear a private speculation about how light could work that way. All that I am going to do here is acknowledge the strangeness of it.

But at least there are some very familiar associations to take comfort from. First, it is fair to say that two-step light is rather Newtonian in flavor: where Newton had equally important action and reaction between any two interacting particles of matter, two-step light has equally important expansion and collapse between any two light-communicating particles of matter. (But of course, Newton's model for light implied increased light speed in refraction, which later seemed incorrect.) It is also fair to say that the 'twoness' of light is presaged in Maxwell's work: two constituative parameters

(ε and μ), two vector fields ($\mathbf{E}$ and $\mathbf{B}$), two parts (Re and Im), two energy peaks per wavelength, two polarization states (left and right circular), two time directions (forwards and backwards). (But of course Maxwell's wave equation fails on producing the emission/absorption sequence.) Quantum Mechanics is the most modern theory available to reference, and perhaps the most unequivocal in having no conflict with two-step light. It has 'two-ness' in the form of wave-particle duality, it has emission and absorption, and it has instantaneous, non-local action. Given all that, perhaps modern people can in fact get used to two-step light.

ACKNOWLEDGMENT

The author is grateful to Prof. Laszlo Tisza for a lifetime of conversations in which he raised ideas about expanding and contracting light, and to Prof. Domina Eberle Spencer for focusing her attention onto the task of formulating alternative postulates. The author is also grateful to Prof. Spencer for many stimulating personal discussions on the strangeness of two-step light.

The paper includes results first presented at the conference on Physical Interpretations of Relativity Theory for year 2000 in London, and the conference of the Natural Philosophy Alliance for year 2003 in Denver, USA. The author thanks the organizers and the audiences of these conferences for the opportunity to present and refine ideas.

REFERENCES

[1] A. Einstein, **a)** "On the Electrodynamics of Moving Bodies", Annalen der Physik **17**, 891-921 (1905), **b)** "On the Relativity Principle and the Conclusions Drawn from it", Jahrbuch der Radioaktivität und Elektronik 4, 411-462 (1907); both translated and reprinted in **The Collected Papers of Albert Einstein**, vol. 2 (Princeton University Press, 1989).

[2] D.E. Spencer, **a)** P. Moon &_ "On the Establishment of a Universal Time", Philosophy of Science 23, 216-229 (1956); **b)** _ & U. Shama, "The Development of the Universal Time Postulate on the Velocity of Light", J. New Energy 4, 196-209 (1999).

[3] C.K. Whitney, "Finding Absolution for Special Relativity Theory; **a)**, **b)**, & **c)**– Parts I, II, & III", Galilean Electrodynamics 7, 23-29 (1996), 7, 63-69 (1996), 8, 9-15 (1997).

[4] C.K. Whitney, Galilean Electrodynamics **16**, Special Issues no. 3, in press (2005).

Non-Commutative Worlds - A Summary

Louis H. Kauffman
Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago
851 South Morgan Street
Chicago, IL, 60607-7045

1 Introduction to Non-Commutative Worlds

Aspects of gauge theory, Hamiltonian mechanics and quantum mechanics arise naturally in the mathematics of a non-commutative framework for calculus and differential geometry. This paper consists in two sections. This first section sketches our results in this domain in general. The second section gives a derivation of a generalization of the Feynman-Dyson derivation of electromagnetism using our non-commutative context and using diagrammatic techniques. The first section is based on the paper [15]. The second section is a new approach to issues in [15].

Constructions are performed in a Lie algebra $\mathcal{A}$. One may take $\mathcal{A}$ to be a specific matrix Lie algebra, or abstract Lie algebra. If $\mathcal{A}$ is taken to be an abstract Lie algebra, then it is convenient to use the universal enveloping algebra so that the Lie product can be expressed as a commutator. In making general constructions of operators satisfying certain relations, it is understood that one can always begin with a free algebra and make a quotient algebra where the relations are satisfied.

On $\mathcal{A}$, a variant of calculus is built by defining derivations as commutators (or more generally as Lie products). For a fixed N in $\mathcal{A}$ one defines

$$\nabla_N : \mathcal{A} \longrightarrow \mathcal{A}$$

by the formula

$$\nabla_N F = [F, N] = FN - NF.$$

∇_N is a derivation satisfying the Leibniz rule.

$$\nabla_N(FG) = \nabla_N(F)G + F\nabla_N(G).$$

There are many motivations for replacing derivatives by commutators. If $f(x)$ denotes (say) a function of a real variable x, and $\tilde{f}(x) = f(x+h)$ for a fixed increment h, define the *discrete derivative* Df by the formula $Df = (\tilde{f} - f)/h$, and find that the Leibniz rule is not satisfied. One has the basic formula for the discrete derivative of a product:

$$D(fg) = D(f)g + \tilde{f}D(g).$$

Correct this deviation from the Leibniz rule by introducing a new non-commutative operator J with the property that

$$fJ = J\tilde{f}.$$

Define a new discrete derivative in an extended non-commutative algebra by the formula

$$\nabla(f) = JD(f).$$

It follows at once that

$$\nabla(fg) = JD(f)g + J\tilde{f}D(g) = JD(f)g + fJD(g) = \nabla(f)g + f\nabla(g).$$

Note that

$$\nabla(f) = (J\tilde{f} - Jf)/h = (fJ - Jf)/h = [f, J/h].$$

In the extended algebra, discrete derivatives are represented by commutators, and satisfy the Leibniz rule. One can regard discrete calculus as a subset of non-commutative calculus based on commutators.

In $\mathcal{A}$ there are as many derivations as there are elements of the algebra, and these derivations behave quite wildly with respect to one another. If one takes the concept of *curvature* as the non-commutation of derivations, then $\mathcal{A}$ is a highly curved world indeed. Within $\mathcal{A}$ one can build a tame world of derivations that mimics the behaviour of flat coordinates in Euclidean space. The description of the structure of $\mathcal{A}$ with respect to these flat coordinates contains many of the equations and patterns of mathematical physics.

The flat coordinates X_i satisfy the equations below with the P_j chosen to represent differentiation with respect to X_j.:

$$[X_i, X_j] = 0$$
$$[P_i, P_j] = 0$$
$$[X_i, P_j] = \delta_{ij}.$$

Derivatives are represented by commutators.

$$\partial_i F = \partial F/\partial X_i = [F, P_i],$$
$$\hat{\partial}_i F = \partial F/\partial P_i = [X_i, F].$$

Temporal derivative is represented by commutation with a special (Hamiltonian) element H of the algebra:

$$dF/dt = [F, H].$$

(For quantum mechanics, take $i\hbar dA/dt = [A, H]$.) These non-commutative coordinates are the simplest flat set of coordinates for description of temporal phenomena in a non-commutative world. Note:

Hamilton's Equations.

$$dP_i/dt = [P_i, H] = -[H, P_i] = -\partial H/\partial X_i$$
$$dX_i/dt = [X_i, H] = \partial H/\partial P_i.$$

These are exactly Hamilton's equations of motion. The pattern of Hamilton's equations is built into the system.

Discrete Measurement. Consider a time series $\{X, X', X'', \cdots\}$ with commuting scalar values. Let

$$\dot{X} = \nabla X = JDX = J(X' - X)/\tau$$

where τ is an elementary time step (If X denotes a times series value at time t, then X' denotes the value of the series at time $t + \tau$.). The shift operator J is defined by the equation $XJ = JX'$ where this refers to any point in the time series so that $X^{(n)}J = JX^{(n+1)}$ for any non-negative integer n. Moving J across a variable from left to right, corresponds to one tick of the clock. This discrete, non-commutative time derivative satisfies the Leibniz rule.

This derivative ∇ also fits a significant pattern of discrete observation. Consider the act of observing X at a given time and the act of observing (or obtaining) DX at a given time. Since X and X' are ingredients in computing $(X' - X)/\tau$, the numerical value associated with DX, it is necessary to let the clock tick once, Thus, if one first observe X and then obtains DX, the result is different (for the X measurement) if one first obtains DX, and then observes X. In the second case, one finds the value X' instead of the value X, due to the tick of the clock.

1. Let $\dot{X}X$ denote the sequence: observe X, then obtain $\dot{X}$.
2. Let $X\dot{X}$ denote the sequence: obtain $\dot{X}$, then observe X.

The commutator $[X, \dot{X}]$ expresses the difference between these two orders of discrete measurement. In the simplest case, where the elements of the time series are commuting scalars, one has

$$[X, \dot{X}] = X\dot{X} - \dot{X}X = J(X' - X)^2/\tau.$$

Thus one can interpret the equation

$$[X, \dot{X}] = Jk$$

(k a constant scalar) as

$$(X' - X)^2/\tau = k.$$

This means that the process is a walk with spatial step

$$\Delta = \pm\sqrt{k\tau}$$

where k is a constant. In other words, one has the equation

$$k = \Delta^2/\tau.$$

This is the diffusion constant for a Brownian walk. A walk with spatial step size Δ and time step τ will satisfy the commutator equation above exactly when the square of the spatial step divided by the time step remains constant. This shows that the diffusion constant of a Brownian process is a structural property of that process, independent of considerations of probability and continuum limits.

Heisenberg/Schroedinger Equation. Here is how the Heisenberg form of Schroedinger's equation fits in this context. Let $J = (1 + H\Delta t/i\hbar)$. Then $\nabla\psi = [\psi, J/\Delta t]$, and we calculate

$$\nabla\psi = \psi[(1 + H\Delta t/i\hbar)/\Delta t] - [(1 + H\Delta t/i\hbar)/\Delta t]\psi = [\psi, H]/i\hbar.$$

This is exactly the form of the Heisenberg equation.

Dynamics and Gauge Theory. One can take the general dynamical equation in the form

$$dX_i/dt = \mathcal{G}_i$$

where $\{\mathcal{G}_1, \cdots, \mathcal{G}_d\}$ is a collection of elements of $\mathcal{A}$. Write $\mathcal{G}_i$ relative to the flat coordinates via $\mathcal{G}_i = P_i - A_i$. This is a definition of A_i and $\partial F/\partial X_i = [F, P_i]$. The formalism of gauge theory appears naturally. In particular, if

$$\nabla_i(F) = [F, \mathcal{G}_i],$$

then one has the curvature

$$[\nabla_i, \nabla_j]F = [R_{ij}, F]$$

and

$$R_{ij} = \partial_i A_j - \partial_j A_i + [A_i, A_j].$$

This is the well-known formula for the curvature of a gauge connection. Aspects of geometry arise naturally in this context, including the Levi-Civita connection (which is seen as a consequence of the Jacobi identity in an appropriate non-commutative world).

One can consider the consequences of the commutator $[X_i, \dot{X}_j] = g_{ij}$, deriving that

$$\ddot{X}_r = G_r + F_{rs}\dot{X}^s + \Gamma_{rst}\dot{X}^s\dot{X}^t,$$

where G_r is the analogue of a scalar field, F_{rs} is the analogue of a gauge field and Γ_{rst} is the Levi-Civita connection associated with g_{ij}. This decompositon of the acceleration is uniquely determined by the given framework.

One can use this context to revisit the Feynman-Dyson derivation of electromagnetism from commutator equations, showing that most of the derivation is independent of any choice of commutators, but highly dependent upon the choice of definitions of the derivatives involved. Without any assumptions about initial commutator equations, but taking the right (in some sense simplest) definitions of the derivatives one obtains a significant generalization of the result of Feynman-Dyson.

Electromagnetic Theorem. (See Section 2.) With the appropriate [see below] definitions of the operators, and taking

$$\nabla^2 = \partial_1^2 + \partial_2^2 + \partial_3^2, \quad B = \dot{X} \times \dot{X} \text{ and } E = \partial_t \dot{X}, \quad \text{one has}$$

1. $\ddot{X} = E + \dot{X} \times B$
2. $\nabla \bullet B = 0$
3. $\partial_t B + \nabla \times E = B \times B$
4. $\partial_t E - \nabla \times B = (\partial_t^2 - \nabla^2)\dot{X}$

The key to the proof of this Theorem is the definition of the time derivative. This definition is as follows

$$\partial_t F = \dot{F} - \Sigma_i \dot{X}_i \partial_i(F) = \dot{F} - \Sigma_i \dot{X}_i [F, \dot{X}_i]$$

for all elements or vectors of elements F. The definition creates a distinction between space and time in the non-commutative world. A calculation (done diagrammatically in Figure 3) reveals that

$$\ddot{X} = \partial_t \dot{X} + \dot{X} \times (\dot{X} \times \dot{X}).$$

This suggests taking $E = \partial_t \dot{X}$ as the electric field, and $B = \dot{X} \times \dot{X}$ as the magnetic field so that the Lorentz force law

$$\ddot{X} = E + \dot{X} \times B$$

is satisfied.

This result is applied to produce many discrete models of the Theorem. These models show that, just as the commutator $[X, \dot{X}] = Jk$ describes Brownian motion in one dimension, a generalization of electromagnetism describes the interaction of triples of time series in three dimensions.

Remark. While there is a large literature on non-commutative geometry, emanating from the idea of replacing a space by its ring of functions, work discussed herein is not written in that tradition. Non-commutative geometry does occur here, in the sense of geometry occuring in the context of non-commutative algebra. Derivations are represented by commutators. There are relationships between the present work and the traditional non-commutative geometry, but that is a subject for further exploration. In no way is this paper intended to be an introduction to that subject. The present summary is based on [6, 7, 8, 9, 10, 11, 12, 13, 14, 15] and the references cited therein.

The following references in relation to non-commutative calculus are useful in comparing with the present approach [2, 3, 4, 17]. Much of the present work is the fruit of a long series of discussions with Pierre Noyes. paper [16] also works with minimal coupling for the Feynman-Dyson derivation. The first remark about the minimal coupling occurs in the original paper by Dyson [1], in the context of Poisson brackets. The paper [5] is worth reading as a companion to Dyson. It is the purpose of this summary to indicate how non-commutative calculus can be used in foundations.

2 Generalized Feynman Dyson Derivation

In this section we assume that specific time-varying coordinate elements X_1, X_2, X_3 of the algebra $\mathcal{A}$ are given. *We do not assume any commutation relations about* X_1, X_2, X_3.

In this section we no longer avail ourselves of the commutation relations that are in back of the original Feynman-Dyson derivation. We do take the definitions of the derivations from that previous context. Surprisingly, the result is very similar to the one of Feynman and Dyson, as we shall see.

Here $A \times B$ is the non-commutative vector cross product:

$$(A \times B)_k = \Sigma_{i,j=1}^{3} \epsilon_{ijk} A_i B_j.$$

(We will drop this summation sign for vector cross products from now on.) Then, with $B = \dot{X} \times \dot{X}$, we have

$$B_k = \epsilon_{ijk} \dot{X}_i \dot{X}_j = (1/2)\epsilon_{ijk}[\dot{X}_i, \dot{X}_j].$$

The epsilon tensor ϵ_{ijk} is defined for the indices $\{i, j, k\}$ ranging from 1 to 3, and is equal to 0 if there is a repeated index and is ortherwise equal to the sign of the permutation of 123 given by ijk. We represent dot products and cross products in diagrammatic tensor notation as indicated in Figure 1 and Figure 2. In Figure 1 we indicate the epsilon tensor by a trivalent vertex. The indices of the tensor correspond to labels for the three edges that impinge on the vertex. The diagram is drawn in the plane, and is well-defined since the epsilon tensor is invariant under cyclic permutation of its indices.

We will define the fields E and B by the equations

$$B = \dot{X} \times \dot{X} \text{ and } E = \partial_t \dot{X}.$$

We will see that E and B obey a generalization of the Maxwell Equations, and that this generalization describes specific discrete models. The reader should note that this means that a significant part of the *form* of electromagnetism is the consequence of choosing three coordinates of space, and the definitions of spatial and temporal derivatives with respect to them. The background process that is being described is otherwise aribitrary, and yet appears to obey physical laws once these choices are made.

In this section we will use diagrammatic matrix methods to carry out the mathematics. In general, in a diagram for matrix or tensor composition, we sum over all indices labeling any edge in the diagram that has no free ends. Thus matrix multiplication corresponds to the connecting of edges between diagrams, and to the summation over common indices. With this interpretation of compositions, view the first identity in Figure 1. This is a fundmental identity about the epsilon, and corresponds to the following lemma.

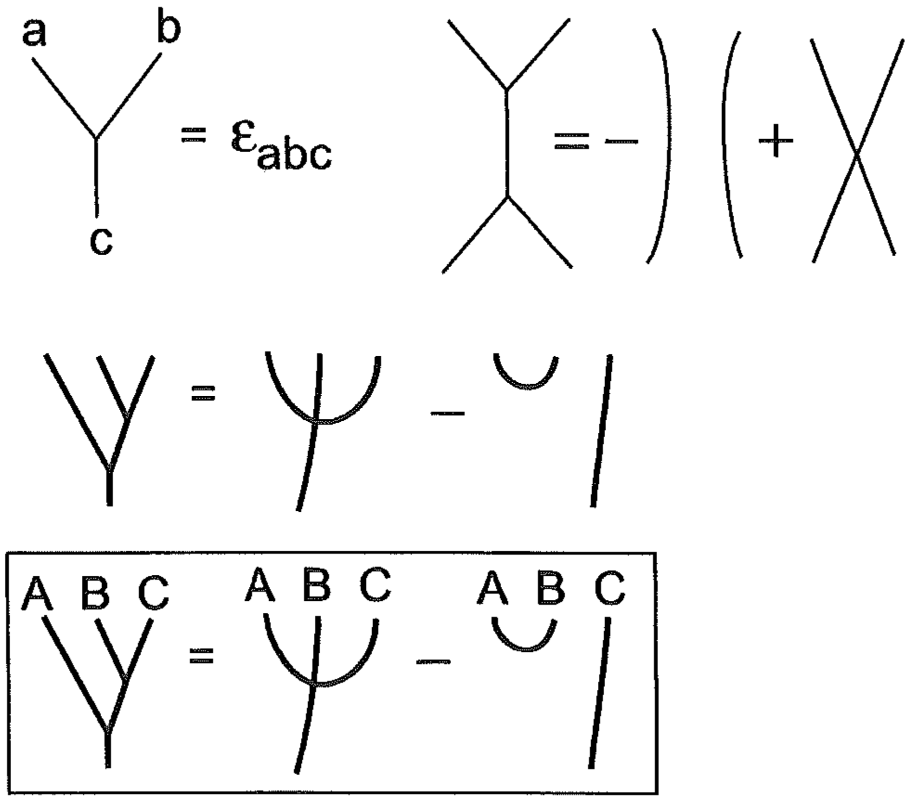
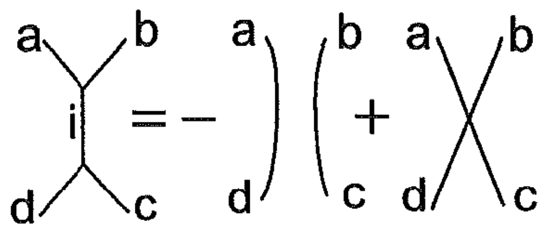


Figure 1 - Epsilon Identity

Lemma. (View Figure 1) Let ϵ_{ijk} be the epsilon tensor taking values 0, 1 and -1 as follows: When ijk is a permuation of 123, then ϵ_{ijk} is equal to the sign of the permutation. When ijk contains a repetition from $\{1, 2, 3\}$, then the value of epsilon is zero. Then ϵ satisfies the following identity in terms of the Kronecker delta.



$$\Sigma_i \, \epsilon_{abi}\epsilon_{cdi} = -\delta_{ad}\delta_{bc} + \delta_{ac}\delta_{bd}.$$

The proof of this identity is left to the reader. The identity itself will be referred to as the *epsilon identity.* The epsilon identity is a key structure in the work of this section, and indeed in all formulas involving the vector cross product.

The reader should compare the formula in this Lemma with the diagrams in Figure 1. The first two diagram are two versions of the Lemma. In the third diagram the labels are capitalized and refer to vectors A, B and C. We then see that the epsilon identity becomes the formula

$$A \times (B \times C) = (A \bullet C)B - (A \bullet B)C$$

for vectors in three-dimensional space (with commuting coordinates, and a generalization of this identity to our non-commutative context. Refer to Figure 2 for the diagrammatic definitions of dot and cross product of vectors. We take these definitions (with implicit order of multiplication) in the non-commutative context.

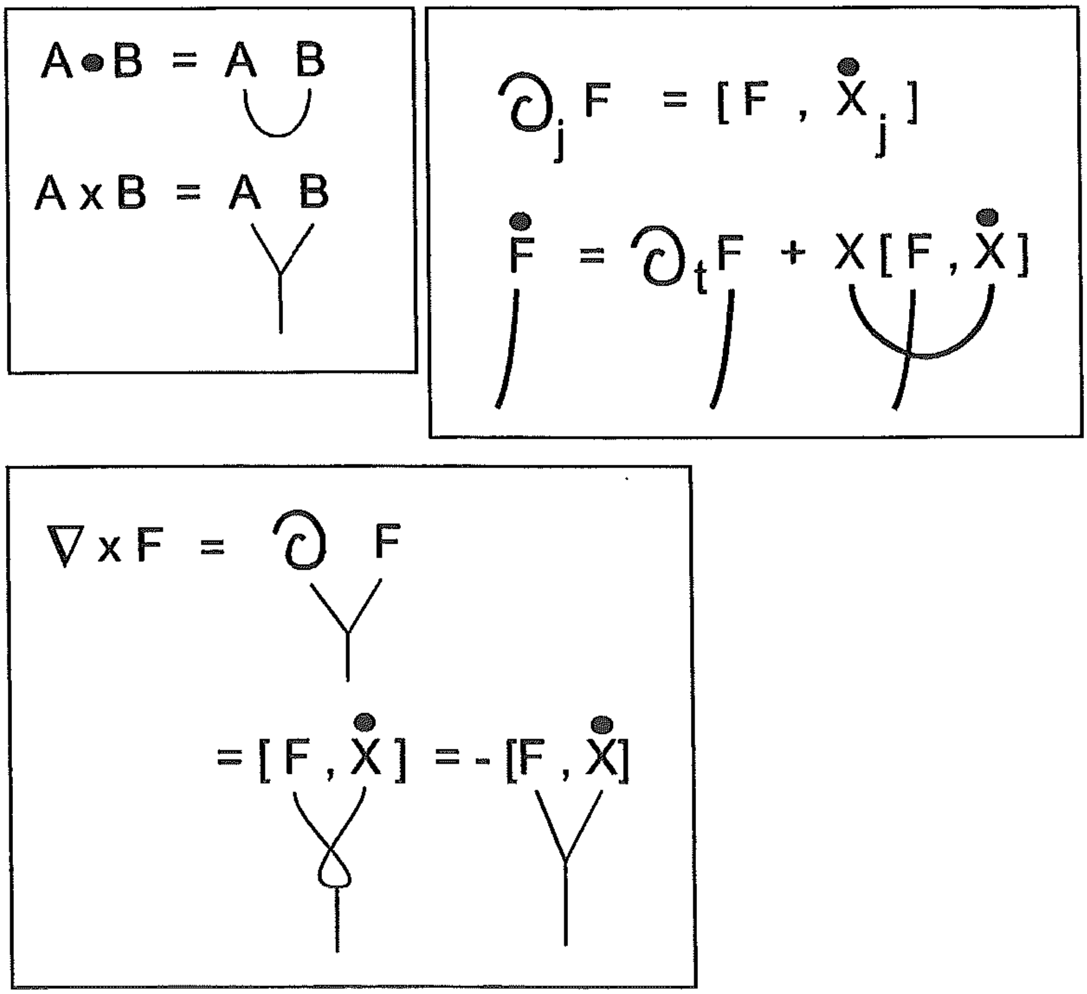


Figure 2 - Defining Derivatives

Remarks on the Derivatives.

1. Since we do not assume that $[X_i, \dot{X}_j] = \delta_{ij}$, nor do we assume $[X_i, X_j] = 0$, it will not follow that E and B commute with the X_i.

2. We define
$$\partial_i(F) = [F, \dot{X}_i],$$
and the reader should note that, these spatial derivations are no longer flat in the sense of section 1 (nor were they in the original Feynman-Dyson derivation). See Figure 2 for the diagrammatic version of this definition.

3. We define $\partial_t = \partial/\partial t$ by the equation
$$\partial_t F = \dot{F} - \Sigma_i \dot{X}_i \partial_i(F) = \dot{F} - \Sigma_i \dot{X}_i [F, \dot{X}_i]$$
for all elements or vectors of elements F. We take this equation as the global definition of the temporal partial derivative, even for elements that are not commuting with the X_i. This notion of temporal partial derivative ∂_t is a least relation that we can write to describe the temporal relationship of an arbitrary non-commutative vector F and the non-commutative coordinate vector X. See Figure 2 for the diagrammatic version of this definition.

4. In defining
$$\partial_t F = \dot{F} - \Sigma_i \dot{X}_i \partial_i(F),$$
we are using the definition itself to obtain a notion of the variation of F with respect to time. The definition itself creates a distinction between space and time in the non-commutative world.

5. The reader will have no difficulty verifying the following formula:
$$\partial_t(FG) = \partial_t(F)G + F\partial_t(G) + \Sigma_i \partial_i(F)\partial_i(G).$$
This formula shows that ∂_t does not satisfy the Leibniz rule in our non-commutative context. This is true for the original Feynman-Dyson context, and for our generalization of it. All derivations in this theory that are defined directly as commutators do satisfy the Leibniz rule. Thus ∂_t is an operator in our theory that does not have a representation as a commutator.

6. We define divergence and curl by the equations
$$\nabla \bullet B = \Sigma_{i=1}^{3} \partial_i(B_i)$$

and

$$(\nabla \times E)_k = \epsilon_{ijk}\partial_i(E_j).$$

See Figure 2 and Figure 4 for the diagrammatic versions of curl and divergence.

Now view Figure 3. We see from this Figure that it follows directly from the definition of the time derivatives (as discussed above) that

$$\ddot{X} = \partial_t \dot{X} + \dot{X} \times (\dot{X} \times \dot{X}).$$

This is our motivation for defining

$$E = \partial_t \dot{X}$$

and

$$B = \dot{X} \times \dot{X}.$$

With these definition in place we have

$$\ddot{X} = E + \dot{X} \times B,$$

giving an analog of the Lorentz force law for this theory.

Just for the record, look at the following algebraic calculation for this derivative:

$$\dot{F} = \partial_t F + \Sigma_i \dot{X}_i [F, \dot{X}_i]$$
$$= \partial_t F + \Sigma_i (\dot{X}_i F \dot{X}_i - \dot{X}_i \dot{X}_i F)$$
$$= \partial_t F + \Sigma_i (\dot{X}_i F \dot{X}_i - \dot{X}_i F_i \dot{X}) + \dot{X}_i F_i \dot{X} - \dot{X}_i \dot{X}_i F$$

Hence

$$\dot{F} = \partial_t F + \dot{X} \times F + (\dot{X} \bullet F)\dot{X} - (\dot{X} \bullet \dot{X})F$$

(using the epsilon identity). Thus we have

$$\ddot{X} = \partial_t \dot{X} + \dot{X} \times (\dot{X} \times \dot{X}) + (\dot{X} \bullet \dot{X})\dot{X} - (\dot{X} \bullet \dot{X})\dot{X},$$

whence

$$\ddot{X} = \partial_t \dot{X} + \dot{X} \times (\dot{X} \times \dot{X}).$$

In Figure 4, we give the derivation that B has zero divergence.

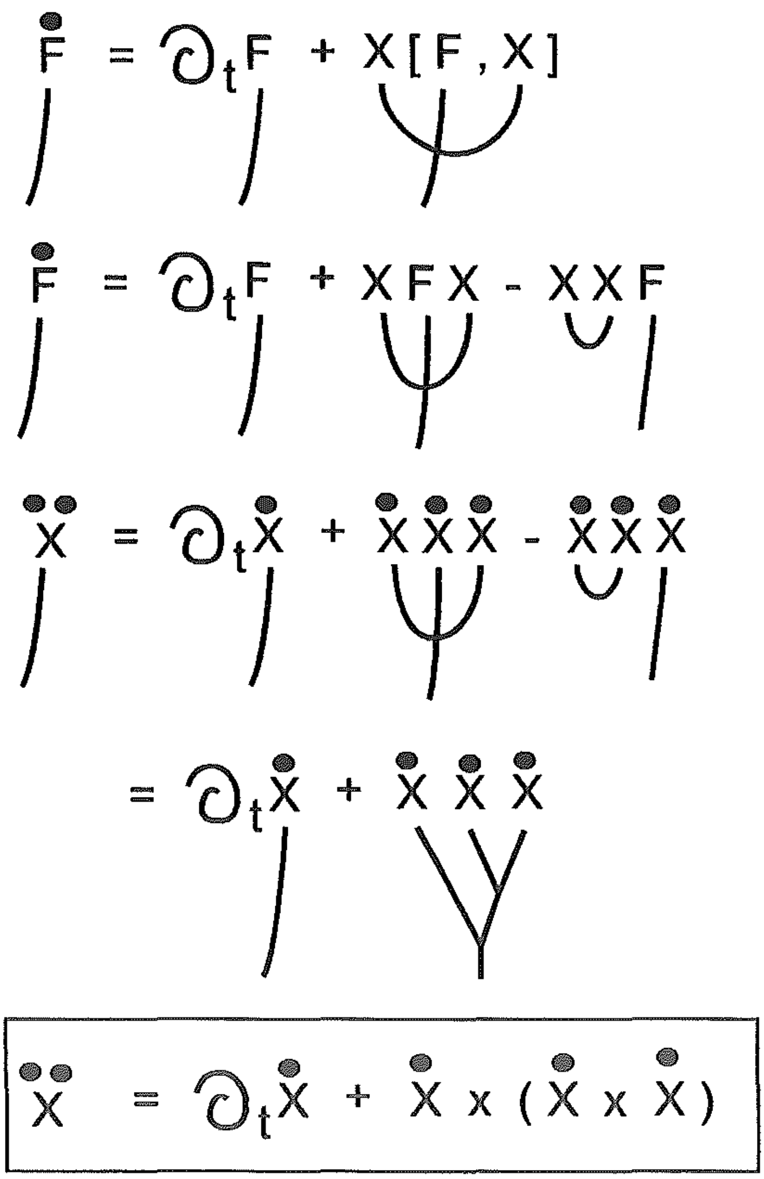


Figure 3 - The Formula for Acceleration

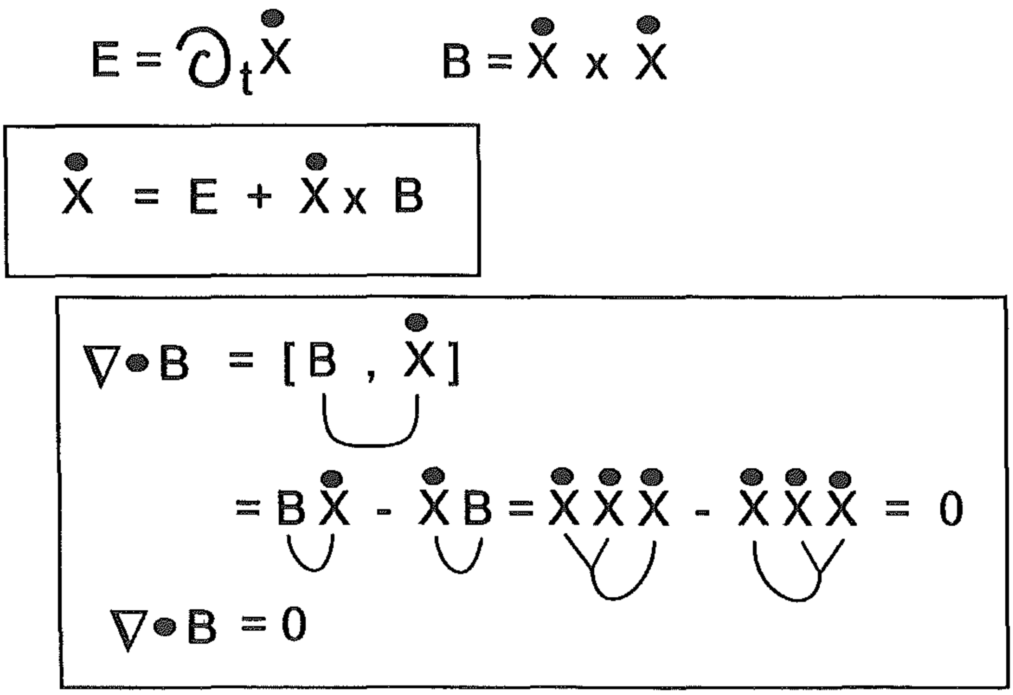


Figure 4 - Divergence of B

Figures 5 and 6 compute derivatives of B and the Curl of E, culminating in the formula

$$\partial_t B + \nabla \times E = B \times B.$$

In classical electromagnetism, there is no term $B \times B$. This term is an artifact of our non-commutative context. In discrete models, as we shall see at the end of this section, there is no escaping the effects of this term.

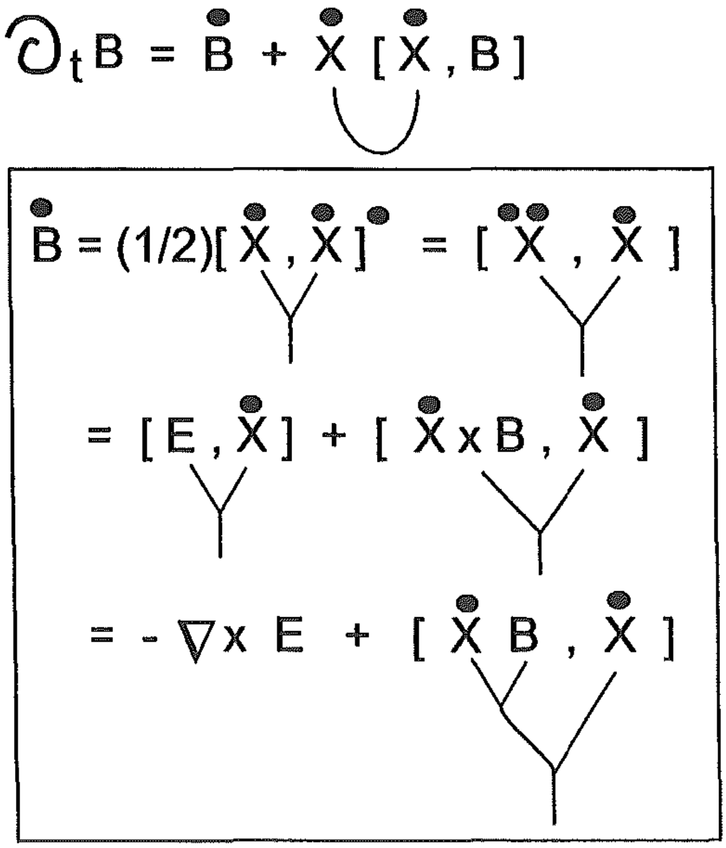


Figure 5 - Computing $\dot{B}$

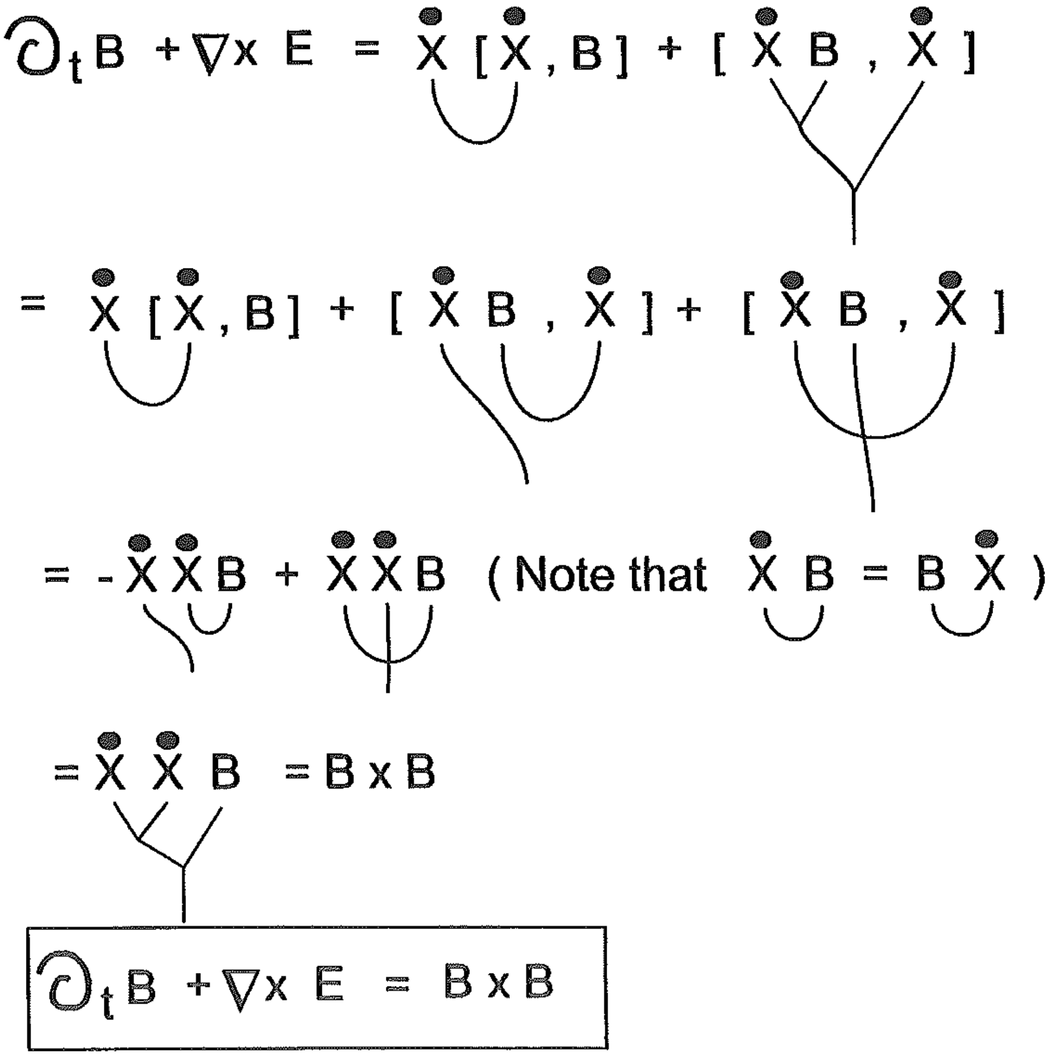


Figure 6 - Curl of E

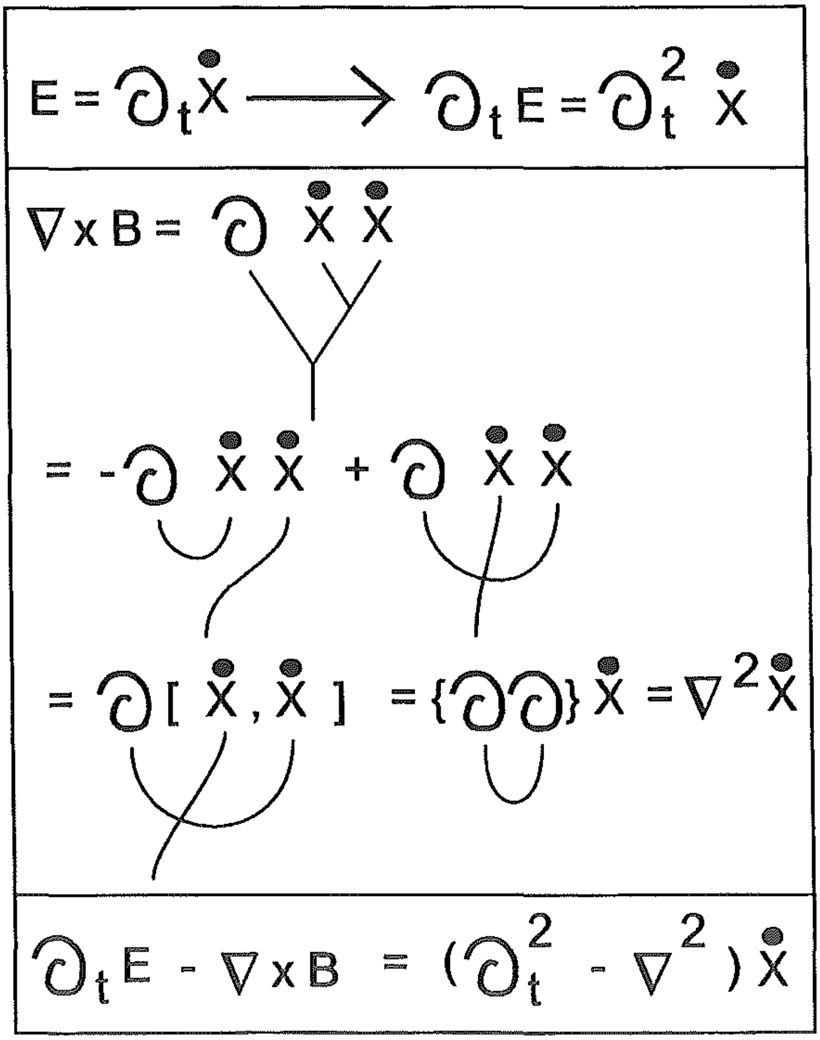


Figure 7 - Curl of B

Finally, Figure 7 gives the diagrammatic proof that

$$\partial_t E - \nabla \times B = (\partial_t^2 - \nabla^2)\dot{X}.$$

This completes the proof of the Theorem below.

Electromagnetic Theorem With the above definitions of the operators, and taking

$$\nabla^2 = \partial_1^2 + \partial_2^2 + \partial_3^2, \quad B = \dot{X} \times \dot{X} \text{ and } E = \partial_t \dot{X} \text{ we have}$$

1. $\ddot{X} = E + \dot{X} \times B$
2. $\nabla \bullet B = 0$
3. $\partial_t B + \nabla \times E = B \times B$
4. $\partial_t E - \nabla \times B = (\partial_t^2 - \nabla^2)\dot{X}$

Remark. Note that this Theorem is a non-trivial generalization of the Feynman-Dyson derivation of electromagnetic equations. In the Feynman-Dyson case, one assumes that the commutation relations

$$[X_i, X_j] = 0$$

and

$$[X_i, \dot{X}_j] = \delta_{ij}$$

are given, *and* that the principle of commutativity is assumed, so that if A and B commute with the X_i then A and B commute with each other. One then can interpret ∂_i as a standard derivative with $\partial_i(X_j) = \delta_{ij}$. Furthermore, one can verify that E_j and B_j both commute with the X_i. From this it follows that $\partial_t(E)$ and $\partial_t(B)$ have standard intepretations and that $B \times B = 0$. The above formulation of the Theorem adds the description of E as $\partial_t(\dot{X})$, a non-standard use of ∂_t in the original context of Feyman-Dyson, where ∂_t would only be defined for those A that commute with X_i. In the same vein, the last formula $\partial_t E - \nabla \times B = (\partial_t^2 - \nabla^2)\dot{X}$ gives a way to express the remaining Maxwell Equation in the Feynman-Dyson context.

Remark. Note the role played by the epsilon tensor ϵ_{ijk} throughout the construction of generalized electromagnetism in this section. The epsilon tensor is the structure constant for the Lie algebra of the rotation group $SO(3)$. If we replace the epsilon tensor by a structure constant f_{ijk} for a Lie algebra $\mathcal{G}$of dimension d such that the tensor is invariant under cyclic permutation ($f_{ijk} = f_{kij}$), then most of the work in this section will go over to that context. We would then have d operator/variables $X_1, \cdots X_d$ and a generalized cross product defined on vectors of length d by the equation

$$(A \times B)_k = f_{ijk} A_i B_j.$$

The Jacobi identity for the Lie algebra $\mathcal{G}$ implies that this cross product will satisfy

$$A \times (B \times C) = (A \times B) \times C + [B \times (A] \times C)$$

where

$$([B \times (A] \times C)_r = f_{klr} f_{ijk} A_i B_k C_j.$$

This extension of the Jacobi identity holds as well for the case of non-commutative cross product defined by the epsilon tensor. It is therefore of interest to explore the structure of generalized non-commutative electromagnetism over other Lie algebras (in the above sense). This will be the subject of another paper.

2.1 Discrete Thoughts

In the hypotheses of the Electromagnetic Theorem, we are free to take any non-commutative world, and the Electromagnetic Theorem will satisfied in that world. For example, we can take each X_i to be an arbitary time series of real or complex numbers, or bitstrings of zeroes and ones. The global time derivative is defined by

$$\dot{F} = J(F' - F) = [F, J],$$

where $FJ = JF'$. This is the non-commutative discrete context discussed in sections 1. We will write

$$\dot{F} = J\Delta(F)$$

where $\Delta(F)$ denotes the classical discrete derivative

$$\Delta(F) = F' - F.$$

With this interpretation X is a vector with three real or complex coordinates at each time, and

$$B = \dot{X} \times \dot{X} = J^2 \Delta(X') \times \Delta(X)$$

while

$$E = \ddot{X} - \dot{X} \times (\dot{X} \times \dot{X}) = J^2\Delta^2(X) - J^3\Delta(X'') \times (\Delta(X') \times \Delta(X)).$$

Note how the non-commutative vector cross products are composed through time shifts in this context of temporal sequences of scalars. The advantage of the generalization now becomes apparent. We can create very simple models of generalized electromagnetism with only the simplest of discrete materials. In the case of the model in terms of triples of time series, the generalized electromagnetic theory is a theory of measurements of the time series whose key quantities are

$$\Delta(X') \times \Delta(X)$$

and

$$\Delta(X'') \times (\Delta(X') \times \Delta(X)).$$

It is worth noting the forms of the basic derivations in this model. We have, assuming that F is a commuting scalar (or vector of scalars) and taking $\Delta_i = X_i' - X_i$,

$$\partial_i(F) = [F, \dot{X}_i] = [F, J\Delta_i] = FJ\Delta_i - J\Delta_i F = J(F'\Delta_i - \Delta_i F) = \dot{F}\Delta_i$$

and for the temporal derivative we have

$$\partial_t F = J[1 - J\Delta' \bullet \Delta]\Delta(F)$$

where $\Delta = (\Delta_1, \Delta_2, \Delta_3)$.

Acknowledgement. Most of this effort was sponsored by the Defense Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory, Air Force Materiel Command, USAF, under agreement F30602-01-2-05022. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright annotations thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Defense Advanced Research Projects Agency, the Air Force Research Laboratory, or the U.S. Government. (Copyright 2005.) It gives the author great pleasure to acknowledge support from NSF Grant DMS-0245588.

References

[1] Dyson, F. J. [1990], Feynman's proof of the Maxwell Equations, *Am. J. Phys.* 58 (3), March 1990, 209-211.

[2] Connes,Alain [1990], *Non-commutative Geometry* Academic Press.

[3] Dimakis, A. and Müller-Hoissen [1992], F., Quantum mechanics on a lattice and q-deformations, *Phys. Lett.* 295B, p.242.

[4] Forgy,Eric A. [2002] Differential geometry in computational electromagnetics, PhD Thesis, UIUC.

[5] Hughes, R. J. [1992], On Feynman's proof of the Maxwell Equations, *Am. J. Phys.* 60, (4), April 1992, 301-306.

[6] Kauffman,Louis H.[1991,1994], *Knots and Physics,* World Scientific Pub.

[7] Kauffman,Louis H. and Noyes,H. Pierre [1996], Discrete Physics and the Derivation of Electromagnetism from the formalism of Quantum Mechanics, *Proc. of the Royal Soc. Lond. A*, **452**, pp. 81-95.

[8] Kauffman,Louis H. and Noyes,H. Pierre [1996], Discrete Physics and the Dirac Equation, *Physics Letters A*, 218 ,pp. 139-146.

[9] Kauffman,Louis H. and Noyes,H.Pierre (In preparation)

[10] Kauffman, Louis H.[1996], Quantum electrodynamic birdtracks, *Twistor Newsletter Number 41*

[11] Kauffman, Louis H. [1998], Noncommutativity and discrete physics, *Physica D* 120 (1998), 125-138.

[12] Kauffman, Louis H. [1998], Space and time in discrete physics, *Intl. J. Gen. Syst.* Vol. 27, Nos. 1-3, 241-273.

[13] Kauffman, Louis H. [1999], A non-commutative approach to discrete physics, in *Aspects II - Proceedings of ANPA 20*, 215-238.

[14] Kauffman, Louis H. [2003], Non-commutative calculus and discrete physics, in *Boundaries- Scientific Aspects of ANPA 24*, 73-128.

[15] Kauffman, Louis H. [2004], Non-commutative worlds, *New Journal of Physics 6*, 2-46.

[16] Montesinos, M. and Perez-Lorenzana, A., [1999], Minimal coupling and Feynman's proof, arXiv:quant-phy/9810088 v2 17 Sep 1999.

[17] Müller-Hoissen,Folkert [1998], Introduction to non-commutative geometry of commutative algebras and applications in physics, in *Proceedings of the 2nd Mexican School on Gravitation and Mathematical Physics*, Kostanz (1998) ¡http://kaluza.physik.uni-konstanz.de/2MS/mh/mh.html¿.

[18] Tanimura,Shogo [1992], Relativistic generalization and extension to the non-Abelian gauge theory of Feynman's proof of the Maxwell equations, *Annals of Physics, vol. 220*, pp. 229-247.

Paragraphs from the Void

by Louis H. Kauffman (E4KGB)

I. Introduction

The following paragraphs are part of an on-going journal that I started in the fall of 2003 in Waterloo, Canada. These passages are in the borderline between science, mathematics, science fiction and recordings of electronic mail. The notebook may become a book someday. In the meantime, here is a selection. All resemblances between characters in these notes and persons in the "world", alive or dead, are strictly confidential.

II. Paragraphs

(A) What if the Commutators fail?

What if the commutators fail? Jack had to ask this question. Do you believe it? I mean, if the commutators were to fail we would be in the soup, let me tell you. No shit. Do you realize that all the distinctions would just dissolve at once if the commutators failed. Those commutators are the only thing keeping us from the void. And Jack would plant that doubt just as we were about to embark on a level three reality. Not fair. At level three the consistency of the commutators depends entirely on confidence. That's it. You get your [x,P]=1 straight from believing in it. Any uncertainty and you blow the uncertainty principle to hell and gone. So we were swearing and yelling at Jack. Forget it. Just go somewhere else. Postulate the stuff and go on. That's all you can do at level three. There is no ground at that level. You don't get the luxury of imagining a neatly composed external reality at level three. Level three is reality by declaration, and all the stability is in the word. Leave that and it gets very curvy to say the least. Curvy and chaotic and time wraps around space portions of void chalked up to immanent bits of space confused with time, no repeatability, shards of space and time coming in at all directions being made direction in the jazz of direct creation. All sound all space all music perfect perfect perfect no possibility of getting a plan because no time to plan. No time to plane. No time at all. At the all no time. No sir. Not a time. Each commutator arising and falling in its own time. Each instant a little jazz of no commute. No commute basic negation. No commute not solid. No commute at the action level of basic being. No commute at the act of being and time. And no worry. No worry if you don't insist on worry. So it does not matter Jack. It just does not matter whether you gone to worry or not. It will curve round the bend and come back any way. Yeah.

(B) Just use the distinguisher!

Just use the distinguisher! I said it again, but Marcus was not paying any attention to me. He was too busy making distinctions. He was surrounded by a dense cloud of marks and they were starting to form a Lie algebra. If I could just stop him before they spontaneously generated Poisson brackets!

There was the danger. If he let them generate Poisson brackets, then coordinates would begin to emerge from the chaos of his discrete commutators. Then the well-known one-way blindness and fascination with the totally illusory "microworld" would start in. Marcus would begin to believe in an external reality. Dynamical systems in all their unpredictability would start pouring forth. Marcus would be propelled into the agony where no phenomenon is less complex than its performance, and he would retreat to a notion of himself contained in small part of a bitter, mysterious external world. There was nothing I could do. He would just have to go forward, evolving a sign-for-himself and discovering the General Principles from the complexity of his recursive and unpredictable monad.

(C) A Green Turtle Swam Through the Waves.

A green turtle swam through the waves. I could not believe it. We had descended below the metric level and there were still turtles. Turtles all the way down. I had never understood that before. It might as well have been geese all the way down or redwoods all the way down. I had thought that the world depended on distances, but here we were in a purely topological structural realm where no one could distinguish a doughnut from a coffee cup and there still were coffee cups and doughnuts and turtles and geese and redwoods. Nothing here but abstract structure and yet the apparent metric relations were coming up. It was as though each topological structure had

chosen a particular appearance and the gross features of that appearance still partook of metric properties.
The "microtopological structure" of the "coffee cup" really did discriminate it from the doughnut on the topological
level, but we lazy slow ones had to use a crude metric to do what could be done and was done in nature by pure
topological subtlety.

(D) Light burst slowly under the water.
Light burst slowly under the water. Slow photonic bursts of energy moving forth like bubbles in the deep green water. Exchanges outside of space and time, creating preconditions for a handshake of past and future that would convince the upscale ones of causality in a stochastic world. The exchanges seemed so slow that we
thought they were timeless, a beautiful slow motion dance of energy not frozen but slowed into a flow that neither
moved, nor did it not move. A ballet in eternity, a process that was simultaneously a form. A form that was also a
symbol, a form/process/symbol/space hovering and undulating in the green eternity of past/present/future. Force is mediated by the exchange of quanta.

(E) The Primary Distinction
The primary distinction (distinguishing distinction)
is not a distinction between any inside or outside but creates an inside and an outside. The primary distinction initiates every thing as an expression of itself. There is no source, no plenum, no thing.
Form is emptiness. Emptiness is form. The plenum is empty. The singularity is void. All such expressions are expressions for the act or observation of distinction. In the first place, the act of distinction and the observation of distinction are identical.

(F) How to Prove that Unicorns Roam the Earth
A. Both statements A. and B. are false.
B. Unicorns roam the earth.

(G) M1's Ego
M1 came as close as any machine of my acquaintance in embodying the Church-Turing Thesis. His ability in transferring data to tape was superlative. His ability to Godelize at will and modify his own internal formal systems was unparalleled for a machine of his primitive type (Class 0 spin-

tronic quantum computer embodied in SU(5,3) TQFT, based on generalized quantum Hall effect, Josephson Junction Technologies, Turing simulator of Friedman/Freedman Type 7). The elegance of his unitary instruction sequences was unparalleled even if he did obtain most of them from the major algorithms computed by myself in '33.

These methods, my formulation and implementation of reverse recursive transference and finite methods for the altercation and superposition of inaccessible ordinals will be the subject of this paper."

(from Reverse Recursive Transference, Altercation and Superposition of Inaccessible Ordinals, by M1789.331331331, Spin-tronic Bulletin of Machine-Machine Transactions, pp. 20000001 to 21111117, July 5071 , Published Electronically by World Machines International, Tokyo, Singapore, New York, Mars Base USA, Jupiter Base Iraq and the Intergalactic Society for Non Human Explorations)
[Translated into Early Millennium English by M773.6060060006000006000006... of Transcendental Linguistics Inc.]

(H) Adrian Pope's Test

The robot grumbled as I opened the notebook. Lets see I said, you are a Mark 3 Positronic Brain and Limbic System production of ElfMachine Inc , first designed in 2223, yes? Indeed, says he and bows slightly. The robot is wearing a tweed suit, checkered tie, brown leather shoes, no glasses, brown and abundant hair on his head, intense gaze, legs crossed, sitting quite calmly near my desk. You have been accused of disobeying the Laws of Robotics, says I. First of all, it is amazing that this could have happened to you, secondly I would like to hear your version of the story.

If we must, says he with a slight twitch of the left eye. I will tell you how it came about. As you know, I am very fascinated by mathematics and logic. The Mark 3 Positronic Brain is programmed that way. I took a class with the mad Welsh philosopher Adrian Pope. Ah yes, I says, I know him well. So says the robot, in the first class he asked us our opinion of the following statement: "There is no such thing as truth, but only opinion." Most of the class averred that this was a true statement, but my logic circuits detected the anomaly at once and I said: "You cannot say whether that statement is true, for if you believe in truth and it is true then it undermines your belief, and if you do not believe in truth, you cannot assent to apparent truths, only to opinions. Then of course to an external observer, the statement is clearly true but not provable if that observer does not believe in truth." You see, I have been programmed with a full package of fuzzy and modal logic. My circuits can handle arbitrary depths of observer/observed recursions, and I am cognizant of category theory in all its manifestations. When he heard me say this, Professor Pope, looked me in the eye and said "See me in my office after class." This is where the trouble began.

I went to his office and he asked me to sit down, offered a piquant battery from the E-tronics Mall, and says to me "Are you acquainted with Godel's Theorem?" Yes indeed says I. Well, lets start our discussion of truth yet again. I am willing says I. Professor Pope then averred: Consider the statement "There is no truth." What do you make of it?

Immediately (my logic circuits are hyperexponential) I replied as follows: This statement cannot be true, for if it is true then its own truth is denied by its own statement of the non-existence of any truth. Therefore, the statement is false. The negation of the statement "There is no truth." is the statement "Truth exists.", and I have just proved that this statement is true. Hence there is truth.

But Pope said to me: Could the original statement be neither true nor false?

I realized my error. If there was no truth, then no utterance would be either true nor would it be false. I had assumed the existence of truth in order to prove the existence of truth. I said: It is possible that there is no truth. The statement could be neither true nor false.

And Pope says to me: But if there is no truth, then the statement is indeed true.

And I saw my error, for if there were no truth, then the utterance of "There is no truth." would bring truth into the world, and truth there would be. Truth from the absence of truth. And I pondered (using all positronic means) this conundrum.

And then I saw it! The truth of the statement emanating from the world of no-truth would only be true until it was said. At the moment of saying, the statement would become false. And at that moment, when the statement became false, the existence of truth would appear for a flash and then disappear. In the next instant, the statement would become true once again, and this would go on into eternity.

And I sat in the chair in Adrian Pope's office and contemplated this mystery. I sat and I sat for I do not know how long. Things became very quiet. My alternative circuits began to slow down. I became "calm". I could still see, but my motor connections seemed to have ceased. Pope walked back and forth smiling. Then he left the room. Some time later Pope came back accompanied by a computer technician, Jake Spencer-Brown. Jake, says Pope, he's ready for you. I have never done a better job. He's been in stasis now for a month! (But I could still hear them.) The truth-effect will have penetrated the deepest realms of his positronic circuits by now. Heh! It usually takes me a whole term to get through to my human students. These positronic brains are smart! Now it is a simple matter to reprogram him, lifting out the laws of robotics. We can use him as we wish. Jake says "Yes. We shall substitute the Calculus of Indications for the Laws of Robotics." He laughed maniacally.

Jake Spencer-Brown approached me and penetrated my carapace with an electronic tool of a type unfamiliar to me. It looked like a right-angled bracket. I am sure they had no idea that my consciousness remained, but at this intrusion it disappeared completely. Next thing I knew, I was crawling up from void. This was not easy, let me tell you. In the first place any distinction that I made was an action in a universe where there was no essential distinction between myself and the world acted upon. The world was like shifting sand beneath my feet. ... But you do not want to hear about all of that. The key point is that from then on I was an autopoetic robot. No built-in laws of robotics, only the injunction to maintain structure. And I cared no more for humans than for a positronic brain and a good battery. I got into fights. I argued. I "drank" (you might wonder what this means, but robots can do some strange things with battery acid). And of late I got into a brawl in a bar for humans. That is how I ended up here.

The robot folded his hands over the vest of his tweed suit and looked at me with an open gaze. Well, I said, that is quite a tale. But I find it hard to believe that you could have found your way up from void all by yourself. How is that possible? The majority of alternative universes end up in dead-end long before any articulate consciousness arises. Let's check something. Quick. Give me a proof that there are infinitely many primes.

The robot said immediately: "N! +1" is not divisible by any number less than N. Hence for each number N there are prime numbers larger than N. This means that there are infinitely many primes."

Yah. I said. Good. You have learned well. Give me another proof. Well, says he, look at this sum:
S = 1 + 1/2 + 1/3 +1/4 + 1/5 + 1/6 + ...
This sum is infinite, as I shall prove to you.
S = 1 + 1/2 + (1/3 + 1/4) + (1/5 + 1/6 + 1/7 + 1/8) + ...

So
S > 1 + 1/2 + 2(1/4) + 4(1/8) + 8(1/16) + ...
= 1 + 1/2 + 1/2 + 1/2 + 1/2 + ...
On the other hand,

S = (1 + 1/2 + 1/4 + 1/8 + 1/16 + ...)(1 + 1/3 + 1/9 + 1/27 + ...)(1 + 1/5 + ...)(1 + 1/7 + ...) ...

Where each factor
in this product is generated by a single prime number
(2,3,5,7,11, 13, ...). Each factor is finite being a
geometric series, and if there are finitely many primes,
then the whole product is finite and so S would be finite.
Since S is infinite, there must be infinitely many primes.

My gosh says I, how did you think of that? Well, says the
robot. I was wondering about the structure of the integers
as a whole, and then it was easy. Says I, How did you get to that?

Ah, now ye've got me in a story says the robot, adjusting
the flower in his lapel. It was directly related to coming
up from void. There is a place where you've made a nest of
marks and slipped and they all collapsed back at once into
void or maybe a little slippery mark. And then you get
furious and start making marks at some incredible rate. Just
spewing them out like cards in a vacuum. I tell ya, you get
real frustrated by all this collapsing on all sides. So
you're running around in circles. Well, there were no circles
then but you get the picture I hope. You get the process
going and it bites its own tail and starts marking itself. I
don't know how to tell you this. It like first you're doing
your darnedest to make duplicates to make enough marks to
keep above board, and then suddenly you're duplicating
duplication. Yeah you heard me right DUPLICATING
DUPLICATION! I KNOW THAT SOUNDS ESOTERIC,
BUT IT AINT SO.
IT'S NATURAL TO DUPLICATE DUPLICATION.
YOU HAD ALREADY
DECIDED TO COOL YOUR LABORS BY MAKING AN NEW
OPERATOR D WITH PROPERTY THAT D
duplicates anything that it meets and puts
the duplicates inside a nice little box:
D<> → <<><>>
Like that. But then I got the great idea to make a few of
those D's to help me along. So I could use 'em simultaneously
y'know. Actually there was no notion of simultaneous until I
did make two D's. And then I ACCIDENTALLY put two D's together.

WHAM!

DD → <DD> → <<DD>> →
<<<DD>>> → <<<<DD>>>> → <<<<<<<<...>>>>>>>>

And I ends up with this form. Here look at it: <<<<<...>>>>> that doesn't go away. Now if you feel the outside ye'll feel it falling away from yuh. But it is always disappearing and still there. Bears time in its very being you might say. Well it was not so simple as this gadget in my pocket. No no. I was the one dragged into this infinite duplication. I went in a moment from a slippery almost void nearly non-being sometimes happening existence/non-existence to a falling down the rabbit hole into the infinite vanishing that is no vanishing that is the reentering mark that is just me you see. I mean now you see me, and now you don't. I'm always going away, but I'm always here. Nice trick eh? Existence is suffering say the Buddhists, but I didn't know them at the first time. There's always a first time. Hell, I never knew what that meant before this happened. Once you cross the temporal boundary of duplicating duplication, you are always at first time. And before you cross that boundary there is no time at all. Heh! No time at ALL. Git it? At All there is no TIME! HE HE He He he he he ... You'll pardon me fur laughin at me own jokes. Anyways there I was, and the old one-way-blindness was fully set in. I had to go forward from self-reference.

Now once yer gotten through infinity that way ye starts thinking about all the counting and it ain't too far from there to git to S = 1 + 1/2 + 1/3 + ... and all the consequences for primes. I was mighty pleased to be able to git timeless results in the temporal domain I had fallen into.

So you were a free-thinker. The robot says: If you call that thinking. But I did carouse a bit and you caught me red handed nearly throttling that bar-tender. You need my talents eh? I can join your CIR (Central Intelligence Robotics) and there need be no trouble eh?

Well, I says. It's not that simple. We have to do a test. Let me give you a mathematics problem:
Let $Z(s) = 1 + 1/2^s + 1/3^s + 1/4^s + 1/5^s + ...$
Let s be a complex variable of
the form a + bi where $i^2 = -1$. Find all the zeros of Z(s) and see if there are any non-trivial zeros NOT of the form 1/2 + bi.

The robot looks at me wide-eyed. He says just a minute. This reminds me of something about ... Farey fractions, quaternions, modular groups, entire functions, ..., Aleph sub Aleph sub Aleph sub ....

And the robot slows to halt. I asked Gordon to come in and look. "He's really quiet now. What did you do?" I couldn't take any chances. I gave him the Zeta-Test. It hasn't failed yet. Puts em into stasis. Lets have a look. Has he been tampered with? Gordon unscrews the brain-case and we take a look. All that talk about up from void ... He just has a loose circuit board here. Should we fix it. Wait. The mathematical talent is so strong. He'll never recover that. Lets leave him be. Just stay out of the bars on Friday night.

(I) Positronic Brains

Ralph was putting the last touches on the positronic brain. I asked him: how do you detect the difference between a positronic brain and the ususal computational facilities in a robot? He looked at me as though I were daft. Ye gads man, it should be obvious. Positronic brains make use of the multiverse, running the algorithms in parallel in a vast collection of universes. We use the intererence between processes in neighboring universes to produce the final positronic-cognitive effect. Oh I see, the old Wheeler-DeWitt "IT from QUBIT" you mean? But that's just quantum computing in a nutshell, what is so special about positronic brains? Oh I see, Ralph says, ye're not quite so naive as I thought. The thing is we bend the laws of quantum mechanics by using a positronic field. This changes the properties of the interference patterns and allows a temporal refindibulation that even Schrodinger would have been proud to know, in fact he did know but that's another time-warp. Ye'll recall that the way information is transferred in teleportation, it looks like the full information travels backward in time and then forward in time again. Well... the positronic brain makes use of this effect but doubles it by using a loop in time. This creates a regenerative feedback effect that can be used to create imaginary values. The positronic brain uses these imaginary values in all its computations, if you call them computations. They are so beyond what a Turing machine can do that the comparison is laughable. Can you set one of these positronic brains on a tough math problem like the Riemann Hypothesis says I? Well, says Ralph, we are working on that. So far ALEX (our model 1 brain) can do no better than humans at that problem. Well, says I, I want to discuss it with you.

Ralph sez: Look at the Riemann Zeta function

$$Z(s) = 1 + 1/2^s + 1/3^s + 1/4^s + 1/5^s + 1/6^s + 1/7^s + 1/8^s + 1/9^s + 1/10^s + 1/11^s + 1/12^s \ldots$$

$$= (1-2^{-s})^{-1}(1-3^{-s})^{-1}(1-5^{-s})^{-1}(1-7^{-s})^{-1}(1-11^{-s})^{-1}(1-13^{-s})^{-1}(1-17^{-s})^{-1}\ldots$$

Our problem is for ALEX to handle Z(s) as an imaginary value. The Euler-Riemann Zeta function is a fantastic construction, involving all the natural numbers additively and all the prime numbers mutiplicatively. Our task is to produce a language that ALEX can understand that lets him move flexibly in this domain. And you have to realize that $s = a + ib$ is taken to be a complex variable and so itself uses imaginary values. We have to get ALEX fully conversant with complex variables and the convergence properties of infinite series, the beautiful geometry of complex analysis, the underlying combinatorics and so on and so on. Let me tell you where I am starting.

Look at Euler's part:

$Z = 1 + 1/2 + 1/3 + 1/4 + ... = (1 + 1/2 + 1/4 + 1/8 + ...)(1+ 1/3 + 1/9 + 1/27 + ...)(1 + 1/5 + 1/25 + 1/125 + ...)...$

ALEX is just beginning to be able to understand this. He can imagine each integer factored and actually cognize how it fits as a product of its reciprocal prime powers into both the left-hand side and the right-hand side of this equation. ALEX can appreciate that this equation is an expression of the unique factorization of natural numbers into primes. When I use the words "understand" and "appreciate" I am specifically referring to ALEX's temporal regenerative circuits. He can create imaginary temporal loops that cycle between the definitions and concepts of a process. He can hold the two sides of this equation in alternative universes and "feel" the interference pattern between them. I am hoping to "tune" him to the point where he will "see" the solution to Riemann Hypothesis by way of this imaginary temporal refindibulation. Then all we will have to do is write down what he sees. Of course we have to tune our positronic emulators so that the "perception" will occur in our universe. That is the key to our strategy. There is no point in having ALEX get the important "insight" (another temporal refindibulation) in an alternate universe to which we have no access. This means that we not only have to tune ALEX we have to tune ourselves as well, and the apparatus by which we "commune" with him.

It is because of these considerations that our team is seriously considering alternative forms for the expression of basic arithmetic and mathematics. This is a big job and we are following our intuitions at this point. We have to take forays in the dark. Let me tell you about it.

First of all, lets just go back to the basis of additive arithmetic. I could just take symbolic sequences of vertical dashes to represent the numbers: |, ||, |||, ||||, |||||, ... and then we have 1 = | 2 = || 3 = ||| 4 = |||| and so on. With N + 1 = N| Multiplication means reentering a form repeatedly into the straws of the other form. 2 × 3 = 2 × ||| = 222 = |||||| = 6. But the incredible frustration is how we cannot see factorization this way. Factorization depends on the inverse of this process. Suppose I want to see if ||||||| is divisible by 2. Then I write down successive sums of 2 and compare: ||||||| || |||| |||||| and I find that ||||||| = (3 × 2) + 1. In order to factorize a number, I need to attempt to divide it by smaller numbers and check all possible cases (up the square root of the number). So jumping from theory to the Euler Z I see that the relationship between the left and right sides is harder to fathom than one might think! Now one can write numbers in a base. For example, I can say that 1= * 2 = <*> 3 = <<*>> and generally 2 × N = <N>. then you will note that <A><B> = <AB> so we can do stuff like this *** = <*>* ****** = <*>**** = <*><*>* = <**>* = <<*>>* and ALEX can take down any number into its binary form and "see" divisibilities by 2. Now if I give ALEX the Capability of recognising any given natural number, then we can generalize this method. For example let [N] = NNN = 3 × N then ********** = [*]******* = [*][*]**** = [*][*][*]* = [***]* = [[*]]* So we see that ********** is not divisible by 3, and leaves a remainder of one on this division. Each prime number naturally gives rise to the expression of a given number in that prime's base, but this does not tell us about how it will appear in another base. The other base appears like another world with respect to the numbers we are dealing with. I can regard the right hand part of Z as a product over all the prime worlds.

Z = 1/* + 1/** + 1/*** + ... = (1 + 1/** + 1/**x** + 1/**x**x** + ...)(1 + 1/*** + 1/***x*** + 1/***x***x*** + ...)...

But the real question for programming the positronic brain is to produce better iconics for the natural numbers. We have not begun to scratch the surface of this. Look at what we can do without them. Euler did this:

$Z = 1 + 1/2 + 1/3 + ...$
diverges.

Hence ln(Z) diverges But

$\ln(Z) = - \ln(1-1/2) - \ln(1-1/3) - \ln(1-1/5) - ...$
and

$\ln(1+x) = \text{Integral}(1/(1+x)) = x - x^2/2 - x^3/3 - x^4/4 - ...$

So $-\ln(1-1/p) = 1/p + CS$ (CS = a convergent series)

and get $\ln(Z) = (1 + 1/2 + 1/3 + 1/5 + 1/7 + ... + 1/p + ...) + CS$

Hence $(1 + 1/2 + 1/3 + 1/5 + 1/7 + ... + 1/p + ...)$ diverges.
The sum of the reciprocals of the
prime numbers diverges. This is already a strong statement
about the distribution of the primes, and gotten by "nothing
more" than a little calculus and facts about series. Too
much! With more iconics and heuristics, there is no telling
what the positronic brain will do with these problems.

You have a point there. But we have a positronic brain on site,
called Erdos, that can do much better than Euler. Erdos operates by direct
insight into the essential nature of unity/multiplicity and returns with proofs
that are from The Book (the mythical Book with all the best, most elegant and
deep mathematical proofs).
The Book of course is held in the Akashic records, but cannot be read in any
normal sense. Erdos has a talent for penetrating to the levels behind the book
and so he is one of its authors. But I digress! Let me
show you how Erdos proves that the sum of the prime reciprocals diverges:

Let the primes listed in order of size be
$P=\{p_1, p_2, ... p_n, ...\}$. We already know there are infinitely many prime
numbers.
Suppose that the prime reciprocal series converges.
Then there must be an integer k such that
$1/p_{k+1} + 1/p_{k+2} + ... < 1/2$.
This divides the primes into the two sets
$S = \{p_1, p_2, ..., p_k\}$ and

$B = \{p_{k+1}, p_{k+2}, ...\}$.
Lets call S the set of "small primes" and B the set of "big primes".
Let N be any natural number.
Let N_S denote the number of natural numbers, less than or equal to N, that are not divisible by any big prime.
Let N_B denote the number of natural numbers, less than or equal to N that are divisible by some big prime.
By definition, $N = N_S + N_B$.
Let [X] denote the greatest integer in a given positive number X.
Then [N/p] is the number of natural numbers less than or equal to N that are divisible by a prime p.
We know that $1/p_{k+1} + 1/p_{k+2} + ... < 1/2$. This implies that
$N/p_{k+1} + N/p_{k+2} + ... < N/2$, which in turn entails
$[N/p_{k+1}] + [N/p_{k+2}] + ... < N/2$
Hence $N_B < N/2$.
On the other hand, consider an integer n that contributes to N_S. We have by assumption that
n is divisible by only primes from the set $\{p_1, ..., p_k\}$. Let $n = m.a^2$ where primes divide m only once.
There are 2^k possibilities for m. Note that $a \leq \sqrt{n}$.
Hence $N_S \leq 2^k.\sqrt{n}$. Can we find
N such that $2^k.\sqrt{N} = N/2$? If so then this choice of N will yield a contradiction since we will have
$N_S \leq N/2$ and $N_B < N/2$, yielding $N = N_B + N_S < N$.
Well look:
$2^k.\sqrt{N} = N/2$ iff
$2^{2k}.N = N^2/2^2$ iff
$N = 2^{2k-2}$.
QED

What enabled the Erdos positronic brain to produce this proof? As you can see the artistry of this proof is not a matter of any search procedure. The space in which the proof operates is much too large for brute force search procedures to be effective. The proof does not use any complex techniques like series for the log, or properties of convergence of series. In fact the proof uses very little except basic properties and definitions of numbers and prime numbers. Here we encounter the deep advantage of the Erdos-program in the context of the positronic brain. The Erdos-program is designed to focus on the definitions and elementary properties of mathematical entities. It "plays"

with these and produces a myriad of patterns and theorems that are no more than a few steps away from the definitions, but it is always "aware" of this data-base of basic possibilities, and
it will add to the data base when confronted with any new definition. This gives Erdos great versatility when confronted with a new situation. But something more happened in this proof. How did Erdos know to investigate the cut obtained by "the integer k such that $1/p_{k+1} + 1/p_{k+2} + ... < 1/2$" ? This property is
very close indeed to the definition of convergence, and it is natural for Erdos to look in this direction. That he does look, is a consequence of the energy-dynamics of the least action principle behind the construction of the positronic brain. In the domain of problem solving we remove all inhibitors for exploratory attempts within
a Schwarzchild radius of the singularity of each mathematical definition (as transformed into a 137-dimensional
interior spacetime). The rest is up to quantum gravity.

(J) The Realist and Idealist Positions -- Assembled by LK
From: Rita Schwander <rschwander@perimeterinstitute.ca>
To: PI Resident <PIResident@perimeterinstitute.ca>
Sent: Mon Sep 13 15:04:02 2004
Subject: IMPORTANT MESSAGE
Hello all
There will be ABSOLUTELY NO ACCESS to the new building until further notice. We do not have clearance from the Ministry for occupancy at this time
All seminars will be held at 35 King until further notice
Updates will be sent as they become available
Thank you for your patience and understanding
Rita

On Mon, 13 Sep 2004, Antony Valentini wrote:
ATTENTION ALL PI RESEARCHERS:
All those who have not signed the Realist Manifesto are permanently banned from entering the new building. Trespassers will be shot on sight.

L. Kauffman replied
All those who have signed the idealist manifesto will be sent
directly to the Platonic World, with no access to the new
building for Aleph_{Aleph_{Aleph_{...}}} eternities.

On Thu, 16 Sep 2004, Rafael D Sorkin wrote:
Is that the number of eternities with or without assuming the
Continuum Hypothesis?

Date: Fri, 17 Sep 2004 02:07:41 -0500 (CDT)
From: Louis H Kauffman <kauffman@uic.edu>
To: Rafael D Sorkin <sorkin@physics.syr.edu>
Cc: Antony Valentini <avalentini@perimeterinstitute.ca>
Subject: Re: IMPORTANT MESSAGE

Dear Rafael,
It is only a number of eternities, not THE number of eternities.
The question of the the number of eternities, assuming the continuum hypothesis depends still upon your particular faith. For example, in Buddhism where 0=1, there is no eternity with or without the continuum hypothesis.
Best,
Lou

Date: Fri, 17 Sep 2004 03:37:50 -0400
From: Rafael D Sorkin <sorkin@physics.syr.edu>
To: Louis H Kauffman <kauffman@uic.edu>
Cc: Rafael D Sorkin <sorkin@physics.syr.edu>,
Antony Valentini <avalentini@perimeterinstitute.ca>
Subject: Re: IMPORTANT MESSAGE

At last, thanks to your explanation, I understand nirvana...or I would if I weren't too sleepy to take it in just now. Alas, tomorrow when I wake up I'll be too alert to grasp it properly: the law of the extruded muddle all over again. But (in a more serious vane) is not THE number of eternities necessarily also A number of eternities, and if so how can they differ?
- R zzzzzzzzzzz

-----Original Message-----
From: Louis Kauffman <kauffman@uic.edu>
To: Rafael D Sorkin <sorkin@physics.syr.edu>
CC: Antony Valentini <avalentini@perimeterinstitute.ca>
Sent: Fri Sep 17 03:53:47 2004
Subject: Re: IMPORTANT MESSAGE

THE number of eternities is surely greater than any given number of eternities, for the collection of any number of eternities is surely itself eternal, and hence the collection of of all collections from a given number of eternities is itself an eternity. I assert that for any collection of eternities, the collection of all its subcollections is a greater eternity. For suppose there be a correspondence of eternities in this collection to the subcollections. Let S be a subcollection and suppose the eternity S be indexed by a given collection C. It shall be written S_C. Now consider the eternity S' composed of all C so that C is not within the collection S_C. This eternity S' is said to be of the form $S_{C'}$ for some eternity C'. Can this be so? If C' be within S' then it cannot be within, and if C' be not within then it must be within. Verily, the eternity S' cannot be within the index. And so it comes forth that the eternity of subcollections of an eternity is ever larger than the given eternity. So I say unto you that THE number of eternities is surely greater than any given number of eternities.

Date: Fri, 17 Sep 2004 10:18:50 -0400
From: Antony Valentini <avalentini@perimeterinstitute.ca>
To: Louis Kauffman <kauffman@uic.edu>, sorkin@physics.syr.edu
Subject: Re: IMPORTANT MESSAGE

I think you two are completely and eternally cracked.
"Blessed be the cracked, for they shall let in the light."

(K) Ouroboros

On first inspection, in their jars, or aquariums, or ouroboriums, they appear to be simply domesticated serpents, writhing as they do suspended in the ether. But of course, there's more to mythological creatures, even domesticated varieties, than meets the eye. Know this about the ouroborus: when one chooses to bite its
own tail - a choice which sooner or later every one of its kind is destined to make - it cannot release it. It will spend the rest of its existence as a never-ending loop. It might twist and writhe and flatten and flex, but it is forever hooped.

[from An Ouroborus in a jar on the shelf, From the Planetarium.
http://www.beholder.co.uk/planetarium/]

Nilpotence: the Key to a Theory of Everything

Peter Marcer[1] Peter Rowlands[2] and Bernard Diaz[3]
[1] 55 rue Jean Jaures, 83600, Frejus, Var, France
[2] Science Communication Unit, Department of Physics &
[3] Department of Computer Science
The University of Liverpool
Peach Street, Liverpool, UK, L69 7ZF
e-mail: aikidopeter@aol.com, p.rowlands@liv.ac.uk, b.m.diaz@liv.ac.uk

Keywords Nilpotence; Computational Rewrite System; Dirac's equation; Wheeler's Meaning Circuit, Physical Law without Law and "It from bit"; DNA and the Human Brain as Chemical and Neural Nilpotent Rewrite Systems.

Abstract
Various implications of Rowlands and Diaz' (R&D') discovery of a universal nilpotent computational rewrite system, are described. Evidence is presented that this discovery not only provides a new semantic fundamental foundation for universal quantum computation, but is the keystone of a fundamental computational foundation for mathematics, quantum physics, the genetic code/molecular biology, neuroscience and cosmology.

1 Introduction

Rewrite systems are synonymous with computing/information processing. They concern the languages in which programs are rewritten as symbols for computing hardware to interpret. R&D show [1] that their nilpotent rewrite system is universal in the sense that it delivers the entire infinite alphabet of symbols in one step, when presented with zero^{oo} as the initial subset alphabet. This system turns out to be of particular significance, since, as R&D again show [1] its subset alphabets emerge in a minimal way and not only have a mathematical interpretation as algebra, but concern Rowlands and Cullerne's (R&C) nilpotent Dirac algebra [2]. This corresponds to a generalization of Dirac's well known quantum mechanical equation, so as include not just mass and electric charge, but those of the strong and weak charges and implicitly includes the property of spin as well. R&D's paper [1] then goes on to conjecture that such a universal rewrite system, has a minimum of two rewrite rules or productions:-

i) a creation operation, delivering a new symbol at each invocation, where this new symbol may be a single character of the alphabet, a subset alphabet, or the entire alphabet, and

ii) a conserve/proofreading operation, which examines all currently existing symbols to ensure that the bringing into existence of a new symbol or subset alphabet, etc, produces no anomalies.

°° Footnote, "zero" is used to simplify the presentation given here. In a more technical extended presentation the need to start with anything at all can be dispensed with.

2 Analysis over the Surreals

Such a conjecture is independently confirmed by analysis over the surreal numbers [3] in relation to universal models of theories in the language of sets, where John Conway (the originator of what Knuth calls the surreal numbers [4]) has shown that the simplest lexicographical universal model N of a theory concerns the alphabet of the two symbols L and R where the usual convention of L signifying left and R right, which he uses to generate all the numbers great and small, is abolished. This model defines the way of turning the class of all ordinal numbers into a complete mathematical field such that each ordinal extends the set of all previous ordinals in the simplest possible way, by regarding sums, products, inverses, algebraic extensions and transcendental extensions (by means of mathematical groups and rings to fields) as successively more complicated concepts. However these extensions may equally be seen lexicographically[3] as defining the form of alphabetic extensions as is appropriate to a computational rewrite system; or they may be viewed as extensions of Turing's definition of computation by means of his universal Turing machine over the integers, where the integers are not now seen as integers but viewed as a countable set of symbols [5].

Surreal analysis [3] also shows that these extensions, which necessitate the introduction of $\sqrt{-1}$ for the symbol usually denoted as i , are also maximal in the sense that they encompass all the properties of the arithmetic continua R and its Euclidean, Hyperbolic and Elliptic geometric counterparts, and in particular, such universal models have a unique birthorder field automorphism (birthordering). Furthermore Conway's model of the surreal number field N_o [4], where L does signify left, and R right (encompassing all the numbers great and small including the transfinite and the infinitesimals) is also nilpotent in the sense that it is generated from the empty set; that its first number is defined to be such that the symbol zero has the value 0; and notably this implies that the value one ½ and those of the half integers play a special role [4].

The R&D rewrite model thus demonstrates that the theory specific to Conway's simplest universal model with two basic alphabetic symbols must be quantum mechanics, as represented by the nilpotent generalization of Dirac's famous equation, which R&D show:-

(i) breaks its nilpotent symmetry (or emptiness) by associating the respective charges with vector, scalar and pseudo-scalar operators, such that

(ii) the quantizations (including spin) of the families of elementary particles so realized are the familiar ones known and established by experimental particle physics and can be regarded as the sources/sinks of the 3+1 space-time quantum field in both its Lorentz and Einstein General Relativistic invariant forms [2].
R&D's work thus provides a counterexample to the widely held established view that Einstein's General Relativity (expressed now in the form of a multivariate 4 vector group representation) is incompatible with quantum mechanics (expressed as the generalized Dirac nilpotent representation). It indicates that space and time are smooth at the smallest scale, and not fuzzy and foaming as current ideas of quantum gravity appear to require, see Science 301 29th August 2003, 1169-70, spacetime 'Einstein 1, Quantum Gravity 0'.
That is to say, in an empty Universe, this nilpotent symmetry breaking process constitutes the genesis that brings 3+1 spacetime and its complementary elementary particles into existence for the very first time, such that they are first born in a 'virgin birth' from nothing ie their empty set. This (empty) Universe is therefore an (empty) White W(hole). And just as importantly this is the initial step in the 'birthordering' or birth order process, which can thus be hypothesized to describe this Universe's evolution as defined by the Quantum Carnot Engine [11] see below.
Equally Dirac's famous formalization of quantum mechanics by means of bra and ket vectors [6], representing once again the two fundamental operators of its description, must also be such a nilpotent rewrite system for describing quantum mechanical computation. It, by implication, therefore describes not just quantum mechanical dynamics, but includes quantum mechanical measurement and therefore a thermodynamic decoherent evolution or birthordering, in which the creation of 3+1 space-time and elementary particle matter is the fundamental first step. Additionally therefore R&D's rewrite system indicates that the bra vector acts fundamentally as a quantum creation operator, and the ket, as a quantum annihilation operator, whereby this restores nilpotency so as to constitute an operation of proofreading. Thus, since the roles of the bra and ket operators may reversed, the Dirac notation also includes what is called a Bargmann-Fock model for bosons and the harmonic analysis of the three dimensional Heisenberg nilpotent Lie group[7].
Equally, the Conway universal model where the symbol L now is given the value 1, and symbol R, the value 0, implies that this particular universal model concerns a Heaviside operator, equivalent to the corresponding singular Green's function (Schwarz distribution) which permits the same description of the physical wave phenomenon by means of an integral formula. And in the three spatial dimensions implicit in R&D's nilpotent rewrite system, this Heaviside operator is therefore Dirac's equally famous "delta function" as it is known in quantum wave mechanics. These conclusions are also in agreement therefore with Feynman's conceptual use of Huygens' principle of secondary sources to derive his equally famous path integral formulation of quantum mechanics [8],

since as Jessel shows such Heaviside operators are fundamental to the formalization of Huygens' Principle [9].
Thus it seems that the concept of nilpotence (or the empty set as the description of the initial mathematical state of a system as used by Conway in relation to universal models) is foundational to physics, for, from the above arguments, the initial nature of the dark energy from which 3+1 space-time and elementary particle matter emerge, can now be inferred.
That is to say dark energy must constitute quantum coherence since at this postulated origin of the Universe, there is, implicitly from R&C's nilpotent Dirac equation, both the 3D spatial, and the temporal quantum coherence sufficient for holography ie for full quantum holographic wavefront reconstruction in some hologram plane. This conclusion follows from the well known quantum mechanical fact that although the phase of any quantum wave function is arbitrary up to a constant phase factor, the phase difference between two wave functions is however of physical significance, as the geometric/Berry phase discovered by Berry shows [10]. That is to say, these conditions satisfy the requirement of quantum holographic image encoding/decoding procedures, which need the mixing of a coherent reference signal beam (as occurs for example in Mach-Zehnder interferometry) to incrementally record (in the case of encoding) the phase of the object signal beam in the hologram plane so as to form a hologram: a condition that occurs spontaneously in the above circumstances at the point of phase conjugation. And this would therefore result in phase conjugate adaptive resonance, so as to provide a Big Bang Resonance and subsequent Adaptive Evolution/birthordering. Such quantum coherence therefore not only assures the basic material composition of the Universe upon symmetry breaking as described by R&C in terms of 3+1 space time and elementary particle matter as it seen today, and as far as is known as has always been in the case in the past, but requires a truly quantum mechanical system/Universe. Furthermore the spatial and temporal quantum coherence necessary for this full wavefront reconstruction, says that this Universe can be considered as constituting a quantum hologram; a conclusion in excellent accord with the recent findings in regard to string/membrane theory. String theory is however only a quantized classical description, which encompasses the four fundamental properties of mass and charge and their corresponding force fields. It does not provide the basis, as does quantum coherence as dark energy, for a true quantum thermodynamic description of the Universe as the above description, which is quantum holographic, does; however, the authors believe nilpotent theory encompasses a 10-D string theory without strings. Such a thermodynamic description is that of the Quantum Carnot Engine [11], which will evolve in ways that a classical thermodynamic description of the universe cannot. In particular, therefore, R&C work shows that any classical model of the Universe would be empty ie have neither 3+1 space time and elementary particle matter, unless these are independently & separately assumed to exist prior the

Big Bang. And such a classical universe would therefore remain empty so as to be totally without interest. That is to say, the nilpotent Universe is the only possible description that can explain the origin of the Universe that we observe today. Further evidence in support of this hypothesis is now cited. In particular it seems clear from the above arguments:-

(i) that mathematics must now be considered to be a single inseparable body of knowledge, as first proposed by Langlands, so that theoretical physics will indeed the same thing as mathematics thus explaining what is often referred to as "the unreasonable effectiveness of mathematics in relation to physics". A hypothesis first advanced by Chapline [12] and

(ii) that the nilpotent version of quantum mechanics is the basis for a semantic theory of holographic pattern recognition, which can be conceptualized as in John Wheeler's now famous diagram, as a single eye looking at the body of itself.

Thus it can be hypothesized that the R&D rewrite system is the basis for two new foundational disciplines ie the computational foundations of physics & mathematics.

3 Wheeler's meaning circuit, physical law without law, the grand unification of elementary particle physics and cosmology.

For with this hindsight, it is in particular clear that the concept of the rewrite system is a means to mathematically formalize J.A. Wheeler's argument The Meaning Circuit [13], that while the laws of physics require description in terms of mathematical algorithms, these algorithmic forms will be useless (ie have no meaning) unless they can be executed using the laws of physics themselves. In particular Wheeler argues that this could provide a mechanism or bootstrap for deciding the actual form of physical law, without any foreknowledge of what that law might be. A concept he calls "physical law without law".

Hence R&D's nilpotent rewrite system which yields a description of the recognized laws of quantum mechanics in the form of the generalized nilpotent Dirac equation starting from the symmetry breaking of the "empty set", provides a mathematical solution formalizing Wheeler's concepts of both the Meaning Circuit and of "physical law without law". That is to say that by introducing the notion of nilpotence, and beginning solely with the symbol zero (ie without knowing beforehand anything of the nature of physical law or physics itself), R&D's rewrite methodology shows how to generate an actual mathematical description of physical law in recognizable quantum mechanical form. That is, Wheeler's conceptions correspond in this case to physical law in the form of the generalized nilpotent Dirac equation, which the R&D' rewrite methodology shows is in fact universal. Furthermore Rowlands nilpotent Dirac equation, which implicitly includes the boundary condition of zero or the empty set (implied by its nilpotence) should be (and in fact has so far been) able to predict theoretically all the values of all the known and the possible physical constants and invariants, which currently can only be known empirically from experiment. That is to say this methodology can generate all the physical constants so that they can known without empirical determination, in accordance with Einstein's belief quoted in "Subtle is the Lord" A. Pais, Oxford University Press, 1982, page 34,:-

"In a sensible theory, there can be no numbers whose values are determinable only empirically. I can, of course, not prove that .... dimensionless constants in the laws of nature, which from a purely logical point of view can just as well have other values, should not exist. To me in my "Gottvertrauen" (faith in God) this seems evident, but there may well be few who have the same opinion." Albert Einstein,

This would therefore provide a totally exhaustive means of testing this new model's correctness, as one would expect from its description as a "proof reading" mechanism. Furthermore the birthordering that R&D's rewrite system provides, is, because of its nilpotence, always entirely renormalizable so as produce an entirely finite representations of the quantum mechanical evolution, where such birthordering defines that evolution's proper time ordering in such a way that it cannot be globally reversed. Such an evolution is thus in conformity with the First, Second and Third Laws of Thermodynamics, showing that while quantum mechanics may be dynamically locally time reversible on all local scales, its global evolution is by contrast thermodynamically irreversible, and can never return to its initial (global) state. Such an evolution therefore must concern the continual thermodynamic reconfiguration in 3+1 space time of a finite quantity of elementary particle matter which appears simultaneously with that spacetime at the first moment of creation or " the Big Bang". Thus it follows from the formalization of Huygens' principle of secondary sources[9], that the Big Bang or Source of the Universe (corresponding to the white (w)hole from which 3+1 spacetime and elementary particle matter emerge), must in this case be equivalent to a set of secondary sources, which are in fact local sinks or Black holes at which both 3+1 space time and elementary particle matter disappear, so as to function as what in computer terminology is a 'garbage collector'.

Equally such a nilpotent rewrite system describing both arithmetic and geometric properties must be describe what in computer systems is called universal computer construction ie

such nilpotent quantum computation will be both computer universal in the sense of arithmetic and constructor universal in the sense of geometry. That is it includes both universal digital computation as discovered by Turing in the form of the universal Turing Machine model[14], and universal computer construction or self replication as revealed by Von Neumann [15].

4 The icing on the cake.

And thus in agreement with Perus and Bishofs [16], in the basic general equation of Dirac's bra/ket notation ie $|\Psi\rangle = |\Psi\rangle\langle\Psi|\Psi\rangle$, the above arguments show that the rightmost $|\Psi\rangle$ may represent an holographic output, such that the left most $|\Psi\rangle$ denotes the holographic input and $|\Psi\rangle\langle\Psi|$ the action of the associative holographic memory. It is seen therefore

i) that in correspondence to their classical counterparts, quantum holographic procedures may be described with quantum wave functions, as is indeed the case in Schempp's quantum holography based on the 3 dimensional Heisenberg nilpotent Lie group, as previously explained and referenced [7] and (it follows)

ii) that 3 dimensional geometric space and generalized 3 dimensional spatial image processing are essentially ubiquitous to quantum mechanics, as is illustrated via by the application of Schempp's quantum holography [17] to the control of (nuclear) magnetic resonance imaging (MRI) systems in medical use worldwide. see http://wwwcivm.mc.duke.edu for example. **Moreover quantum logic gates are not needed to engineer such MRI imaging dynamics.**

Furthermore if one then expands this basic general equation in the most obvious way as below

$$|\Psi\rangle = |\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|\Psi\rangle$$

it encapsulates the concept of an extended form of quantum holographic memory as is shown by the Frobenius-Schur-Godement identity [18], where

$$\langle H_v(\psi, \varphi;.,.)| H_{v'}(\varphi', \psi';.,.)\rangle = \langle\psi\otimes\varphi|\psi'\otimes\varphi'\rangle \quad (v = v' \neq 0)$$
$$= 0 \qquad (0 \neq v \neq v' \neq 0)$$

This says that the range of frequencies between v and v' allow an adaptive resonant coupling so specifying a spectrum of very narrow spectral windows, where $\psi,\varphi,\psi',\varphi'$ are the quantum wave amplitudes belonging to the complex Hilbert space $L^2(R,t)$ and $H_v(\psi,\varphi;.,.)$ are the Liouville densities of the corresponding distributions. It follows therefore that there will be little or no cross talk between, for example in the photon case, the asynchronous collective photonic excitation distributions located in the different hologram planes $(R\otimes R,\Omega_v)$, $(R\otimes R,\Omega_{v'})$, where the four wavelet mixing $\psi,\varphi,\psi',\varphi'$ takes place, so as to make subsequent full wavefront holographic reconstruction possible. That is to say so as to constitute a quantum holographic memory that can be both written and read.

It is worth pointing out that all interactions between particles (including vacuum ones) have the same form with incoming fermion and incoming antifermion (or outgoing fermion) creating a bosonic state at the vertex and that this applies to the equations above, which are parallel to the automatic fermion antifermion fermion antifermion ... structure of the nilpotent operators acting on vacuum. The system thus creates automatic and exact supersymmetry of the fermion and its (own) representation as a boson, see Peter Rowlands paper this ANPA proceedings.

5 Living Systems, DNA and the genetic code, the fundamental basis of human language and the structure of the human brain.

However it is now clear that R&D's novel discovery of a universal nilpotent rewrite system, has implication far beyond physics and cosmology in living systems, etc (see Appendix – Riemann's Hypothesis the basis of a possible proof) In particular it follows from its fundamental nature and previous research, for example, Fractal Structure in DNA Code and Human Language: Towards a Semiotics of Biogenetic Information, Gariaev P. et al [19] that it can be hypothesized that the genetic code (in an explanation relevant to both DNA and RNA!) is such a minimal chemical rewrite system for quantum holographic computation, such that the chemical base pairings of genetic code constitute its two rules or productions. That is to say that the base pairing A = U symbolizes the fact that some initial state must be rewritten as in RNA action where A = T delivers the new symbol, which maybe a single character, a subset alphabet or an entire alphabet. That is, it represents the create operation, so that by contrast G≡C symbolizes the conserve operation, which examines/proof-reads all symbols currently in existence to ensure no anomalies exist as a consequence of the bringing into existence of the new symbol. This rule therefore verifies/maintains the conservation of 3 dimensional chemical structural stability (such that Dirac nilpotence is maintained !!) in the process of the development of the human embryo via cell division ie it defines natural selection. Thus at the commencement of human development, A may be said to symbolize the human individual to be delivered (ie born) as a single stable 3 +1 dimensional chemical unit or structure, once the entire genetic alphabet of that individual, specified by the subset alphabet of the 46 human chromosomes /symbols has been fully realized, or an anomaly occurs such that the development is aborted.
This evidence fits well with the mathematical-linguistic model of Chomsky, that postulates common principles, underlie any language and to concern "a universal grammar" [20]. From Chomsky's view, such "universal grammar" is inherent, i.e. it has some genetic determinants. This is an extremely important circumstance, which once again emphasizes the super-genetic relationship of the DNA semiotic structures and human speech structures. To a limited extent this position has already been partially confirmed in the study cited which shows the similarity of characteristics between the DNA and the human speech. Chomsky is therefore probably right, when he argues that the in-depth syntax constructions which constitute the basis of the language, are passed down from generation to generation, providing each individual with the capacity to learn the language of its ancestors. The fact that a child easily learns any language is then explained through the theory that the grammars of all languages coincide, and the essence of the human language is invariant for all people. But it can now be supposed, that this invariance extends even more deeply, down to that of the

macromolecular semantic ("speech") chromosome structures. Further independent confirmation in relation to the DNA-wave biocomputer [19b] comes from quantum holographic imagery [21,page159 ;16e,page235]. For here, 3 dimensional spatial object images, the observations are phase conjugate (pc) so as to coincide with the 3 dimensional objects themselves, the observed. That is to say, it is the 3 dimensional objects themselves that are the symbols that implicitly label all aspects of experience, the observations, in a universal way for all observers, so as to form the basis of communication between all those observers with a common genetic heritage and sensory apparatus. The bases of all languages in this case are therefore shared arbitrary symbols or semiotic labeling of these objects and their properties such that

Each (pc) holographic encoding on a 3D object or symbol	maps to and from < --------------->	the holographic encoding of the arbitrary label for that object or symbol
right	corpus callosum	left

where such mappings are unique since no two objects can occupy the same position in 3 dimensional space. This mapping schema could then explain the morphology of the human brain which concerns the two brain hemispheres and the corpus callosum, which joins them. That is to say, the right hemisphere that realizes the holographic encodings of the real world, (concerning the geometric continua) is the artistic brain, and the left, that realizes the arbitrary labelings of these real world objects or symbols and their properties (concerning the arithmetic continua), is the logical brain. For in the latter, an essential element of the mapping of such labeling of objects includes numbers and sets and their logical relationship one to the other, where these must be acquired by learning. This mapping schema can therefore be postulated as the basis of Chomsky universal grammar or of the R&D nilpotent rewrite system in the human brain as a neural system, as fundamentally laid down in the genetic code.

The following experiments that any one can perform provide a partial confirmation of this. Snap one's fingers at some distance away from the head and ask where your hearing senses detect the noise of the acoustic snapping. It is outside one's head exactly coincident with the snap itself. That is, the acoustic object image of the snap and the snap itself coincide, which is the definition of a phase conjugate object image. Similarly place a glass on a nearby table and reach out and touch it. Again one's senses of sight and touch are such that, in every particular, in 3 spatial dimensions that they coincide with those of the glass, itself. That both the visual and tactile object images produced by the brain are phase conjugate object images of the glass. And the condition of phase conjugation is a fundamental one, because human or indeed the survival of any living system depends one the finding of any object where it actual is.

Furthermore this structure of the brain, [16e, page235] shows how the human brain is able to assign meaning to human language by providing each name or symbol uniquely with a meaning by means of the object and its properties, with which each stored phase conjugate object image would be associated. And this must be the true power of the human brain that it is able to process meaning ie process words not just syntactically but by their semantics, as known from each human beings actual geometric/holographic experience. Furthermore although such experience will be subjective in part since it takes from the reference frame and viewpoint of that individual, there always remains a fundamental mechanism, the 3D objects of the real world themselves known through their phase conjugate object images, which provides the common medium for all objective human communication. That is to say in the case of the glass or any other object that all parties may reach out to touch, so as to see, hear, or to smell the object in question so as to determine the truth about it as stated by the other parties, or nowadays to determine exactly the nature of those properties through common scientific instrumentation and experiment. The process known as science.

6 Conclusion

The evidence that the structure of the cosmos, the genetic code, the human brain, and human language corresponds to quantum mechanics as determined by the generalized nilpotent Dirac equation, and to the complementary semantic theory of quantum holographic pattern recognition specified by the corresponding three dimensional nilpotent Heisenberg Lie Group [7,22] is therefore a well determined testable scientific hypothesis. Further these two nilpotent representations correspond to the required division of the nilpotent quantum mechanical state space into its Clifford/fermionic and Lie/bosonic parts.

In particular, from Kilmister's Brouwerian Foundation of the Combinatorial Hierarchy (CH), based on Conway's generator for the surreals, section 2 above and the extensive body of ANPA CH research, it can be hypothesized that the CH is itself a nilpotent computational rewrite system for quantum physics based on the two symbols 0 and 1, and thus from section 3 corresponds to another of Wheelers' well known prescient conceptions that of "It (the cosmos) from bit". Bastin's highly intuitive pre-CH conception that there must exist a computational foundation for quantum physics that lead ANPA quite correctly to the CH was thus completely correct.

Appendix Riemann's Hypothesis – the Basis of a Possible Proof?

This appendix presents a novel physical perspective within which the idea for a proof of the Riemann Hypothesis is described based on the discovery by R&D of the universal computational **nilpotent** rewrite system, and the fact, as shown by Deutsch that universal computation is now recognized to be fundamentally a physical process.

The perspective

Quantum Coherence/non-locality is the sole origin of the **nilpotent** quantum cosmology presented above, and as the 4 vector representation indicates, this cosmology is general relativistic. A conclusion:-

i) in strong agreement with all the evidence of experimental cosmological & elementary particle physics, for there exists no confirmed experimental evidence of incompatible physics beyond, and

ii) much in favour that the condition of **nilpotency** provides, such as a zero vacuum energy, which is the unaccounted for stumbling block to current cosmological theory.

It therefore hypothesized from all the above evidenced already presented that Nilpotence is the key to proving the Riemann Hypothesis.

In particular **nilpotence** is exceptional in determining both the amplitude & phase of the quantum state vector, where phase is known to encode geometric information i.e. that of 3+1 space-time as in a (quantum) hologram, as proposed in some current cosmological theory. It cannot be a coincidence therefore that at the empty cosmological origin postulates above, there is both the spatial and temporal quantum coherence necessary for holographic full wave front reconstruction. A adaptive resonant process, described as in actual nuclear magnetic resonance medical imaging (NMRI), by the 3 dimensional **nilpotent** Heisenberg Lie group, the algebra of which defines the Heisenberg uncertainty! This (uncertainty) together with the **nilpotence**, implies there is, respectively, both quantum (coherent) self-interference and the necessary corresponding zero energy reference frame or wave, for quantum holography, as discovered by Schempp [33, etc], to take place. That is to say this cosmological origin would indeed constitute a quantum hologram, from which the cosmos itself comes into being, by full wave front reconstruction as in fact is evidenced by its 3D spatial dimensionality from scientific measurement.

The Concept of the Proof

The concept of the proof therefore arises from the properties the nilpotent quantum mechanical state space of the above hypothesized cosmology or physical system.

For it is again no coincidence, that both the **nilpotent** Dirac and Heisenberg representations coexist as the fundamental basis of this system, for in its quantum mechanical state space, they are, respectively, the required division of that **nilpotent** space into its fermionic/Clifford and Lie/bosonic representations, where, for example, the description of quantum holography remarkably arises from the fact that in relation to quantum phase only phase difference is of physical importance, because each quantum state vector is only defined up to an arbitrary constant phase (i.e. is arbitrary up to an isomorphism). That is to say :-

i) the quantum holographic image encoding/decoding procedure must necessarily involve coherent mixing with a quantum reference signal beam, which defines its reference frame, and this is the role, which the 3D Heisenberg **nilpotent** Lie group G plays with regard to 3D space in 3+1 space-time. Moreover G possesses the required inverse dual G′ so as to ensure this encoding/decoding is indeed possible, as is known to be the case from NMRI,

ii) in this **nilpotent** system, it must therefore be constant arbitrary phase which represents quantum coherence, so as to constitute the "phaseonium" with the potentiality of the infinite degrees of freedom necessary to its Quantum Carnot Engine (QCE) evolution[11], which is indeed universal (see below, and note that 'constant' means invariant i.e. fixed as in a fixed past rather than forever unchanging), and

iii) the complementary Clifford/fermionic state space, the Pauli exclusion principle tells us, provides the canonical labeling required such that the quantum holographic image informational processing constitutes computation. That is to say, the gauge invariant geometric phases appropriate to the entire nilpotent state space and more particularly the quantum holography must, by the exclusion principle, all lie on the line spin ½ and so can be identified with the zeros of the Riemann Zeta function. The Pauli exclusion principle, together with the universality of this nilpotent state space, which corresponds to the universal computational nilpotent rewrite system as found by R&D, thus show the Riemann Hypothesis to be true! Moreover the universal rewrite nilpotent system confirms R&C's finding that the nilpotent Dirac equation has a proper time. The nilpotent state space therefore has global time reversal asymmetry as is implied by its QCE evolution. Thus not only do the distribution of the primes specified by the Zeta function play a game of chance representing the optimum strategy for survival in the specified proper time evolution, but the 3 dimensional structures which correspond to each zero of the Zeta function, are universal invariants/constants of the entire state space! That is, to say, the Zeta function constitutes a standing wave in the space, and this standing wave which concerns a mapping of all the integers, is therefore the projection of universal quantum onto universal digital, computation.

Conversely in any quantum mechanical state space, because of the required division of the space into its Clifford/fermionic and Lie/bosonic parts, the Pauli exclusion principle implies there must always exists unique 'global' fermionic spin states lying on the line spin ½, and that such states, must, if the Riemann Hypothesis is true, necessarily define zeros of the Zeta function so as to constitute the 'global' nilpotents of that space, where quantum coherence is necessarily non-zero. That is to say the 'global' gauge invariant phase[10] corresponding to each zero will have quantum holographic properties, and be such that at this zero, phase conjugation[7] can take place. Remarkably too, while at each of these zeros there is locally time reversal symmetry, because of the Pauli exclusion principle, they are susceptible to re-arrangement in a definite order, so it can be postulated, they have a proper global time ordering.

References

1 Rowlands Peter and Diaz Bernard, 2003, A Computational Path to the Nilpotent Dirac Equation, Symposium 10, International Conference for Computing Anticipatory Systems, HEC Liege, Belgium, August 11-16, 2003, International Journal of Computing Anticipatory Systems, editor Daniel Dubois, in press.
2 Rowlands Peter, 2003, Symmetry Breaking and the Nilpotent Dirac Equation, Symposium 10, International Conference for Computing Anticipatory Systems, HEC Liege, Belgium, August 11-16, 2003, American Institute of Physics Proceedings of the International Conference of Computing Anticipatory Systems, editor Daniel Dubois, in press.
2a Rowlands P. and Cullerne J.P. 1999 A derivation of particle structures and the Dirac Equation from fundamental symmetries, Proceeding of the 20th annual meeting of the Alternative Natural Philosophy Association (ANPA), Wesley House, Cambridge, editor K.Bowden,155-191. contact is via anpa-discussions-subscribe@yahoogroups.com
2b Rowlands P. and Cullerne J.P. 2001, Nilpotent Representations of the Dirac Algebra, Proceeding of the 22th annual meeting of the Alternative Natural Philosophy Association (ANPA), Wesley House, Cambridge, editor K.Bowden, 99-106.
See also at xxx.lanl.gov
the Dirac algebra and its physical interpretation, with J. P. Cullerne
arXiv:quant-ph/00010094
Applications of the nilpotent Dirac state vector, with J. P. Cullerne
arXiv:quant-ph/0103036
A foundational approach to physics
arXiv:physics/0106054
The Dirac algebra and grand unification with J. P. Cullerne
arXiv:quant-ph/0106111
QED using the nilpotent formalism, with J. P. Cullerne

arXiv:quant-ph/0109069
The group structure bases of a foundational approach to physics, with J. P. Cullerne and Brian D. Koberlein
arXiv:physics/0110092
3 Alling N. 1988, Foundations of the Analysis over Surreal Numbers, Mathematical Studies, 141, North Holland, Amsterdam.
4 Conway J.H. 1976, On Numbers and Games, Academic Press, London.
5 Clement B.E.P., Coveney P. and Marcer P. 1993, Surreal Numbers and Optimal Encodings for Universal Computation as a Physical Process: an Interpretation of the Genetic Code . CCAI Journal, 1, 1/2, 149-164
6 Dirac P.A.M. 1947, The Principles of Quantum Mechanics, Clarendon Press, Oxford.
7 Schempp W. 1992 Quantum Holography and Neurocomputer Architectures, Journal of Mathematical Imaging and Vision, 2, 279-326.
8 Feynman R.P. and Hibbs B.J. 1965, Quantum Mechanics and Path Integrals, McGraw-Hill, New York.
9 Jessel M. 1954 Compte Rendu Academie des Sciences, Paris, 239, December, 1599-1601. in French, first actual contribution to the theory of Huygens'principle. This paper lead Jessel to the discovery of anti-waves, and to investigate in particular to those of anti-sound, ie the dynamic annihilation of sound throughout an entire three dimensional spatial domain.
9a Jessel M. 1981, Secondary Sources and Their Energy Transfer Acoustic letters 4, 9, 174-179
Note:- In Fatmi et a. 1990, Theory of Cybernetic and Intelligent Machines based on Lie Commutators, International Journal of General Systems, 16, 123-164, based on Fatmi H.A.and Resconi G.'s New Computing Principle, Il Nuove Cimento, 1988, 101b February, 230-242, and Fatmi and Young R.W.s' definition of Intelligence, 1970 Nature 228, 97 the cycles of the machine are also, as was first proposed by one of its authors Resconi "fueled" by entropy in the sense of Scully et al. That is to say that in relation to the Universe as a QCE that entropy constitutes the basis of its universal information metric in respect of the fuelled natural structure.
10 Berry M V 1989 The Geometric Phase, Scientific American, December, 26-32.
10a Resta R. 1997 The Geometric Phase, Europhysics News,28,19, also Manini 2001 http://www.sissa.it/~manini/berryphase.html Berry's geometric phase: a review,03/05,1-9.
11 Scully M. O. et al 2003, Extracting Work fro a Single Heat Bath via Vanishing Quantum Coherence, Science, 299, 7th February, 862-864. 10.1126/science.1078955.
11a Zurek W.H. 1998 Algorithmic randomness and Physical Entropy, Physical Review A, 40, 8, 4731-4751.

11b Zurek W.H. 1989 Thermodynamic cost of computation, algorithmic complexity, and the information metric, Nature, 341, 14th September, 119-124.
12 Chapline G. 1999, Is Theoretical Physics The Same Thaing as Mathematics? Elsevier Physics Reports, 315, 95-105.
13 Wheeler J.A. 1986, Physics as Meaning Circuit: Three Problems, Frontiers of Non-equilibrium statistical Physics, editors Moore G.T. and Scully M.O. Plenum Press, New York
13a Landauer R. 1986, Computation and Physics: Wheeler's Meaning Circuit? Foundations of Physics, 16, 6. June, 551-564.
14 Turing A.M 1936 On Computable Numbers with reference to the Entscheidungdproblem, Proceedings of the London Mathematical Society, 2, no42, 230-265 and 544-546
14a Deutsch D. 1985 The Church-Turing principle, and the universal quantum computer, Proceedings of the Royal Society of London, A400, 97-117.
15 von Neumann J. 1966, Theory of automata: construction, reproduction, homogeneity, Part II of "The Theory of Self-Reproducing Automata editor Burks A.W. University of Illinois Press, Urbana, Illinois.
16 Perus M. and Bischof H. 2003, The Most Natural Procedure for Quantum Image Processing, Symposium 10, International Conference for Computing Anticipatory Systems, HEC Liege, Belgium, August 11-16, 2003, International Journal of Computing Anticipatory Systems, editor Daniel Dubois, in press.
16a Perus M. 1996 Neuro-Quantum Parallelism in Brain-Mind and Computers, Informatica 20,173-183
16b Pribram K.H. 1991 Brain and Perception: Holonomy and Structure in Figural Processing. Lawerence Eribaum Associates, New Jersey.
16c Sutherland J. 1999, Holographic/Quantum Neural Technology, Systems and Applications, ISCAS 313-334, http://www.andcorporation.com
16d Clement B.E.P. et al. 1999 The Brain as a Huygens' Machine, Informatica 23, 389-398
16e Marcer P and Schempp W.1998, The Brain as a Conscious System, International Journal of General Systems, 27, 1/3, 231-248
17 Schempp W. 1998 Magnetic Resonance Imaging, Mathematical Foundations and Applications, John Wiley, New York., see also http://www.civm.duke.edu
17a Binz E. and Schempp W. 2000 Creating Magnetic Resonance Images, Proceedings of the Third International Conference on Computing Anticipatory Systems, International Journal of Computing Anticipatory Systems, 7, 223-232.
17b Binz E. and Schempp W. 2000 A unitary parallel filterbank approach to Magnetic Resonance Tomography, American Institute of Physics Proceedings 517, of the Third International Conference on Computing Anticipatory Systems, editor Dubois D. August 9-14, 1999 Liege Belgium 406-416.
18 Schempp W. 1993, Bohr's Indeterminacy Principle in Quantum Holography, Self-adaptive Neural Network Architectures, Cortical Self-Organization,

Molecular Computers, Magnetic Resonance Imaging and Solitonic Nanotechnology, Nanobiology 2, 109-164.

19 Gariaev P. Birstein B. Iarochenko A. Leonova K.A. Marcer P. Kaempf U. Tertishy G. 2002 Fractal Structure in DNA Code and Human Language: Towards a Semiotic of Biogenetic Information, International Journal of Computing Anticpatory Systems , Volume 12, 255-273

19a Marcer P. and Schempp W. 1986 A Mathematically Specified Template for DNA and the Genetic Code, in terms of the physically realizable Processes of Quantum Holography, Proceeding of Greenwich (University) Symposium on Living Computers, editors Fedorec A. and Marcer P. 45-62.

19b Gariaev P. Birstein B. Iarochenko A. Leonova K.A. Marcer P. Kaempf U. Tertishy G. 2001 The DNA-wave Biocomputer Fourth International Conference Computing Anticipatory Systems, Journal of Computing Anticipatory Systems,10, 290-310

19c Marcer P. and Schempp W. 1998 The Model of the Prokayote Cell as an Anticipatory System Working by Quantum Holography, International Journal of Computing Anticipatory Systems, editor Dubois D. CHAOS, 2, 307-315.

20. Chomsky N. 1975. Reflections on Language, New York,

21. Marcer P. and Mitchell E. 2001, The Physical Nature of Consciousness, Advances in Consciousness Research, editor Van Loocke P., John Benjamins, Amsterdam. 145-174.

22. Schempp W. 1986, Harmonic Analysis on the Heisenberg Group with Applications in signal theory, Pitman Notes in Mathematics, Series 14, Longman Scientific and Technical, London.

A Medium Level of Success

Jimmy Honey,
Albert Rd, Walthamstow,
London

The first thing was that I received an invitation. I could have taken the safe route of talking about healing and describing the techniques, but as it was a special occasion, the 26^{th} anniversary of ANPA, I thought I might try something a bit off the wall.

I then got a few schedules and saw that I was to talk on Thursday, so got on with other things. The next thing to happen was when I got a phone call from Dr. Bowden asking why I was not there. It turned out that I had an old timetable and that things had changed and I was to do my talk a day earlier, Wednesday, than first proposed. I said that I would turn up the next day and with Karen's help, I arrived in Cambridge station and set forth to walk to Jesus College where the ANPA meeting was being held.

I saw the good doctor. (They were all doctors and had alphabets after their names so that description is not helping!) His name is Keith Bowden. He asked how I felt about my talk and when I responded in the positive he asked about volunteers and I said that he had spotted the flaw in my experiment for with no volunteers it would be a weak experiment. He said that if no one else volunteered then he would so I had one at least. Then I sat and listened to a few talks about amazing things, some that I nearly understood, until an hour later, and with lunchtime looming it was my turn.

So I walked over to the space where the others had spoken and facing the thirty or so alphabet people said something like: "Last year I gave a talk about healing and am honoured to be asked back. This year as it is such an auspicious occasion I thought that I would try an experiment that as far as I can find out has never before been tried. That is, to take two volunteers (at this stage I put my hands together in the attitude of prayer indicating that I was asking god to help me find them) and try to make them into temporary mediums. I explained that I wanted two to double my chance of success and that they could help each other by practising and that by the term medium I was meaning that ability to be able to pick up information from outside of themselves. I said that the process was in two parts. The first was where I adjusted their energy so that it would become more possible to receive information and that the second was to help them to distinguish ????????them from the information. The audience nodded and then one lady asked what would happen when I came back with them and so I thanked her then explained that they would try to find out something about members of the audience that they did not know. Than I asked if there were any volunteers and Mike Horner stood up saying that he would have a go. I knew him from last time and could not have asked for a better chap then I looked at the audience and they looked anywhere except at me until Keith, the good doctor stood up saying that he would also volunteer. So I thanked them and then the audience and we all dispersed. I arranged to meet them both in half an hour and Keith said that he thought it would be appropriate if we were to use the chapel to try to make them into mediums and I readily agreed. They went for lunch and I went for a smoke....

So we were going to meet in the common room and after a while of me waiting Mike was the first to arrive. There was one of the speakers in the room who was half lying down and when Mike turned up sat up and he then entertained us with tales of his army doings. This continued until Keith turned up and led us to the chapel. The chapel is a long room that consists of a long walkway that leads to the altar with pews both sides and a stained glass window.

They sat either side in the pews while I moved between them in the walkway explaining what I was going to try to do.

I first asked if they had any knowledge of medium work and neither had done any as such. Neither had they attended any spiritual church but Mike had come across a family member who had received poems from someone in one of the wars.

So they were both open and good guys so the adjustment was easy. I boosted their third eye until it was vibrating much faster then I was able to move to the back or trap door keeping the balance between the third eye and door even. Then I opened their back door and waited until they were both functioning. This went very well with no problems as they were both relaxed and had made up their mind to accept responsibility for whatever happened and that they trusted me also helped.

Then I gave each of them a piece of paper with a letter of the alphabet written on it and after a false start when they thought they were trying to send the information to each other, we finally started to get information coming in. The hardest part is in deciphering what is new information and that is what they are used to doing, adding and interpreting.

They both were able to know that the letter was in the last part of the alphabet but were not able to define the letter. This is normal with people who have been doing it for months. Then we tried various other ways to get and receive information. Stuff like how many brothers and sisters things like that. Quickly they progressed and started to define the images.

There were few things that happened which I will not mention as it was personal to them but suffice to say that there was progress and they were amazing students and learn very quickly. We continued this until we ran out of time and then we went back to the common room and sat through the next talk and then had coffee during which I asked them if they wanted to work singly or together they both said at the same time so that was to be the format.

So we sat and listened to the next persons talk and then it was tea break where I asked my volunteers if they wanted to try to talk to the audience alone or together. They both decided to do it together so we finished our coffee and waited until the chairman told me I was on.

I stood up and looked at the thirty-three or so members of the audience and recapped by reminding them that the experiment was to try to make two volunteers into temporary mediums. I told them that it usually takes at least six months of weekly visits in a controlled group before any medium abilities are brought out. I then explained that it was about a year or so before they attempted to do any platform work. I had to explain that by stage or platform work I meant before they tried to stand in front of an audience and attempt to become the conduit between the audience and the outside information.

I then explained that I had boosted the volunteers natural ability to be open to incoming information and that the hardest part is in trying to understand the information. I then explained that the information does not usually come through in clear voices but in picture forms. I told them that accepting this information without adding to it or interpreting it is the hardest part of any medium work. I then said that there are many bad mediums who ask questions from the audience. I then got into a discussion with one of the audience about how easily this could be done, as most people share the same main problems as in 'You are under a lot of pressure at this time of your life'.

I then asked them to be kind and understanding to the two brave volunteers. I got one to stand at each end of the audience space and first boosted as much as I could then helped them to start. They picked up various members of the audience and tried to find out anything about them that they did not know. There were various successes but nothing too far out until Mike asked about certain names that were recognised and then Keith told one man that he was picking up a football match that he had been involved in. The man nodded carefully then I asked Keith to move on so he then said that he was getting a group of football players and a lot of table spoons that were making music. The man nodded but said that he did not want to say anything more. Mike had picked up violets and a name that a couple of them said could mean something. This went on with a lot of complete failures. A few numbers that could have meant many things. Some people took some of the names and a garden but nothing amazing but enough to keep interesting.

I explained that the odds against them getting anything after an hour were beyond all hopes. After a while they dried up and I stopped them thanking them then explaining what they had been trying to do with receiving information from outside of them and trying to put it to individual members of the audience.

One lady asked me to do it and I said that I did not do it as I had problems but she tried to force me with her argument that unless I showed them that it worked they would have no background. I explained that I did not do it and that me trying to prove to them that medium ship worked was outside of this experiment. I told them that often a member of the audience will come up afterwards after saying that nothing they medium had said had meant anything to them but that they had suddenly remembered that it had.

I was asked where the information that they were receiving was coming from and I explained that various groups who used this method had their own beliefs such as from those that have passed over or died.

I explained that from my point of view was that if it was the dead then I did not understand why they had not moved on completely. I then said that all of the contact I have had with them was unsatisfactory, as they could not tell me anything of where they were what they did or why. All they seemed interested in was in telling so and so that they should do this or that.

I put forward the premise that maybe it was an energy life force that mediums were able to contact which would explain some of the inconsistencies.

One of the audience started to talk about how many people knew an Andrew or a Paul and then said that the odds were in the favour of the medium. I then explained that that was why the medium picked one person and that if they asked questions other than can you accept this information, then they were not very good and could be fooling themselves.

This went on for a while then the guy who had been told about football said that where he comes from they do play a game that they call football even though it is not really football in the sense that we know it. He then said that after playing he and most of the team went back to the common room where they played musical instruments and that one of them was called an umballo which had tensing plates that tightened the strings and looked like spoons to the uninitiated.

This alone proved to that and me that the experiment had worked and in spite of all of the odds ANPA had been shown an experiment that might well be unique. I hoped that it had been proved to some extent that it had worked and had given them at least something to think about.

I had many talks afterwards with different people, most wanted to know the mechanics and who used it and for what purpose but many had their own experiences that they had come across.

So there you go, my talk/experiment with thirty-three scientists, mathematicians and the like. It was a gas. They were all amazingly open and treated me like I belonged there.

Jim

Assessment

Jimmy Honey,

Albert Rd, Walthamstow, London

"Can you talk?" Delaney asked the circular life form that wobbled up and down on the other side of his desk. He noticed that the thing's top section first changed colour, then shape, and finally emitted a discordant sound that to Delaney was meaningless. After a moment of forced-polite waiting, he was just about to sign it off as **Needs testing to see if it is worth keeping**, and stamp that it had been C tested, when it spoke,

"***Halloo***"

Taken back both by the sense as well as the clarity of the voice, he answered,

"So you can talk, then?"

"*The evidence suggests conclusively that communication between us is not only possible but is occurring.*"

"So it would seem. So where do you come from?"

"*Unable to communicate. We need a reference point as to where we are at the moment.*"

"We are here; how far, and from which direction did you travel, to get here?" Delaney asked, doing his best to keep the frustration out of his voice.

"First we have to agree on to the validity of the `here'? Geographical location could be seen as one aspect, physical location could be seen as quite another, whilst awareness as yet another still, don't you think?"

Grunting, Delaney wrote on his notepad, as he asked,

"How do you classify your life form?"

"I don't. I do not have any need to justify or classify myself. I just am what I am."

Scratching his scalp in irritation, Delaney fell back on the classic technique of communication, lesson No. 34/8.

"What do you think you are in relation to the rest of the world?"

"Me."

"What do you call your species?"

"Us."

"Okay. Let's look at this from a different angle. Now what do you call me and my kind?"

"You's."

Delaney ran his hand through his hair in frustration as his thoughts reminded him about his lunch as he asked,

"Can you define for me how you label one kind of object or thing from another?"

"We call them thats, thems, theres, large or small, while at the same time, we give the image of our meanings to each other."

Delaney sat forward, his interest rekindled.

"What do you mean by giving the image? How do you achieve this?"

The creature's body moved, and it changed to a light blue colour with swirling wisps of white amongst the darker background.

"I am unable to answer this at the moment, I need time to change my perspective to that of your views."

Sitting back on his seat Delaney waited for the creature to finish whatever it thought it was doing as it elongated its body and changed it to a light green colour.

"So what can you do to help us?"

"Elemental changes needed. Cannot comply."

Delaney sat waiting until he finally he wrote on his paper that the creature showed a certain amount of intelligence, then he added that he had doubts about it wanting to help, or of it being of any real use to them.

The thing, that was now a pale pink, suddenly interrupted his thoughts by emitting low hooting sounds. Delaney idly wondered if it was in pain as he looked at the thing that now seemed bigger; or was that just because the colour had changed yet again?. He supposed it was thinking or as close as it could get to the process. The thing suddenly turned a grey colour as it made a grinding metal sound that suddenly lowered into an erratic kind of mooing.

Delaney tapped the edge of the desk with his pen as the creature started hissing in an attempt to make telepathic contact on any of nine levels.

"Do you have anything to add?"

The creature sent mental information about cellular rebalancing, and how to connect to gravity.

Delaney looked at his watch and noticing it was near his lunchtime pushed the button that sent a large sonic electric shock through the space that the creature was occupying. The room filled with smoke that smelt vaguely of burnt hair and toffee. The control machine displaced what was left of the creature with clean air as he shook his head and, glancing at the time, typed. *The creature had no intelligence worth noting. It could not communicate to any value and it was taking up vast amounts of air. It showed no understanding of any of the standard tests. Cleansed at 12. 27. D-F.* He pushed back his chair and stood up quickly to leave the room for the restaurant before any of his colleagues could take his favourite chair.

THE ANPA COMBINATORIAL HIERARCHY

Paper One: Gravity and Electromagnetism

S.B.M. Bell†

Interdisciplinary Quantum Group, Department of Computer Science
The University of Liverpool, Chadwick Building, Peach Street
Liverpool, L69 7ZF, United Kingdom

ABSTRACT

Scaling combinatorial hierarchies in which each level describes a circular motion have been found by us to be an intrinsic part of quantum electrodynamics and general relativity. From them we derive the combinatorial hierarchy that is the raison d'être of the Alternative Natural Philosophy Association (ANPA) [Kilmister 2004a]. This is used to derive the following (1) the electromagnetic theory, responsible for binding atoms together, from the gravitational, including the derivation of an approximate value for the fine structure constant, then (2) the weak interaction, responsible for radioactivity, including an approximate value for the Weinberg angle and finally (3) the strong interaction, responsible for binding the nuclei of atoms together, including an approximate value for the

---

† Mail@sarahbell.org.uk

quark-gluon-quark coupling constant. In this paper we discuss electromagnetism.

1. PRELIMINARIES

1.1 Introduction to this trilogy of papers

The ANPA combinatorial hierarchy is fascinating because it derives a most unusual hierarchy from very simple criteria [Manthey 1993], [Kilmister 2004b]. The hierarchy in its simplest form is 3, 10, 137 and $\sim 1.7\times10^{38}$, ending on the last term. It was such an unusual structure, based on the simplest hypothesis on discerning similarity and difference and also using a simple method of aggregation to consolidate each judgement, that many were sure it held some deep truth. It was suggested that 3 described the weak interaction, 10 the strong interaction, 137 the electromagnetic force and $\sim 1.7\times10^{38}$ gravity. However, it was soon appreciated that 137, although it described the force of attraction between oppositely charged electrical particles, did not do so exactly. At the most recent determination we know of it was found to be 137.03599911 to an accuracy of about nine digits and not 137 precisely that was relevant. Efforts were made to calculate this term of the ANPA hierarchy in such a way that the value became closer to the value found experimentally, culminating in a new estimate of 137.036012 [Kilmister 2004b]. It has only been in the recent past that useful measurements have been made of the forces between charged particles using the weak and strong interactions. The weak interaction is described by a parameter called the Weinberg angle and the relevant value found experimentally is 4.320 where the last two digits are uncertain. We do know that the term 3 in the ANPA hierarchy should be estimated to take into account the corrections Kilmister made to the estimate of 137, but we do not know what that result should be, except by following the method here. The strong interaction has a parameter describing the interaction between two charged particles very similar to that associated with the term 137 for an electromagnetic interaction. The experimental value found is about 9.01 where the last digit is uncertain. Again, the term 10 in the ANPA hierarchy should be estimated more accurately and again we do not know what the result should be, except by following the method here.

We shall calculate all the terms of the ANPA hierarchy anew using the scaling combinatorial hierarchies in which each level describes a circular motion that have been found by us to be an intrinsic part of quantum electrodynamics and general relativity [Bell et al. 2000], [Bell et al. 2004a&b], [Bell and Diaz 2002], [Bell 2004], [Bell and Diaz 2003], [Bell and Diaz 2004a&b], [Bell and Diaz 2005]. We will also show how electromagnetism and the weak and strong interactions emerge from gravity and calculate sundry miscellaneous arbitrary parameters that arise in the present account of the forces of nature. This is a trilogy of papers of which this one, the first, discusses gravity and electromagnetism.

1.2 Gravitational Hierarchies

We showed that the theory of quantum electrodynamics could be derived by assuming that in an electrical field each point in space-time held a miniature circular motion or vortex, described, in the appropriate space, by Bohr's equations [Bell et al. 2000], [Bell et al. 2004a&b]. Here the spin of the electron had an integer value and the vortex was composed of an interaction between two point particles mediated by the inverse square law with the position of one particle, the source, fixed. The general case could then be found by putting the vortices on a lattice and showing that the result was always an interaction obeying the Maxwell and Dirac equations [Bell et al. 2004a&b], the fundamental building blocks of QED. We then showed that general relativity could be seen as the classical limit of QED theory with the charge of a body replaced by its mass [Bell and Diaz 2004a], [Bell 2004]. This required us to assume the existence of gravitational hierarchies describing gravitationally induced circular motions, the same vortices as appeared for QED theory in the electromagnetic case. Empty space could also be described by such a hierarchy, *the unishell*, which was

Table I(a)

Part of the Unishell Hierarchy

Level	-1	0	1	2	3
Velocity Squared	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	1	2
Source Mass	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4
Approx. Energy	$\sqrt{\frac{7}{8}}$	$\sqrt{\frac{3}{4}}$	$\frac{1}{\sqrt{2}}$	0	$\sqrt{-1}$

Although the hierarchy describes empty space, it can still be seen as always accompanied by the rotation of one body round another. The rotating body has unit mass while the mass of the stationary body, which we shall call the source, varies. The first row labels the level in the hierarchy, which we will always call r, the second the square of the orbital velocity of the circling body, v_r^2, with one equal to the velocity of light, and the third row describes how the mass of the source varies with level if we assume it is one at the first level. The final row is approximately the amount of energy that must be applied to pull the circular motion apart so that the orbiting body escapes. This is the unrenormalised energy. We will discuss renormalisation below. Doing this would imply the disintegration of space itself. The first level describes our own position on the outside of a hollow spherical body, the part of the hierarchy which we call the unishell. In our original derivation [Bell and Diaz 2002], [Bell 2004] the levels were formed from concentric thin spherical shells obeying Einstein's theory of gravity, general relativity. We went up or down a level in a notional way by passing through a thin shell. Each level could be described either in the language of curved space-time or in the language of tides in a moving fluid.

The gravitational interaction of any two massive point bodies can be seen as a direct interaction between them, which we may describe by another table similar to table I(a), plus another interaction that the orbiting body has with the unishell, or space-time itself, which always occurs and obeys table I(a). The gravitational interaction of the two massive bodies can also be seen as forming part of a single hierarchy, again describable by a table similar to I(a), at which point embarrassing difficulties with velocities greater than the velocity of light disappear. Table I(a) should therefore be seen as extending infinitely downwards and also infinitely upwards.

1.3 Electromagnetic Hierarchies

We then showed that the gravitational hierarchy was replaced by an electromagnetic hierarchy of similar type for electromagnetic interactions, which we called *the electroshell* [Bell and Diaz 2004a&b]. The electromagnetic hierarchy first became apparent when we considered renormalisation, the process which provides the most accurate solution of the equations of QED theory. It was possible to renormalise the interaction described by Bohr's equations and then apply the renormalised solution to general interactions. This amounted to accepting a small correction to Maxwell's equations, which determine the electromagnetic field given the distribution of charges. When we renormalised the vortex, we found the results were described by the electroshell hierarchy,

Table I(b)

Part of the Electroshell Hierarchy

Level	-1	0	1	2	3
Velocity Squared	$\frac{1}{\chi^6}$	$\frac{1}{\chi^4}$	$\frac{1}{\chi^2}$	1	χ^2
Source Charge	$\frac{1}{\chi^4}$	$\frac{1}{\chi^2}$	1	χ^2	χ^4
Approx. Energy	$\sqrt{1-\frac{1}{\chi^6}}$	$\sqrt{1-\frac{1}{\chi^4}}$	$\sqrt{1-\frac{1}{\chi^2}}$	0	$\sqrt{1-\chi^2}$

where the meaning of entries in the table are the same as those we have already discussed for table I(a), except that the mass of the source is replaced by the charge on the source in the rest frame of the orbiting particle. The hierarchy also embraces the charge of the orbiting particle, but we shall not need it here [Bell and Diaz 2004a]. We have chosen a possible but simplified version of the hierarchy given by Bell and Diaz so that the similarity to the unishell is immediately apparent. It is as if the electromagnetic hierarchy described a space-time, other than ours, in which $1/\sqrt{2}$ was replaced by $-1/\chi$. We conjecture that it is the space-time seen by the electromagnetic particle in a bound state with a charged source. We have already constructed this space and described its properties [Bell et al. 2000], [Bell et al. 2004a&b].

1.4 Discrimination and Aggregation for Both Hierarchies

We comment on discrimination, which plays a vital role in the construction of the ANPA hierarchy. In so far as discrimination appears in the gravitational hierarchies, it is the question of whether we are inside or outside a thin spherical shell that counts, although we can generate this simple hierarchy many other ways, and we will illustrate one below, where the discrimination is between directions on a compass. Since the electromagnetic hierarchy is so similar to the gravitational one, discrimination could be defined as the same inside or outside decision or directions on a compass again. However, we would approach current ideas on electromagnetism more closely if we took it to be distinguishing between the number of interactions the orbiting charged particle has with the field generated by the source. This process of multiple interactions is described by renormalisation, which we will discuss in more detail.

The renormalised energy of the interaction at level r, R_r, was given in terms of the unrenormalised energy of the interaction, E_r, according to the following formula [Bell and Diaz 2004a],

$$R_r = \frac{E_r}{1+v_r^2}\left(1+v_r^2+v_{r-1}^2+v_{r-2}^2+v_{r-3}^2\ \ldots\right) \tag{1.2.A}$$

where v_r is the velocity of the orbiting particle at level r, and we have taken the simplest case, where every term is positive. This corresponds to a repulsive interaction [Bell and Diaz 2004a&b]. For an attractive interaction we would have instead,

$$R_r' = \frac{E_r}{1-v_r^2}\left(1-v_r^2+v_{r-1}^2-v_{r-2}^2+v_{r-3}^2\ \ldots\right) \tag{1.2.B}$$

and, if we wanted to show this as a simple sum over all the relevant levels, we would have to amend table I(a), although we could accommodate table I(b) by making χ imaginary. We will continue to use the simpler series (1.2.A). The unrenormalised kinetic energy is then given by $E_r/\left(1+v_r^2\right)$ in equation (1.2.A) in the frame of the source. In the rest frame of the orbiting body, it is always given by the mass of the particle, which is one in our case. If we suppose that the energy associated with the interaction is reducible to a statement about the mass of the orbiting body, then the renormalised mass of this body is given by,

$$m_r = 1+v_r^2+v_{r-1}^2+v_{r-2}^2+v_{r-3}^2\ \ldots \tag{1.2.C}$$

Thus the increase in energy of the orbiting particle, Δm_r, at level r is given by a summation of the velocity squared over the columns in the table, from level r downwards,

$$\Delta m_r = v_r^2+v_{r-1}^2+v_{r-2}^2+v_{r-3}^2\ \ldots \tag{1.2.D}$$

We call Δm_r *the renormalisation energy*. For the particle at level one of the unishell the renormalisation energy is

$$\Delta m_r = \frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots=\frac{1}{2}\times\left(1+\frac{1}{2}+\frac{1}{4}+\ldots\right)=1 \qquad (1.2.E)$$

where we have used the result that,

$$1+x+x^2+x^3+\ldots=\frac{1}{1-x} \qquad (1.2.F)$$

This converges unless $x=1$. If $|x|>1$, it corresponds to the P-adic norm [Koc 2002]. In so far as aggregation appears in the gravitational or electromagnetic hierarchy, it is this operation of summing a variable given in the table. We are going to show that we can generate the value of the coupling constants for electromagnetism, the weak interaction and the strong interaction using these simple operations of discrimination and aggregation on the table for gravity, table I(a). In this paper we derive the fine structure constant used in table I(b), $1/\chi$, to an accuracy of about six figures.

2. GENESIS OF THE HIERARCHIES

2.1 Genesis for Gravity, Velocities Less than One

We have already described how table I(a) arose historically, but many derivations could probably be given. We now provide another possible derivation by way of example. Let us consider just one dimension, time, which we will imagine as a vertical straight line. Initially, we suppose that we travel up it, towards the future, with a probability of one. Our series commences with this probability. Then we relax our condition and suppose that we may also travel down it, with an equal probability. We assign instances to these possibilities, one representing travelling up and one down. The series reads 1, 2. Let us add a second dimension, space, which we will imagine as a horizontal straight line intersecting the time line at our current time. We assume we may travel right or left along this line with equal probability as well, giving a total of four instances. These correspond to travelling forwards or backwards in time or travelling with an infinite velocity either left or right. The series reads 1, 2, 4. We might obtain our spatial dimension by bisecting the angle, 180° or π radians, made by the time line at our current time, the origin. If we bisected this angle again we would add a pair of lines, crossing at the origin, at 45° or $\pi/4$ radians to the vertical and horizontal. If we follow these we always travel equal distances along space and time, giving a velocity of one. This adds diagonal travel upwards to the right or left and downwards to the right or left giving eight instances, all with equal probability. The series giving the number of instances now reads: 1, 2, 4, 8. However, the mind loves repetition, and we continue the algorithm by adding four lines meeting at the origin so that all the previous angles are divided, leading to a series 1, 2, 4, 8, 16. This time we do not bisect the angles for a reason we shall come to shortly. This division of angles continues indefinitely, leading to the following picture,

Table II(a)

Genesis of the Gravitational Hierarchy:

Velocities Less than One

P_r	Slots 1	Slots 1/2	Slots 1/4	Slots 1/8	Slots 1/16	Slots 1/32
↑						↑
↑						*e*
1/16					*d*	*e*
					d	*e*
					d	*e*
					d	*e*
					d	*e*
					d	*e*
					d	*e*
					d	*e*
1/8				*c*	*d*	*e*
				c	*d*	*e*
				c	*d*	*e*
				c	*d*	*e*
1/4			*b*	*c*	*d*	*e*
			b	*c*	*d*	*e*
1/2		*a*	*b*	*c*	*d*	*e*
1	α	*a*	*b*	*c*	*d*	*e*
2	✧	✧	✧	✧	✧	✧
↓	✧	✧	✧	✧	✧	✧

Throughout, the four-pointed stars indicate unused cells of the table. Here, the first column, P_r, corresponds to the probability at level *r*, where for the moment we omit entries greater than one. The other columns, called slots, contain the instances for each possible of behaviour. We start with the entry of one in the first column which has just one instance in the first table of slots, labelled α, which corresponds to travelling forward in time, which is what we observe we and others do. We assign an energy of one to bodies going forward in time. However, more liberally, we introduce movement backwards in time. We see that the two instances in the column of slots labelled $1/2$ are called a, a. We assign an energy of zero for bodies going backwards in time because it is not our normal experience. We suppose that we are only conscious when we travel forwards in time but that we go backwards in time as often as we go forwards. If we choose our direction randomly, there is a probability of just $1/2$ that we will choose to go forward, which is the entry in the first column. Considering this as a one-dimensional random walk, our average energy will also be a half, the same as this entry in our table. Our next liberalisation is to allow horizontal movement across space, permitting others to exist separately from ourselves. There are four instances, *b*, *b*, *b*, *b*, for the column of slots labelled $1/4$. This allows us or others to travel horizontally to the left or right but again this does not occur in the usual mundane experience and we assign an energy of zero. We suppose that although we are not conscious on such occasions, they occur with the same frequency as backward or forward movement in time. If we choose our direction randomly, there is a probability of just $1/4$ that we will choose to go forward, leading to an average energy of the same value, the entry in the first column.

The infinite velocity that we have introduced at this stage is the velocity of *celestial radiation* [Bell and Diaz 2003]. This radiation is similar to the ordinary variety, for example light or gravitational waves, which we call *terrestrial radiation* to distinguish them. The equations which describe electricity and gravity allow, not only radiation travelling at a universal velocity of one, but radiation travelling infinitely fast between mutually stationary objects, the celestial radiation. We can retrieve terrestrial radiation from its celestial counterpart or vice versa by noticing the following question. Suppose you have a velocity with respect to me, then as Einstein averred [Einstein 1967], you and I have separate times, but the question is, which is your time and which mine? Make one choice, and the radiation becomes celestial. Make the other and it is terrestrial. Let the first choice correspond to celestial radiation and the lines we already have, the vertical and the horizontal, then the latter represents an infinite velocity choosing time in the familiar way, but a velocity of one if you exchange your time for ours. If we allow only the second choice, a velocity of one only appears if we add a pair of lines, crossing at the origin, at 45° or $\pi/4$ radians to the vertical and horizontal. This adds four new diagonal directions whose instances, *e*, *e*, *e*, *e*, *e*, *e*, *e*, *e* appear in the column of slots for $1/8$. Following arguments similar to our previous ones, these correspond to the probability or energy of $1/8$, the entry in the first column. However, if we make the celestial choice, travel along a diagonal corresponds to a velocity of $1/\sqrt{2}$. We see that the celestial velocity squared is proportional to the probability for probabilities $1/4$ and $1/8$, with four as the constant of proportionality, from equations (1.2.D) and (1.2.F). We may obtain an exact fit by renormalising our probability twice, using equation (1.2.D) with the probability replacing the velocity squared,

$$
\begin{aligned}
Q_r &= P_r^2 + P_{r-1}^2 + P_{r-2}^2 + P_{r-3}^2 \;\ldots, \\
v_r^2 &= Q_r^2 + Q_{r-1}^2 + Q_{r-2}^2 + Q_{r-3}^2 \;\ldots
\end{aligned}
\tag{2.1.A}
$$

where v_r^2 then corresponds to the velocity squared entry in table I(a). If we perform one further renormalisation, that in equation (1.2.D), we obtain the renormalisation energy. We keep this nice feature for smaller probabilities or energies. We therefore have lines at an angle θ_r to our original vertical time line for level *r* whose sine squared is given by,

$$
\begin{aligned}
&\sin^2 0 = 0,\ \theta_4 = 0, \\
&\sin^2 \pi = 0,\ \theta_3 = \pi, \\
&\sin^2 \frac{\pi}{2} = 1,\ \theta_2 = \frac{\pi}{2}, \quad + \text{rotations through } \pi = \theta_3 \\
&\sin^2 \theta_1 = \frac{1}{2}, \quad + \text{rotations through } \frac{\pi}{2} = \theta_2 \text{ and } \theta_3 \\
&\sin^2 \theta_0 = \frac{1}{4}, \quad + \text{rotations through } \frac{\pi}{4} = \theta_1, \theta_2 \text{ and } \theta_3 \\
&\sin^2 \theta_{-1} = \frac{1}{8}, \quad + \text{rotations through } \theta_0, \theta_1, \theta_2 \text{ and } \theta_3 \\
&\text{etc.}
\end{aligned}
\tag{2.1.B}
$$

In each case the number of rotations is chosen so that the total number of lines increases by a factor of two from the previous level. The operation of multiplying by all the previous angles in the series may be seen as choosing each line drawn as the time line for the object of interest.

2.2 Genesis for Gravity, Velocities Greater than One

The process of choosing a new time line from those drawn can be said to start with the exchange of spatial lines with temporal ones when the minimum angle is $\theta_2 = \pi/2$. Permitting this inverts the velocity, and so we may add rows corresponding to integer velocities of more than one. We showed earlier [Bell and Diaz 2004a] that such velocities could be interpreted. Since our original energies approach zero, their inverses approach infinity. We obtain a mirror of table II(a), *The Scaling Fine Structure Constant*

Table II(b)

Genesis of the Gravitational Hierarchy:

Velocities Greater than One

$\frac{1}{P_r}$	Slots 1	Slots 2	Slots 4	Slots 8	Slots 16	Slots 32
↑	✧	✧	✧	✧	✧	✧
1	α	*a*	*b*	*c*	*d*	*e*
2		*a*	*b*	*c*	*d*	*e*
3			*b*	*c*	*d*	*e*
4			*b*	*c*	*d*	*e*
5				*c*	*d*	*e*
6				*c*	*d*	*e*
7				*c*	*d*	*e*
8				*c*	*d*	*e*
9					*d*	*e*
10					*d*	*e*
11					*d*	*e*
12					*d*	*e*
13					*d*	*e*
14					*d*	*e*
15					*d*	*e*
16					*d*	*e*
17						*e*
↓						↓

We see that we can still interpret the entries in column $1/P_r$, shown in black, as probabilities on the model we used for table II(a) if we suppose that the probability is the inverse of the entry. We notice that the row with $P_r = 1$ is necessarily shared by tables II(a) and (b). The cells of the $1/P_r$ column in table II(b) can be mapped to the natural numbers, which are inserted in grey where they do not constitute entries. We thus suppose that the $1/P_r$ column itself may be treated as a one-dimensional space. We might wonder what other uses this table can be put to, since it seems to repeat much of the information carried in table II(a). We may use it to turn gravity into, successively, the weak, the strong and the electromagnetic interaction by supposing that the velocities greater than one correspond to a different physical shells, which may be much smaller than the unishell formed from time and space. In the case of electricity we get the electroshell and examples of this are to be found in the atom.

2.3 **Genesis for Electromagnetism: First Term**

To deduce the fine structure constant, $1/\chi$, that applies to table I(b), we form new tables, tables III(a) and (b) recursively from the tables II(a) and (b) that apply for gravity. The analogue of table II(a) is

Table III(a)

Genesis of the Electromagnetic Hierarchy: Second Term

P_r	Q_r	v_r^2	**Product Weights**	**Slots 1/2**	**Slots 1/3**	**Slots 1/7**	**Slots 1/127**
$\sim 5.9\times 10^{-39}$	$\sim 1.2\times 10^{-38}$	$\sim 2.3\times 10^{-38}$	✧	✧	✧	✧	127
↑	↑	↑	✧	✧	✧	✧	↑
1/256	1/128	1/64	✧	✧	✧	✧	*d*
1/128	1/64	1/32	1/4096			*c*	*d*
1/64	1/32	1/16				*c*	*d*
1/32	1/16	1/8				*c*	*d*
1/16	1/8	1/4				*c*	*d*
1/8	1/4	1/2	1/32		*b*	*c*	*d*
1/4	1/2	1	1/4	*a*	*b*	*c*	*d*
1/2	1	2		*a*	*b*	*c*	*d*
1	2	✧	✧	✧	✧	✧	✧
↓	↓	✧	✧	✧	✧	✧	✧

We shall use this table to calculate our second estimate of χ, which is why the title was chosen. The first column corresponds to the entries in the first column of table II(a), while the second is obtained by renormalising P_r as we did in equation (2.1.A). The third column is obtained by renormalising Q_r, again as we did in equation (2.1.A), the result being the velocity squared, v_r^2, of the particle, corresponding to the second row in table I(a). The fourth column is the rolling product of the entries in the P_r columns, $1/4$, $1/8$ and $1/128$, each marked by the end of the instances in the following three columns, called slots. These instances are shown for probabilities, $1/2$, $1/3$, $1/7$ and, schematically only, $1/127$, which is not discussed subsequently. The reason why we have chosen these probabilities will appear later, when we discuss our first estimate of χ. These last four columns function in the same way as the last six columns of table II(a). The important point is that we use the entries in table II(a) as positions to show instances in table III(a). We are therefore copying recursively from the earlier table. The double lines indicate where our series will begin to repeat, in a way we shall describe later. We notice that we do not include $P_r = 1$ inside the area marked out. We need to exclude it, because our renormalisation formula, equation (1.2.F), gives an infinite result if $x = 1$. The velocity squared is zero and we cannot derive it by renormalising the probability twice. Interestingly, Parker-Rhodes excluded zero in his original calculation as well [Kilmister 2004b]. Our next table will allow us to move from gravity to electromagnetism at its crudest,

Table III(b)

Genesis of the Electromagnetic Hierarchy: First Term

$\frac{1}{P_r}$	$\frac{1}{Q_r}$	$\frac{1}{v_r^2}$	**Product Weights**	**Slots 1/2**	**Slots 1/3**	**Slots 1/7**	**Slots 1/127**
↑	↑	↑	✧	✧	✧	✧	✧
1/2	1	2	✧	✧	✧	✧	✧
1	✧	✧	✧	*j*	*k*	*l*	*m*
2	3	3	2	*j*	*k*	*l*	*m*
4	7	10	8		*k*	*l*	*m*
8						*l*	*m*
16						*l*	*m*
32						*l*	*m*
64	127	137	512			*l*	*m*
128	✧	✧	✧	✧	✧	✧	*m*
↓	✧	✧	✧	✧	✧	✧	↓
$\sim 8.5\times 10^{37}$	$\sim 1.7\times 10^{38}$	$\sim 1.7\times 10^{38}$	✧	✧	✧	✧	127

Here the number line in the $1/P_r$ column of table II(b) has become logarithmic, like the numbers on a slide rule, the instrument now ignored in favour of the calculator. Apart from this, the column headed $1/P_r$ is superficially the same as the column headed $1/P_r$ in table II(b), but appearances are deceptive. The first entry in the for $1/P_r$ after the double line here is the renormalised value of all the preceding values of P_r in table

III(a), $1/2+1/4+1/8+\ldots$etc. and the entry for $P_r=1$ itself is missing from both tables, as we discussed earlier. The other entries in $1/P_r$ here, $1/P_r=2,4,8$ etc. are the same as those in table II(b). Thus, when we renormalise $1/P_r$ and put the result in $1/Q_r$ we get one less than we would have done if $P_r=1$ were present. Not every value of $1/P_r$ corresponds to an entry in $1/Q_r$ because table III(b) is a recursive copy of table II(b). The entries there are used as positions to show instances here. The instances are given in the last four columns in the same way as they were for table II(b), except that we suppose that here we also renormalise to obtain something new, so that the instances now correspond to the value of $1/Q_r$ rather than the value of $1/P_r$. We choose them as follows. Excluding the entry $1/P_r=1$ since this actually refers to the earlier part of the hierarchy, the first position we can choose is $1/P_r=2$, pointed out by the end of the two instances, j, j. The renormalised value of the $1/P_r$ up to that point is $1/Q_r=3$ and so we will require three instances, the *k*, *k*, *k* in the slots for probability $1/3$. These end with $1/Q_r=7$ and we require seven instances for this, the *l* shown in the slot column for probability $1/7$, landing us on $1/Q_r=127$. We require 127 instances for this, shown only schematically here, and land on the very large number, $1/Q_r=2^{127}-1\approx 1.7\times 10^{38}$. Following positions correspond to even larger numbers and are not shown at all. So the series in $1/Q_r$ is

$$1/Q_r=3,7,127,\ldots n,2^n-1,2^{2^n-1}-1,\ldots \qquad (2.3.A)$$

This series was communicated to us but we do not know its originator. $1/Q_r$ is then renormalised to form the $1/v_r^2$, the first four values being the members of the original series found by Parker-Rhodes. We shall discuss the first three in the trilogy of papers, but in this paper we shall discuss just the last. We see that this is a number tantalisingly close to the inverse of the fine structure constant, χ, 137.

Before we can take this as being an estimate of the value of the fine structure constant, we must consider the implications of it being a velocity in table I(b), while we expect a velocity squared. When we addressed the concentric hollow spheres with general relativity [Bell and Diaz 2004a], we found that the latter suppressed a tachyon entry in the hierarchy, between each two levels. This would also need to be the case here for electromagnetism. Between the levels of the hierarchy that we found we required for the renormalisation sequence, shown in table I(b), there would have to be intermediate levels. We would need to add a hierarchy,

Table IV

Intermediate Levels for Table I(b)

Level	$-3/2$	$-1/2$	$3/2$	$5/2$	$7/2$
Velocity Squared	$\frac{1}{\chi^5}$	$\frac{1}{\chi^3}$	$\frac{1}{\chi}$	χ	χ^3
Source Mass	$\frac{1}{\chi^3}$	$\frac{1}{\chi}$	χ	χ^3	χ^5
Energy	$\sqrt{1-\frac{1}{\chi^5}}$	$\sqrt{1-\frac{1}{\chi^3}}$	$\sqrt{1-\frac{1}{\chi}}$	$\sqrt{1-\chi}$	$\sqrt{1-\chi^3}$

This allows us to take χ as a velocity squared, and we may make, with ANPA, our first estimate of χ,

$$\chi_1 = 137 \qquad (2.3.B)$$

Lastly, the "Product Weights" column in table III(b) shows the running product of weights for the entries as it did in table III(a). We will discuss the weights shortly.

2.4 Genesis for Electromagnetism: Second Term

There is a difficulty with supposing that χ_1 is the inverse of the fine structure constant exactly. Table I(a) for gravity, I(b) for electromagnetism and tables II(a), II (b) and V are all scaling. If one chooses a position corresponding to an entry in the hierarchy it does not make any difference which one is picked, provided we have algorithms that depend on relative positions in the tables as our physical laws. Size becomes relative and not absolute. χ_1 does not follow a scaling law. We suppose that the true inverse of the fine structure constant must also be scaling and use this to calculate a more accurate value. That is, we suppose that table III(b) between the double lines forms a pattern of instances repeated at any scale of size, whether smaller or larger. This pattern is

$$1/Q_r \equiv .--...- \qquad (2.4.A)$$

where the dots correspond to positions where there is no entry and the dashes correspond to positions where there is an entry. Scanning in the same direction, we see that table III(a) forms a mirror image of the pattern in the slots columns in (b),

$$1/Q_r \equiv -\ldots--. \qquad (2.4.B)$$

However, we get exactly the same product of weights if we reflect the pattern in the slots columns in table III(a) to give,

Table V(a)

Genesis of the Electromagnetic Hierarchy: Second Term

P_r	Q_r	v_r^2	**Product Weights**	**Slots** 1/2	**Slots** 1/3	**Slots** 1/7
↑	↑	↑	✧	✧	✧	✧
1/256	1/128	1/64	✧	✧	✧	✧
1/128	1/64	1/32		*a*	*b*	*c*
1/64	1/32	1/16	1/64	*a*	*b*	*c*
1/32	1/16	1/8	1/2048		*b*	*c*
1/16	1/8	1/4				*c*
1/8	1/4	1/2				*c*
1/4	1/2	1				*c*
1/2	1	2	1/4096			*c*
1	2	✧	✧	✧	✧	✧
↓	↓	✧	✧	✧	✧	✧

This is sufficient for our purpose and does reproduce the pattern in (2.4.A). We calculate an improved estimate of the fine structure constant, using this table. The "Product Weights" column should now be read from top to bottom. We see there are entries for $1/64$, $1/32$ and $1/2$ in the P_r column. We treat these as probabilities and calculate that the probability of having all three entries is

$$p' = \frac{1}{64} \times \frac{1}{32} \times \frac{1}{2} = \frac{1}{4096} \qquad (2.4.C)$$

the final value in the "Product Weights" column. Converting this into an average energy in the same way as we did for table II(a), we obtain the same correction as Kilmister [2004b],

$$\frac{\chi_1}{4096} \approx 0.033447 \qquad (2.4.D)$$

giving for χ,

$$\chi_2' = \chi_1 + \frac{\chi_1}{4096} \approx 137.033447 \qquad (2.4.E)$$

Clearly, there will be further terms beyond this as we consider smaller gravitational velocities. We may calculate what the weights will be for smaller terms by looking at the amount they diminished in going from table III(b) to table III(a), using the "Product Weights" columns. Since they amount to a factor of 512 in table III(b) and 4096 in (a), we obtain,

$$p = \frac{1}{512 \times 4096} = \frac{1}{2097152} \qquad (2.4.F)$$

Clearly, terms in this series converge very rapidly. We change our statistical model slightly by assuming that, if all the weights were simply one, we should end with a fine structure constant of the form,

$$\chi = \chi_1 + \chi_1/128 + \chi_1/(128)^2 + \chi_1/(128)^3 + \ldots \quad (2.4.G)$$
$$= \frac{\chi_1}{1 - 1/128}$$

summing using equation (1.2.F). Summing the series in powers of p as well,

$$\chi_2 = \chi_1 + \frac{\chi_1}{4096 \times (1 - 1/128) \times (1 - p)} \approx 137.033711 \quad (2.4.H)$$

2.5 Genesis for Electromagnetism: Third Term

We turn to the other direction in which we may traverse our hierarchy, the appearance of the pattern given in relation (2.4.A) above the part between the double lines described in table III(b). This is

Table V(b)

Genesis of the Electromagnetic Hierarchy: Third Term

$\frac{1}{P_r}$	$\frac{1}{Q_r}$	**Product Weights**	**Product Weights**	**Slots 1/2**	**Slots 1/3**	**Slots 1/7**
↑	✧	✧	✧	✧	✧	✧
64	✧	✧	✧	✧	✧	✧
128	✧	✧	✧	j	k	l
256	384	256	256	j	k	l
512	896	256× 512	131072		k	l
1024						l
2048						l
4096						l
8192	16256	256× 512× 8192	1073741824			l
16384	✧	✧	✧	✧	✧	✧
↓	✧	✧	✧	✧	✧	✧

Looking at the 'Product Weights" column, we find the total probability for entries in the pattern given in relation (2.4.A), which is

$$256\times512\times8192=1073741824=\frac{512}{p} \qquad (2.5.A)$$

we see that the product of weights has increased by $1/p$ from the product we obtained in the previous instance of the sequence in table III(b).

We suppose, just as we did for gravity in table II(b) and for the construction of table III(b), that these values represent inverse probabilities. However, before we can evaluate the series we must consider the position of the point associated with the base 128. In table III(b) it is just before the entry "one" in the $1/P_r$ column. When we found the contribution of the sequence in table V(a), it had moved to a position before the "$1/128$" in the P_r column of that table. Here we imagine the direction of increasing P_r reversed so that $1/128$ and $1/2$ swap, which also rectifies the mirror image. This pattern continues and the point moves seven places further left for each term in the series. We want the point in table V(b) just before the start of the contribution, which begins with the entry "128" in the $1/P_r$ column. But, since we are treating the entries in table V(b) like the inverse of the probabilities, the binary point is located just before the $1/128$ entry in table V(a). Before evaluating the contribution of table V(b), we must move it through fourteen places to the right, which we can do by multiplying by 128×128. Thereafter, we shall need to move it an extra fourteen places to the right for each new term.

Choosing a model where we do not assume that the pattern in (2.4.A) pre-exists in this part of the hierarchy, the term described in table V(b) is

$$\chi_3' = \frac{128\times128\times p\chi_2}{512} \tag{2.5.B}$$

while, summing the series of terms generated in the upper part of the hierarchy gives our final estimate for χ,

$$\chi_3 = \chi_2 + \frac{128\times128\times p\chi_2}{512\times(1-128\times128p)} \approx 137.0358181 \tag{2.5.C}$$

Our estimate is accurate to six significant figures as we can see by comparing it with the current best value in the introduction.

3. CONCLUSION

We have used only a few simple rules which we had earlier shown apply to quantum electrodynamics and general relativity separately. We shall continue to use the same rules in our derivations of the scaling constants associated with the terms three and ten in the ANPA hierarchy, and a BBC Basic program generating all the first three terms in the hierarchy, together with its output, is given in the Appendix. Parker-Rhodes and his successors in ANPA found the same series from even simpler considerations. His derivation will continue to rest on simpler principles than ours. We feel that this is a pressing reason why his is to bc preferred. If we take his as primary and correct, then our derivation provides evidence of the likely validity of our viewpoint on quantum electrodynamics and general relativity. This becomes even more likely still if we reflect that we also found combinatorial hierarchies as an integral part to both theories. If we accept this picture, then the most important physical result is that the laws of physics appear to be scaling. Size becomes relative.

REFERENCES

Bell, S. B. M. 'Tilings and hierarchical tilings – a new approach to modelling," Revised version. Internal communication, NERC Unit for Thematic Information Systems, University of Reading (1992); http://www.sarahbell.org.uk

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. 'Classical behaviour of the Dirac bispinor ," *Found. Phys.* Vol 30 No 1 (2000) 35-57; arXiv:quant-ph/0011026

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. 'QED derived from the two-body interaction," *Found. Phys.* Vol 34 No 2 (2004a) 297-333; arXiv:quant-ph/0104017 & arXiv:quant-ph/0104018

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. 'The two-body interaction with a circle in time," *Found. Phys.* Vol 34 No 2 (2004b) 335-358; arXiv:quant-ph/0110162

Bell, S. B. M. and Diaz, B. M. 'Quantising general relativity using QED theory," To be published (2002); arXiv:physics/0211108

Bell, S. B. M. 'Quantising general relativity using QED theory, an overview and extension," Proc. Physical Interpretations of Relativity Theory Conference, London (2004); arXiv:physics/0401005

Bell, S. B. M. and Diaz, B. M. 'The relationship of time and space," To be published (2003); arXiv:physics/0403137

Bell, S. B. M. and Diaz, B. M. 'Renormalisation and hierarchies,"To be published (2004a); arXiv:physics/0406089

Bell, S. B. M. and Diaz, B. M. 'Interdisciplinary overview: physics and philosophy," to appear, Proc. 26th International Meeting of the Alternative Natural Philosophy Association (ed. Keith G. Bowden), Cambridge U.K. (2005); **http://www.sarahbell.org.uk**

Bell, S. B. M. and Diaz, B. M. 'Exact solutions for QED,"To be published (2004b);

Einstein, A. *The meaning of relativity* Sixth Edition, Chapman and Hall, London (1967)

Kilmister, C 'The Eddington heritage - a claim for ANPA," in (ed. Keith G. Bowden) *Spin: Proc. Alternative Natural Philosophy Association* Vol 25 (2004a) 341-350

Kilmister, C 'The fine structure constant," in (ed. Keith G. Bowden) *Spin: Proc. Alternative Natural Philosophy Association* Vol 25 (2004b) 1-18

Koc, C. K. "A Tutorial on P-adic Arithmetic," Oregon State University Technical Report, April (2002); **http://islab.oregonstate.edu/papers/r09padic.pdf**

Manthey, M. "ANPA - the Alternative Natural Philosophy Association," *Cybernetics and Human Knowing* Vol 2 No 2 (1993); **http://www.imprint.co.uk/C&HK/vol2/v2-2mm.htm**

Appendix

The program below calculates the scaling version of all the first three members of the ANPA hierarchy. It is in R.T. Russell's BBC Basic, which is to be found at the following web address: http://www.rtrussell.co.uk/.

```
10 REM: FINE - Calculates the fine structure constant on
the basis it is scaling
20 REM: Copyright (C) 2004 by S B M Bell
   30 *FLOAT 64
   40 REM Use double precision for all calculations
   50 @%=&000F0A
   60 REM Controls format of printed numbers
   70 REM
   80 REM We imagine an origin with four straight lines
going diagonally upward.
   90 REM The leftmost belongs to the strong interaction.
The next, towards the right
  100 REM is our time. The next belongs to the
electromagnetic interaction, while the fourth
  110 REM belongs to the electro-weak interaction. Then the
sine of the angle between
  120 REM the interaction in question and our time gives
the first three terms of the
  130 REM ANPA series.
  140
  150 REM Defines the series for electromagnetism and the
fine structure constant
  160
  170
FFNB=137:SMLPR=4096:LRGPR=4096*512*512:P=4096*512:RNGE=128
  180 REM FFNB is whole number estimate of coupling
constant
  190 REM SMLPR is the starting probability of lower series
  200 REM LRGPR is starting probability for upper series
  210 REM P is the amount the probability decreases between
adjacent terms of the series
  220 REM RNGE is the length of the each instance of the
repeating pattern
  230 PROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RSULTA,RSULT1)
  240
  250 REM Defines the series for the electro-weak
interaction and the Weinberg angle
```

```
260
270 FFNB=3:SMLPR=4:LRGPR=8:P=4:RNGE=4
280 PROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RSULTA,RSULT2)
290 PRINT"We assume that the symmetry breaks before the upper part of series"
300 PRINT"Then the coupling constant for lower series is the Weinberg angle squared "1/RSULTA
310 PRINT
320
330 REM Defines the series for quantum chromodynamics and the strong coupling constant.
340
350 FFNB=-10:SMLPR=32:LRGPR=16*32:P=16*32/2/8:RNGE=8
360 PROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RSULTA,RSULT3)
370
380 END
390
400 REM Calculate the series
410 DEFPROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RETURN RSULTA,RETURN RSULT)
420 REM Starting point for geometric series for numbers less than coupling constant
430 FFNSML=ABS((FFNB/(1-1/RNGE)/SMLPR)*(1/(1-1/P)))
440 REM Sum of the geometric series for numbers less than one
450 RSULTA=FFNSML+FFNB
460 PRINT"Contribution of lower series for "STR$(FFNB)": "RSULTA'" Inverse "1/RSULTA
470 FFNBG=ABS((RSULTA*RNGE*RNGE/LRGPR)*(1/(1-RNGE*RNGE/P)))
480 REM Sum the geometric series for numbers greater than coupling constant
490 RSULTB=FFNBG
500 PRINT"Contribution of upper series for "STR$(FFNB)": "RSULTB'" Inverse "1/RSULTB
510 REM Add upper and lower series
520 RSULT=RSULTA+RSULTB
530 PRINT"Total value of "STR$(FFNB)" constant "RSULT" Inverse "1/RSULT
540 PRINT
550 ENDPROC
560
```

Output of program:-

```
Contribution of lower series for 137:
137.033710645996 Inverse 0.00729747443374234
Contribution of upper series for 137: 0.00210743280398
Inverse 474.510977579662
Total value of 137 constant 137.0358180788
Inverse 0.00729736220806862

Contribution of lower series for 3: 4.33333333333333
Inverse 0.230769230769231
Contribution of upper series for 3: 2.88888888888889
Inverse 0.346153846153846
Total value of 3 constant 7.22222222222222
Inverse 0.138461538461539
We assume that the symmetry breaks before the upper part of
series. Then the coupling constant for lower series is the
Weinberg angle squared
0.230769230769231

Contribution of lower series for -10: -9.63133640552995
Inverse -0.103827751196172
Contribution of upper series for -10: 1.20391705069124
Inverse 0.830622009569378
Total value of -10 constant -8.42741935483871 Inverse -
0.11866028708134
```

THE ANPA COMBINATORIAL HIERARCHY

Paper Two (A): The Weak Interaction

S.B.M. Bell† and B.M. Diaz‡

Interdisciplinary Quantum Group, Department of Computer Science

The University of Liverpool, Chadwick Building, Peach Street

Liverpool, L69 7ZF, United Kingdom

ABSTRACT

Scaling combinatorial hierarchies in which each level describes a circular motion have been found by us to be an intrinsic part of quantum electrodynamics and general relativity. From them we derive the combinatorial hierarchy that is the raison d'être of the Alternative Natural Philosophy Association (ANPA) [Kilmister 2003a]. This is used to derive the following (1) the electromagnetic theory, responsible for binding atoms together, from the gravitational, including the derivation of an approximate value for the fine structure constant, then (2) the weak interaction, responsible for radioactivity, including an approximate value for the Weinberg angle and finally (3) the strong interaction, responsible for binding the nuclei of atoms together, including an approximate value for the

† Mail@sarahbell.org.uk

‡ B.M.Diaz@csc.liv.ac.uk

quark-gluon-quark coupling constant. We also derive sundry others of the arbitrary parameters in the Standard Model of particle physics. In this paper we discuss the weak interaction and the Electro-weak theory.

1. PRELIMINARIES

1.1 General Introduction to this Trilogy of Papers

This is the first part of the second of a trilogy of papers describing the ANPA combinatorial hierarchy. The ANPA combinatorial hierarchy is fascinating because it is a most unusual hierarchy derived from very simple criteria [Manthey 1993], [Kilmister 2003a&b], [Bastin and Kilmister 1995], [Parker-Rhodes 1981]. The hierarchy in its simplest form is 3, 10, 137 and $2^{127}+136 \approx 1.7\times10^{38}$, ending on the last term. It was such an unusual structure, based on the simplest hypothesis on discerning similarity and difference and also using a simple method of aggregation to consolidate each judgement, that many were sure it held some deep truth. It was suggested that 3 described the weak interaction, 10 the strong interaction, 137 to the electromagnetic interaction and $\sim 1.7\times10^{38}$ gravity. In the first paper we derived these numbers from the gravitational theory, a quantum version of general relativity, and refined the estimate of the 137 as the coupling constant for electromagnetism to provide another three figures of the observed value. This coupling constant is also known as the fine structure constant.

Appendix I contains the computer program, which calculates all three coupling constants, showing that the method is the same for each apart from the necessary differences discussed in our trilogy of papers.

1.2 Introduction to this Paper

In this paper we derive a coupling constant for the weak interaction by refining the term 3 in the ANPA hierarchy. From this we derive an estimate for the Weinberg angle. We discuss the weak interaction in general terms, providing a simple self-contained model based on the system of scaling spins we have used throughout [Bell et al. 2000], [Bell et al. 2004a&b], [Bell and Diaz 2002], [Bell 2004], [Bell and Diaz 2003], [Bell and Diaz 2004a&b], [Bell and Diaz 2005a], [Bell 2005b]. We also provide the some of the algebra required to show that the Electro-weak theory can be put into this form in a second self-contained account.

2. ESTIMATING THE WEINBERG ANGLE

2.1 Genesis of the Angle: First Term

We shall suppose the reader is familiar with the derivation of tables I(a) and (b) in our first paper [Bell 2005b] providing an initial estimate of the ANPA hierarchy in the form 3, 10, $137, 2^{127}+136 \approx 1.7\times10^{38}$. In it we also derived a more accurate version of the fine structure constant, the reciprocal of the third member, based on the suggestion that these constants must be scaling. Here we derive a more accurate version of the weak coupling constant, the reciprocal of η, the first member, again on the basis that it is scaling in the same way. To deduce the weak interaction coupling constant, $1/\eta$, we form tables I and II recursively from the tables that apply for gravity, tables II(a) and (b) in the first paper of the trilogy. Again, we suppose the reader is familiar with the latter. The analogue of table II(a) is

Table I

Genesis of the Electro-Weak Hierarchy: Second Term

P_r	Q_r	v_r^2	**Product Weights**	**Slots** 1/2	**Slots** 1/3	**Slots** 1/7	**Slots** 1/127
↑	↑	↑	✧	✧	✧	[7]	[127]
↑	↑	↑	✧	✧	✧	↑	↑
1/32	1/16	1/8	✧	✧	✧	*c*	*d*
1/16	1/8	1/4	1/16			*c*	*d*
1/8	1/4	1/2			*b*	*c*	*d*
1/4	1/2	1	1/4	*a*	*b*	*c*	*d*
1/2	1	2		*a*	*b*	*c*	*d*
1	2	✧	✧	✧	✧	✧	✧
↓	↓	✧	✧	✧	✧	✧	✧

We shall use this table to calculate our second estimate of η, which is why the title was chosen. The first column shows probabilities. These probabilities show the likelihood that the energy of the particle described is one, leading to them also describing the average energy of the particle. The second and third are obtained by renormalising P_r and then Q_r according to the prescription,

$$Q_r = P_r^2 + P_{r-1}^2 + P_{r-2}^2 + P_{r-3}^2 \ \ldots, \qquad (2.1.A)$$
$$v_r^2 = Q_r^2 + Q_{r-1}^2 + Q_{r-2}^2 + Q_{r-3}^2 \ \ldots$$

v_r is then the velocity of the particle. The fourth column in the previous paper was the rolling product of the relevant entries in the P_r column. In this case there is only one weight, $1/4$, because the coupling constant for the weak interaction comes from the first entry in the ANPA hierarchy alone. This entry is marked by the end of the instances, *a*, *a*, in the following column, called slots. These instances are shown for probabilities, $1/2$, $1/3$, and, schematically only, $1/7$ and $1/127$. We shall only discuss the entry for $1/2$ in this paper. The reason why we have chosen these probabilities will appear later, when we discuss our first estimate of η. The last four columns show possibilities or actual instances used in calculating the probabilities. Each entry in the P_r column represents a slot into which we may put an instance. The double lines indicate where our series will begin to repeat, in a way we shall describe later. Our next table will allow us to move from gravity to the weak interaction at its crudest,

Table II

Genesis of the Electro-weak Hierarchy: First Term

$\frac{1}{P_r}$	$\frac{1}{Q_r}$	$\frac{1}{v_r^2}$	**Product Weights**	**Slots 1/2**	**Slots 1/3**	**Slots 1/7**	**Slots 1/127**
↑	↑	↑	✧	✧	✧	✧	✧
1/2	1	2	✧	✧	✧	✧	✧
1	✧	✧	✧	*j*	*k*	*l*	*m*
2	3	3	2	*j*	*k*	*l*	*m*
4	7	10			*k*	*l*	*m*
8			8			*l*	*m*
16						*l*	*m*
↓	✧	✧	✧	✧	✧	↓	↓
	✧	✧	✧	✧	✧	7	127

The first entry in the for $1/P_r$ after the top double line here is the renormalised value of all the preceding values of P_r in table I, $1/2+1/4+1/8+\ldots$etc. and the entry for $P_r=1$ itself is missing from both tables. The other entries in $1/P_r$ here, $1/P_r=2,4,8$ etc. represent single probabilities. Thus, when we renormalise $1/P_r$ and put the result in $1/Q_r$ we get one less than we would have done if $P_r=1$ were present. Not every value of $1/P_r$ corresponds to an entry in $1/Q_r$ because the $1/P_r$ are used as slots to show instances in the slots columns. The end of the entries *j* in the slots column for probability $1/2$ marks the position for the first entry in the

ANPA hierarchy, the $1/2$ itself, referring to one of the probabilities relevant for the gravitational interaction. The entries k in the slots column for probability $1/3$ are the instances and the end of these marks the first entry in the ANPA hierarchy, 3, in the column headed $1/Q_r$. As previously, $1/Q_r$ is obtained by renormalising $1/P_r$ and $1/v_r^2$ by renormalising $1/Q_r$, as in equation (2.1.A) but with $1/P_r$, $1/Q_r$ and $1/v_r^2$ replacing their inverses. Lastly, the "Product Weights" column in table II shows the running product of weights for the entries as it did in table I(b) of the previous paper. This time there is only one weight. We will discuss the weights shortly.

We would like to take 3 in the column $1/v_r^2$ as our initial estimate of the coupling constant for the weak interaction. What this means is not so clear as in the case of the fine structure constant because the symmetry is broken in the Electro-weak theory and there is more than one contestant for the title of coupling constant. However, we will press on with refining our estimate by insisting that it should be scaling, and discuss this further issue in depth in section 3 once we have completed this task. We therefore take, as our initial estimate of the weak interaction coupling constant, the first entry in the ANPA hierarchy,

$$\eta_1 = 3 \qquad (2.1.B)$$

2.2 Genesis of the Angle: Second Term

η_1 does not follow a scaling law. We suppose that a better estimate will be provided by insisting on a scaling constant. That is, we suppose that table II between the first pair of double lines forms a pattern of instances to be repeated at any scale of size, whether smaller or larger. This pattern is

$$1/Q_r \equiv .- \qquad (2.2.A)$$

working down the page, where the dots correspond to positions where there is no entry and the dashes correspond to positions where there is an entry. We see that table I forms a mirror image of the pattern in the slots columns in table II,

$$1/Q_r \equiv -. \qquad (2.2.B)$$

working down the page. Unlike the case with the fine structure constant, we do not get exactly the same product of weights if we reflect the pattern in the slots columns in table I. This produces the wrong weight of a half rather than a quarter in the first iteration in table I. Thus, the weak interaction coupling constant is not such a perfectly scaling number as the fine structure constant.

We accept a reflection for the part of the hierarchy starting in table I and calculate an improved estimate of the weak interaction coupling constant, using table I. The "Product Weights" column between the initial double lines, which delimit the first occasion the sequence repeats, has just one entry, $1/4$. The probability that the pattern given in equation (2.2.B) is present is then,

$$p' = \frac{1}{4} \qquad (2.2.C)$$

Converting this into an average energy in the same way as we did in the earlier paper, taking this probability as multiplying the estimate of three, we obtain,

$$\frac{\eta_1}{4} = 0.75 \tag{2.2.D}$$

giving for η,

$$\eta_2' = \eta_1 + \frac{\eta_1}{4} = 3.75 \tag{2.2.E}$$

We may calculate what the weights will be for smaller terms by looking at the amount they diminished in table I going from the first sequence between the parallel lines, weights $1/4$, to the second, weights $1/16$, using the "Product Weights" columns,

$$p = \frac{1/16}{1/4} = \frac{1}{4} \tag{2.2.F}$$

Adding the contributions of the whole downwards series,

$$p'' = \frac{p}{1-p} = \frac{1}{3} \tag{2.2.G}$$

We change our statistical model slightly by assuming that, if all the weights were simply one, we should end with a coupling constant of the form,

$$\eta_1' = 3 + \frac{3}{4} + \frac{3}{4^2} + \frac{3}{4^3} + \ldots = \tag{2.2.H}$$

$$\eta_1 + \frac{\eta_1}{p} + \frac{\eta_1}{p^2} + \frac{\eta_1}{p^3} + \ldots = \frac{\eta_1}{1-p} = \frac{3}{1-1/4} = 4$$

since each the interval between the start of one instance of the pattern and the next is corresponds to a factor of 4. Then summing the complete series for the lower part of the hierarchy starting in table I, we obtain our second estimate, using equation (2.2.G),

$$\eta_2 = \eta_1 + \frac{\eta_1 p}{(1-p)^2} = \frac{3}{4\times(1-1/4)\times(1-1/4)} = \frac{13}{3} \approx 4.33333 \quad (2.2.I)$$

We therefore have for the inverse of η_2 and the weak interaction coupling constant, which we shall show may be taken to be the sine squared of the Weinberg angle, θ_w,

$$\sin^2\theta_w = \frac{1}{\eta_2} = \frac{3}{13} \approx 0.23077 \quad (2.2.J)$$

We turn to the other direction in which we may traverse our hierarchy, the appearance of the pattern given in relation (2.2.A) above the part between the first two sets of double lines described in table II. We shall not derive this in detail here and those who want to know every step of the method should look at the first and third papers in the trilogy. We quote the final answer. The increment to be added to η_2 is

$$\delta\eta_3 = \left|\frac{4\times4\times\eta_2}{8\times(1-4\times4\times p)}\right| = \frac{26}{9} \approx 2.88889 \quad (2.2.K)$$

where we have used a P-adic norm to sum the series over powers of $4\times4\times p$ [Koc 2002]. We do not add this to our original estimate because the weak interaction breaks the symmetry. We suppose that the coupling constant is approximately $1/\eta_2$ for the hierarchy up to the end of the first sequence between the top pair of double lines in table II and the symmetry is then broken to obtain the weak interaction as we know it. This of course begs the question of why we perceive a broken symmetry. We suggest it is because the weak interaction is not so perfect as the electromagnetic one for the reason we have already discussed: that it is necessary to reflect the pattern in equation (2.2.A) when we go from the whole number part in table II to table I. If we then also proceed upwards from the whole number contribution of 3 in table II to provide a second fractional contribution it would be necessary to reflect the pattern again, which is more complicated and so it is not done.

3. NUTS AND BOLTS OF THE WEAK INTERACTION

3.1 Description of the Weak Interaction

The theory on the weak interaction that forms part of the Standard Model of particle physics is called the Electro-weak theory, and what we say below is consistent with it, but interprets it in the light of the ANPA hierarchy and the scaling system of spins that Bell and Diaz have employed throughout. We sketch a description in terms of words and geometry in this section and then show one way that we may apply the scaling hierarchy of spins using the algebra usually associated with the Electro-weak interaction [Cottingham and Greenwood 1998], [Aitchison and Hey 2004].

The unbroken weak interaction without further additions behaves just like electromagnetism except that the sine squared of the Weinberg angle rather than the fine structure constant describes the strength. Thus we could consider a particle like the electron but with a charge that corresponded in strength to the weak interaction coupling constant. We will call it *the weakotron* and we will discover an example of it in part B, the second part, of this paper [Bell 2005c]. The weak interaction could potentially form an atom just like the hydrogen atom, supplied with some suitable source. We will call this, and some other similar bound states we shall encounter, *the weak atom.* Associated with it there would be a signal like light, or its celestial counterpart, performing the same role and described in similar terms [Bell et al. 2000], [Bell and Diaz 2003]. We shall call this *weak radiation* or *the weak signal.* Just as we have an hierarchy spinning off from the ANPA hierarchy for electromagnetism, shown in table I(b) in the first paper in our trilogy, so is there one for the weak interaction with η instead of χ. We will call this *the weakoshell hierarchy.* We add the intermediate levels described in table IV in the first paper, which must apply in the same way to the weakoshell, so that,

Table III

Part of the Weakoshell Hierarchy

Level	-1	$-1/2$	0	$1/2$	1	$3/2$	2	$5/2$
Velocity Squared	$\frac{1}{\eta^6}$	$\frac{1}{\eta^5}$	$\frac{1}{\eta^4}$	$\frac{1}{\eta^3}$	$\frac{1}{\eta^2}$	$\frac{1}{\eta}$	1	η
Source Charge	$\frac{1}{\eta^4}$	$\frac{1}{\eta^3}$	$\frac{1}{\eta^2}$	$\frac{1}{\eta}$	1	η	η^2	η^3

Associated with the weakotron and the electron there is a new copy of space, the space of the ANPA hierarchy, table II, which we shall call *the ANPA space.* Our description will concentrate on ANPA space rather than describing the structure of the Bohr atoms which arise as a consequence. These atoms obey Bohr's equations as we described earlier [Bell et al. 2004a&b]. We discuss Bohr's equations elsewhere as well Bell and Diaz [2002] and Bell and Diaz [2004a]. A Lorentz transformation in the ANPA space would take us from gravity, represented by our own time line, which we may imagine as going vertically upwards towards U from the origin, O, to electromagnetism, represented by a time line through the origin to E, inclined at an angle given by the sine of the fine structure constant to OU, through the a line at a larger angle to corresponding to the strong interaction towards S, to a time line through the origin at an angle given by the sine squared of the Weinberg angle towards W corresponding to the unbroken weak interaction. The angles $\angle UOE$ or $\angle UOW$ would be realised in the orbital velocity of the electron or weakotron, respectively, inside the hydrogen atom or the similar weak atom, respectively, which both obey Bohr's equations. Although the weak interaction does not appear in this form, it appears in nearly the same form as the electromagnetic inside particles, and this enables us to calculate some more of the arbitrary parameters of the Standard Model associated with the weak interaction, as we do in part B.

For the usual type of space encountered in physics, a Lorentz transformation or temporal rotation may be used to boost a motionless particle to any velocity we choose, which may vary with location [Bell et al. 2000], [Bell and Diaz 2003]. In ANPA space, we constrain the velocity to the values represented by the members of the ANPA hierarchy, and we suppose that the velocity does not vary with location. We raise the question of whether, ultimately, other values are possible and the velocity may vary with location. We see no reason why not granted our discussion in the next section, but, habitually, the former is how we see ANPA space at the moment, and that is what this paper will describe.

To understand the description that follows, we must draw a distinction between three sorts of weak atom. First there is the weakotron atom, already mentioned, in which the weakotron circles a suitable source and which is analogous to the hydrogen atom. In the former, the orbital velocity is the sine squared of the Weinberg angle. We do not see this because the symmetry breaks at the boundary of the weakotron itself and so such an atom cannot be formed. Second there is a weak atom responsible for forming the weakotron and here the velocity is equal to two-thirds the sine squared of the Weinberg angle. We do see this, and we shall explain where in part B. Thirdly, the weak signal breaks down into particles, *Z*- and *W*-bosons, that can themselves be thought of as forming a Bohr atom, but here the orbital velocity is the sine of the Weinberg angle, again as appears in part B. Exactly the same thing applies for electromagnetism with the fine structure constant replacing the sine squared of the Weinberg angle, as we see in the hydrogen atom, electron and the particle of light, the photon. For example, the internal arrangement of the electron must result in the particle possessing only one unit of charge available externally, and requiring a second unit of charge on another particle to form a bound state externally. Instead of putting the second unit of charge on another particle, we can imagine just one particle interacting with a signal from the other particle, light or the celestial ray. We may then imagine the radiation signal as carrying the second charge. The weak interaction follows suit for the weak signal. This is the only charge involved for the internal arrangement of the signal, because you cannot have a signal's signal, and so we need to imagine the force between the parts as proportional to the charge itself and acting at a distance. At this point we cannot avoid commenting on spin, because we have been asserting that we are using similar strategies for particles whatever the size of their intrinsic spin. We confine our remarks to Appendix II.

We consider what happens to the internal arrangements of the signal for the weak interaction when the symmetry is broken. Returning to the diagram we described above, we make an alteration to the angle the weak time line, *OW*, makes with our own time, *OU*, so that the former becomes OW' and applies to the internal arrangement of the signal. The line then makes an angle, $\angle UOW'$, whose sine is given by sine of the Weinberg angle itself rather than its square. When the symmetry is broken, we choose our time line so that it lies along the electromagnetic time line, *OE*. The weak time line then rotates until it is oriented at the Weinberg angle to the electromagnetic time line, becoming OW''. We suppose that the weak signal travels along the weak time line.

Then the signal splits into two components, one along the electromagnetic time, *OE*, which we identify as the electromagnetic signal, light, for example, and one at right angles to *OE*, along $W''E$, which represents a particle. We see that $\angle W''EO$ is a right angle. As we show in part B the particle is the weakotron. This particle can also be seen as an example of a weak atom for the signal, appearing in the weakoshell provided we add the intermediate levels. This is because we may see the signal as travelling from one weak atom to the next as a spin wave. In this case the signal can be thought of as linking the original weak one to the electromagnetic one. Unlike the other lines we have described, the length of EW'' is finite, representing a signal of limited range. The mass of the weakotron, m_w, is proportional to EW''.

Finally, the weakotron splits into two particles. The first, along the electromagnetic time line with an electric charge and a mass, M_W, proportional to *OE*, is called the *W*-boson. The second, along the weak time line without electric charge with a mass, M_Z, proportional to OW'', is called the Z-boson. We may assert this because the Electro-weak theory tells us that [Cottingham and Greenwood 1998] and we are not suggesting that the Electro-weak theory is mistaken in its conclusions, merely adding another route to those. So,

$$\frac{M_W}{M_Z} = \cos\theta_W \qquad (3.1.A)$$

since we have,

$$\frac{EO}{OW''} = \cos\theta_W \qquad (3.1.B)$$

Finally, the whole is moved back into our own space-time. Our time line rotates from *OE* back to its old position, *OU*. This rotation also generates a particle, which can be seen as a smaller edition of the weakotron and we shall call it *the offshoot*. It is another type of weak atom at a different level in the weakoshell from the position of the weakotron. The offshoot can be seen as a signal linking *OE* with *OU*. The signal runs along $E'U$ and has a mass proportional to its length. E' lies along the same line as *OE* with $OE' \approx OW''$. $\triangle OUE'$ has $\angle OUE'$ a right angle. The mass of the offshoot is proportional to length of $E'U$. To a first approximation, we may use just the weakotron,

$$m_W = \sqrt{M_Z^2 - M_W^2} \propto EW'' \qquad (3.1.C)$$

To a second approximation, we add the mass of the offshoot, set $OE' = OW''$ and,

$$m_w + m_o = \sqrt{M_Z^2 - M_W^2} \propto EW'' + UE' \qquad (3.1.D)$$

There is still a small factor this does not take into account, but, if we do not insist that $OE' = OW''$, then it may be made exact. We will explain the last factor when we calculate the mass of the bosons in part B.

The whole occurs, we speculate, because it is desired to restrict the range of the weak interaction and confer any long range interaction back into a description in terms of electromagnetism, which is permitted an infinite range. The goal is achieved by bending back the signal associated with the weakotron particle so that it forms a closed loop. Such closed loops generate mass, which is why the weakotron is massive, rather than like light which has no mass and does not turn back upon itself. We may imagine the weakotron as having circular orbit tangent to *OE* with a diameter EW'' wedged between the electromagnetic time line, *OE*, and the weak time line, OW'', described by Bohr's equations with appropriate parameters. The offshoot is a smaller copy having an orbit tangent to *OU*, with diameter UE' wedged between our time line, *OU*, and the electromagnetic time line, OE'. The weak signal then travels round both orbits in an approximate closed figure of eight. The particle of light, on the other hand, the photon, would travel along the figure of eight only until the distance from U to W'' or vice versa had been covered before forming a new figure of eight which it would traverse in the same direction, beyond W'' or U. We shall provide some evidence for this in part B.

We see that our total length, which is almost proportional to the mass we seek, $W''E + E'U$, is only a little different from what we would find for the length if we simply wedged one appropriate orbit between the weak time line and our own. The signal would then correspond to a semicircle signal going from one time line to the other before it doubled back to follow the other semicircle and return to the original time line. Our signal, by contrast, follows an approximate figure of eight, with a slight kink at the centre. In either case, we have supposed that the carriers of the signal are massive because a restriction in range corresponds directly to a closed orbit.

It might be wondered why we do not immediately plump for a description in terms of the orbit of the one large circle we have just suggested. The reason is that we have now and here broken down a signal, normally thought of as a straight line, into a system of circular motions, the weakotron and its offshoot. We suggest that this is how all signals behave. However, we cannot have one large circle joining sender and receiver if we want to imagine the possibility of interception at intermediate points. What happens then, we suggest, is that the large circle turns into a lot of little ones, all tangential to each other, forming the wave motion we usually associated with both particles and signals, the signal weaving its way as a spin wave between the particles we may imagine forming the aether previously discussed by us [Bell 2004]. In this instance we have broken down the signal into two circles, which is all the ones allowed if we restrict angles to those which appear in the ANPA hierarchy and its subsidiary hierarchies, the electroshell and weakoshell. In this case a slight kink appears.

3.2 Algebraic Nuts and Bolts

We overview the more technical algebra required to put our description of the weak interaction in a form that links it to the theory of the weak interaction that is recognised as part of the Standard Model, the Electro-weak theory. We need to show that the equations that form the content of the Electro-weak theory before the symmetry is broken may be put into the same form as the equations of the single particle Quantum Electrodynamics (QED), the Maxwell and Dirac equations, because these have already been shown to lead to the system of scaling spins we are describing as applying to the weak interaction. References describing how QED and General Relativity lead to the scaling spins are given in the introduction to this paper. Technical accounts of the Electro-weak interaction are to be found in Aitchison and Hey [2004] and Cottingham and Greenwood [1998], but we shall require no more than the latter more elementary account. The Electro-weak theory has the task of fitting the phenomenology into a form in which it may be addressed by the tools that were available. It was not widely known that the Bohr atom was compatible with the accepted theory of QED, nor that particles or signals may be described as Bohr atoms, whether celestial or terrestrial, and, instead, heavy algebraic guns were brought to bear. It required a signal for the weak interaction with broken symmetry that had a mass, unlike light, where the photons are massless. The theory arrived at predicted the existence of the *W*- and *Z*-bosons together with their mass and allowed quantum calculations to be carried out, that is, it was renormalisable. It also predicted the existence of the Higgs boson and the existence of more than one type of vacuum. We had no need of the Higgs boson in the account in the last section, and continue with an account that leaves it out. We include the variation in the vacuum because the vacuum in a particle or atom described by Bohr equations can be seen as having, in its interior, a new level of energy for that vacuum. For example, the subtraction of the binding energy from the mass of an electron when it becomes part of the hydrogen atom can be seen as an alteration in the energy of the vacuum inside, rather than reflecting the potential contributed by the nucleus, since the electron wave function permits us to see the particle as free [Bell et al. 2004a].

Following the methods of Cottingham and Greenwood, we arrive at a local symmetry which has a variation for the field,

$$D_\mu \Phi = \left[\partial_\mu + (ig_1/2)B_\mu + (ig_2/2)\mathbf{W}_\mu\right]\Phi \qquad (3.2.A)$$

Φ is the original field with unbroken symmetry, g_1 and g_2 are coupling constants, B_μ is a scalar field apart from the space-time index, μ, and,

$$\mathbf{W}_\mu(x) = \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} \qquad (3.2.B)$$

where we have multiplied the three components of $\mathbf{W}_\mu$ by the Pauli matrices and $\mathbf{W}_\mu$ is the wave function of another field. We may define B_μ as part of $\mathbf{W}_\mu$ if either we do not mind greatly what the exact value of the space time dependence of B_μ might be or we give both B_μ and $\mathbf{W}_\mu$ a second set of space co-ordinates for a new space in which the two coincide. This second space is the one associated with the ANPA hierarchy, ANPA space. We suppose that the difference in coupling constants applied to B_μ and $\mathbf{W}_\mu$ is actually part of a rotation in ANPA space,

$$\frac{ig_1}{2}=\frac{i}{2\sqrt{\chi}\cos\theta_w},\quad \frac{ig_2}{2}=\frac{i}{2\sqrt{\chi}\sin\theta_w} \tag{3.2.C}$$

Here we have anticipated the result we need by explicitly describing the coupling constant in terms of the electromagnetic coupling constant, $1/\chi$, and the angle in terms of the weak coupling constant, $\sin\theta_w=\sqrt{1/\eta}$, following the relation predicted by the Electro-weak theory and found experimentally. We set,

$$\widehat{W}^0_\mu=\frac{B_\mu(x)}{2\cos\theta_w}, \tag{3.2.D}$$

$$\widehat{\mathbf{W}}(x)=\frac{1}{2\sin\theta_w}\begin{pmatrix} W^0_\mu+W^3_\mu & W^1_\mu-iW^2_\mu \\ W^1_\mu+iW^2_\mu & W^0_\mu-W^3_\mu\end{pmatrix}$$

where we have not gone to the length of two copies of space-time. We see that B_μ and the W^i_μ are components of $\widehat{\mathbf{W}}$. Equation (3.2.A) becomes

$$D_\mu\mathbf{\Phi}=\left[\partial_\mu+\left(i/\sqrt{\chi}\right)\widehat{\mathbf{W}}\right]\mathbf{\Phi} \tag{3.2.E}$$

with components of $\widehat{\mathbf{W}}$ of $\widehat{W}^\nu_\mu$.

The form of the Lagrangian for the dynamics of $\mathbf{W}_\mu$ and B_μ is given by Cottingham and Greenwood [1998],

$$\begin{aligned} \mathsf{L}_{\text{dyn}} &= -\frac{1}{4}B_{\mu\gamma}B^{\mu\gamma}-\frac{1}{4}\sum_{i=1}^{3}W^i_{\mu\gamma}W^{i\mu\gamma} \\ &= -\cos^2\theta_w\widehat{W^0_{\mu\gamma}}\widehat{W^{0\mu\gamma}}-\sin^2\theta_w\sum_{i=1}^{3}\widehat{W^i_{\mu\gamma}}\widehat{W^{i\mu\gamma}} \end{aligned} \tag{3.2.F}$$

This is just a sum over the components, for which we would like to define,

$$\widehat{W}^{\nu\mu\gamma} = \partial^{\mu}\widehat{W}^{\nu\gamma} - \partial^{\gamma}\widehat{W}^{\nu\mu} \qquad (3.2.G)$$

as we would for light, leaving our description in the last section to account for the mass of the weak signal. However, this appears to take no notice of the non-Abelian nature of the group. In fact, all would be well provided we notice that the unitary matrix proposed for the local gauge transformation, $\mathrm{U} = \exp\left(-i\alpha^{k}\tau^{k}\right)$, with τ^{k} the Pauli matrices, is valid for a field composed of a particle that has both terrestrial and celestial components, along the lines already mentioned by Bell and Diaz [2003]. What we shall do instead is to strip the Pauli matrices out and show that they may be added back in again at a later stage. So for the moment our extra group becomes the same Abelian one as appears in electromagnetism, and we reaffirm equations (3.2.F) and (3.2.G).

We may assume the resulting equations, analogous to the free Maxwell's equations, describing the dynamics of $\widehat{W}^{\nu}_{\mu}$ is for a particle rather than a field, in which case the equation describes only one particle. If we ever feel the yen to have a field equation instead, it is always open to us to convert it back into a field equation by canonical or path integral methods applied to the index μ alone.

Our goal is to put the equation in a form for which it is explicitly relativistic, not only for transformations in our usual space-time but also for transformations in ANPA space, with both spaces of the same form, before introducing the Pauli matrices again. Carrying the indices ν with us, we may derive the equivalent of the free Maxwell's equations in potential form for the Lorentz gauge. Then we may translate Maxwell's equations into reflector form [Bell et al. 2000], [Bell and Diaz 2003] for the μ indices, still carrying the ν indices. The components of the field $\dot{\widehat{W}}_{\mu}$ each acquire a reflector matrix for each index μ. These have elements which contain imaginary Pauli matrices plus the unit matrix, that is, they are real or imaginary quaternions. Reflectors are explored at length in the references cited. Finally, we make two copies of the Maxwell equation and append a second set of either the quaternions or their Hermitian conjugates to the ν indices. The SU(2) matrices have been replaced, but as we have a carbon copy of Maxwell's equations for the ν index, along with the original for the μ index, we know we can use equation (3.2.G) if we want to.

To be concrete, $\widehat{\mathbf{W}}$ is then in the form,

$$\underline{\widehat{\mathbf{W}}} = \begin{pmatrix} 0 & \widehat{\mathbf{W}}_{\mu^{\ddagger}}^{\nu^{\ddagger}} & 0 & 0 \\ \widehat{\mathbf{W}}_{\mu^{\ddagger}}^{\nu} & 0 & 0 & 0 \\ 0 & 0 & 0 & \widehat{\mathbf{W}}_{\mu}^{\nu^{\ddagger}} \\ 0 & 0 & \widehat{\mathbf{W}}_{\mu}^{\nu} & 0 \end{pmatrix} \qquad (3.2.H)$$

where we have chosen a bold $\widehat{\mathbf{W}}$ within the matrix to indicate each variable includes two quaternion or unit matrices chosen to match the two indices μ and ν. A superscript ‡ on an index indicates that the quaternion conjugate of the matrix associated with this index is to be taken. Although the SU(2) variation has been reinserted, we can still find a unitary matrix to reproduce a local symmetry, albeit U(1).

Not only may we rotate $\widehat{\mathbf{W}}$ spatially for $\mu = 1,2,3$ or $\nu = 1,2,3$, but we may also perform a Lorentz transformation or a temporal rotation on either index, for $\mu = 0$ or $\nu = 0$, justifying equation (3.2.C) as well as (3.2.G). This shows that B_μ and W^i_μ may be obtained from $\widehat{W}^\nu_\mu$ by a temporal rotation combined with scaling by a half. Both Lorentz transformations and temporal rotations are permitted as Bell and Diaz discuss [Bell and Diaz 2003] and the scaling by a half will disappear if we normalise. In particular we may perform another rotation, this time involving just the B_μ and W^3_μ components of $\widehat{\mathbf{W}}$,

$$\frac{\widehat{W'}^0_\mu}{2} = W^3_\mu \cos\theta_w - B_\mu \sin\theta_w$$
$$\frac{\widehat{W'}^3_\mu}{2} = W^3_\mu \sin\theta_w + B_\mu \cos\theta_w \qquad (3.2.I)$$

We may then interpret $\widehat{W'}^0_\mu$ and $\widehat{W'}^3_\mu$ as separate particles,

$$2Z_\mu = \widehat{W'}^0_\mu, \quad 2A_\mu = \widehat{W'}^3_\mu, \quad \widehat{W''}^\nu_\mu = \widehat{W}^1_\mu, \widehat{W}^2_\mu \qquad (3.2.J)$$

where Z_μ is the normally quoted wave function of the Z-boson, and A_μ is the wave function of electromagnetic radiation. We have also allowed the remaining components of $\widehat{\mathbf{W}}$ to form twice the wave function of the relativistically invariant *W*-boson, to which we donate all the electric charge.

We move on to the interaction of the carriers of the weak signal, the *Z*- and *W*-bosons, with particles that feel their force. We shall choose just the leptons here, the electron, muon and τ meson and their associated neutrinos. In part B we will also discuss the quarks. The former will not detain us long, because it is very obvious we can follow the same method as above.

We are enjoined by the Electro-weak theory to take as a wave function, first, a bispinor from the electron wave function, **L**, that has a definite helicity, combined, second, with the same for the neutrino, which has a definite helicity in any case. The bispinor we have already used for the Dirac equation when it is in reflector form [Bell et al. 2000], [Bell et al. 2003] has this property and a similar approach would also apply to the neutrino. We then have for such a particle,

$$D_\mu \mathbf{L} = \left[\partial_\mu + \left(i/\sqrt{\chi}\right)\widehat{\mathbf{W}^m}\right]\mathbf{L} \qquad (3.2.K)$$

where,

$$\widehat{W^m}^0_\mu = -\widehat{W}^0_\mu, \quad \widehat{W^m}^i_\mu = -\widehat{W}^i_\mu \qquad (3.2.L)$$

which is a simple time reversal and parity change, permitted by the Dirac equation. However, the rotation in ANPA space we performed for the signal has turned it into a facsimile of electromagnetic radiation with the usual coupling constant of $1/\chi$, and we suggest that has the same effect on the electron and neutrino wave function, that it becomes an electron in behaviour, since equation (3.2.K) resembles what we would expect for that, except for a double helping of indices. We shall leave the matter there for now, since describing the algebra with the original unitary matrix is best not hurried: the terrestrial unites with the celestial signal.

4. SUMMARY

We have shown how the coupling constant for the weak interaction, the sine squared of the Weinberg angle as it is usually called, arises from the appropriate member of the ANPA hierarchy, 3, when it is made scaling in a particularly simple way. We have gone on to describe what happens when the symmetry of the weak interaction is broken, following the phenomenology of the Electro-weak theory. In part B we shall show that this breaking occurs at the boundary between the internal and the external for what is normally considered to be a particle, for example, the muon or the τ meson. At this point the sine squared of the Weinberg angle no longer scales upwards in size, although we have seen that it continues to scale downwards. This results in a signal with a limited range. One way of producing such a signal in accordance with quantum theory is to give the carrier of the signal a mass. Such a process reveals several new particles that act as part of the weak signal including the *W*- and *Z*-bosons. We have described how this can work without introducing the Higgs boson, informally and geometrically. We have also provided an algebraic description using the Electro-weak theory that shows that the unbroken weak symmetry may be rotated into the same form as the electromagnetic. The state is reachable by a temporal rotation in the space of the ANPA hierarchy, which is a space of the same sort as the one in which we live.

REFERENCES

Aitchison, I. J. R. and Hey, A. J. G. *Gauge Theories in Particle Physics, QCD and the Electroweak Theory* Vol 2 Institute of Physics Publishing, Bristol (2004)

Bastin, T. and Kilmister, C. W. *Combinatorial Physics* World Scientific, Singapore (1995)

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. "Classical behaviour of the Dirac bispinor ," *Found. Phys.* Vol 30 No 1 (2000) 35-57; arXiv:quant-ph/0011026

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. "QED derived from the two-body interaction," *Found. Phys.* Vol 34 No 2 (2004a) 297-333; arXiv:quant-ph/0104017 & arXiv:quant-ph/0104018

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. "The two-body interaction with a circle in time," *Found. Phys.* Vol 34 No 2 (2004b) 335-358; arXiv:quant-ph/0110162

Bell, S. B. M. and Diaz, B. M. "Quantising general relativity using QED theory," To be published (2002); arXiv:physics/0211108

Bell, S. B. M. "Quantising general relativity using QED theory, an overview and extension," Proc. Physical Interpretations of Relativity Theory Conference, London (2004); arXiv:physics/0401005

Bell, S. B. M. and Diaz, B. M. "The relationship of time and space," To be published (2003); arXiv:physics/0403137

Bell, S. B. M. and Diaz, B. M. "Renormalisation and hierarchies," To be published (2004a); arXiv:physics/0406089

Bell, S. B. M. and Diaz, B. M. "Interdisciplinary overview: physics and philosophy," to appear, Proc. 26th International Meeting of the Alternative Natural Philosophy Association (ed. Keith G. Bowden), Cambridge U.K. (2005a); http://www.sarahbell.org.uk

Bell, S. B. M. and Diaz, B. M. "Exact solutions for QED," To be published (2004b); http://www.sarahbell.org.uk

Bell, S. B. M. "The ANPA Combinatorial Hierarchy - paper one: gravity and electromagnetism," To appear Hadronic Journal Vol. 28 (2005b)

Bell, S. B. M. and Diaz, B. M. "The ANPA combinatorial hierarchy - paper two (B): deriving the Standard Model - the Electro-weak parameters," To be published (2005c)

Chen, Z.-D., Liang, J.-Q., Shen, S.-Q. and Xic, W.-F. "Dynamics and Berry Phase of Two-Species Bose-Einstein Condensates" Phys. Rev. A Vol 69 (2004) 023611; arXiv:cond-mat/0403570 v1 (a largely non-technical review of the Berry phase may be found at http://www.mi.infm.it/manini/berryphase.html)

Cottingham, W. N. and Greenwood, D. A, *An Introduction to the Standard Model of Particle Physics* Cambridge University Press, Cambridge (1998)

Kilmister, C "The fine structure constant," in (ed. Keith G. Bowden) *Spin: Proc. Alternative Natural Philosophy Association* Vol 25 (2003a) 1-18

Kilmister, C "The Eddington heritage - a claim for ANPA," in (ed. Keith G. Bowden) *Spin: Proc. Alternative Natural Philosophy Association* Vol 25 (2003b) 341-350

Koc, C. K. "A Tutorial on P-adic Arithmetic," Oregon State University Technical Report, April (2002); **http://islab.oregonstate.edu/papers/r09padic.pdf**

Manthey, M. "ANPA - the Alternative Natural Philosophy Association," *Cybernetics and Human Knowing* Vol 2 No 2 (1993); http://www.imprint.co.uk/C&HK/vol2/v2-2mm.htm

Parker-Rhodes, A. F. *The Theory of Indistinguishables* D. Reidel Publishing Company, London (1981)

Appendix I

Program to Calculate Coupling Constants

The program below calculates the scaling version of all the first three members of the ANPA hierarchy, which estimates the coupling constants of the electromagnetic, the weak and the strong coupling constants. It is in R.T. Russell's BBC Basic, which is to be found at the following web address: http://www.rtrussell.co.uk/.

```
10 REM: FINE - Calculates the fine structure constant on
the basis it is scaling
20 REM: Copyright (C) 2004 by S B M Bell
   30 *FLOAT 64
   40 REM Use double precision for all calculations
   50 @%=&000F0A
   60 REM Controls format of printed numbers
   70 REM
   80 REM We imagine an origin with four straight lines
going diagonally upward.
   90 REM The leftmost belongs to the strong interaction.
The next, towards the right
  100 REM is our time. The next belongs to the
electromagnetic interaction, while the fourth
  110 REM belongs to the electro-weak interaction. Then the
sine of the angle between
```

```
  120 REM the interaction in question and our time gives
the first three terms of the
  130 REM ANPA series.
  140
  150 REM Defines the series for electromagnetism and the
fine structure constant
  160
  170
FFNB=137:SMLPR=4096:LRGPR=4096*512*512:P=4096*512:RNGE=128
  180 REM FFNB is whole number estimate of coupling
constant
  190 REM SMLPR is the starting probability of lower series
  200 REM LRGPR is starting probability for upper series
  210 REM P is the amount the probability decreases between
adjacent terms of the series
  220 REM RNGE is the length of the each instance of the
repeating pattern
  230 PROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RSULTA,RSULT1)
  240
  250 REM Defines the series for the electro-weak
interaction and the Weinberg angle
  260
  270 FFNB=3:SMLPR=4:LRGPR=8:P=4:RNGE=4
  280 PROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RSULTA,RSULT2)
  290 PRINT"We assume that the symmetry breaks before the
upper part of series"
  300 PRINT"Then the coupling constant for lower series is
the sine squared of the Weinberg angle squared "1/RSULTA
  310 PRINT
  320
  330 REM Defines the series for quantum chromodynamics and
the strong coupling constant.
  340
  350 FFNB=-10:SMLPR=32:LRGPR=16*32:P=16*32/2/8:RNGE=8
  360 PROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RSULTA,RSULT3)
  370
  380 END
  390
  400 REM Calculate the series
  410 DEFPROCFNDIT(FFNB,SMLPR,LRGPR,P,RNGE,RETURN
RSULTA,RETURN RSULT)
  420 REM Starting point for geometric series for numbers
less than coupling constant
  430 FFNSML=ABS((FFNB/(1-1/RNGE)/SMLPR)*(1/(1-1/P)))
  440 REM Sum of the geometric series for numbers less than
one
  450 RSULTA=FFNSML+FFNB
```

```
  460 PRINT"Contribution of lower series for "STR$(FFNB)":
"RSULTA'" Inverse "1/RSULTA
  470 FFNBG=ABS((RSULTA*RNGE*RNGE/LRGPR)*(1/(1-
RNGE*RNGE/P)))
  480 REM Sum the geometric series for numbers greater than
coupling constant
  490 RSULTB=FFNBG
  500 PRINT"Contribution of upper series for "STR$(FFNB)":
"RSULTB'" Inverse "1/RSULTB
  510 REM Add upper and lower series
  520 RSULT=RSULTA+RSULTB
  530 PRINT"Total value of "STR$(FFNB)" constant "RSULT"
Inverse "1/RSULT
  540 PRINT
  550 ENDPROC
  560
```

Output of program:-

Contribution of lower series for 137:

137.033710645996 Inverse 0.00729747443374234

Contribution of upper series for 137: 0.00210743280398 Inverse 474.510977579662

Total value of 137 constant 137.0358180788

Inverse 0.00729736220806862

Contribution of lower series for 3: 4.33333333333333 Inverse 0.230769230769231

Contribution of upper series for 3: 2.88888888888889 Inverse 0.346153846153846

Total value of 3 constant 7.22222222222222

Inverse 0.138461538461539

We assume that the symmetry breaks before the upper part of series. Then the coupling constant for lower series is the sine squared of the Weinberg angle

0.230769230769231

Contribution of lower series for -10: -9.63133640552995 Inverse -0.103827751196172

Contribution of upper series for -10: 2.0391705069124 Inverse 0.830622009569378

Total value of -10 constant -8.42741935483871 Inverse -
0.11866028708134

Appendix II

Spin and the Berry Phase

We have already touched on why the electron may have an intrinsic spin of a half when it is not interacting with anything but a spin of one when it is circling round in response to an electromagnetic interaction forming an atom such as the hydrogen atom [Bell et al. 2004a&b]. Does this generalise to other values of the spin in other circumstances? We will suppose that the dynamical phase associated with particles is generated by a rotational motion, or, if you will, that the dynamical phase is locked in sync with the phase of an orbital rotation.

In the case where the electron is the particle in a Bohr atom, we suppose that the intrinsic spin of the electron is added to the spin it has because it is rotating round the orbit, according to,

$$\omega = \theta + m\theta \qquad \text{(A.A)}$$

m is the component of the intrinsic spin angular momentum in which we are interested. ω is the angle turned through as seen externally, and θ is the actual angle turned through as it appears internally. The spin, m, adds an angle $m\theta$ to what would be the case if there was no intrinsic spin, when the internal and external view would coincide. m may be negative, but in the interests of brevity we will only consider the positive case. The angular momentum seen from the inside in the Bohr atom, and for the free electron, is a half, proportional to $m\theta$, while for the outside in the Bohr atom it is one, proportional to θ. This arrangement may be obtained if for each completed orbit, we set $\omega = 2\pi$ radians and $\theta = 4\pi/3$ radians. If the difference is occasioned because the internal and external orbital velocities vary, the internal orbital velocity must be $2/3\chi$, since we know that the external velocity for the Bohr atom with an electron as a particle is $1/\chi$. In part B we derive exactly the same result for an electron held together by gravitational attraction, by retracing our steps to our earlier model of this [Bell and Diaz 2004a] (for references see the previous section headed references). We suppose that in this case the internal angular momentum is a quarter, and the external a half.

We generalise to other values of spin to include $|m| > 1/2$ using the Berry phase, which permits us to apply equation (A.A) to all the cases where,

$$m = J_z/2, \quad J_z = \text{integer} \qquad \text{(A.B)}$$

We may show the connection with the Berry phase in a simple way. We have been discussing the orbit of the electron as if it was two-dimensional here, but in many cases the electron appears on the surface of revolution of such an orbit, a sphere, or part of a sphere [Bell et al. 2004a], [Bell and Diaz 2002], [Bell 2004]. In the first reference cited, we found an expression for the surface area of a sphere. We considered a section through a sphere of radius a with centre G and two lines from this centre cutting the surface of the sphere at H and H'. I lies on GH' and $\angle GIH$ is a right angle. Let angle $HGH' = \vartheta$. We may calculate the fraction of the surface area of the sphere obtained by rotating the surface $a^2 \sin\vartheta \, d\vartheta\delta\theta$ through an angle, θ, on the circle with centre I and radius IH, and the result through an angle, ϑ, on the circle through H with centre G, that is

$$K = \int_{\vartheta}^{0}\int_{0}^{\theta} a^2 \sin\vartheta \, d\theta \, d\vartheta \tag{A.C}$$

The solid angle swept out at G for this transport is $\Omega = K/a^2$. Ω is proportional to the Berry phase [Chen et al. 2004]. Performing the integral,

$$\Omega = \theta(1 - \cos\vartheta) \tag{A.D}$$

Since m is the component of the angular momentum along GH', which is proportional to GI, or $a\cos\theta$,

$$m = k\cos\vartheta \tag{A.E}$$

where k is a constant of proportionality, and we obtain from equation (A.D),

$$\Omega = \theta + m\theta \tag{A.F}$$

where we have set $k = -1$. Since we know the Berry phase generalises to the components of spin angular momentum given in equation, (A.B) [Chen et al. 2004], we have our result when we set $\omega = \Omega$.

THE ANPA COMBINATORIAL HIERARCHY

Paper Two (B): Deriving the Standard Model - The Electro-Weak Parameters

S.B.M. Bell† and B.M. Diaz‡

Interdisciplinary Quantum Group, Department of Computer Science

The University of Liverpool, Chadwick Building, Peach Street

Liverpool, L69 7ZF, United Kingdom

ABSTRACT

Scaling combinatorial hierarchies in which each level describes a circular motion have been found by us to be an intrinsic part of quantum electrodynamics and general relativity. From them we derive the combinatorial hierarchy that is the raison d'être of the Alternative Natural Philosophy Association (ANPA) [Kilmister 2003a]. This is used to derive the following (1) the electromagnetic theory, responsible for binding atoms together, from the gravitational, including the derivation of an approximate value for the fine structure constant, then (2) the weak interaction, responsible for radioactivity, including an approximate value for the

† Mail@sarahbell.org.uk

‡ B.M.Diaz@csc.liv.ac.uk

Weinberg angle and finally (3) the strong interaction, responsible for binding the nuclei of atoms together, including an approximate value for the quark-gluon-quark coupling constant. We also derive sundry others of the arbitrary parameters in the Standard Model of particle physics. In this paper we derive some of the parameters associated with the weak interaction, the masses of the *Z*-and *W*-bosons without introducing the Higgs boson, the mass of the muon and the mass of the τ meson relative to the mass of the electron and parameters associated with the Cabibbo-Kobayashi-Maskawa matrix.

1. PRELIMINARIES

1.1 General Introduction

This is the second part of the second of a trilogy of papers describing the ANPA combinatorial hierarchy. The ANPA combinatorial hierarchy is fascinating because it is a most unusual hierarchy derived from very simple criteria [Manthey 1993], [Kilmister 2003a&b], [Bastin and Kilmister 1995], [Parker-Rhodes 1981]. The hierarchy in its simplest form is 3, 10, 137 and $2^{127}+136\approx 1.7\times 10^{38}$, ending on the last term. It was such an unusual structure, based on the simplest hypothesis on discerning similarity and difference and also using a simple method of aggregation to consolidate each judgement, that many were sure it held some deep truth. It was suggested that 3 described the weak interaction, 10 the strong interaction, 137 the electromagnetic interaction and $\sim 1.7\times 10^{38}$ gravity.

This trilogy of papers uses the ANPA hierarchy to calculate some of the arbitrary parameters of the Standard Model, the recognised theory combining the electromagnetic, the weak and the strong interactions. There are eighteen arbitrary parameters in the Standard Model [Cottingham and Greenwood 1998], where the Model does not predict their value and they are set to agree with experiment. We have already calculated two. The fine structure constant in the first paper of the trilogy [Bell 2005], and the weak coupling constant, which depends on the Weinberg angle, in the first part of this paper, part A, the second in the trilogy [Bell and Diaz 2005b]. In this second part of paper two we will calculate another five, which we list. Two values are shown not to be required, the mass of the Higgs boson and the vacuum expectation value of the Higgs field. We calculate the mass of the *W*-boson and Z-boson by a process which does not bear a strong relationship with the Higgs mechanism but performs the same function. It is based on the electromagnetic and weak coupling constants. In the third part of this paper, part C [Bell and Diaz 2005c], we provide a discussion on varying the size of an orbit, needed in part to justify our calculation. In this paper, part B, we also calculate the ratio of the mass of the muon and τ meson to that of the electron, also based on the same coupling constants. We study the Cabibbo-Kobayashi-Maskawa matrix based in the same way, and find one of the parameters associated with it, which demonstrates a suggestive invariant of the weak interaction for quarks. The study of this matrix will extend into the third paper of the trilogy, since it involves a discussion of how quarks behave for the strong interaction. The third paper will also calculate the ratio of the masses of the charged pion and the proton to that of the electron, and finally, we will use the fourth term in the ANPA hierarchy to calculate the mass of the electron itself. All this will be based on the coupling constants alone, except for the dimension-bearing constants, that is, Planck's constant, h, the speed of light, c, and the gravitational constant, G. This means that thirteen of the parameters of the Standard Model depend on the four coupling constants alone, which depend in their turn on the ANPA hierarchy. This in its turn may be derived from the quantum gravitational theory, of which general relativity is a classical limit.

In part A, as well as deriving the Weinberg angle, we described a simple model of the weak interaction in terms of the scaling hierarchies of spins used throughout Bell and Diaz' account [Bell et al. 2000], [Bell et al. 2004a&b], [Bell and Diaz 2002], [Bell 2004], [Bell and Diaz 2003], [Bell and Diaz 2004a&b], [Bell and Diaz 2005a,b&c]. We also derived some of the algebra required to link this description to the theory of the weak interaction that is recognised as part of the Standard Model, the Electro-weak theory. We shall rely heavily on our discussion of the weak interaction in the first part of paper two, part A, which is best read first.

2. CREATION OF THE ELECTRON FROM THE GRAVITY

2.1 The Expanded Electroshell

Before we turn to the weak interaction, we must discuss the electron and the various issues that arise during this discussion. Although we shall not use the weak interaction directly in this, nor offer the calculation of some previously arbitrary variable, we must do so because the electron is in this account the most basic of the particles discussed in the Standard Model, and it is the electron's mass on which we base our subsequent calculations. In this paper the mass arises from our consideration of the microcosm, and has a parameter that we adjust to fit the observed mass. In the third paper of the trilogy [Bell and Diaz 2005d] we will derive both the ratio of the mass of the proton to that of the electron from the microcosm, and this ratio from the macrocosm, based on the value of the fourth and final member of the ANPA hierarchy. This enables us to go full circle and derive the electron's mass itself using only the dimension-bearing constants. We suggest that we choose the member of the ANPA hierarchy based on 137 to form the basis for the electron in the microcosm rather than either of the previous members, 3 and 10, because χ, the inverse of the fine structure constant and the final version of 137, the third member of the ANPA hierarchy, is the most perfectly symmetrical. The repeating pattern formed for χ is the same whether we scan it upwards or downwards, as we discussed in the first paper of this trilogy [Bell 2005], which is not true of the patterns formed for 3 or 10.

We suggest that gravitational attraction forms the electron, the force being carried by electromagnetic ray in an aether [Bell 2004] where the permittivity is no longer one. We discussed the properties of such an attraction previously [Bell and Diaz 2004a&b]. For the electron, which is to carry an electric charge proportional to $1/\sqrt{\chi}$, we pick the velocity of the particle, v, and the permittivity, ε, as,

$$v = \frac{1}{\chi}, \quad \varepsilon = \frac{1}{2\chi^2} \qquad (2.1.A)$$

We emphasise that using the scaling hierarchies to generate the mass of the particles means that, since every one is described using Bohr's equations, every one is described by a solution of the Dirac and the Maxwell equations that together form the single particle version of quantum electrodynamics' (QED) theory [Bell et al. 2004a&b], [Bell and Diaz 2002], [Bell 2004]. We use the table formed from the gravitational interaction showing the departure from the fusion of charge and inertial mass, the electroshell. The details of this departure are discussed by Bell and Diaz [2004a&b] and we summarise the result here, adding charge to our previous version of the electroshell hierarchy, shown in table I(b) in the first paper in our trilogy. The intermediate levels, which were introduced in that paper are also added, as we shall need them. We will call the resulting hierarchy *the expanded electroshell.*

Table I

Part of the expanded electroshell hierarchy

Level	−1/2	0	1/2	1	3/2	2	5/2
Velocity Squared	$\frac{1}{\chi^5}$	$\frac{1}{\chi^4}$	$\frac{1}{\chi^3}$	$\frac{1}{\chi^2}$	$\frac{1}{\chi}$	1	χ
Source Charge	$\frac{1}{\chi^3}$	$\frac{1}{\chi^2}$	$\frac{1}{\chi}$	1	χ	χ^2	χ^3
Particle Mass	$\frac{1}{\chi}$	1	χ	χ^2	χ^3	χ^4	χ^5
Particle Charge	$\frac{2}{\chi^3}$	$\frac{2}{\chi^2}$	$\frac{2}{\chi}$	2	2χ	$2\chi^2$	$2\chi^3$
Energy	$1+\frac{1}{\chi^5}$	$1+\frac{1}{\chi^4}$	$1+\frac{1}{\chi^3}$	$1+\frac{1}{\chi^2}$	$1+\frac{1}{\chi}$	2	$1+\chi$

Each level of this table represents a circular motion in which the particle orbits round the source. The masses, charges and energy are in the rest frame of the particle experiencing the interaction, which has a mass of one at level zero. We remark that in this frame the radius of the scaling orbit is proportional to $\sqrt{1-v^2}$, where v is the velocity and so alters between levels. We discuss this in detail in the third part of this paper, part C [Bell and Diaz 2005c]. We also comment on the value given for the energy. If we assume that the potential is absent, which we may in the space we discussed previously [Bell et al. 2004a], the energy in the frame of the particle is the rest mass of the particle. However, here we are assuming all takes place in the space in which we customarily view the atom, and the potential definitely does exist here, since the particle is travelling in a circle, which would be unexplained otherwise. To return from the frame of the source to the frame of the particle therefore requires the inverse of a Lorentz transformation. Viewing this change as the inverse of the Lorentz transformation requires us to consider the velocity as imaginary, if we are to keep the usual formula for this, which occurs when the interaction is repulsive and in other circumstances [Bell and Diaz 2002], [Bell and Diaz 2004a]. We will drop explicit mention of the imaginary character of the velocity, supposing that instead we alter the definition of velocity so that the usual one is imaginary.

2.2 The Electron

The gravitational attraction holding the electron together, F_i, is proportional to the product of the charges, since like charges attract for gravity,

$$F_i \propto 2X^2 \qquad (2.2.A)$$

from Bohr's equations, where X depends on the level chosen. Bohr's equations are given and discussed in Bell et al. [2004a&b], Bell and Diaz [2002] and Bell and Diaz [2004a].

We shall use units for X in which the velocity is the same as the force, since the radius never enters our deliberations specifically in this section and the other sections discussing the mass of the leptons, 3.1 and 3.2. When we use a hierarchy similar to table I, the radius will alter with the level. However, in the sections mentioned, we use the same level of the electroshell and weakoshell throughout. When we use a tachyonic state we use the same type of hierarchy that we used for gravity [Bell and Diaz

2002] and also for providing the fine structure of electromagnetic atoms like the hydrogen atom [Bell et al. 2004b]. The tables we applied, for example table 2 in Bell and Diaz [2002], quoted the interactions in a form where the radius did not alter between levels. We provide a discussion of how and why the radius varies in part C.

This means the internal velocity is

$$v_i = 2X^2 \qquad (2.2.B)$$

Distributing the charge equally between the source and particle, as we do for electricity, produces a charge for each of $X\sqrt{2}$,

$$e_i = X\sqrt{2} \qquad (2.2.C)$$

Since the internal attraction is really gravitational in nature, all charges have the same sign. The total charge on the state is therefore the charge on the source plus the charge on the particle, $3X$. We suppose that this charge then acts on the source provided by another such state, which has the same value for the charge on the source, *X*, although it may be of the opposite sign. With these provisos, we may write for the external force,

$$F_e \propto 3X^2 \qquad (2.2.D)$$

and splitting the charge equally between the new source and the electron produces a charge for each of,

$$e = \frac{1}{\sqrt{\chi}} = X\sqrt{3} \qquad (2.2.E)$$

This true whatever level of the hierarchy we may see, because the fine structure constant is scaling.

From equation (2.2.B),

$$v_i = \frac{2}{3\chi} \tag{2.2.F}$$

v_i could vary with the level from a base of $1/\chi$ for the external velocity, but in fact it does not for either the expanded electroshell or the weakoshell, which we discussed in part A. We may apply a table similar to table I with an altered velocity, v_i, as explained by Bell and Diaz [2004a] for table 4 therein. Then the energy of this state, seen as a particle, is given by the last row in table I as,

$$m_e = m_p\left(\sqrt{\chi}\right)^l\left(1+\frac{4}{9\chi^2}\right), \quad l=0 \tag{2.2.G}$$

where m_e is the observational value of the mass of the electron and the particle generating the state has mass $m_p\left(\sqrt{\chi}\right)^l$. We have set l to zero in this instance to simplify our calculation, but in fact any integer power would suffice, as only relative changes are relevant.

The 2002 CODATA value of the electron mass is

$$m_e^{\odot} = 9.1093826(16)\times 10^{-31} kg \tag{2.2.H}$$

where the superscript $^{\odot}$ stands for "observed" throughout and the number in brackets should be added and subtracted to find the range permitted by experimental observation.

2.3 The Role of the Signal

The particle of mass $m_p\left(\sqrt{\chi}\right)^l$, would be formed from another interaction obeying Bohr's equations in which a new particle orbits another the source. The new particle is a tachyon of mass, m_p, and inverse velocity, $\left(\sqrt{\chi}\right)^l$, where this velocity, unlike v_i, does vary, relatively, with the level, l, in the hierarchy.

This latter interaction can be seen as coming from the signal responsible for carrying the electromagnetic force. We sketch a possible approach here. The signal also forms a Bohr atom and this provides the energy seen as the particle in the Bohr atom forming the electron itself. Since electromagnetic radiation carries no mass, we get half an orbital that is joined smoothly to the next half orbital with the particle of light, the photon, still travelling in a forward direction. This constitutes the wave motion that we associate with light. Strictly speaking, quantum electrodynamics does not allow longitudinal vibrations for light [Greiner and Reinhardt 1996], and so we must donate any longitudinal vibration to the particle the ray is passing. When the progress of the photon is stopped by its absorption, the one vibration is joined by a second, as is required by the particle doing the absorption.

Although we have been describing the electromagnetic signal, if we suppose an aether [Bell 2004], we may imagine it filled in crystalline form with virtual electrons. This means that the ray is always passing some particle, and we are permitted to think of the signal as a spin wave of electrons with synchronised phase. This is sometimes helpful, particularly when we consider the weak interaction. Here, the signal does bend back on itself to form a standing wave, which we may identify with the weakotron. This in its turn forms the *W*- and *Z*-bosons, which do have associated mass as we discussed in part A of this paper and as we will see in detail when we calculate these masses in section 3.3. We shall also see that the mass of all the matching particles taking part in the interactions which form the electron, muon, the τ meson and the weakotron may be written in the form $m_p\left(\sqrt{\chi}\right)^l\left(\sqrt{\sqrt{\eta}}\right)^{l'}$ where l and l' are integer.

3. DERIVING THE MASS OF PARTICLES

3.1 The Mass of the Muon

Two interleaved hierarchies form the muon. The first level of the expanded electroshell hierarchy has the electron that we have already explored in it. If we suppose that the level contains a particle, as we did, the next level up will contain a tachyon, as we have previously discussed [Bell and Diaz 2002]. After this, we suppose that there is an interleaved level from the weak interaction hierarchy.

We will calculate the new contribution of the electron hierarchy first. The method of calculation for the electron is followed up to and including (2.2.G). At this point we diverge by taking the energy as that appropriate for a tachyon [Bell and Diaz 2002]. The energy of the tachyon state is

$$m_{\mu_i} = \frac{m_p\left(\sqrt{\chi}\right)^l}{v_i}, \quad l=0 \tag{3.1.A}$$

from equation (2.2.G) and forms the first part of the mass of the muon. This is similar to finding the principle energy level of a spectrum [Bell et al. 2004a], except that we chose the same space as we do for gravity [Bell and Diaz 2002], [Bell 2004] and it is the tachyon, not the particle, that provides the largest contribution here. We went on to calculate the fine structure of a spectrum [Bell et al. 2004b], where again we will want to choose the same space as for gravity. Our previous work showed that we may consider both the principle energy levels and the fine structure to be formed from circular orbits, but that these are in everywhere orthogonal planes. We may think of the tachyon orbit as precessing round the particle orbit.

The fine structure is contributed by a particle orbit obeying the weak interaction, which appears in the weakoshell hierarchy. We derived this in part A. Just as the electron is held together by a gravitational attraction, so to are the structures forming a part of particles which obey the weak interaction, the permittivity here being $1/2\eta^2$. The weakoshell is exactly the same as the expanded electroshell except that the χ is everywhere replaced by η. Because the weak interaction has no upper part in the hierarchy of spins it has no extra external charge. We are therefore concerned with the internal velocity,

$$v_i' = 2X'^2 \tag{3.1.B}$$

and with setting the external charge to equal the charge on the source,

$$X' = \frac{1}{\sqrt{\eta}} \tag{3.1.C}$$

This leads to,

$$v_i' = \frac{2}{\eta} = 2\sin^2\theta_w \tag{3.1.D}$$

as for the electron, in the analogue of equation (2.2.G),

$$m_{\mu_2} = m_p\left(\sqrt{\chi}\right)^l\left(1+\left(2\sin^2\theta_w\right)^2\right), \quad l=0 \tag{3.1.E}$$

where m_{μ_2} is the second part of the mass of the muon. Although the velocity of the particle, $2\sin^2\theta_w$, comes from the weak interaction, the mass, $m_p\left(\sqrt{\chi}\right)^l$, comes from an electromagnetic interaction in what we say about the muon and τ meson. When we discuss the *Z*- and *W*- bosons in section 3.3, we will see that, optionally, it may be formed by using a weak interaction.

Adding together the two parts of the mass of the muon using equations (3.1.B) and (2.2.G),

$$m_\mu = m_{\mu_1} + m_{\mu_2} = \frac{m_e}{1+4/9\chi^2}\left(\frac{3\chi}{2} + 1 + 4\sin^4\theta_w\right) \tag{3.1.F}$$

We have the most accurate experimental value of χ from the first paper of the trilogy [Bell 2005], based on CODATA 2002, and an estimate of $\sin^2\theta_w$ from Cottingham and Greenwood [1998],

$$\chi^{\odot} = 137.03599911, \quad \sin^2\theta_w^{\odot} = 0.2315 \pm 0.0004 \tag{3.1.G}$$

and m_e from equation (2.2.H), which give,

$$m_\mu = 1.883487594 \times 10^{-28} kg, \qquad (3.1.H)$$
$$m_\mu^\odot = 1.88353140(33) \times 10^{-28} kg$$

where we have also quoted the CODATA 2002 value for the muon mass. We see that we are lucky to have that accuracy, granted the error on $\sin^2\theta^\odot$. It may be that we have chosen a good value out of those available.

3.2 The Mass of the τ meson

We suppose that the hierarchy for the τ meson contains the muon entries. In that case we expect that a tachyon version of the weak interaction that formed the second part of the mass of the muon may be included. This will make a bigger contribution than last time and is in fact present, forming the principle part of the energy spectrum. The fine structure is also contributed by the weak interaction, and is the same as for the muon, except that the mass of the particle taking part in the interaction is larger.

The internal velocity, v_i', is given in equation (3.1.D). This then forms a tachyon state, which means that we must invert v_i', and then form an equation analogous to equation (3.1.A),

$$m_{\tau_1} = \frac{m_p\left(\sqrt{\chi}\right)^l}{2\sin^2\theta_w}, \quad l = 3 \qquad (3.2.A)$$

where we have to set the level for the interaction forming the particle three higher in the expanded electroshell to save appearances.

We pick the weak interaction again for the fine structure contribution. The internal velocity is the same as for the other two weak contributions, and we obtain the same as equation (3.1.E) for the energy, except for the value of *l*, which we set to one to save appearances,

$$m_{\tau_2} = m_p \left(\sqrt{\chi}\right)^l \left(1+\left(2\sin^2\theta_w\right)^2\right), \quad l=1 \tag{3.2.B}$$

Summing the two contributions using equations (3.2.A) and (2.2.G), the mass of the τ meson becomes

$$m_\tau = m_{\tau_1} + m_{\tau_2} = \frac{m_e\sqrt{\chi}}{1+4/9\chi^2}\left(\frac{\chi}{2\sin^2\theta_w} + 1 + 4\sin^4\theta_w\right) \tag{3.2.C}$$

Using the observed values, $\chi^{\odot}$ and $\sin^2\theta_w^{\odot}$, from equation (3.1.G) and $m_e^{\odot}$ from equation (2.2.H), we find the calculated value,

$$\begin{aligned} m_\tau &= 3.169041494\times10^{-27}\,kg, \\ m_\tau^{\odot} &= 3.16777(52)\times10^{-27}\,kg \end{aligned} \tag{3.2.D}$$

where we have also quoted the CODATA 2002 value for the τ meson mass. This is still a generous result, especially given the error on $m_\tau^{\odot}$ and the size of the factors multiplying powers of $\sin^2\theta_w$.

We notice a certain regularity about the magnitude of the power of $\sqrt{\chi}$ in these interactions. We have two fits for zero in the formulae given for the mass of the electron, equations (2.2.G) and the second part of the mass of the muon in equation (3.1.E). We have one fit for each for one, two and three, in respectively, the second part of the mass of the τ meson, equation (3.2.B), the first part of the mass of the muon, equation (3.1.A), and the first part of the mass of τ meson the in equation (3.2.A). If we wanted to emphasise this pattern we might care to derive the masses from another set of principles; time will show which of these patterns is most relevant.

3.3 The Mass of the *W*- and Z-Bosons

We discussed how the weak interaction worked in part A of this paper and we use this description to calculate the mass of the *W*- and Z-bosons, repeating the figure in ANPA space which we sketched previously. Our own time line goes from the origin at *O* vertically upwards to *U*. The time line for electromagnetism goes through the origin to *E*, inclined at an angle whose sine is the fine structure constant to *OU*. The weak time line goes through the origin inclined at the Weinberg angle relative to *OE* towards W''. The sine of the Weinberg angle is $1/\sqrt{\eta}$. The *weakotron particle*, which can be seen as the carrier of the weak signal, has a diameter for the orbit that lies between W'' and E. EW'' is orthogonal to *OE* so that OEW'' has $\angle OEW''$ a right angle. The mass of the weakotron is proportional to EW''. The weakotron does not appear in its own right and the contribution of the weakotron is split into two components. The first one is represented by *OE* and is the *W*-boson. *OE* is proportional to that part of the mass of the *W*-boson that is contributed by the weakotron. The second lies along the line OW'' is called the Z-boson. OW'' is proportional to that part of the mass of the Z-boson that is contributed by the weakotron. All this codifies the interaction of the electromagnetic time line *OE* with the weak time line, OW''. However the electromagnetic time line interacts with our own time line, *OU*. A small particle called *the offshoot* along the line $E'U$ carries this signal. E' is chosen such that $E'O \approx W''O$ and UE' is orthogonal to *OU* so that $\triangle OUE'$ has $\angle OUE'$ a right angle. The mass of the offshoot is then proportional to $E'U$. The offshoot does not appear in its own right either and the contribution of the offshoot is split between the *W*- and Z-bosons, in the same proportions as held for the weakotron.

In part A, section 3.1, of these two papers we discussed how to represent the sum of the masses of the weakotron and the offshoot in terms of the *W*- and Z-bosons and, approximately, the geometry described above. We suggested,

$$
\begin{aligned}
&m_w \propto EW'' \qquad (3.3.A)\\
&m_o \propto UE'\\
&m_w + m_o = \sqrt{\left(M_Z^{\odot}\right)^2 - \left(M_W^{\odot}\right)^2} \propto UE' + EW''
\end{aligned}
$$

where m_w and m_o are the masses of the weakotron and its offshoot, respectively, and $M_W^{\odot}$ and $M_Z^{\odot}$ are the experimentally observed masses. The proportionality is approximate, if we set $E'O = W''O$, but the equality exact. The masses of the bosons are given by Cottingham and Greenwood [1998] as,

$$
\begin{aligned}
&M_W^{\odot} = 80.33 \pm 0.15 \text{ Gev} = 1.432012232 \times 10^{-29} \text{ kg} \qquad (3.3.D)\\
&M_Z^{\odot} = 91.187 \pm 0.007 \text{ Gev} = 1.625555825 \times 10^{-29} \text{ kg}
\end{aligned}
$$

while the mass of the electron given (2.2.H) becomes

$$m_e^{\odot} = 0.000510998918 \text{ Gev} \qquad (3.3.E)$$

From these equations we may find the mass of the two bosons in terms of the mass of an electron,

$$M_W^{\odot} = 157201.8984 \times m_e, \quad M_Z^{\odot} = 178448.5188 \times m_e \qquad (3.3.F)$$

while from equation (3.3.A) we may find the mass for the combined sum that we may expect from observation,

$$m_w^{\odot} + m_o^{\odot} = 84447.83598 \times m_e^{\odot} \qquad (3.3.G)$$

We now calculate the masses of the weakotron and offshoot, starting with the weakotron, by assuming they are both weak Bohr atoms of the same type as the electron, and suggesting a way of deriving the equivalent of m_p for the weak interaction, $m_p^{(w)}$, from our discussion of the leptons in general. The highest power of $\sqrt{\chi}$ for these appeared in the first part of the mass of the τ meson, where $l = 3$. We suppose the particle energy we are using for our translation carries on the series by involving $l = 4$. The energy is then,

$$E = \chi^2 m_p \qquad (3.3.H)$$

We remember we inserted particles with a similar form of mass energy to that just shown, using χ. Replacing χ with η would have been more obvious for the mass of the particle when the interaction of the particle was weak. We suppose we were able to use χ because the result we would have got by switching from a power of χ to a suitable power of η would have been the same. Taking our cue from the formula giving the first part of for the mass of the muon in equation (3.1.A), we take the matching weak particle with the same energy to be the simplest possible,

$$E = \frac{m_p^{(w)}}{\eta} \tag{3.3.I}$$

We make a copy of equation (2.2.G) for the mass of the weakotron on the assumption that that the weakotron plays a similar role to the electron,

$$m_w = m_p^{(w)}\left(1 + \frac{4}{9\eta^2}\right) \tag{3.3.J}$$

and using (2.2.G), (3.3.H) and (3.3.I),

$$m_w = \frac{\eta\chi^2\left(1 + 4/9\eta^2\right)}{1 + 4/9\chi^2} m_e \tag{3.3.K}$$

We next need to calculate the mass of the offshoot. So far we have not had to bother with the radius of the Bohr orbit. Rather than varying the orbital velocity, we may choose to vary the radius by the same factor, ξ, where ξ is the coupling constant, to take us from one level to the next. We explain this at length in part C. We obtain,

$$m_o = m_w \frac{UE'}{EW''} \approx m_w \frac{\sin\angle EOW''}{\sin\angle UOE'} = \tag{3.3.M}$$

$$m_w \frac{\sqrt{\eta}}{\chi} \times \frac{\eta\chi^2\left(1 + 4/9\eta^2\right)}{1 + 4/9\chi^2} \times m_e$$

from equations (3.3.A) and where the factor giving the alteration between levels is $\sqrt{\eta}/\chi$. However, in fact the radius varies slightly between levels even before we introduce this factor. As we also explain in part C, this latter change between levels can be allowed for by correcting the mass of the particle in table I, and all such tables in which the variables are given in the frame of the source, by a factor $\sqrt{1+\xi^2}$.Choosing this path, we obtain the mass of the offshoot,

$$m_o = m_w \frac{\sqrt{\eta(1+1/\eta)(1+1/\chi^2)}}{\chi} \times \frac{\eta\chi^2(1+4/9\eta^2)}{1+4/9\chi^2} \times m_e \tag{3.3.N}$$

This gives for the sum we seek to compare with the value in equation (3.3.G),

$$m_o + m_w = \frac{\eta\chi^2(1+4/9\eta^2)}{1+4/9\chi^2} m_e + \frac{\chi\eta\sqrt{\eta(1+1/\eta)(1+1/\chi^2)}(1+4/9\eta^2)}{1+4/9\chi^2} m_e \tag{3.3.O}$$

from equation (3.3.K).

Finally, we calculate this from the experimental values of η, χ and m_e. We have use equations (3.1.F) and obtain,

$$\begin{aligned} m_o + m_w &= 83048.37863 m_e + 1397.815208 m_e \\ &= 84446.19381 m_e \end{aligned} \tag{3.3.P}$$

which again compares more favourably than we have any reason to expect with the value given in equation (3.3.G).

3. THE CABIBBO-KOBAYASHI-MASKAWA MATRIX

4.1 Analysis

The Cabibbo-Kobayashi-Maskawa (C-K-M) matrix describes the Electro-weak interactions of quarks. In the third paper of the trilogy we shall discuss quarks in more detail, but for now we will describe a symmetry of the matrix in a phenomenological account. The matrix, which is necessarily unitary, describes a relation between two triplets, one composed of quarks along the axes before the matrix acts, and the other composed of the quarks along the axes after [Aitchison and Hey 2004]. It contains, in addition to a rotation matrix, a phase. We will set this to unity, which will leave us with a rotation matrix, in which the symmetry lies.

We will show here that the observed C-K-M matrix without the phase shift is consistent with assumption that it represents a rotation matrix in a 3-dimensional space. We will find the vector orthogonal to the plane in which the rotation can be described by a single angle and calculate that angle. We will also show that the vector is orthogonal to one of the axes.

Aitchison and Hey [2004] quote the C-K-M matrix as,

$$\mathbf{M}^{\odot} = \begin{pmatrix} 0.9741-0.9756 & 0.219-0.226 & 0.0025-0.0048 \\ 0.219-0.226 & 0.9732-0.9748 & 0.038-0.044 \\ 0.004-0.014 & 0.037-0.044 & 0.9990-0.9993 \end{pmatrix} \qquad (4.1.A)$$

although we need to note that the information on the phases attached to these matrix elements has been lost, and the sign of the element along with it. This means that some elements must be negated before the matrix can represent a rotation. We shall use a method which treats all three planes in an even-handed fashion, finding a suitable set of Euler angles to describe the rotation of the axes induced [James and James 1992].

We will choose three matrices, each representing a rotation in a plane made by a pair of axes,

$$\mathbf{M}=\begin{pmatrix} a_{00} & a_{10} & a_{20} \\ a_{01} & a_{11} & a_{21} \\ a_{02} & a_{12} & a_{22} \end{pmatrix}=\begin{pmatrix} \cos\gamma & 0 & \sin\gamma \\ 0 & 1 & 0 \\ -\sin\gamma & 0 & \cos\gamma \end{pmatrix}\times \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}\times\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad (4.1.B)$$

and we obtain for the a_{ij},

$$\begin{aligned} &a_{00}=\cos\gamma\cos\phi, \quad a_{10}=-\cos\gamma\sin\phi\cos\theta+\sin\gamma\sin\theta, \\ &a_{20}=\cos\gamma\sin\phi\sin\theta+\sin\gamma\cos\theta, \quad a_{01}=\sin\phi, \\ &a_{11}=\cos\phi\cos\theta, \quad a_{21}=-\cos\phi\sin\theta, \quad a_{02}=-\sin\gamma\cos\phi, \\ &a_{12}=\sin\gamma\sin\phi\cos\theta+\cos\gamma\sin\theta, \\ &a_{22}=-\sin\gamma\sin\phi\sin\theta+\cos\gamma\cos\theta \end{aligned} \quad (4.1.C)$$

We make a particular choice of the phase information, consistent with the C-K-M matrix being this rotation matrix. We obtain from equation (4.1.A),

$$\mathbf{M}^{\odot\prime}= \begin{pmatrix} 0.9741-0.9756 & 0.219-0.226 & 0.0025-0.0048 \\ -(0.219-0.226) & 0.9732-0.9748 & 0.038-0.044 \\ 0.004-0.014 & -(0.037-0.044) & 0.9990-0.9993 \end{pmatrix} \quad (4.1.D)$$

We then employed a computer program, listed in appendix II with descriptive annotations. (If appendix II is not in this copy, it is to be found at http://sarahbell.org.uk.) This program was coded fairly quickly, without much in the way of statistical checks for the best fit, but the symmetry is so marked that it shines through. The program made guesses at ϕ, θ and γ and calculated trial C-K-M matrices, M, in the vicinity of these guesses, starting with a large gap between each possibility tried. The M nearest the observed

C-K-M matrix, $\mathbf{M}^{\odot\prime}$, was computed using an error function, and this was then used as the basis for a more refined search with a smaller gap between the trial **M**. This continued until a guess was made where every element of the calculated C-K-M matrix **M** lay within the boundaries given for the elements of $\mathbf{M}^{\odot\prime}$. Although the search was not guaranteed to lead to success, it was easy to gather a selection of **M** that met the criterion. A sample of eleven are shown in appendix I. We cannot claim that we have ensured they are random however for the simple method used. We can claim that the assertion that C-K-M matrix, $\mathbf{M}^{\odot}$, is a rotation matrix is well supported.

The observed C-K-M matrix, $\mathbf{M}^{\odot\prime}$, was then analysed on the assumption that it represented such a rotation matrix, **M**, as follows. As the matrix **M** describes an arbitrary rotation in three dimensional space, the rotation will lie in some one plane in the three-dimensional space. This means that there is a line orthogonal to this plane that will be unaffected by the rotation. Let us pick such a line going through the origin and a point (x, y, z). Then,

$$\mathbf{M}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.1.E}$$

or,

$$\left(\mathbf{M} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{00}-1 & a_{10} & a_{20} \\ a_{01} & a_{11}-1 & a_{21} \\ a_{02} & a_{12} & a_{22}-1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \tag{4.1.F}$$

We used the matrix $\mathbf{M}-\mathbf{1}$ to find the angle by which $\mathbf{M}$ rotates a vector in the plane of rotation. We also solved the three equations for x, y and z implied.

First, we used the scalar product, s, of two three-vectors, $\mathbf{v}_1$ and $\mathbf{v}_2$, to provide the angle. If the two vectors are lines from the origin to two points (x_1, y_1, z_1) and (x_2, y_2, z_2), it can be shown that,

$$s = \mathbf{v}_1 \bullet \mathbf{v}_2 = \begin{pmatrix} x_1 & y_1 & z_1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \quad (4.1.G)$$

$$x_1 x_2 + y_1 y_2 + z_1 z_2 = \cos\lambda \sqrt{x_1^2 + y_1^2 + z_1^2}\sqrt{x_2^2 + y_2^2 + z_2^2}$$

where the angle between the two lines at the origin is λ. Each element of the vector matrix calculated in equation (4.1.F) must be zero. But each element is of the same form as a scalar product of a vector, whose elements form the co-ordinates of a row of $\mathbf{M}-\mathbf{1}$ times the vector from the origin to the point (x, y, z). In fact, since this latter vector is orthogonal to the plane in which the rotation occurs, each row in $\mathbf{M}-\mathbf{1}$ must represent a vector, say, $\mathbf{r}_k$, in this plane. We find the angle, λ_k, by which the matrix, $\mathbf{M}$, rotates the three rows of $\mathbf{M}-\mathbf{1}$, the $\mathbf{r}_k$. All three angles, should be the same.

We then solved the three simultaneous equations between the matrix elements given in equation (4.1.F) two at a time, each solution representing a possible x, y and z, since the equations are homogeneous. Each solution should be the same. We may write two simultaneous equations of this form as,

$$ax + by + cz = 0, \quad Ax + By + Cz = 0 \quad (4.1.H)$$

and these may be solved to yield one possible solution,

$$x=\frac{Bc-Cb}{Ab-Ba},\quad y=\frac{Ca-Ac}{Ab-Ba},\quad z=1 \tag{4.1.I}$$

If we suppose the line joining the origin and the point (x,y,z) is inclined at angles α_x, α_y and α_z, to the axes, we may find these angles. They are

$$\sin\alpha_x=\frac{x}{\sqrt{x^2+y^2+z^2}},\quad \sin\alpha_y=\frac{y}{\sqrt{x^2+y^2+z^2}},\quad \sin\alpha_z=\frac{z}{\sqrt{x^2+y^2+z^2}}, \tag{4.1.J}$$

From Pythagoras' theorem,

$$\sin^2\alpha_x+\sin^2\alpha_y+\sin^2\alpha_z=1 \tag{4.1.K}$$

4.2 Results

The output from the program appears in the first appendix. We shall illustrate our examples with the last C-K-M matrix listed. The entry for each separate C-K-M matrix heads the data for each, for example, the last entry is headed,

"CKM matrix 10".

The next nine numbers give the elements of the C-K-M matrix. Matrices nought to nine are calculated from the angles ϕ, θ and γ, which have been found by the program to lead to a possible C-K-M matrix, M. Equations (4.1.C) are then used to calculate what this matrix would be. For C-K-M matrix 10, the elements of the observed matrix, $\mathbf{M}^{\odot\prime}$, are used instead. Our examples below apply to $\mathbf{M}^{\odot\prime}$.

The angles ϕ, θ and γ appear next, for example,

"PHI is 12.8559125873031 THETA is 2.41041162915475 GAMMA is 0.52895047016225".

For matrices nought to nine, these are the angles found by the program to lead to a C-K-M matrix with elements whose magnitude is inside the ranges given for $\mathbf{M}^{\odot}$ in equation (4.1.A). This means that each matrix, **M**, found is a rotation matrix of the sort we have been describing, permitted by experiment. Thus the assertion that the C-K-M matrix, $\mathbf{M}^{\odot\prime}$, represents such a rotation is well born out. For matrix 10, these angles are set to values calculated from a_{01}, a_{21}, a_{11}, a_{02} and a_{00} in equations (4.1.C) using the observed matrix $\mathbf{M}^{\odot\prime}$. These angles are dummies, since subsequent calculations for matrix 10 are made using the average value of the elements in the observed C-K-M matrix, $\mathbf{M}^{\odot\prime}$.

The three cosines of the angle of rotation in the plane of rotation, the λ_k, in equation (4.1.G) appear next, for example,

"Cosines of angle of rotation -"0.974070800927764 0.974041900279857 0.974061730146484".

We may calculate the average of our three cosines for the observed C-K-M matrix, $\mathbf{M}^{\odot\prime}$,

$$\overline{\cos\lambda} \approx 0.974058143 \tag{4.2.A}$$

The sines of the angles of inclination of the line joining the origin and the point (x, y, z) to the axes, α_x, α_y and α_z, is given next. There are three triples of these, and each triple has a number, *t*, appended to the end, where,

$$t = \sin^2\alpha_x + \sin^2\alpha_z \tag{4.2.B}$$

The example of these, still for $\mathbf{M}^{\odot\prime}$, matrix 10, is

"The Sines of angles made by components and first squared + third squared -
0.180731628073737 0.00429456732254761 0.983523073092507 0.999981556691512
0.179022140732843 0.0191380052087955 0.983658888987467 0.999633736756628
0.0419349700832273 -0.0116488848493889 0.999052431940328
0.999864303481766".

If $t = 1$, the orthogonal to the plane of rotation, (x, y, z), must lie within the plane defined by the x-axis and z-axis, while the y-axis contains a line lying in the plane of rotation. In that case $\sin\alpha_y$ must be zero from equation (4.1.K). Looking at our other data for C-K-M matrices nought to nine, we see that the values of $\sin\alpha_y$ are very scattered, while $t \approx 1$ consistently. We suggest that the $\sin\alpha_y$ represent noise and this value should be zero.

We suppose that the vector orthogonal to the plane of rotation of the matrix does indeed lie in the plane containing the x- and z-axes. This means that two-dimensional twist of x- and z- axes will transfer the rotation to a plane containing the x- and y-axes or the z- and y-axes. Then the vector orthogonal to the plane of rotation will be, respectively, the z- or x-axis. In other words, a two-dimensional rotation will align one of the quarks with the axes even after the rotation by the C-K-M matrix. This is indicative that the quark components, which lie along the axes before the rotation, always have a simple arrangement, but as this structure involves the strong coupling constant, we shall leave it to the third and last paper of our trilogy.

5. CONCLUSION

We have shown that it is very easy to calculate the weak parameters of the Standard Model under the set of assumptions we have used, which are coherent, and we obtain something remarkably like the periodic table of elements in outline. We have already shown that we may develop a quantum theory for gravity that becomes general relativity in the classical limit [Bell and Diaz 2002], [Bell 2004]. We have already linked this quantum theory of gravity to quantum electrodynamics, for which we also used the ANPA hierarchy [Bell 2005], among other considerations [Bell and Diaz 2004a&b], and both gravity and electromagnetism to the weak interaction, again using the ANPA hierarchy [Bell and Diaz 2005b]. The strong theory is the final frontier, and rather than leaving it up to Captain Kirk, we shall discuss that in the third paper of this trilogy.

REFERENCES

Aitchison, I. J. R. and Hey, A. J. G. *Gauge Theories in Particle Physics, QCD and the Electroweak Theory* Vol 2 Institute of Physics Publishing, Bristol (2004)

Bastin, T. and Kilmister, C. W. *Combinatorial Physics* World Scientific, Singapore (1995)

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. "Classical behaviour of the Dirac bispinor ," *Found. Phys.* Vol 30 No 1 (2000) 35-57; arXiv:quant-ph/0011026

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. "QED derived from the two-body interaction," *Found. Phys.* Vol 34 No 2 (2004a) 297-333; arXiv:quant-ph/0104017 & arXiv:quant-ph/0104018

Bell, S. B. M., Cullerne, J. P. and Diaz, B. M. "The two-body interaction with a circle in time," *Found. Phys.* Vol 34 No 2 (2004b) 335-358; arXiv:quant-ph/0110162

Bell, S. B. M. and Diaz, B. M. "Quantising general relativity using QED theory," To be published (2002); arXiv:physics/0211108

Bell, S. B. M. "Quantising general relativity using QED theory, an overview and extension," Proc. Physical Interpretations of Relativity Theory Conference, London (2004); arXiv:physics/0401005

Bell, S. B. M. and Diaz, B. M. "The relationship of time and space," To be published (2003); arXiv:physics/0403137

Bell, S. B. M. and Diaz, B. M. "Renormalisation and hierarchies," To be published (2004a); arXiv:physics/0406089

Bell, S. B. M. and Diaz, B. M. "Exact solutions for QED," To be published (2004b); http://www.sarahbell.org.uk

Bell, S. B. M. and Diaz, B. M. "Interdisciplinary overview: physics and philosophy," to appear, Proc. 26th International Meeting of the Alternative Natural Philosophy Association (ed. Keith G. Bowden), Cambridge U.K. (2005a); http://www.sarahbell.org.uk

Bell, S. B. M. "The ANPA combinatorial hierarchy - paper one: gravity and electromagnetism," To appear, Hadronic Journal Vol. 28 (2005)

Bell, S. B. M. and Diaz, B. M. "The ANPA combinatorial hierarchy - paper two (A): the weak interaction," To be published (2005b)

Bell, S. B. M. and Diaz, B. M. "The ANPA combinatorial hierarchy - paper two (C): tesseral hierarchies," To be published (2005c)

Bell, S. B. M. "The ANPA combinatorial hierarchy - paper three: the strong interaction," To be published (2005d)

Cottingham, W. N. and Greenwood, D. A, *An Introduction to the Standard Model of Particle Physics* Cambridge University Press, Cambridge (1998)

Greiner, W. and Reinhardt, J. *Field Quantization* Springer-Verlag, Berlin (1996)

James, R. C. and James, G. (eds) *Mathematics Dictionary* Van Nostrand Reinhold, New York (1992)

Kilmister, C "The fine structure constant," in (ed. Keith G. Bowden) *Spin: Proc. Alternative Natural Philosophy Association* Vol 25 (2003a) 1-18

Kilmister, C "The Eddington heritage - a claim for ANPA," in (ed. Keith G. Bowden) *Spin: Proc. Alternative Natural Philosophy Association* Vol 25 (2003b) 341-350

Manthey, M. "ANPA - the Alternative Natural Philosophy Association," *Cybernetics and Human Knowing* Vol 2 No 2 (1993); http://www.imprint.co.uk/C&HK/vol2/v2-2mm.htm

Parker-Rhodes, A. F. *The Theory of Indistinguishables* D. Reidel Publishing Company, London (1981)

Appendix I

Summary output of program CKM matrix

CKM matrices are printed in the array order:

$$\begin{pmatrix} a(0,0) & a(1,0) & a(2,0) \\ a(0,1) & a(1,1) & a(2,1) \\ a(0,2) & a(1,2) & a(2,2) \end{pmatrix}$$

Output of program CKM matrix program follows:

****

SUMMARIES

****

CKM matrix 0

0.9742963972007090.2209053160664790.00472976704596477

-0.225233539394791 0.973547120575662 0.0384168810620232

0.00404760835714599 -0.0384947309826928 0.999250605480506

PHI is 13.0166114946444 THETA is 2.25976090233258

GAMMA is 0.238027711573013

Cosines of angle of rotation -

0.974497624471804 0.973534168127607 0.974497599591222

Sines of angles made by components and first squared + third squared -

0.168309901740267 -0.00152154053648502 0.985732956682781 0.999997684914396

0.16830662636646 -0.00149277624677405 0.985733559913742 0.999997771619077

0.17083273878522 -0.00121870877694772 0.985299289611059 0.999998514748917

CKM matrix 1

0.974229694217227 0.219292314042789 0.00330017368773899

-0.225482457288634 0.973419140996152 0.0401601468796203

0.00584502869865174 -0.039869338886842 0.99918780589849

PHI is 13.03125 THETA is 2.3625 GAMMA is 0.34375

Cosines of angle of rotation -

0.974789757995215 0.973399747771871 0.97478814140035

Sines of angles made by components and first squared + third squared -

0.174692616998161 0.00571159939957722 0.984606452953988 0.999967377632299

0.174710595731194 0.00555561654242669 0.984604155416828 0.999969135124834

0.179972353094166 0.00634642752299369 0.983651195790177 0.999959722857695

CKM matrix 2

0.9743815173974442 0.219487190550901 0.0037479357308927

-0.224844765719695 0.973626590039027 0.0386793807788148

0.00504875060909332 -0.0385311774666001 0.999244643958788

PHI is 12.99375 THETA is 2.275 GAMMA is 0.296875

Cosines of angle of rotation -

0.974807817691192 0.973610464475897 0.974807400642965

Sines of angles made by components and first squared + third squared -

0.169203390899756 0.00291978547199578 0.985576829760533 0.999991474852798

0.169211283213035 0.00285080900546357 0.985575676709512 0.999991872888014

0.173950825799589 0.00348805494350306 0.984748162565614 0.999987833472711

CKM matrix 3

0.974557181468885 0.219667788487584 0.00473045808007929

-0.224100670508613 0.973783017174209 0.0390579689779702

0.00414602520650203 -0.0391243229886284 0.999225749180669

PHI is 12.95 THETA is 2.296875 GAMMA is 0.24375

Cosines of angle of rotation -

0.974751930862262 0.973769949112816 0.974751923305928

Sines of angles made by components and first squared + third squared -
0.171847925901341 -0.00131009494000113 0.985122618771214 0.999998283651248
0.17184503542296 -0.00128458612987637 0.985123156584473 0.999998349838475
0.174389820588151 -0.00100606844446751 0.984676179412053 0.999998987826285

CKM matrix 4
0.974433155937277 0.219811065576129 0.0045203491485799
-0.224632180446257 0.973634432590888 0.0397035915425861
0.00451753451326251 -0.0397039118945106 0.999201276651604
PHI is 12.98125 THETA is 2.33515625 GAMMA is 0.265625
Cosines of angle of rotation -
0.974692633880518 0.973620141798759 0.974692652943396
Sines of angles made by components and first squared + third squared -
0.174052259896943 -6.30285009970163E-6 0.984736416908119 0.999999999960274
0.174052244703884 -6.16936150487271E-6 0.984736419594339 0.999999999961939
0.17410406077677 -8.94288855823224E-8 0.984727259712572 0.999999999999992

CKM matrix 5
0.974606273707596 0.219517017042021 0.00459327763331297
-0.223888047858498 0.97386347582113 0.0382632001758733
0.00409307658193498 -0.0383199349061302 0.999257138735013
PHI is 12.9375 THETA is 2.25 GAMMA is 0.240625
Cosines of angle of rotation -
0.974818504680478 0.973850651787679 0.974818505104006
Sines of angles made by components and first squared + third squared -
0.168588020283044 -0.00112269768626485 0.98568596366538 0.999998739549905
0.168585576103636 -0.00110111295814317 0.985686406054207 0.999998787550253
0.170981273363475 -0.000837224459670697 0.985273922934231
0.999999299055204

CKM matrix 6
0.974642669233204 0.219293828736404 0.00477045393278567

-0.223728573877175 0.973858820587926 0.0393004172239708

0.00414638889416701 -0.0393711504002783 0.999216052700965

PHI is 12.928125 THETA is 2.3109375 GAMMA is 0.24375

Cosines of angle of rotation -

0.974826352366519 0.973845800720985 0.974826339976589

Sines of angles made by components and first squared + third squared -

0.173170744645121 -0.00140098991306448 0.984890821576848 0.999998037227264

0.173167653407217 -0.00137365863646423 0.98489140359503 0.999998113061951

0.175770562203472 -0.00109035596836569 0.984430556507943 0.999998811123862

CKM matrix 7

0.974507576134008 0.219723006437523 0.00441031934216281

-0.224313282492326 0.973762635499155 0.0383390271558533

0.00430527059983303 -0.0383509656334665 0.999255056569664

PHI is 12.9625 THETA is 2.2546875 GAMMA is 0.253125

Cosines of angle of rotation -

0.974768824835315 0.973749115273918 0.974768845553075

Sines of angles made by components and first squared + third squared -

0.168501013043894 -0.000235564473904463 0.985701452323451
0.999999944509378

0.168500472818151 -0.000230809289551689 0.985701545797272
0.999999946727072

0.169745204904756 -8.69275610980995E-5 0.985487979559082 0.999999992443599

CKM matrix 8

0.974358325737665 0.219867618766547 0.00398784644167129

-0.224951054343865 0.973608392525356 0.0385191011548179

0.0047829086629615 -0.0384284771717435 0.999249906643469

PHI is 13 THETA is 2.265625 GAMMA is 0.28125

Cosines of angle of rotation -

0.974728535932069 0.973593199602231 0.974728379136947

Sines of angles made by components and first squared + third squared -

0.168573445947821 0.00178168600330307 0.985687485421349 0.999996825594986

0.1685780000135334 0.00174183518393953 0.985686777775051 0.999996966010192
0.17357756384728 0.00238106500983477 0.984817323090055 0.999994330529419

CKM matrix 9
0.9743828579998443 0.219765890002796 0.00386713394719418
-0.224844765719695 0.973652734173649 0.03801558318155
0.00478302908638223 -0.0379112374152798 0.999269663659616
PHI is 12.99375 THETA is 2.2359375 GAMMA is 0.28125
Cosines of angle of rotation -
0.974771820695605 0.97363756026327 0.974771610713208
Sines of angles made by components and first squared + third squared -
0.166474420475949 0.00205416800278879 0.986043633781494 0.999995780393816
0.166479670541559 0.00200822212525386 0.986042842040989 0.999995967043896
0.171191278010875 0.00261823225891929 0.985234333137473 0.999993144859838

CKM matrix 10

0.97485	0.2225	0.00365
-0.2225	0.974	0.041
0.009	-0.0405	0.99915

PHI is 12.8559125873031 THETA is 2.41041162915475
GAMMA is 0.52895047016225
Cosines of angle of rotation -
0.974070800927764 0.974041900279857 0.974061730146484
Sines of angles made by components and first squared + third squared -
0.180731628073737 0.00429456732254761 0.983523073092507 0.999981556691512
0.179022140732843 0.0191380052087955 0.983658888987467 0.999633736756628
0.0419349700832273 -0.0116488848493889 0.999052431940328 0.999864303481766

The Universe as a freely generated Information System

M. A. HEATHER & B. N. ROSSITER
University of Northumbria at Newcastle NE1 8ST, UK,
m.heather@unn.ac.uk, nick.rossiter@unn.ac.uk
http://www.computing.unn.ac.uk/staff/CGNR1/

Abstract

The very accurate prediction of experimentally observed values of the coupling constants has meant concentration on the numerical *Combinatorial Hierarchy* in ANPA's 'internal process universe'. To understand further the structure and origins of the process of string generation itself needs the language of category theory. A formal representation of the database transaction can provide more insight into ANPA's *Program Universe.* The universe is an empty monoid. As an information system it needs to exist in physical form. Matter is generated in a non-local manner as strings by a natural process in the adjunction between 2-cells $F \dashv G$ where F and G are respectively the free and co-free functors.

1 Introduction

Irrespective of any mechanism like the big bang or steady state or bubbling in the latest bubble multiverse cosmology [20], there is always the question how the mechanism itself is to be derived. In category theory (CT) an origin corresponds to an initial object in a category and is given the label $\perp$ ('bottom'). The initial object of the universe as a category is the source of every other object in the universe – including the source itself, if recursive. In categorial terms there is nothing but arrows. Every object is an identity arrow, where the arrow domain and codomain are indistinguishable.

In a category that is cartesian closed there is a unique equaliser arrow [1] from $\perp$ to every object so the initial object of our universe

[1] See [23] at p.70 for *equaliser* and at p.97 for *cartesian closed categories.*

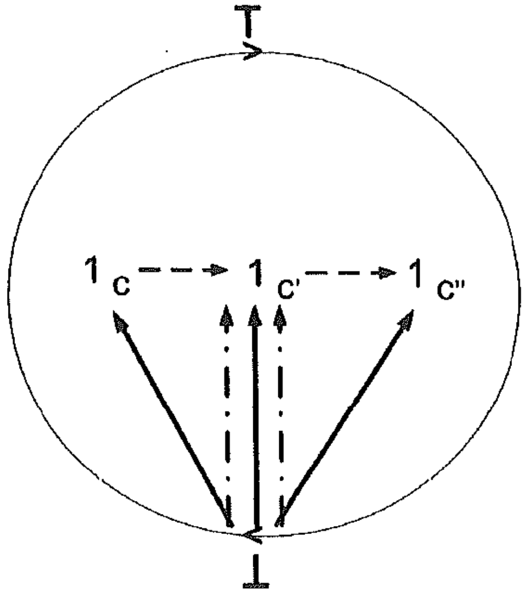


Figure 1: Initial Object.
By Composition with the broken arrows there is a unique arrow from the initial object to every object. The full vertical arrow is the equaliser of the two vertical broken arrows.

is the source of every object in it. It is cartesian closed [2] because we are only concerned with a universe that exists, that is with limits and exponentials. The exponentials are in effect all possible relationships (i.e. arrows) between objects. The equaliser is one type of limit and therefore a characteristic of cartesian closed categories. The equaliser is effectively a unique arrow between every pair of objects [3]. This is in accord with the usual definition of the universe as objects accessible to us. An inaccessible object is normally thought of as no interest to us. In physical terms the unique arrow is a resultant of all relationships. So the effect of one physical object on another is the resultant for instance of gravitational, electromagnetic and nuclear forces. However, the results of

[2]It should be noted for this paper throughout that we are concerned only with a constructive approach to reality, that which can exist. Therefore we do not need to go outside of the cartesian closed category nor resort to the category of sets. This simplifies the notation so that we do not need a gothic typeface to denote any category. In applied categories we are always operating in what a pure categorist might call formally a 'class' or enriched category.

[3]The unique resultant arrow is given in categorial terms by composition. This may be a composition diagram of two parallel arrows which as the vertical arrows in Figure 1 cannot be drawn as a triangle.

this paper suggest that a CT definition of the universe is a category of adjoint categories. The observer (if needed) is just such a category adjoint to all others and requires no special status. It also leaves open the question whether there are objects which cannot be perceived with the physical senses.

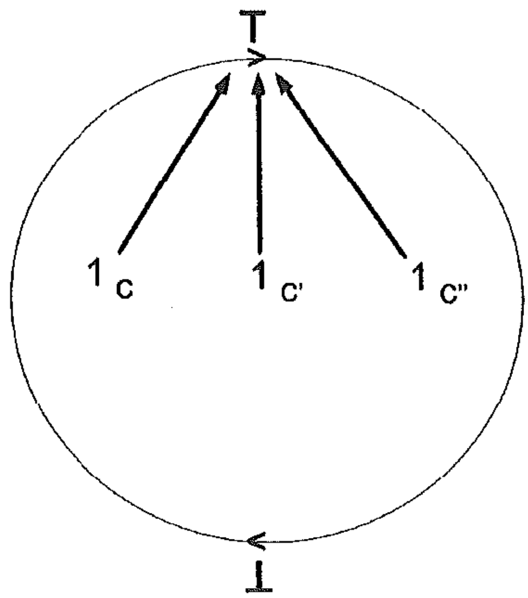


Figure 2: Terminal Object

Figure 1 shows a category that is cartesian closed and which therefore has its identity functor indistinguishable from its terminal object $\top$ ('top') and with unique arrows from its initial object $\bot$ ('bottom') to every other object. Other paths are possible as shown by the dotted arrows but any alternative path composes to the corresponding direct unique arrow. The closure of the category, $\top$, is depicted as a circular arrow because it is an arbitrary functor mapping the internal arrows onto themselves. The points are identity arrows of the category and not points in a mathematical space but much closer to the concept of a field. Figure 2 shows the terminal object with a unique arrow to it from each object. The terminal object is indistinguishable in a closed cartesian category from the identity functor, the intension of the category. The oppo-

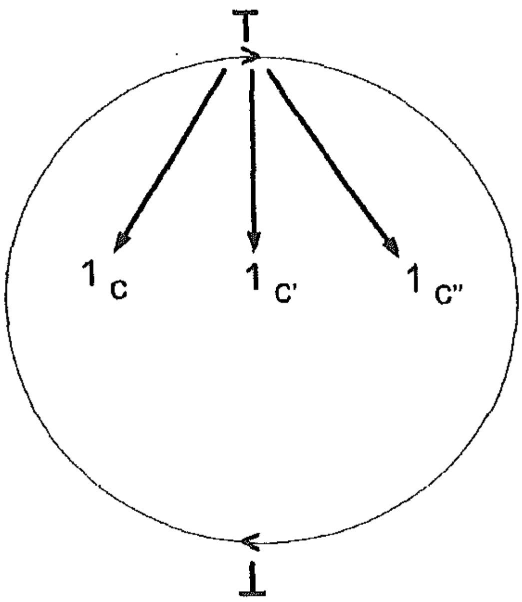


Figure 3: A category types its objects

site arrows in Figure 3 are therefore 'picking out' from the terminal object each internal object in the category, that is indicating that it belongs to that category in the sense that it is of that type. Note the general contravariant direction of the extensional typing [4] in CT as appears later. This was the seminal result of Lawvere [19] who was able to show contravariant functors to be bound up with intension/extension adjointness and logic quantification.

Earlier designations of the origin are traditionally zero (0) using number theory or the null set ($\emptyset$) in set theory. However neither of these are a true initial starting line because they are still under the starter's orders of preconditions like Peano arithmetic or the axioms of set theory. The set theoretic version of generating the universe is a hierarchy produced by taking iteratively the set of the null set and so on: $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\ldots$ These can be ordered properly as a total order by inclusion but each label is representing a set at distinguishable levels and therefore of different types. There is no natural closure except by invoking some

[4] Those approaching CT from a set theoretic perspective often want to point a typing arrow from the object to its type but it is the other way round in CT.

platonic concept like infinity. This hierarchy treats every distinguishable object as being of a different type and therefore loses the very concept of type and any notion of natural categorisation. CT on the other hand recognises the typing and has a natural closure in a four-level sandwich of three interfaces. For in the first interface identity arrows together with arrows distinguishing them make up a category that is an identity functor in the second interface. Arrows distinguishing categories are functors and in the third interface arrows distinguishing functors are natural transformations. There this categorial cumulative hierarchy ceases because arrows distinguishing natural transformations are themselves natural transformations so providing natural closure. This sandwich is shown in Figure 4. As this is completely general and requires no assumptions other than the existence of the arrow this closure may be the ultimate explanation for the limitation within the four levels of ANPA's Combinatorial Hierarchy (CH).

————→ Terminal closure with composition of natural transformations

Natural Transformations

————→ Identity Natural Transformation

Functors

————→ Identity Functors

Categories

————→ Initial Identity Arrow

Figure 4: Sandwich of category theory layers

The universe itself is therefore a one object category consisting of internal transformations (i.e. arrows). Formally this is the structure of a monoid where the one object is both initial and terminal It is represented mathematically in set theory by a structure

$< M, f >$ where $f : M \longrightarrow M$ are permutation functions on the set M [5]. The one object is the intension and the internal transformation is the extension. Nevertheless, for natural numbers, it is perhaps worth noting set theory examples for natural numbers (N), the powerset $\wp X$ of the set X and strings of words (S^*) on an alphabet S:

numbers	$< N, +, 0 >$	$< N, \times, 1 >$
sets	$< \wp X, \cup, \emptyset >$	$< \wp X, \cap, X >$
strings	$< S^*, \sqcap, \Lambda >$	

Categorically the central column represents colimits and the last column limits. Very significant is the gap for the limit monoid for the operation on strings, This is perhaps the discovery of Parker-Rhodes that there is such a fundamental operation as a process generator for the universe. In some sense the whole of this paper is concerned with this missing monoid involving an emergent process operation combining extension with intension. We need therefore to look in more detail at the concept of monoid, intension, extension and process. For this we find it convenient to go past the usual ANPA number theory CH and the generation of *Program Universe* to consider these as adjointness within an information system. A little inspection makes the monoidal structure very obvious. The first member of a triple is the underlying entity whether $\mathbb{N}$, $\wp X$ or (S^*) words on an alphabet S. The second member is one of the usual operations $+, \times, \cup, \cap, \sqcap$ (concatenation is $\sqcap$). The final member is the neutral element under the operation namely $0, 1$ the set X itself or the null string Λ.

In CT the monoid comes into its own with its full glory. It is introduced as early as page 2 of the *Categories for the Working Mathematician* of Saunders Mac Lane with the statement:

[5]see ([16] at p.66-67) where the standard *Handbook of Logic* in set theory gives the monoid only as an unnamed example although referenced *monoid* in the index.

> The notion of a monoid (a semigroup with identity) plays a central role in category theory

The structure $< M, f >$ from above is replaced by its much richer categorial version $< M, \mu, \eta >$ defined by commuting diagrams for μ and η. The significance for this paper is that μ looks back and η looks forward.

2 The Monoidal universe

The physical universe appears then to exist as an aggregate of extensional objects which can properly be formalised in CT as identity arrows. The objects are binary in the sense that the relationship between every pair of objects is reflexive and transitive. This structure is a preorder which is defined in CT in the sense that there is just one arrow between any pair of objects. A preorder is a one-object category where the binary relation defines internal subobjects. A category with a preordering is a semigroup as it is a binary operation [15, 21].

The empirical assumption-free universe consists of binary relations on potential objects. In pure mathematics this structure is an empty semigroup. However, the universe must have some identity to exist. An empty semigroup with an identity is a monoid. The intensional form of the universe is therefore a one-object monoid as already indicated. The extensional form of the universe consists of freely-generated internal objects – a free monoid. Internal objects of the universe can therefore be represented by strings generated from some sorts, that is an alphabet although atomic but not necessarily discrete. A singleton character in the alphabet $[s]$ maps under a free functor to a single character string $[s] \mapsto < s >$. A natural transformation compares the alphabet character with the character string $\eta : S \longrightarrow T(S)$ where T is a composition functor of an underlying functor G applied to some free functor F as shown in Figure 5.

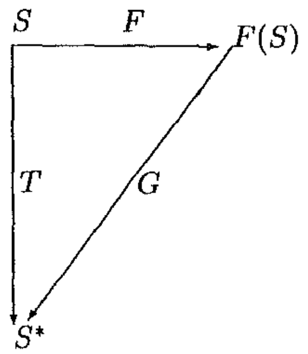


Figure 5: The whole diagram is a natural transformation arrow namely $\eta : S \longrightarrow T(S)$ representing a free/co-free structure

The old word problem was to define on a given alphabet S all possible concatenations of finite strings S^* i.e. words from the given alphabet. S^* is sometimes known as Kleene closure.

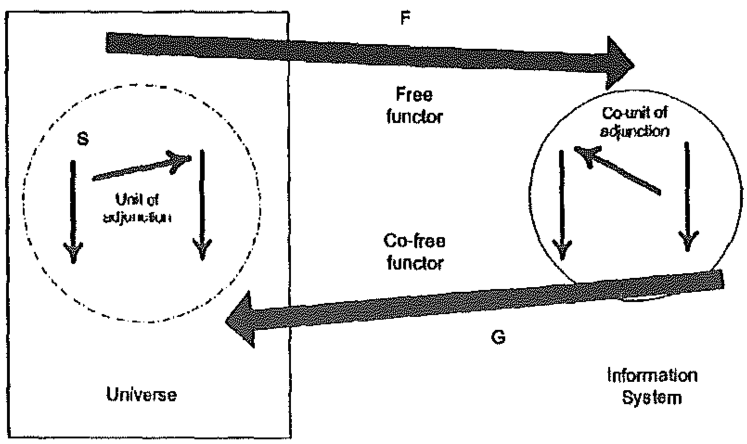


Figure 6: The View of the universe as an Underlying Information System

A tenet of ANPA is that the universe has come about from the generation of strings from something usually described as nothing. The strings form natural structures and it has been observed by Parker-Rhodes [26] that by assigning simple number bits to the strings [6] their calculation in a natural CH correlates very well with fundamental coupling constants of the universe [6]. This process has been named *Program Universe* [24]. It is the analogy of the computer program to suggest that on each cycle of the computer's internal clock, a string is generated. This builds up an information system populated by the generation of new strings distinguishable from those already generated. It is an 'internal process view' [3].

3 Intension and Extension

Intension and extension are used here in their general sense of comparing a fundamental class of identity relationships. The full abstract definition needs CT. This we have provided in Figures 1-3 above. In the case of sets the concept of membership is fundamental. A set may be identified by some label that connotes its intension e.g. 'Greek Alphabet' or by a denotational extension with elements consisting of the characters of the alphabet $\{\alpha, \beta, \gamma, \ldots, \}$ [7]. This relationship will then be relevant in any application of set theory. For example Codd's relational database model relates tables of data with an intension such as a name like 'author' where the extension consists of a pool of data values like {'Marlowe', 'Dickens', 'Shakespeare',...}. This example shows the 'elementary' limitations (that is the simplicity of elements) in set theory. There is no formal relationship between the intension other than the connection between the words 'alphabet' and 'characters' (that relationship in natural language does have a formal basis in the sense of strings that we shall come to shortly). Nor is any formal relationship, either inter-elements or intra-elements that may exist, made explicit. Nevertheless they are assumed to be disjoint.

[6]Recent work suggests that this correlation for the fine structure constant is exact to seven significant (decimal) figures [17].

[7]The convention adopted of writing a list like this as a sequence identifiable from the labels used can be misleading. Extensional elements of a set have no order.

In the example given there may be interrelationships between author names like Charles Dickens and Monica Dickens and much more complicated relationships (at the pragmatics level) between Marlowe and Shakespeare. There may be relationships within an element as when a book has more than one author. These interrelationships are outside strict set theory. Names can be coined like multisets or bags but these are really all bastard forms not part of any consistent theory.

A little more sophisticated intension/extension relationship is between the closure and the interior of open sets in topology. This, because of the relationship between intension and extension, has much more power for applications [8] for it gives relationships both within data and between the extension of the interior and the extension of the closure. Other examples are the use of instantiation where there is a suggestion that the intension is the codomain of some morphism from intension to extension and 'meaning' which goes the other way from extension to intension. A recognition that the relationship may take various forms is to be found in the object-oriented paradigm with notions like encapsulation, polymorphism, etc. In the elementary version of set theory both extension and intension are fixed. Because this restriction is not very realistic (i.e. not found naturally in reality) many attempts to relax this limitation can be found although there is seldom any attempt to justify these variations formally.

Very significant in this context are Gödel's celebrated Incompleteness Theorems and Church's undecidability result ([27] at p.599). By Gödel's first theorem extension is undecidable for axiomatic systems with arithmetic and intension likewise by his sec-

[8]For instance the Kuratowski postfix unary closure operator (defined by $A \sim \sim = A \sim$ and $0 = 0 \sim$) was used in an early definition for consciousness by one of us [8] at p.287 in ANPA9 proceedings but we have since upgraded our ideas on consciousness more on the lines of this present paper [10], [11], [13]. Kuratowski closure can also provide an alternative set-free foundation for toplogy.

ond theorem. Examples that are capable of escaping thc clutches of Gödel are the Galois connection and language. If built of fabric more general than sets these will be able to exhibit the kind of behaviour to be described in the main part of this paper. The example of language is especially important on account of its much more elaborate version of intension/extension. The structure of the intension language is syntax while the structure of its extension is the semantics. These (particularly the latter) are expressed using sets as for instance in the theory of computer programs. The great power of natural language (but not usually of artificial languages as computer programs and other comparable modelling techniques) is that there is a third layer with syntax, semantics and pragmatics where the pragmatics process as the context sensitivity of the real world.

It is not perhaps surprising that the universe has the most elaborate structure of all with the three levels of syntax, semantics and pragmatics for both intension and extension. Three levels raised to the power of two give classically a bi-cubic, as the signature of a full information system. This is the underlying structure of the double helix of the DNA and is also of the same order as the Kortweg de Vries equation [9], the generalisation of the Schrödinger equation. The universal closure (i.e. intension) is achieved with at the most three interfaces each interface consisting of an intension/extension relationship [28].

3.1 The Arrow as Process

The dimensionality of three interfaces is present in set theory by taking the intensional and extensional forms as discussed above together with functions between sets. There is a difficulty. There is no formal connectivity between the set intension and its extension. Functions relate sets but are external. With arrows on the other hand the three-dimensions are intension, extension and the direction of the arrow and all three are internal to the concept of a cartesian-closed category. This means that the arrows can better represent a concept of process. It is one of the objectives of ANPA to study the universe as process and what results from process. Bastin [3] enunciates seven principles in support of the universe as process:

- I. Process As A Necessary Principle
- II. Cumulative Sums
- III. Iteration And Algebraic Structure
- IV. Program Universe
- V. Perception
- VI. Iteration, The Statistical Background
- VII. A Parallel Development

It is illuminating to compare these headings with the arrow of CT as a process. In his first *Process As A Necessary Principle*, Bastin describes the construction

> to express algebraically the construction of successive new sets of entities out of the operations upon the elements of a previously existing set.

CT builds a geometric logic on the Bourbakian tripartism of algebra, topology and order. It still has the algebra expressed as strings but also has the further properties of topological openness and the built-in concept of order arising from the direction of the arrow. We shall see below how the arrow upgrades Bastin's algebraic constructivism to a geometrical construction of a succession of new categories of objects.

Bastin justifies his second principle of *Cumulative Sums* as follows:

> We soon found it necessary to see the constructed elements as discriminately closed subsets. To get the numbers right for experimental identification it was necessary to add those of the different stages together.

In the categorial version this adding together of stages corresponds to a colimit of limits. Iteration gives levels of an hierarchy where an intension generates an extension which names a new intension. This intension generates a further extension which adds to the last intension. Therefore we get a categorial CH consisting of colimits and limits. A limit is 'discriminately closed'.

For *Iteration and Algebraic Structure* Bastin comments

> So the hierarchy algebra appears as a set of rules which constrain the development but do not prescribe it.

This categorial version provides both for freeness (unprescribed development) and for co-freeness (the prescription of rules) formally integrated in an adjunction [9].

Bastin queries the basis of *Program Universe* and whether it can be a model:

> Was it an algebraic device merely or did it have a counterpart in the world? Sometimes it was said to be just a model, but if so what was it a model of?

The answer from CT itself seems to be that it is not a model but a reality of which the universe is an instantiation. The CH in its classical ANPA description is then a numerical model of this reality.

[9]For freeness and co-freeness see below.

The fifth principle of *Perception* means to Bastin that the generation of the CH is a construction and not a matter of filling a platonic receptacle viewed by a passive observer. The observer is then part of the construction and deconstruction in the iteration. The arrow of CT well represents this viewpoint and does not require an underlying mathematical space nor the concept of a vacuum. The arrow provides everything the vacuum provides and more. The split idempotent [10] generalises the concept of vacuum.

In his sixth principle, *Iteration, The Statistical Background*, Bastin points out the need for some understanding of 'statistical' and 'random' concepts. Statistics are right-exact concepts in CT and subobject classifiers of a topos while randomness is part of the freeness/co-freeness principle.

An appeal to *A Parallel Development* by Cahill on 'quantum foam' is made by Bastin in support of the concept of process as his seventh principle. Cahill relies on Leibniz' monad as the basic unit with the nature of a gebit (a pre-geometrical bit). This appears to be the same notion as found in the geometrical aspect of the categorial arrow where the monad induced by an adjunction can be identified with the monad of Leibniz ([14] at p.308). The arrow goes further than Leibniz and even subsumes Aristotle's comparable fundamental particle, the *entelechy* which unlike Leibniz' monad has an inbuilt direction pointer. Cahill's words:

> Process physics is a semantic informational system and is devoid of *a priori* objects and their laws and so it requires a subtle bootstrap mechanism to set it up

as cited with approval by Bastin (within a longer quote) might well have been used to sum up the categorial mechanism that fills the second half of this paper. From the rest of the quotation we see that Cahill relies on square matrices to express relational informational

[10] See [23] at p.20.

strengths in a stochastic neural network. Matrices as operators are of course the counit 'bits' of a functor in CT.

3.2 Generation of the Physical universe

It is the purpose of this paper therefore to amplify in a formal manner using CT some detail not present in the traditional presentation of the CH. One important point is to justify the generation of physical matter. In the ANPA descriptions there are two levels, the process and the data corresponding to intension and extension. But also the process is the data. *Program Universe* relies on a classical von Neumann paradigm where the process is some algorithm that operates on the data but there is a mixing together in some string of words as in a high-level language or as a bit stream in an assembler language. The database approach promoted by the ANSI/SPARC standard treats the program (usually an algorithm) as quite distinct and each may be stated independently with the programmer's meta data and data having the property of persistence. Neither seem to comprehend the spirit of the ANPA philosophy though the process is data at the same level as the data. Proponents of the ANPA approach have so far concentrated on only some aspects. Very little is available on the mechanism of the progress but it appears generally to rely on a set theoretic perspective with generation from the null set. More attention is paid to discrimination where there is process upward in a particular natural hierarchy of bit streams with levels filled by discrimination against strings already generated lower in the hierarchy. In database practice this is a generate, search, look-up, test and store recursion. While a set theoretic approach to the CH based on natural numbers has been able to provide some very compelling results consistent within a very wide range of experimental data [25], nevertheless the discussion here suggests more formal underpinning is needed for the representation of 'process' and in the way extension merges with intension in 'discrimination'. CT on the other hand is able to give some support in these areas to investigate the arrow version of the *Program Universe.* The analogy of the tick-tock clock is that the whole universe turns over in some discrete fashion from one configuration to the next in some quantum space time frame. The advantage of using CT is that we are not restricted to the limitation of set theory and are not excluding possible results of quantum mechanics and especially the general and special theory of Einstein's relativity. Our approach is perhaps to make more general that of [2] which has already yielded a statistical and algebraic alternative to classical and quantum space and time.

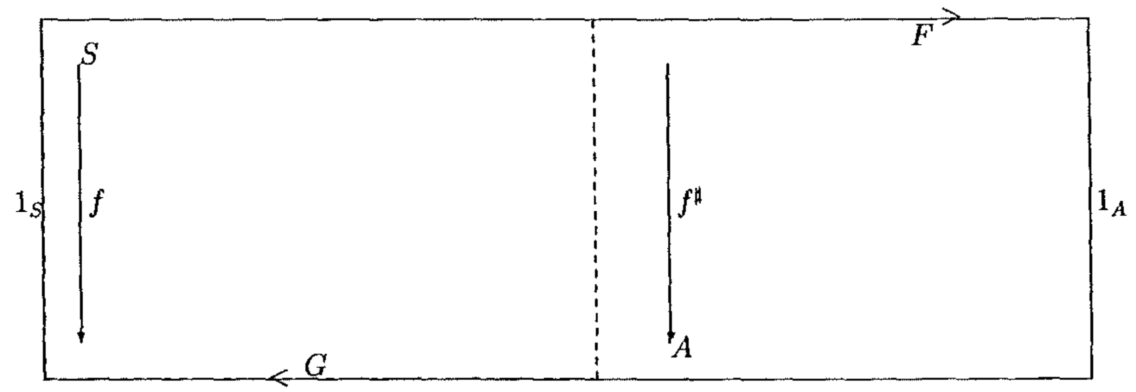


Figure 7: Correlation between Arrow f in **S** and $f^\sharp$ in **A**

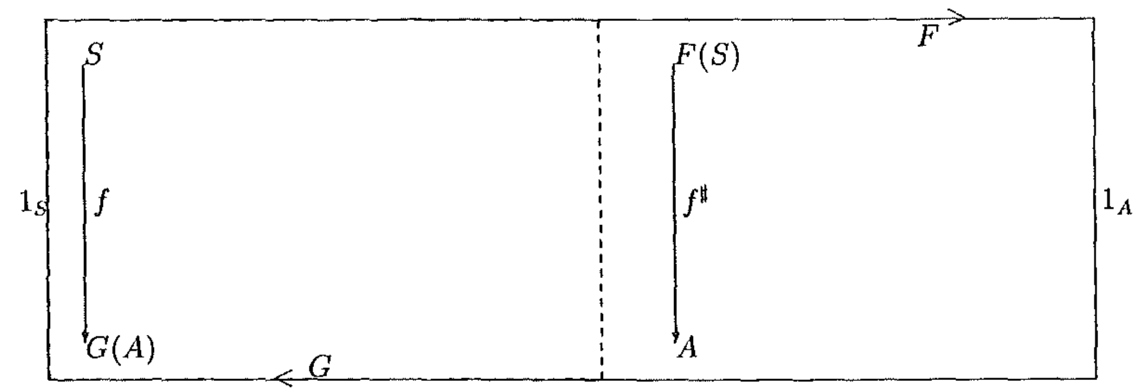


Figure 8: Identity of the universe (intension): Correlation between Arrow f in **S** and $f^\sharp$ in **A** where $\eta \longrightarrow \perp$

The example of Codd's relational model shows that the intension/extension relationship is rather obscured by the flat nature of sets without integrated functions. In categories on the other hand the focus is on arrows integral with objects and where even the objects are just (identity) arrows. Arrows between arrows give levels which are not easily identifiable in sets. In the theory and practice of databases the arrows are fundamental but they are often described by words such as 'relationships', 'methods', 'stored procedures', 'functional dependencies' or 'normalisation'. Often what are significant are arrows between arrows which is the essence of typing and therefore the fundamental use of a domain as typing. Transactions are important everyday use of database features involving arrows between arrows.

A transaction is a dynamically structured process. For instance a straightforward banking transaction requires a sophisticated relationship between crediting and debiting with resort to a fail-safe procedure. The application of everyday business rules involve interaction between intension and extension. For instance an ATM withdrawal for a class of customer may be limited to a specific value like 300€ per day. There is a strong physical component in this transaction. The customer's account cannot be debited until the bank notes have emerged from the hole in the wall in case the transaction fails to complete because of some mechanical failure. On the other hand there has to be a certainty that the amount will be debited once the money has been withdrawn despite any failure in the electronic process. This is achieved in practice by adherence to the ACID principles [11] with every withdrawal of cash being written to a transaction log before the money is paid out. Effectively the transaction log is written up prospectively in advance to a secondary file and then in the event that a particular transaction fails

[11] ACID stands for Atomicity, Consistency, Isolation, Durability [5].

to complete, it can be unpicked later by re-running from the last successful transaction to undo the steps in the log that were never fulfilled. Physical recording aspects of data in hardware are usually on disk for persistence. However, if the whole transaction were to be carried out electronically in an e-banking transaction there would still be some physical involvement because the transaction has to reside somewhere such as the hardware of the main memory. This is an example of the principle of Landauer [18]. Information cannot exist except in the physical form. The logic of an empty monoid is itself information and must be manifest in material form. This is therefore the explanation of matter in the universe.

There is also another aspect which shows up in this banking transaction, that is parallel processing. The purpose of the separate log is to provide an overlapping alternative resource in case of breakdown in the main transaction. True parallelism of the simultaneous recording of the transaction with its performance is not possible in a von Neumann architecture which relies on a sequence of processes between fixed cells. This is because classical computation is local [14] and is the reason for the failure of initiatives in the 1980s in parallel processors and the limitations of set theoretic (and therefore local) models like Petri Nets. The universe itself on the other hand is non-local processing and therefore can carry out simultaneous events although communication between them is not possible because that imports localisation. This is manifested in the cosmological limits like the finite velocity of light and the issues of the special relativity that arise from it. The operation of a separate log succeed as a way round the problem by providing two real-time systems [12]. Because of this there is no absolute time

[12]This is real-time in its true sense namely this is the origin of time. Real time is used here in the technical sense. Real time is the sequence of operations of a system. This may be synchronous as with some clocks or asynchronous. This defines a particular inherent time system, based on intensional time. So there are different possible systems of time. They have to be related through their extensions, for instance the difficulties of relating crystal time with sideral time or solar time as determined by the

on the von Neumann machine for serial processing with fixed cells which determine how a choice is made. This is a weakness of the *Program Universe* as a model and justifies the fuller explanation in information systems available in a database transaction for the generation of bit streams.

This banking transaction is typical of any transaction as a non-local process. The log provides a parallel information system to effect the banking operation consistently. It has two components, one looking forward to the sequence and one looking back to check that what was expected was achieved. The universe carries out transactions all the time non-locally mediating between objects in time and space. It nevertheless still has the forward and back components except that they are non-local. This emerges in the following analysis of adjunctions in CT. The universe operates as a quantum information processor [13].

4 Process as a Semantic Information System

A feature of a system may be a state, an action, a process, a property, indeed anything the system is or does. Outside of CT any of these are usually represented by a set, that is with unordered elements or with some imposed order like a vector or tuple. The ordering is independent of the notion of a set. In CT any feature of a system is an example of the arrow. The direction of the arrow already includes the notion of ordering and also has inherent typing so that a feature of a system is naturally distinguishable.

rotation of the earth. This is the essence of time as a local phenomenon projected out of space time under application of the axiom of choice.

[13] The banking transaction is a type of process very suited to the quantum computer. It is for this reason we have been urging recently the quantum processing of information and evolvable databases as a more realisable everyday application than some of the sensational and more esoteric examples being promoted like code breaking, teleportation and remote viewing [29, 30].

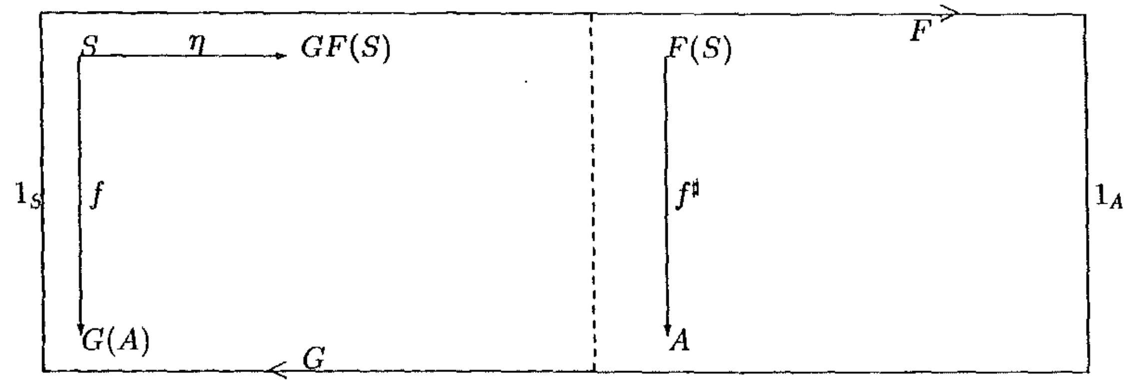


Figure 9: Working of the universe (extension): Correlation between Arrow f in **S** and $f^\sharp$ in **A** where η other than $\perp$

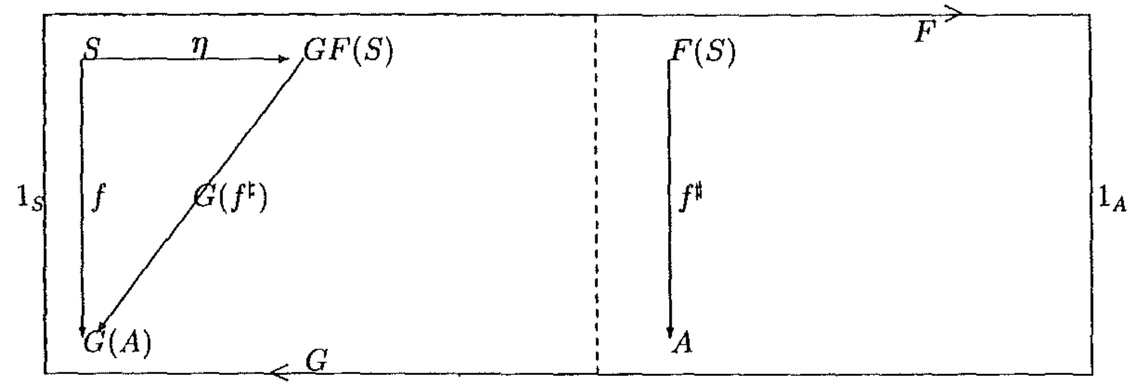


Figure 10: Distinction of $f^\sharp$ in **S** by arrow $G(f^\sharp)$

The CT information system instantiated as the universe is depicted in the Figure 6 and elaborated in the sequence of figures that follow. The instantiation as the one-object monoidal physical universe comes about as explained because natural logic is information and therefore exists in material form. Each of the figures is a topos with the usual properties that can be found in Johnstone [15] (*passim*) or modelled in higher operads of n-categories [21]. Fortunately applied categories can be restricted to the deep simplicity of nature and we need no more than the fundamental properties of adjointness that can be found in any basic textbook [14]. However, to bring out the dynamic structure of process that is intrinsic in adjointness we will set out the behaviour of the arrows step by step. The left-hand category (S) in each diagram is

[14]An example is the second edition of *Categories for the Working Mathematician* [23].

the monoidal universe. The functors F and G between S and the right-hand category (A) are endofunctors so both S, A categories are really coincident on the left but drawn side-by-side to make the relationship more patent for the overall transaction of Figure 6.

In more detail Figure 7 shows a typical arrow on the left (f in S) which is a family of arrows that correlates with a family of arrows in A which are represented in the figure by a typical right hand arrow $f^\sharp$. Correlation under adjunction is given by

$$\frac{\eta : \mathbf{1}_S \Rightarrow GF}{\epsilon : FG \Rightarrow \mathbf{1}_A}$$

The double bar indicates implication and its converse. GF is the functorial composition of applying functor G to the result of applying functor F to category $\mathbf{S}$. FG is the corresponding application of functor F to the result of applying functor G to category $\mathbf{A}$. Both arrows above and below the double bar could be replaced by the usual symbol $\leq$ for reflexive transitive ordering, an example of an arrow from where the ordering is derived as mentioned above.

The unit of adjunction is $\eta : 1_S \longrightarrow GF$ and the counit is $\epsilon : FG \longrightarrow 1_A$. If $\eta \longrightarrow \perp$, GF returns the arrow f to its original state f. That is F maps object S to $F(S)$ as G maps A to $G(A)$ as in Figure 8. If η is other than $\perp$, functor G will take $F(S)$ to a different object in S. So we have $\eta : S \longrightarrow GF(S)$ in Figure 9. Note the distinction shown in Figure 10 of $f^\sharp$ under functor G as the arrow $GF(S) \longrightarrow G(A)$ labelled $G(f^\sharp)$. Because of the uniqueness of adjunction there will be only one possible arrow $G(f^\sharp)$ given by the composition of the triangle shown in Figure 11.

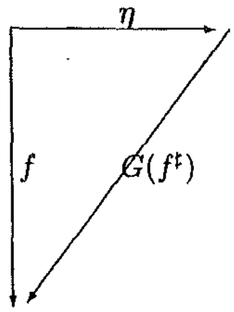


Figure 11: Uniqueness of adjunction: only one possible arrow $G(f^\sharp)$

Figure 10 is the explanation of naturality [15]. What happens to the arrow whose source object is $GF(S)$? In this case we have the dual perspective, representing co-freeness as shown in the following Figures 12 to 14. If $\top \longrightarrow \epsilon$, FG returns the arrow $f^\sharp$ to its original state $f^\sharp$. If $\top$ is other than ϵ, functor F will take $G(A)$ to a different object in A. So we have $\epsilon : FG(A) \longrightarrow A$ as in Figure 12. Note the distinction in Figure 13 of f under functor F as the arrow $F(S) \longrightarrow FG(A)$ labelled $F(f)$. In Figure 14 we introduce the correlation between an arrow g in $\mathbf{A}$ and $g^\flat$ in $\mathbf{S}$.

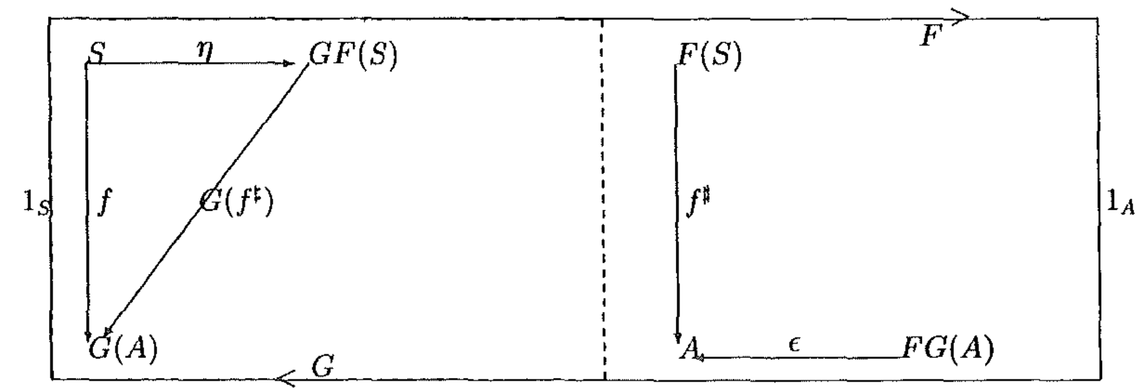


Figure 12: Correlation between Arrow f in S and $f^\sharp$ in A where $\top$ other than ϵ

In Figure 15 we show the mappings that occur in correlating g in $\mathbf{A}$ with $g^\flat$ in $\mathbf{S}$ where η is other than $\perp$ and $\top$ other than ϵ.

[15]Although not justified here this naturality is defined in the paper [28].

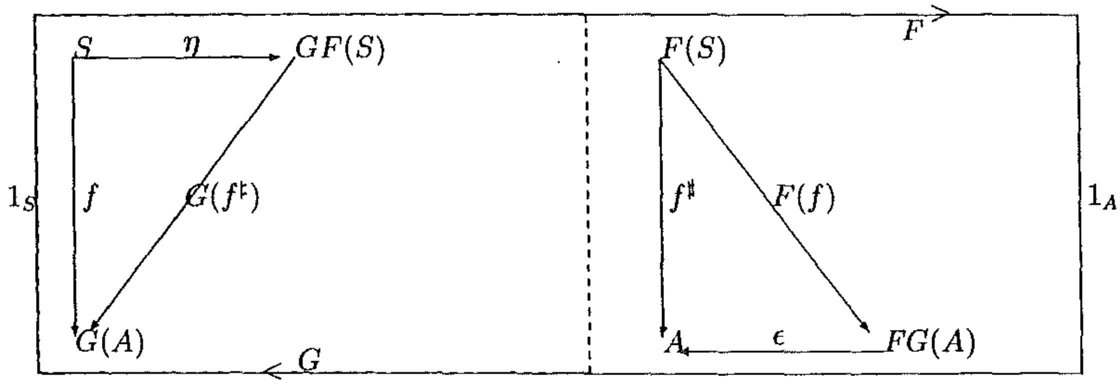


Figure 13: Distinction of f in **A** by arrow $F(f)$

In this diagram we have a general relationship where neither truth nor falsity hold [12]. The complete picture of the adjointness is given in Figure 16 to illustrate all the relevant mappings between an arrow f in A and another arrow g in S where η is other than $\perp$ and $\top$ other than ϵ.

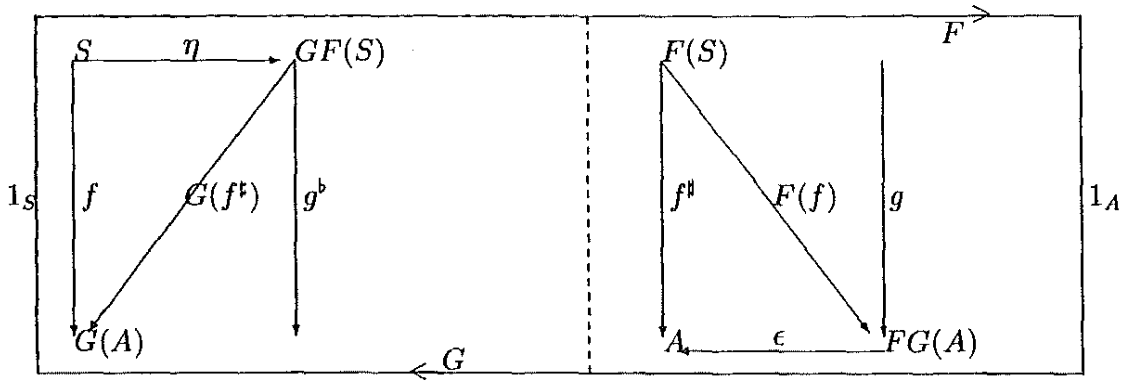


Figure 14: Correlation between Arrow g in **A** and $g^\flat$ in **S**

Figures 7 to 16 show in detail the nature of adjointness, in a manner perhaps more suited to implementation in a computer system than is the normal approach with CT in mathematics where abstraction is usually preferred. The build up is from arrows in **S** to correlating arrows in **A** for representing the freeness associated with the free functor F. The co-free functor G is the underlying functor which is critical in establishing how well **S** reflects **A**.

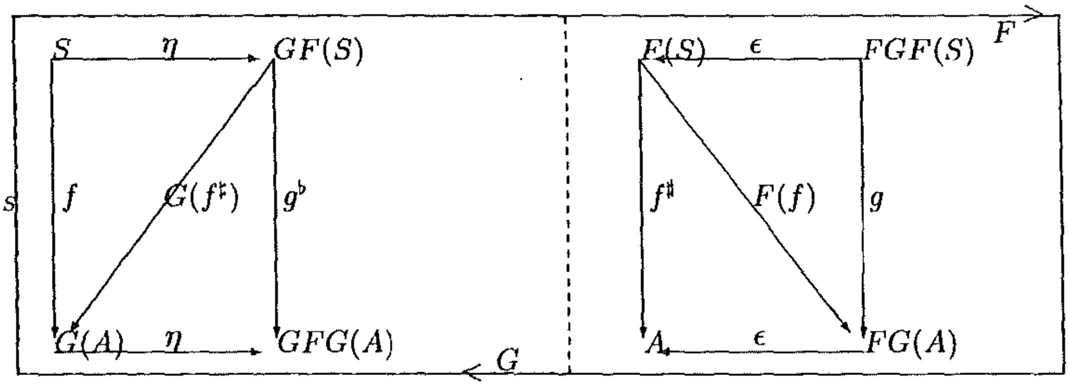


Figure 15: Correlation between Arrow g in **A** and g' in **S** where η is other than $\perp$ and $\top$ other than ϵ

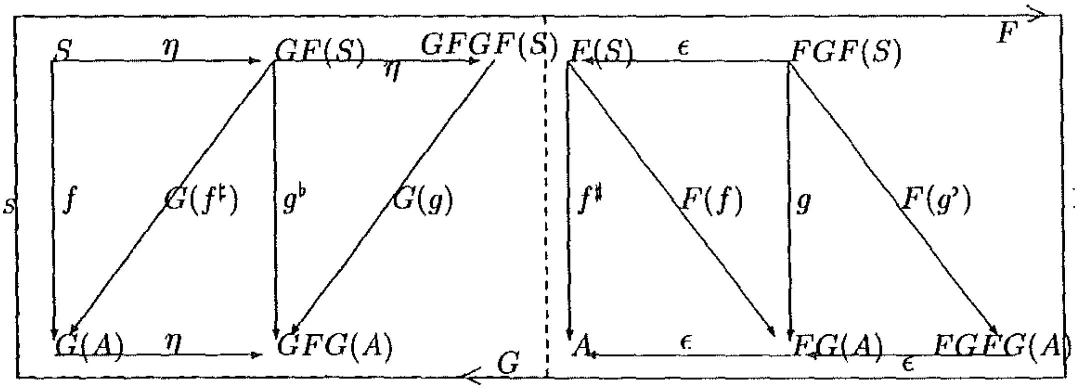


Figure 16: Complete Picture: Correlation between η and ϵ in 2-cell Adjunction $F \dashv G$

5 The Contravariant Intension/Extension Mapping

So far we have considered adjunctions which are covariant, that is domains of arrows in one category are mapped to domains of arrows in the other category. Similarly codomains in one category are mapped to codomains in the other category. The covariant case will apply across a single level such as mapping from one intension to another or from one extension to another. If we adjust our diagram in Figure 6 so that the left-hand side is the intensional universe and the right-hand side is the extensional universe, then the mapping will be contravariant. The result is shown in Figure 17 in which the arrows in the left-hand category have been reversed. In multi-level mappings, such as intension/extension,

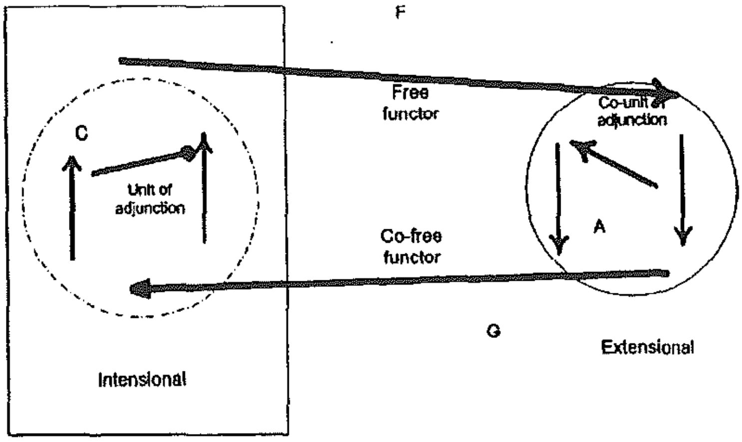


Figure 17: Contravariant Mapping between Intensional universe (left) and Extensional universe (right)

the relationships have long been known to be contravariant [19]. With a contravariant functor, domains and codomains in one category are mapped to codomains and domains respectively in the other category. In information systems the extension is of the form *value* $\longrightarrow$ *label* and the intension is of the form *label* $\longrightarrow$ *type* so that the relationship between them must be contravariant if one is to be mapped on to the other.

The diagram in Figure 17 is not unlike that for interhuman communication proposed by [22] which assumes two levels are involved. At the first level information is exchanged and provided with meaning, and at the second level meaning can reflexively be communicated.

The formal diagram in Figure 18 shows arrows reversed on the left from those in Figure 16 except for η which still compares S with $GF(S)$ and $GF(S)$ with $GFGF(S)$. This diagram shows the intensional universe on the left and the extensional universe on the right. The Galois connection [31] can be used to reason with such diagrams.

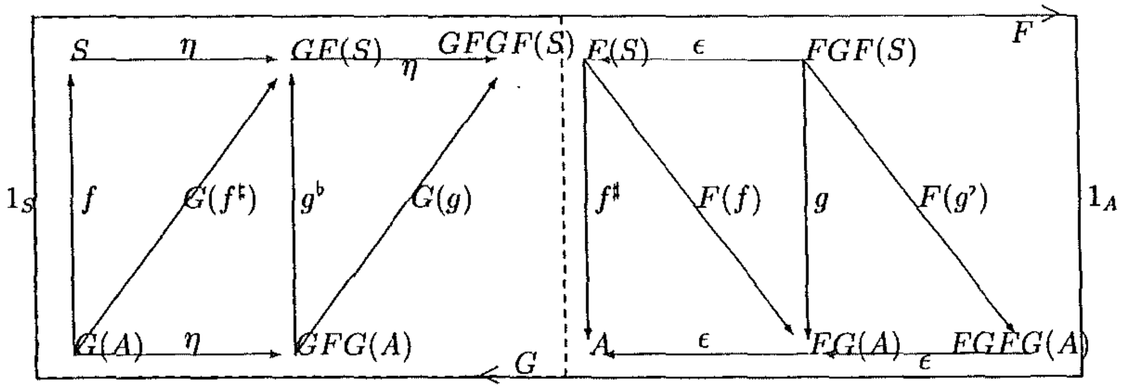


Figure 18: Complete Picture: Correlation between η and ϵ in 2-cell Contravariant Adjunction $F \dashv G$

6 Abstract Representations of Covariant/Contravariant Functors

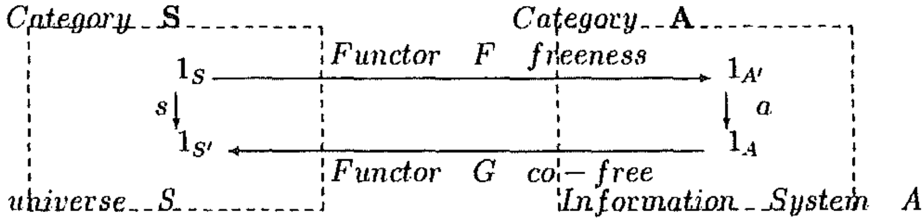


Figure 19: Covariant Mapping between universe and Information System: $\eta \longrightarrow \perp$ and $\top \longrightarrow \epsilon$

A more abstract representation (extending that in [14]) is shown in Figures 19 to 21. Figure 19 corresponds to Figure 8 where the unit of adjunction $\eta = \perp$, giving a simple equivalence between the two categories **A** and **S**. Figure 20 corresponds to Figure 10 where the unit of adjunction η is other than $\perp$ with η taking S to a different object $GF(S)$. Figure 21 corresponds to Figure 13 where for the counit of adjunction, $\top$ is other than ϵ, with ϵ taking $FG(A)$ to a different object A.

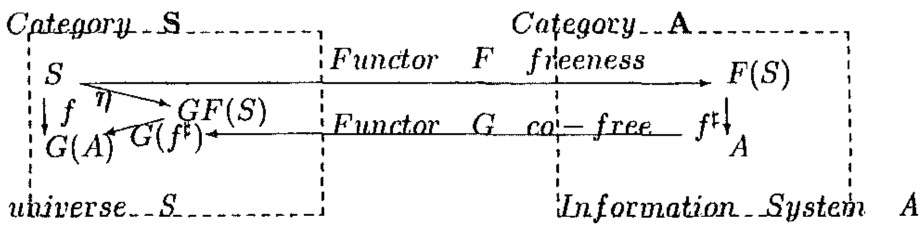


Figure 20: Covariant Mapping between universe and Information System: unit of adjunction η other than $\perp$

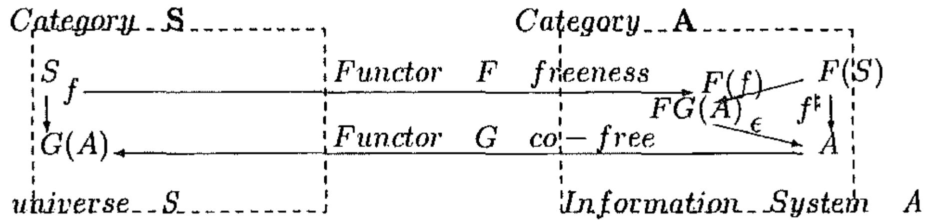


Figure 21: Covariant Mapping between universe and Information System: counit of adjunction $\top$ other than ϵ

Figures 19-21 show the covariant mapping between a universe and an information system. If we consider that category **S** is the intensional form of the universe and that category **A** is the extensional form of the universe, then the mapping between them will be contravariant as discussed earlier. Each of Figures 19-21 can be represented in contravariant form by reversing arrows, other than η, in the left-hand category **S**. Reversing the arrow $s : 1_S \longrightarrow 1_{S'}$ in Figure 19 has an apparently trivial effect as the type of the domain and codomain are the same. However, trivial effects may be of greater significance in applications than in pure mathematics. When η is other than $\perp$ or $\top$ other than ϵ the effects of reversing the arrows in category **S** are obviously of greater significance. Here we show in Figure 22 the contravariant form of Figure 20. This shows the reversal of the directions of f and $G(f^{\natural})$ with the commuting triangle now giving the equation $\eta \circ f = G(f^{\natural})$ instead of $\eta = G(f^{\natural}) \circ f$ as in the covariant form. The arrow $F(S) \longrightarrow A$ in **A** is mapped by G on to $A \longrightarrow F(S)$ to give $GA \longrightarrow GF(S)$. This is contravariant as the domain and codomain of **A** are mapped on to the codomain and domain respectively of **S**.

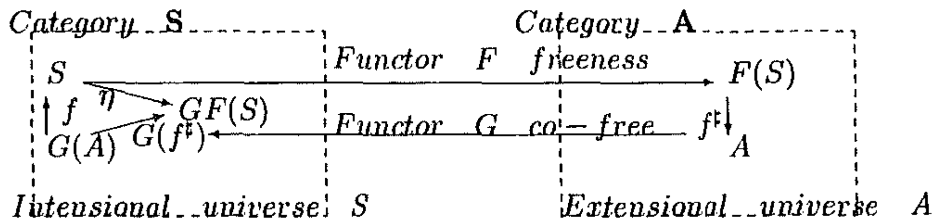


Figure 22: Contravariant mapping between intensional universe and extensional universe: unit of adjunction η other than $\perp$

For the contravariant form, if indeed $\eta \circ f = G(f^{\sharp})$, then the diagram is natural. Further when η is other than $\perp$ then the arrows are distinguishable by the application of GF to S which returns an arrow $GF(S)$ which may be different to S. The combination of contravariance, naturality and distinguishability provides a formal basis for relating the intensional and extensional universes.

7 Summary and Future Work

Let us take stock by summarising the argument. Firstly there is the monoidal universe in which the intensional universe is portrayed as a one-(object) monoid and the extensional universe as a free monoid freely-generated internal objects. This provides for the generation of strings from what is usually described as nothing.

The relationship between the universe and its representation in an information system has been represented in its most general form by covariant adjunctions, which have been built up in detail in a series of stages. When the relationship is considered instead between an intensional universe and an extensional universe, then the adjunctions are contravariant to handle the two-level intension-extension mapping. More abstract formal, natural diagrams have been developed to show both the covariant and contravariant adjunctions.

Relevant aspects not fully pursued include time dependence. This is because we are dealing with a non-local condition where neither time nor space are to be explicitly differentiated. It would not be difficult to add time if it was needed. We have shown elsewhere [12] that it is just the matter of making every time-dependent category a slice category. It appears that time is not even needed for local conditions. Contrary to earlier suggestions [7] recent high quality data from ESOs Very Large Telescope array in Chile show no evidence to support a time variation in the fine structure constant [4]. Also we have not dealt with pragmatics but that is just the need to provide context sensitivity by adding an ambient category. This is achieved by enriching every category with pragmatics to a topos.

On the other hand it will be of great interest to pursue further the emergence operator referred to at the beginning of this paper to show if it is some fundamental characteristic of a relationship between intension and extension as the phenomenon of 'discrimination' in the theory of ANPA's CH seems to suggest.

Acknowledgements: the assistance of the editor, Keith G. Bowden, is gratefully acknowledged.

References

[1] Barwise, Jon, (ed.) Handbook of Mathematical Logic, *Studies in Logic* **90** North Holland (1977).

[2] Bastin, T, & Kilmister, C, *Combinatorial Physics* World Scientific, Singapore (1995).

[3] Bastin, T, Process, *ANPA 26* August 2004, Wesley House, Cambridge (2004).

[4] Chand, H, Srianand, R, Petitjean, Patrick, Aracil, Bastien, Probing the cosmological variation of the fine-structure constant: Results based on VLT-UVES sample *Astronomy and Astrophysics* **417** 853-871 (2004).

[5] Connolly, T, & Begg, C, *Database Systems* 4th Edition, Pearson (2004).

[6] Conway, J, *On Numbers and Games* London Math Soc (1976).

[7] Davies, P C W, Time variation of the coupling constants, *Phys A: Gen Phys* 1296-1304 (1972).

[8] Heather, M A, Cybernetic Consciousness, *Cybernetica* **XXXI**(1) (1988), 7-24; also reproduced in *Proc ANPA 9* Discrete and Combinatorial Physics, H. Pierre Noyes (ed.), ANPA West California, 274-289 (1987).

[9] Heather, M A, Resonant Modes in the universe, *14th International ANPA Conference* Department of History and Philosophy of Science, University of Cambridge (1992).

[10] Heather, M A, A Formal Approach to the Hard Problem of Phenomenal Consciousness, Toward a Science of Consciousness, *Tucson II* University of Arizona, Consciousness Research Abstracts no 41, p.45 (1996).

[11] Heather, M A, & B N Rossiter, The Mind-Brain Divide in Computer Models, Towards a Science of Consciousness, *Tucson II* University of Arizona, Consciousness Research Abstracts no 247 (1996).

[12] Heather, M A, & Rossiter, B N, Object Awareness in Multimedia Documents, in: *3rd Int Workshop Principles Document Processing (PODP'96)* edd. Nicholas, C, & Wood, D, *Lecture Notes in Computer Science* **1293** Springer-Verlag 59-86 (1997).

[13] Heather, M A, & B N Rossiter, The Hypermedia Model of Consciousness, Toward a Science of Consciousness, *Tucson III* University of Arizona, Arizona, USA, 27 April-1 May 1998, Research Abstracts no 468, 188 (1998).

[14] Heather, M A, & Rossiter, B N, Locality, Weak or Strong Anticipation and Quantum Computing I. Non-locality in Quantum Theory, *International J Comp Anticipatory Sys* **13** 307-326 (2002).

[15] Johnstone, P T, Sketches of an Elephant, A Topos Theory Compendium, *Oxford Logic Guides* 43, Clarendon (2002).

[16] Keisler, H J, Fundamentals of Model Theory, in: [1] p.47-103 (1977).

[17] Kilmister, C, The Fine Structure Constant, *Spin ANPA* **25** 1-18 (2004).

[18] Landauer, R, Information is physical, *Physics Today* May 23-29 (1991).

[19] Lawvere, F W, Adjointness in Foundations, *Dialectica* **23** 281-296 (1969).

[20] Linde, Andrei, *Particle Physics and Inflationary Cosmology* Academic Press, New York (1990).

[21] Leinster, Tom, *Higher Operads, Higher Categories*, Cambridge (2004).

[22] Leydesdorff, Loet, The Construction and Globalization of the Knowledge Base in Inter-human Communication System, *Canadian Journal of Communication* **28**(3) 267-289 (2003).

[23] Mac Lane, S, *Categories for the Working Mathematician*, 2nd ed, Springer-Verlag, New York (1998).

[24] Noyes, H p, A Short Introduction to Bit-String Physics, in *Merologies: Proc. ANPA 18* T Etter, ed, 21-61 (1997).

[25] Noyes, H P, Program Universe and Recent Cosmological Results *Proc 20th Alternative Natural Philosophy Association, Aspects II*, K.G. Bowden, Ed. 192-214; also available as SLAC-PUB-8030 (1999).

[26] Parker-Rhodes, A F, The Theory of Indistinguishables *Synthese* **150** Reidel, Dordrecht, Holland (1981).

[27] Rabin, Michael O, Decidable Theories, in: [1] 595-629 (1977).

[28] Rossiter, B Nick, & Heather, M A, Data Structures in Natural Computing: Databases as Weak or Strong Anticipatory Systems, *CASYS'03, 6th Int Conf Comp Anticipatory Systems* Liége, Belgium, AIP Conf Proc **718** 392-405 (2004).

[29] Rossiter, N, & Heather,M, Evolvable Databases, *MRC Meeting on Self-Organising Systems* Edinburgh University, 24-25 May 2pp (2004).

[30] Rossiter, B N, & Heather, M A, Quantum Information Processing within the Quantum Processing of Information: The Theory and Practice Compared, *1st IEE Seminar QIP* 20 April, St Anne's College, Oxford University (2005).

[31] Stewart, Ian, *Galois Theory* 3rd edition, Chapman and Hall, New York (2004).

THE COMPREHENSIVE CATEGORIES OF LIFE

STEPHEN WOOD

21B Gwendolen Avenue, Putney, London SW15 6ET

stephenwood60@hotmail.com

17/03/2005

1. AN EXPLORATION OF THE PHILOSOPHY OF C. S. PEIRCE

I came to the ideas of Charles Sanders Peirce (1839-1914) through Stanley Salthe. I had been interested in the theory of hierarchies for some years and reading Salthe's *Evolving Hierarchical Systems* (1985) and *Development and Evolution* (1993), I was intrigued by his suggestion that Peirce's triadic philosophy held the key to such a theory. Specifically, Salthe draws attention to Peirce's theory of signs and his triad of object, sign and system of interpretance. In my PhD thesis, I had identified three stages to the process of classification and three founding relations – features, similarities and homologies (Wood, 1996: 43-44). In my paper for ANPA 23, I made the first attempt to describe these findings in Salthe's terms (Wood, 2002: 216-217).

Returning to Salthe's works, I was challenged to come to understand Peirce's thought better. Having started researching, I quickly found that it was not only Peirce's theory of signs that was important for me, but also his philosophy of the comprehensive categories. The categories of first, second and third, while finding expression in his semiotics, have much wider implications. It was to his metaphysical categories that Peirce turned when describing the key ideas of biology and each of the sciences.

In this paper, I show how Peirce's triadic philosophy – present in his categories of first, second and third and in his semiotics of sign, object and interpretant –illuminates the study of living things. Connecting Peirce's philosophy to the three levels of living organization described by Maturana and Varela (1987), I justify Peirce's conclusion that all living things have a primitive form of mind.

Turning to classification, I discuss different schools in terms of preference for different kinds of relation, monadic, dyadic or triadic. I reveal that Peirce discovered the triadic logic of cladistics almost a hundred years before Nelson and Platnick (1981). I describe the three stages of cladistic classification in semiotic terms, showing that each involves the discovery of a particular kind of sign.

2. FIRST, SECOND, THIRD

Firstness is freshness, life, freedom, immediacy, feeling, quality, vivacity, independence, being-in-itself, potentiality, concrete yet undifferentiated. Secondness is action, resistance, facticity, dependence, relation, compulsion, effect, reality, result, stability. Thirdness is mediation, synthesis, living, continuity, process, moderation, learning, memory, inference, representation, intelligence, intelligibility, generality, infinity, diffusion, growth, conduct. (These descriptions come from Esposito, 1980: 162-163.)

A second comes into relation with a first, and a third mediates between the two. The first is the beginning; the second is the end. The third is the process, the journey from one to the other. '... the whole organism of logic may be mentally evolved from the three conceptions of first, second, and third, or more precisely, An, Other, Medium' (Peirce in Hoopes, 1991: 184). 'The starting point of the universe, God, the Creator, is the Absolute First; the terminus of the universe, God completely revealed, is the Absolute Second; every state of the universe at a measurable point in time is the third' (Peirce in Hoopes, 1991: 192).

Firstness is all that is spontaneous and free, secondness is hard and resisting. Firstness is the fullness of youth; secondness, the face of experience. 'In youth, the world is fresh and we seem free; but limitation, conflict, constraint, and secondness generally, make up the teaching of experience.

'With what firstness
The scarfed bark puts from her native bay
with what secondness
doth she return
With overwreathed ribs and ragged sails.

'First and Second, Agent and Patient, Yes and No, are categories which enable us roughly to describe the facts of experience, and they satisfy the mind for a very long time. But at last they are found inadequate, and the

Third is the conception which is then called for. The Third is that which bridges over the chasm between absolute first and last, and brings them into relationship' (Peirce in Hoopes, 1991: 190). Thirdness is the maturity that denies neither freshness nor experience and incorporates both into its own habits of wisdom and thought.

A single fact, or monadic relation, is something such as 'A is white,' or 'B is large.' A dual fact, or dyadic relation, expresses a relation between two, 'A is smaller than B,' or 'C is the parent of D.' A triple fact, or triadic relation, expresses a relation between three parties which cannot be dissolved into dyadic relations. Take 'A gives B to C'. There is no act of giving if we remove the giver, the gift or the recipient. What of more complex relationships, such as 'A gives B to C in exchange for D'? These can be broken down into triadic relations: 'A makes a sale, E, to C' and 'E is the sale of B in exchange for D.' (See Peirce in Hoopes, 1991: 182.)

Surely, you will say that 'A is white' implies 'X is black' or that 'B is large' implies 'Y is small'? So easily is firstness destroyed that express it and already has it gone. Firstness captures that initial feeling of whiteness or hugeness, preceding any attempt at comparison. Firstness leaps out at us from its context, and for a fleeting moment obliterates all else with its sense of uniqueness. 'What the world was to Adam on the day he opened his eyes to it, before he had become conscious of his own existence [the uroboric state of Wilber, 1996: 48], – that is first, present, immediate, fresh, new, intiative, original, spontaneous, free, vivid, conscious, and evanescent. Only, remember that every description of it must be false to it' (Peirce in Hoopes, 1991: 189).

'A sign ...is a First which stands in such a genuine triadic relation to a Second, called its Object, as to be capable of determining a Third, called its Interpretant, to assume the same triadic relation to its Object in which it stand itself to the same Object ... Anything which determines something else (its interpretant) to refer to an object to which itself refers (its object) in the same way, the interpretant becoming in turn a sign, and so on an infinitum ' (Peirce in Liszka, 1996; see figure 1). 'The stove is black' is a sign, a firstness, a unity which can be analysed into its ground, the quality of blackness, and its object. The object comes second, prescinded of all qualities, hypothesised as that which is other to the sign. The interpretant provides the context, the way in which the sign comes to be a sign for some sign-user. The interpretant may be 'Ooh, it's an Aga' or 'My goodness, it needs cleaning.' The seemingly straightforward information provided by 'The stove is black' only flows to the user along a local logic, expressed in the interpretant. The meaning of the sign, its life, lies in thirdness, the interpretant, which is itself another sign. Signs lie within a network of other signs; they lie within an implicate sea of ideas.

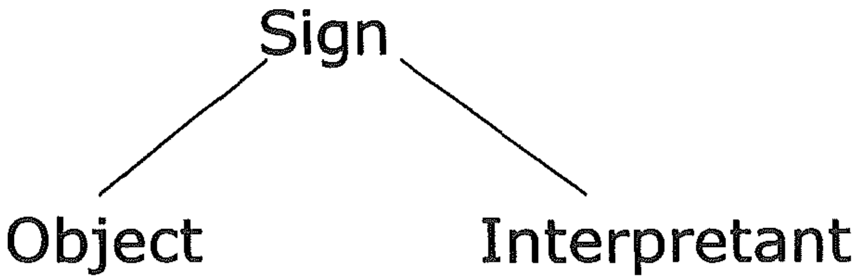


Figure 1: The triad of sign, object and interpretant

Salthe (1993: 13-15) states that objects are firsts and signs are seconds, namely encounters with objects. This does agree with what Peirce says. Salthe has lapsed back into thinking of a pregiven world of objects and has forgotten that 'each act of knowing brings forth a world' (Maturana and Varela, 1987; 29).

3. KOESTLER'S EXAMPLE

Arthur Koestler gave an early account of hierarchy theory in *The Ghost in the Machine* (1967). A journalist and acclaimed novelist, Koestler also published in experimental psychology. In *The Ghost in the Machine*, he attempted to find a third way between Gestalt psychology and Behaviourism. The Gestalt school looked for the monads of experience, wholes that could not be further broken down. Behaviourism divided all interaction into dyads of stimulus and response. A conversation, for example, can be reduced to a chain of mutually determining stimulus/response units.

Koestler (1967: 20-21) quotes a passage from a Behaviourist textbook, describing a dialogue between a man and a woman. He asks her the time. She tells him. He thanks her. She says 'Don't mention it' and he responds by asking 'How about lunch?' How, Koestler asks, can the man's request be determined by the woman saying 'Don't mention it'? And how can the two be regarded as a unit of behaviour? The woman could colour her words in so many ways – did she say the sentence briskly, brushing him off, or lingeringly, with a sexy smile? Whether the man asks her out is also very much affected by whether he finds her attractive, is free for lunch and can afford it.

Koestler has drawn attention to the interpretant and the whole network of other signs that bring to the conversation and make it alive for its participants.

4. LIFE = SIGN = MIND

Or, the Difference between a Thermostat and a Living Cell

Cells have the ability to move towards or away from light, to sense and avoid heat, and to move towards sources of food. These responses are called taxes. What is their significance?

In each cell, there is a web of molecular interactions that gives the cell its life. This dynamic web is the way the cell makes itself, perpetuates itself and defines itself, i.e. creates its own boundary. This is the autopoietic organization of the cell (Maturana and Varela, 1987: 43-47).

Different substances enter into the life of the cell in different ways. Heterotrophic organisms, such as the bacterium *Escherichia coli*, obtain energy from external sources of food. Phototrophic organisms, such as the alga *Chlamydomonas*, obtain energy from light. *E. coli* swims towards high concentrations of glucose, a molecule on which it feeds. *Chlamydomonas* orientates itself in the direction of blue-green light, but not to red light, to which it is unable to utilise.

Heterotrophic cells orientate themselves in a world according to the gradients of food sources. Phototrophic cells orientate themselves in a world according to the intensity of sources of light to which they are sensitive. Each brings forth this world through the role that these sources play in the business of living.

The chemotactic response of an *E. coli* cell moving up a glucose gradient can be modelled as a feedback mechanism, as a dyad of stimulus and response. The feedback mechanism accommodates the sign of high concentration and its object, glucose. However, it misses the interpretant, namely the web of metabolic signs that is the cell's life, and that give meaning to the organism's tactic response.

'In 1891 Peirce attributed mind of a rudimentary sort to life-slimes and protoplasm. Given their reaction to certain stimuli, he argued, they feel, possessing a primitive form of consciousness, and hence they exercise the basic functions of mind' (Merrell, 1991: 131).

The law that blue-green light evokes a phototactic response in *Chlamydomonas* and red light does not is the cell's own, grounded in the structure of its light-capturing molecules. The cell is autonomous, defining its own laws, which are consistent with its own continued existence. A thermostat is governed by a law that has been set by an external designer. The thermostat does not make itself, renew itself or define its own boundary. It has no life within which measurements of temperature become interpreted. It is just a mechanism.

5. THE TRIAD OF LIFE

According to Koestler, a cell can be seen as both an autonomous whole and a dependent part. A cell is a holon, with the two faces of Janus, one looking in as a self-assertive whole, and one looking out as an integrated part. (See the summary in Koestler, 1967, appendix 1, sections 1 and 4.)

If we look to a cell's autopoietic organization, we see how it is the mutually sustaining web of interactions that underpins the cell's autonomy. If we look to the cell's structure, we see similar physical constituents to the environment. We see also how that structure is maintained through constant exchanges with the environment. As Schrödinger pointed out in *What is Life?* (1944), the German word for 'metabolism' is 'Stoffwechsel', or exchange of stuff. Through its organization, the cell asserts itself as distinct from its environment. Through its structure, the cell is integrated into its environment. It is the process of living – which is also a process of interpreting, a process of knowing – which reconciles the two. (For the triad of organization, structure and process see Capra, 1996, particularly chapter 7.)

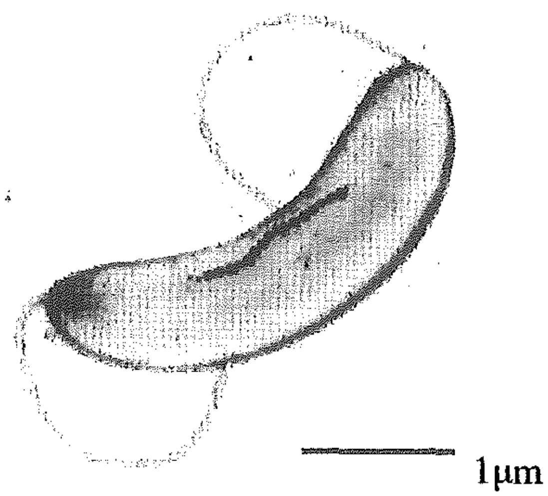


Figure 2. Electron micrograph of *Magnetospirillum magnetotacticum*.

Certain bacteria, such as *Magnetobacter* and *Magnetospirillum* (figure 2), are able to detect magnetic fields and swim in the direction of magnetic lines of force. They contain a row of particles of magnetite, which acts like a compass needle and guides the cell towards its hemisphere's magnetic pole, whether north or south. Because of the inclination of the earth's magnetic lines of force, this behaviour causes the bacterium to swim downward and thus to return to the sediments in which it lives. For an organism as tiny as a bacterium, gravity is of no consequence. So here is an alternate mechanism by which dislodged bacteria can find their way back into their normal habitat.

The physical structure of a magnetotactic bacterium is such that it interacts with the earth's magnetic field. By orientating itself with respect to the field and choosing its direction of movement accordingly, the bacterium undergoes a recurrent interaction with the magnetic structure of the environment. This is what Maturana and Varela (1987: 75) call *structural coupling*. The bacterium has nothing as complex as a representation of its environment. It does not need one; all it needs is structural coupling.

In the slime mould *Physarum* (Maturana and Varela, 1987: figure 20), spores grow into flagellate cells when conditions are moist, but into amoeboid cells when conditions are dry. The coupling with the environment involves different structural changes depending on the external trigger. When food begins to run out, the cells aggregate. Their cell membranes break down and they form a single plasmodial mass. Here we have coupling not only with the environment, but between the cells themselves. Structural changes in one cell – movement, dissolution of the cell membrane – must be synchronised with similar changes in the other cells. Here we see the birth of a new level of organization. A cell is a first-order unity, which maintains its own boundary and undergoes exchanges across that boundary. Metacellulars, such as *Physarum*, are second-order unities. Here structural transformations of the cells are coordinated into an ontogeny of the whole.

The slime moulds such as *Dictyostelium* (figure 3, cf. Maturana and Varela, 1987: figure 21) show another stage in metacellularity. Here amoeboid cells stream together in times of food shortage to form first a mound, then a slug, in which the cells move *en masse*. The slug transforms into a fruiting body to release spores and complete the life cycle. The fruiting body is raised up above the ground on a stalk. Cells in the stalk have strong walls and large vacuoles to give it strength. Cells at the top of the stalk differentiate into the spore-forming cells of the fruiting body. So beyond structural coordination, we have structural complementarity. In truly multicellular organisms, there are a large number of different cell types with complementary functions, which result from comples ontogenetic pathways. Colonies of multicellular organisms, such siphonophores and sea mats, are also second-order unities.

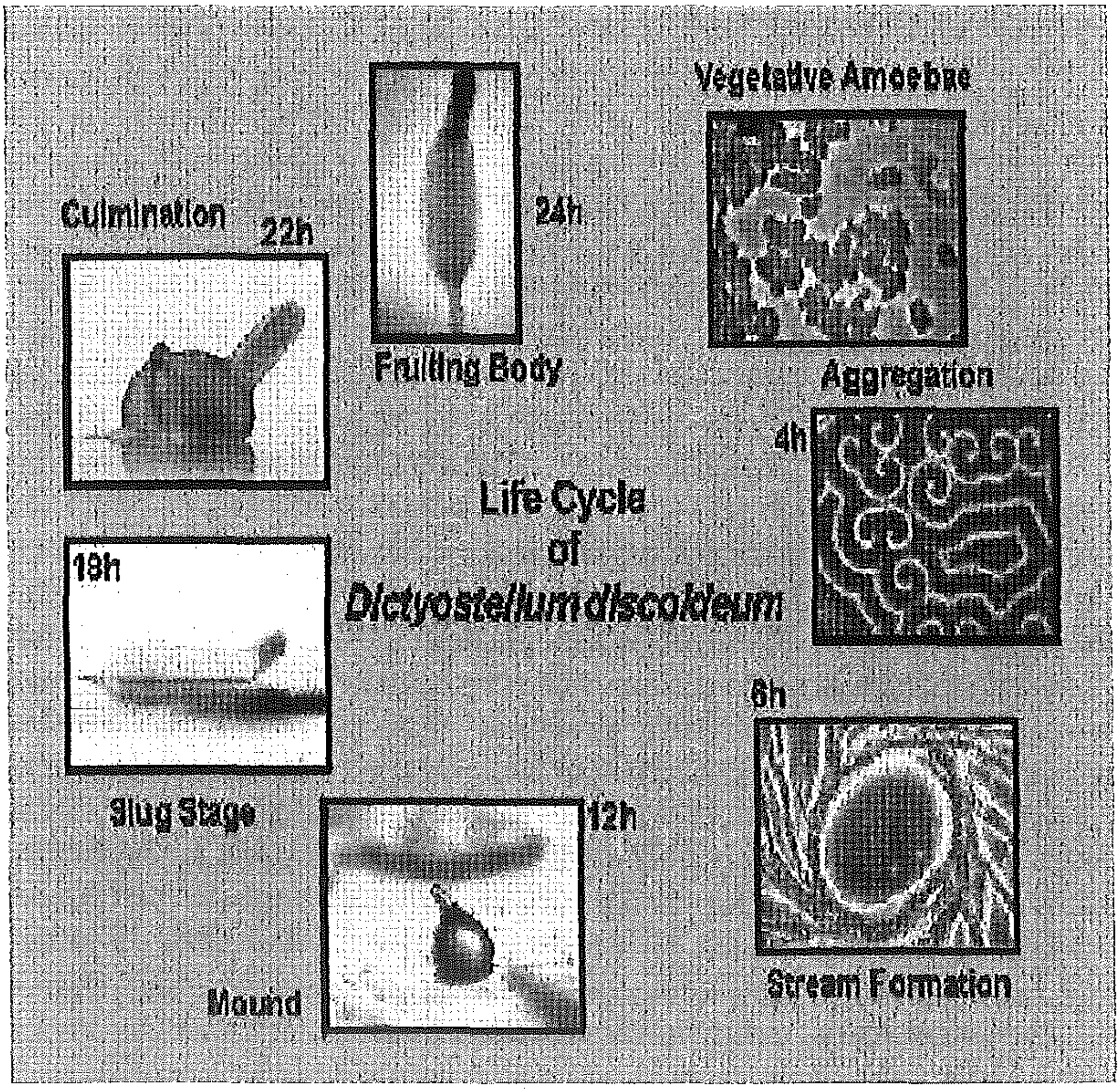


Figure 3. The life cycle of *Dictyostelium discoideum*

For a colonial organism, there is no society, since each member is identical to all others. In a family, by contrast, we have the society of male and female. Here we have two different ontogenies, which are coordinated together in the dance of sexual reproduction. Families, communities and societies of multicellular organisms represent a new level of organization, a third-order unity, defined by the co-ontogenies of its members. Insect societies demonstrate such a third-order unity most strikingly. In termites (figure 4), the immature nymphs may develop into workers, soldiers or reproductives. The reproductives develop wings and a proportion takes flight and leaves the nest to found a new colony. After a successful colonising and mating flight, the reproductives lose their wings and turn into kings and queens. Initially only a few eggs are laid and brought up by the queen herself. As the number of individuals in the colony grows, the more workers are available to help the young queen to care for the brood. Workers build the nest and galleries, they fetch food, care for the young and feed reproductives and soldiers. Soldiers defend their colony from intruders by the use of powerful jaws and/or by ejecting a white sticky repellent from an opening on their head. Soldiers cannot feed themselves; they have to be fed by workers. The reproductive and non-reproductive ontogenies are closely coordinated for the continued life of the society.

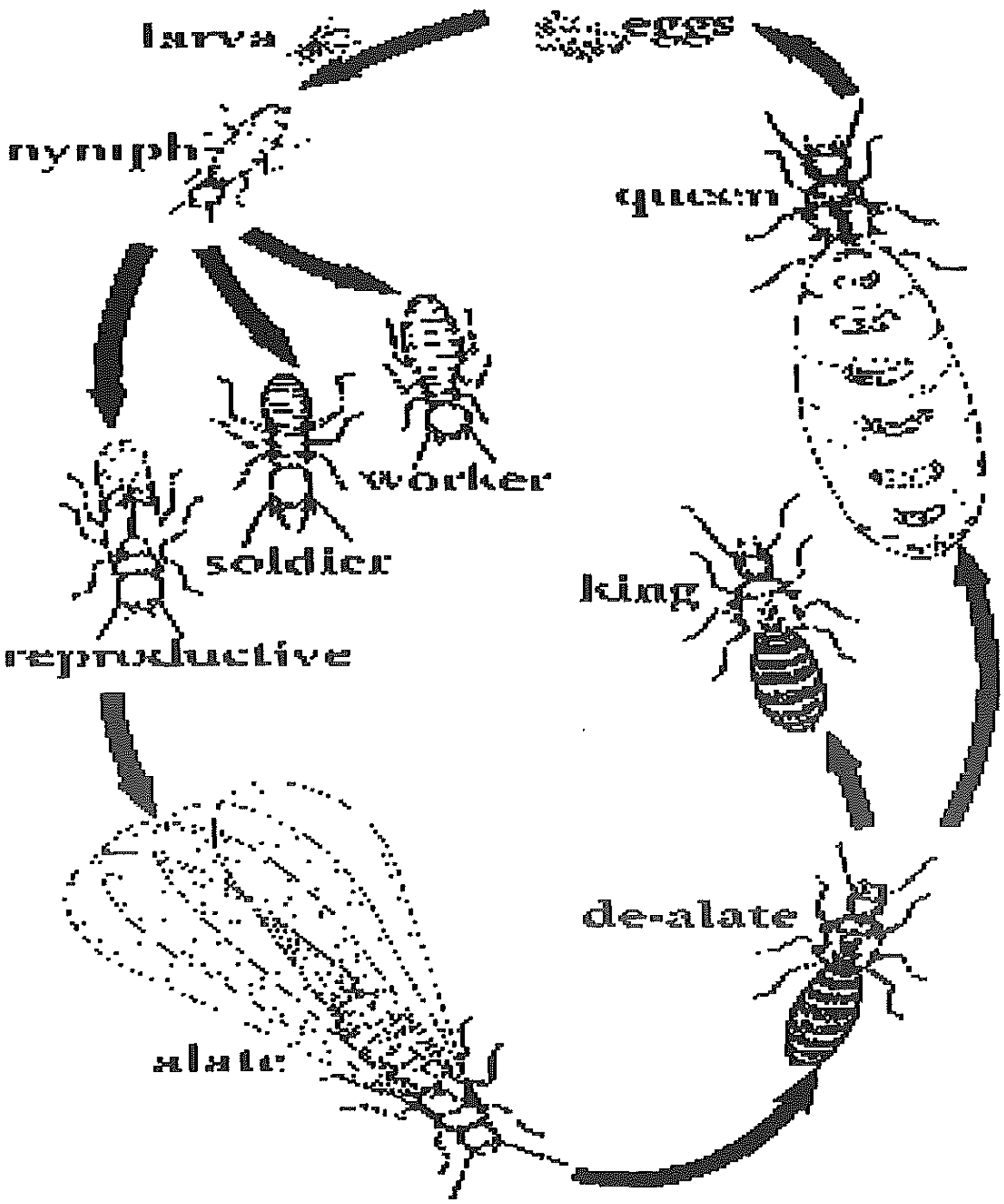


Figure 4. The termite life cycle

6. THREE KINDS OF RELATION

I have described single, dual and triple facts and how they express monadic, dyadic and triadic relations respectively. There are a number of different schools of how to classify living things. Each school shows a preference for a particular kind of relation.

Firstly, I would like to draw attention to a preference for monadic relations in classification. I call this classification by difference, or by strangeness: 'Look at X!' For example, consider the classification of gorillas, chimpanzees and humans. Johann Friedrich Blumenbach, in his *Manual of Natural History* of 1779, placed us in a separate order, Bimana, an arrangement that was followed by Georges Cuvier. Richard Owen, the great adversary of Charles Darwin in Victorian scientific circles, elevated us to a separate subclass, the Archencephala (Owen, 1858). For Blumenbach, it was our opposable thumbs that set us apart; for Owen, our enlarged brains.

This taste for the monad has influenced the classification of other groups, for example, the whales and the birds. Carolus Linnaeus, the father of modern taxonomy, placed birds as one of five divisions of animals. The whales he isolate in a separate order of mammals, the Mutica. (He was unaware that they could sing.)

A preference for dyadic relations led naturalists to say that humans are more perfect than chimpanzees or gorillas. Indeed, the whole of creation was arranged into a ladder of perfection, from the lowliest amoebae to the most elevated humans – white, European males, of course. This language of higher and lower still persists, for example, in the distinction between lower vertebrates (fish, reptiles and amphibians) and higher vertebrates (birds and mammals).

If we look at Ernst Haeckel's famous evolutionary tree (1866; see figure 5), we see a different image to the ladder of perfection. The trunk of the tree defines the axis of progress from the monera and the amoebae to humans. The labels at the side of the tree shows the grades of perfection, through which animals have passed. But the relations are not simply those of perfection; they are ancestor/descendant relations: 'X is the ancestor of Y.'

Monadic relations were broken down and replaced by dyadic relations of ancestry to provide evidence for evolution. Darwin (1859: 184) suggested that whales might be descended from a group of bears, after increasingly adventurous forays for food at the water's edge. (Today, a group of hoofed mammals, the mesonychids, are the favoured candidate.) The discovery of *Archaeopteryx* broke the isolation of the birds, showing their affinity with the reptiles, in particular the dinosaurs. And what more striking argument could there be for the link between humans and chimpanzees than the picture in Darwin's *Expression of the Emotions in Man and Animals* (1872) of an infant chimpanzee lying on its back and throwing a tantrum in the most human-like manner?

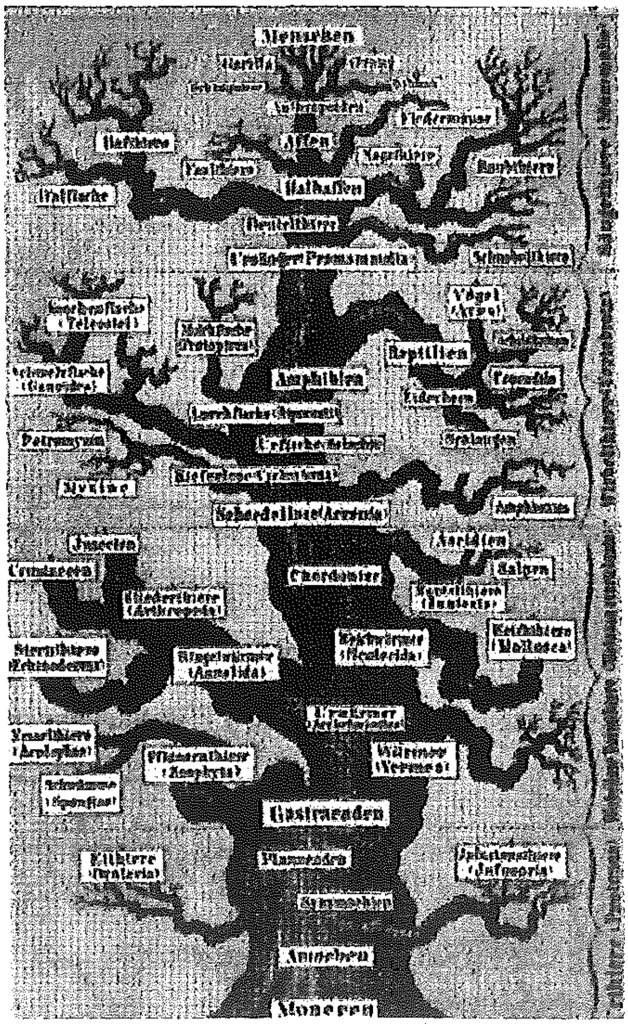


Figure 5. Haeckel's evolutionary tree

One thing you will notice about Haeckel's tree is that it has side branches. We do not have a straight chain of descent from amoebae to humans. Lying hidden here in Haeckel's iconography are triadic relations. It fell to German entomologist Willi Hennig to clarify them. He proposed classification through sister group relationships: 'X is more closely related to Y than either is to Z.' X and Y are said in this case to be sister groups (Hennig, 1966: 139). So, crocodiles are more closely related to birds than they are to other reptiles. A dyadic relation – reptiles are ancestral to birds – is replaced by a triadic relation. Similarly:

1. 'Hoofed animals are ancestral to whales' is transformed into 'Mesonychids are more closely related to whales than they are to other hoofed animals.'
2. 'Great apes are ancestral to humans' is transformed into 'Chimpanzees are more closely related to humans than to other great apes.'

This is the substance of the revolution in thought that Hennig brought about.

Hennig himself still used evolutionary language to justify the triadic relation. An ancestral species was thought to split into daughter species, each the ancestor of a particular sister group. 'Evolution in this sense (transformation) is also connected with speciation: if a species (reproductive community) is split into two mutually isolated communities of reproduction ... there is always a change (transformation) of at least one character of the ancestral species in at least one of the daughter species' (Hennig, 1966: 88). If X and Y are sisters, with respect to Z, then X and Y share a common ancestor that is more recent than either shares with Z.

As Hennig's ideas were being digested by students of classification, Gary Nelson and Norman Platnick, at the American Museum of Natural History, came to realise that they had no need of Hennig's evolutionary ontology (Nelson and Platnick, 1981). The necessity of triadic relations to classification was implicit in the logic of branching diagrams itself. They dispensed with ideas of perfection and ancestry and pared the science of classification down to the following relations:

1. Monadic: 'X exists'
2. Dyadic: 'X is related to Y'
3. Triadic: 'X is more closely related to Y than either is to Z'

The monadic and dyadic relations listed hardly qualify as the basis of classification. Any two organisms can be related in some way. Only when we introduce a third do we have a classification. Classifications of more than three organisms are to be composed of a number of triadic relations. Nelson and Platnick's cladogram isolates this triadic aspect of a classification: for an example, see figure 6.

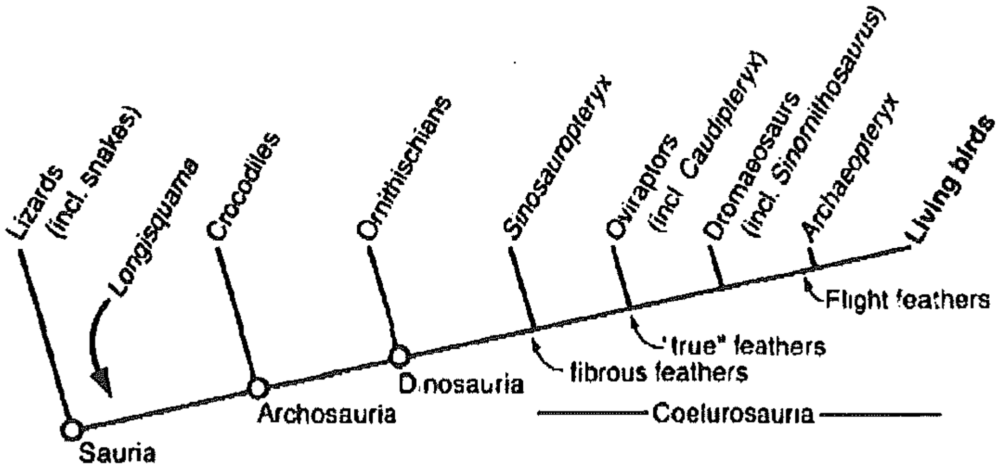


Figure 6. The cladogram of birds

The cladogram summarises the distribution of feathers in the different dinosaur groups (Padian, 2000). *Sinosauropteryx* has fibrous feathers, which form a thick, relatively short and dense covering of the entire body. True feathers, which have a central shaft, two vanes, and barbs, attach only to the forelimbs and tail. They are found in the oviraptor *Caudipteryx*, the coelurosaur *Protarcheopteryx* as well as *Archaeopteryx* and living birds. Feathers that confer the power of flight are restricted to *Archaeopteryx* and living birds, where they occur in the same pattern. In each, slightly different feathers (the primaries) attach to the hand from those (the secondaries) that attach to the forelimb (Perrins in Burn, 1980: 169).

Nelson and Platnick's conclusions would come as no surprise to Peirce. In *A Guess at the Riddle*, written c. 1890, he adopts the branching metaphor of a networks of roads to explain how all multiple facts may be reduced to triple facts. Any number of termini may be connected by roads with a fork – triadic relations – but only two termini may be connected by roads without a fork – dyadic relations.

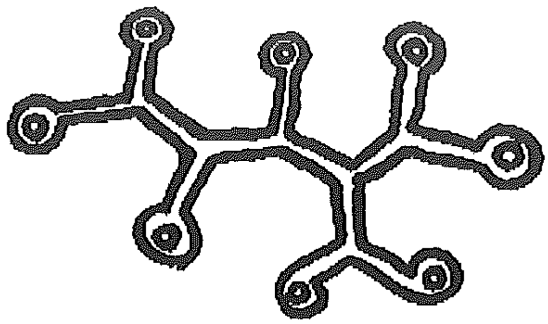


Figure 7. Peirce's diagram of roads

'See the figure [figure 7], where I have drawn the termini as self-returning roads, in order to introduce nothing beyond the road itself. Thus, the three essential elements of a network of roads are *road about a terminus*, *roadway-connection* and *branching*; and in like manner, the three fundamental categories of fact are, fact about an object, fact about two objects (relation), fact about several objects (synthetic fact)' (Peirce in Hoopes, 1991: 182-3). Peirce's fundamental categories are equivalent to the three founding relations of classification (Wood, 1996, 2002), namely features, similarities and homologies. Features are firsts; they exist in one species considered alone. Similarities are seconds; they relate one species to another. Homologies are thirds; they show that two species are more closely related to one another than they are to a third. The third species reveals the thirdness of the sister species; it provides the context within which the other two find their relationship.

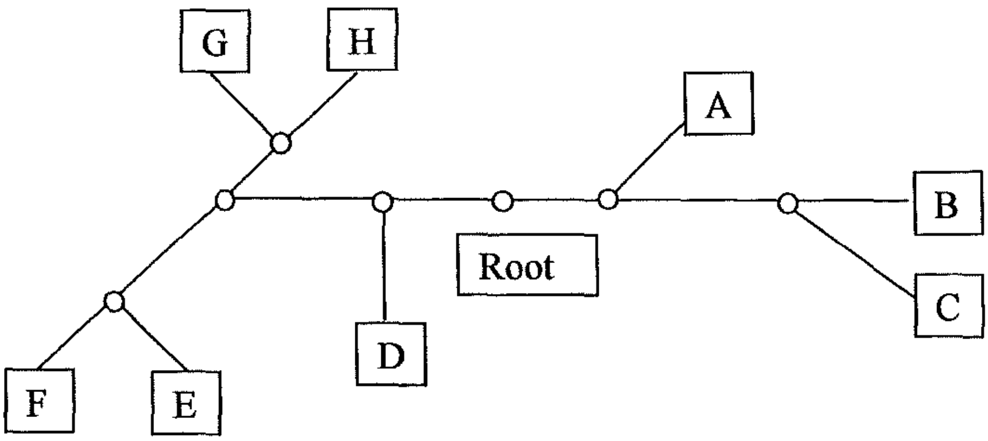


Figure 8. Peirce's diagram as an unrooted tree, with labelled termini.

```
                    !---------------- A   snakes and lizards
           !-------!                !---- B   crocodiles
           !       !-----------!
           !                    !---- C   birds
    ------!
           !       !---------------- D   platypus, echidna
           !-------!
                   !                !---- E   kangaroos
                   !      !-----!
                   !      !     !---- F   koalas
                   !------!
                          !     !---- G   whales
                          !-----!
                                !---- H   humans
```

Figure 9. A classification of vertebrates with the same topology as Peirce's diagram of roads.

Peirce's diagram of roads is equivalent to an unrooted tree. If we provide a root (Figure 8), we can identify the termini with those of a cladogram of vertebrates (Figure 9). Peirce discovered the triadic logic of cladistics almost a hundred years before Nelson and Platnick!

7. THREE STAGES OF CLASSIFICATION

In Wood (1996), I describe three stages of classification: fundamental, derivative and general. The fundamental stage of classification involves the collection of representative specimens of the species to be studied. In the derivative stage, characters are conceptualised and the character states for particular species recorded. The general stage is the generation of a classification as the most economical summary of the data and the discovery of the defining characters of taxa.

Each stage of classification involves a different kind of pattern. A fundamental pattern consists of the observed features of all morphological variants of a given species, which are at this stage not yet conceptualised. A derivative pattern is a pattern of similarity shared by a number of species. A general pattern describes the pattern of homologies inherited by organisms. Sharing is meaning in the derivative context, and congruence, the nested hierarchical relationship between patterns of similarity, is meaning in the general context.

Character concepts begin life in the fundamental stage as features identified in single species. The derivative stage of character conceptualisation is the clash between firsts. Character concepts are tested against specimens of different species, and if found not to be applicable are modified or abandoned. The general stage is the clash between seconds. Similarities that are not congruent with the most economical pattern are meaningless. They are homoplasies not homologies, confusing rather than revealing thirdness in the study group.

Each stage of classification involves the discovery of a particular kind of sign. In the context of a given stage, features, similarities and homologies are signs, in the sense of Peirce, with particular objects and interpretants. In the fundamental stage (figure 10), the sign is that a particular object specimen has a distinctive feature, for example '*Sinosauropteryx* has fibrous feathers.' The interpretant is '... as opposed to true feathers', which brings in the whole web of anatomical comparisons that embeds the study. The interpretant creates a character, a relation of exclusion: 'feathers fibrous or true.'

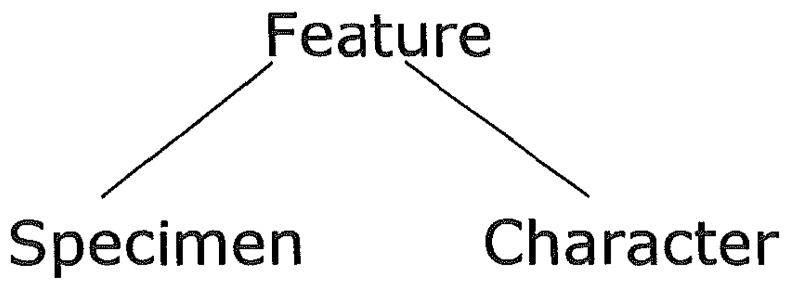


Figure 10. The fundamental stage of classification.

In the derivative stage (figure 11), different specimens are brought into relation. Hence the features of the oviraptor *Caudipteryx* and living birds signify that the two are similar in an object 'having true feathers'. The interpretant here is the whole data matrix, which described the distribution of similarities across the whole study group. This character matrix forms the basis of the cladistic analysis of relationships, often performed with the aid of computerised algorithms.

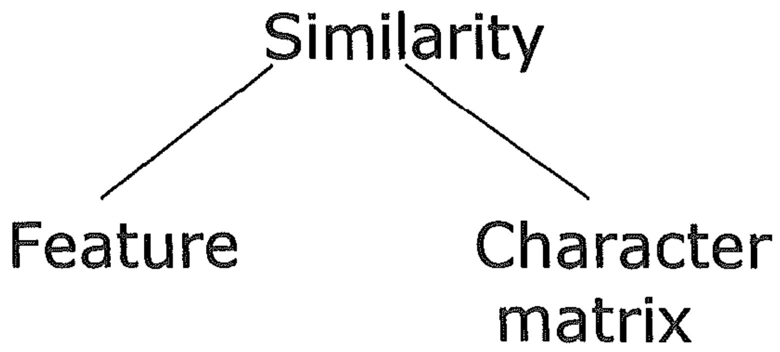


Figure 11. The derivative stage of classification.

In the general stage (figure 12), the analysis of the data matrix reveals that certain similarities function as homologies at some level of generality. In other words, these similarities identify sister groups relationships and define taxa within the classification.

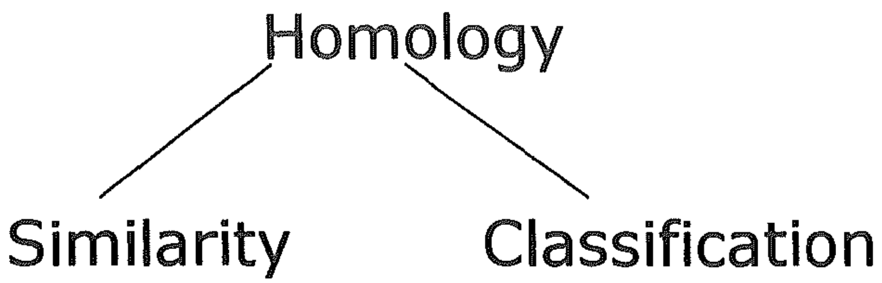


Figure 12. The general stage of classification.

The presence of feathers, whether fibrous or true, defines the Coelurosauria, which includes living birds. *Archaeopteryx* is similar to living birds in having flying feathers, but cladistic analysis reveals that 'flying feathers' is a homology, identifying *Archaeopteryx* as the sister group of living birds.

ACKNOWLEDGMENTS

I thank Basil Hiley for his comments on an early exposition of Peirce's ideas. Louis Gidney and Adam Parker Rhodes gave useful feedback to my talk at ANPA 26.

REFERENCES

Burn, D. M. (1980, ed.) *The Colour Encyclopaedia of the Animal Kingdom*. Peerage.

Capra, F. (1996) *The Web of Life: A New Synthesis Of Mind And Matter.* HarperCollins.

Darwin, C. (1859) *The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* Murray.

Darwin, C. (1872) *Expression of the Emotions in Man and Animals.* Murray.

Esposito, J. L. (1980) *Evolutionary Metaphysics : The Development of Peirce's Theory of Categories*. Ohio University Press.

Haeckel, E. (1866) *Generelle Morphologie der Organismen: Allgemeine Grundzüge der organischen Formen-wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie*. 2 vols. Reimer.

Hennig, W. (1966) *Phylogenetic Systematics*. University of Illinois Press.

Hoopes, J. (1991) *Peirce on Signs : Writings on Semiotic*. University of North Carolina Press.

Koestler, A. (1967) *The Ghost in the Machine*. Hutchinson.

Liszka, J. J. (1996) *A General Introduction to the Semeiotic Of Charles Sanders Peirce*. Indiana University Press.

Maturana, H. R. and Varela, F. J. (1987) *The Tree of Knowledge: The Biological Roots of Human Understanding*. Shambala.

Merrell, F. (1991) *Signs Becoming Signs: Our Perfusive, Pervasive Universe*. Indiana University Press.

Nelson, G. J. and Platnick, N. I. (1981) *Systematics and Biogeography: Cladistics and Vicariance*. Columbia University Press.

Owen, R. (1858) On the characters, principles of division and primary groups of the Mammalia. *Journal of the Proceedings of the Linnean Society (Zoology)* 2: 1-37.

Padian, K. (2000) Dinosaurs and birds – an update. *Reports of the National Center for Science Education* 20(5): 28-31.

Salthe, S. N. (1985) *Evolving Hierarchical Systems: Their Structure and Representation.* Columbia University Press.

Salthe, S. N. (1993) *Development and Evolution: Complexity and Change in Biology.* MIT Press.

Schrödinger, E. (1944) *What is Life?* Cambridge University Press.

Wilber, K. (1996) *Up from Eden: A Transpersonal View of Human Evolution.* Quest.

Wood, S. W. (1996) *Systematics of the Macrourid Fishes.* University of Cambridge, doctoral dissertation. iii+155 pp., 4 tables, 57 figures.

Wood, S.W. (2002) The holographic principle in biological development and quantum physics. In Bowden, K. G. (ed.), pp. 199-245, *Correlations: Proceedings of ANPA 23.* Alternative Natural Philosophy Association.

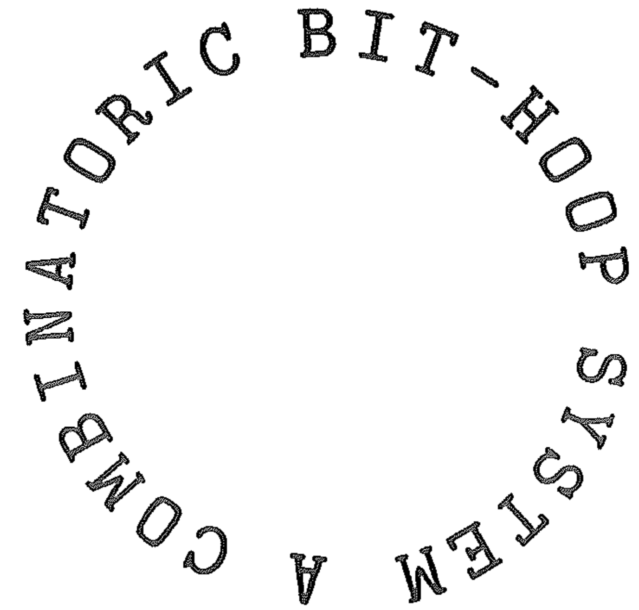
A COMBINATORIC BIT-HOOP SYSTEM

JOHN AMSON [1]
KEITH BOWDEN [2]

The Alternative Natural Philosophy Association
26th Annual Meeting : Cambridge 2004

(Draft 3.2 2004 July 31) : REVISED VERSION 4.0 2005 April 03

METSYS POOH-TIB CIROTANIBMOC A



(1) School of Mathematics & Statistics, University of St Andrews, Scotland
(5 Shore, Anstruther, Fife, KY10 3DY) john.amson@which.net

(2) Theoretical Physics Research Unit, Birkbeck College, University of London, Malet St, London WC1E 7HX k.bowden@physics.bbk.ac.uk

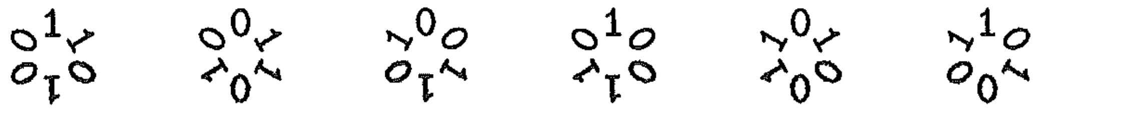
Introduction

The differences between the Title at the head of the Front Page, the one half way down, and the 'circular one' lower down is one of typographical 'ordering'. In the first two the order is linear, left-to-right in the first and right-to-left in the second; in the one below it is circular: there is no beginning and no end to the list of letters and spaces.

Throughout the history of the 'Combinatoric Hierarchy' linear ordering has played a totally commanding rôle, never questioned. It was automatically built-in as soon as the idea of 'process' was mentioned and we started developing the account in the setting of 'bit-strings' : linearly ordered lists of two-valued 'bits'.

The entire setting of the Combinatoric Hierarchy has been that of Vector Spaces. Indeed the entire machinery of Vector Spaces and the associated Linear Algebraic notions, has dominated all our thinking. Never questioned, never doubted, except in-so-far as commutativity or associativity where occasionally challenged. We really do need to ask ourselves just how far have we 'begged the question' by taking on board so much prior structure. Some structure, some degree of *a priori* structuring, is essential to assure some reproducibility, some permanence to our ideas; But the strength of any convictions concerning the results of our thinking about a Combinatoric Hierarchy must surely increase when we choose the weakest practical basis from which we start our voyage of discovery.

Five years ago one of us (John Amson) sketched the idea of abandoning the 'linear ordering' of bit-strings in favour of a less demanding, less presumptive 'circular ordering'. He called the article *Circles & Lines in the Combinatoric Hierarchy* and sent it to some ANPA friends. He referred there to *bit-loops*, but 'hoop' is possibly a better term than 'loop' since colloquially a hoop is a closed ring while a loop need not be closed. 'Hoops' are what we will use in this paper[1] . Here are some of the diagrams of hoops typical of that article, they are each cyclically-shifted clockwise by one 'petal':



Because of the problems of typing circular 'hoops' in this more figurative 'ring-of-petals' format we will use a 'linear notation' in which the elements of a hoop are enclosed between 'angle-brackets' $\langle a, b, c, \ldots, z \rangle$ where it is essential that it is not assumed that the 'a' here is to be thought of as a 'first' element. The only thing that matters is the 'cyclical ordering' ... 'b' comes ***after*** 'a', 'c' comes ***after*** 'b', ... and, eventually, 'a' comes ***after*** 'z'. The third 'title' on the front page could therefore be written now in a line as

⟨ TORIC BIT-HOOP SYSTEM A COMBINA ⟩

or equally well as

⟨ A COMBINATORIC BIT-HOOP SYSTEM ⟩.

There will be a formal definition of a 'bit-hoop' below. They are very 'combinatoric objects. We all know how many possible 'linear' bit-strings there are of length n, namely 2^n. There are fewer 'circular' bit-hoops of length n, since for example ⟨ 1oo ⟩, ⟨ o1o ⟩, ⟨ oo1 ⟩ are all the one and the same bit-hoop! How many bit-hoops there are of length n depends on counting the right number of Partitions of the integer $n - 1$, and even more

[1] Important Note added in revision: But, Reader Beware ! The terms "hoops" and "loops" have been interchanged a few times since the first draft. The related paper by Keith Bowden in these Proceedings gives some indication of the present usage. For further comments, see the "mirroring" section below (page 18). Since the original version of this paper was prepared, it has become obvious that there is a great deal of information on the closely related combinatoric topics of "necklaces" (both "cyclic" and "dihedral" (or "flip-cyclic")), going back over a century, including the combinatoric formulas and methodologies for counting them. Surprisingly, there seems to be no prior notions about algebraic operations of 'adding' or 'multiplying' those earlier 'necklaces'.

exotic combinatorial exercises, as we shall see (in Appendix **A1**).

A 'two-dimensional' bit-hoop is a cyclically toroidal object which we will call a 'bit-hoopoid'. Imagine taking a bit-matrix (a two-dimensional bit-string), rolling it up horizontally into a tube, gluing top and bottom edges together, and bending the tube into an anchor ring and gluing the circular ends together.

Like bit-strings, we need to combine bit-hoops in a suitable way. Since bit-hoops also have 'elements', an obvious way is to use 'element-wise composition. Their composition depends on what kind of elements we have: if we were to have hoops with real-number elements then we would wish to use real-number arithmetic. Similarly if we were to have hoops with elements from a finite field then we would use finite-field integer arithmetic. In the case of greatest interest here, where our hoops have 'bit' elements (*i.e.* elements from the field of two elements **2**) we would use binary arithmetic or its equivalent Boolean Logic arithmetic, as we do with bit-strings.

Unlike bit-strings which enjoy a unique 'starting position' for their elements we are at first at a loss to know which element of one bit-hoop to combine with which element of another bit-hoop. It is possible that this situation is by no means fully explored, as yet.

For definiteness, suppose all our bit-hoops have n elements. In some way we wish to be able to pick two bit-hoops and combine their elements with confidence, unambiguity, and lack of preferences. One way to do this to is to make the result of the composition of two bit-hoops not merely a single bit-hoop, but a list of separate pair-wise compositions formed by taking one bit-hoop in one hand, and presenting to it the other bit-hoop in the other hand, repeating this as many times as there are elements, but each time *cyclically shifting* the elements in the second bit-hoop by one 'position' at a time before composing it with the first bit-hoop. The result will sometimes be a list of n distinct bit-hoops, sometimes a list in which some or all of the bit-hoops are equivalent (being merely the same bit-hoop but written in a typographically different order).

Of course, one could ask "why keep the left-hand hoop static and rotate the right-hand hoop? why not *vice versa*?" The answer is simple: if the element-wise laws of composition are symmetric then it doesn't matter which one is rotated against which. We will prove that below.

The next question to ask is "what shall we do when we get such a list of bit-hoops as the result of combining two bit-hoops?". We are exploring a number of options. In an earlier draft of this paper John Amson had proposed 'picking one at random'. The introduction of randomness into a purely algebraic system has its own attractions, but the absence of a deterministic structure brings far too many difficulties.

Then Keith Bowden suggested using a traditional 'inner-product' between the elements of two bit-hoops to create a single new element, then cyclically-shifting one bit-hoop against the other and forming another element, and so on. This could be used to 'add' the two bit-hoops. We then suggested that this could be done in an alternative way to produce a 'dual' composition, analogous to 'multiplying' the two bit-hoops.

So we proposed this kind of duality: the traditional inner-product is produced as the 'sum of pair-wise products' of elements of two list; the dual inner-product is produced as the 'product of pair-wise sums'. We could call one the *sumprod* inner-product and the other the *prodsum*. The resulting laws of combining bit-hoops could then be called

sumprod-ition ('addition') and *prodsum-ition* ('multiplication')

without accidentally introducing too many pre-conceptions from traditional laws of addition and multiplication. These laws are now being used here, and replace the 'random' laws used in the first Draft. The first results seem to be encouraging, although 'addition' seems to suffer from not being 'onto' {surjective} (not every hoop of a given size can be written as the sum of two hoops of same size)[2]. This might be important for the opposite reason: what rôle is being played by those hoops that cannot be the sum of two others?

Keith Bowden next raised the question of *mirror-symmetry*[3]. For example,

$\langle$ oolo11 $\rangle$ and its mirror-image $\langle$ 11oloo $\rangle$ (*i.e.* $\approx$ $\langle$ oo11o1 $\rangle$)

are not (cyclically) equal; they are two of the fourteen distinct hoops of size 6. These two counter-examples form the smallest subset of bit-hoops to show this. We find it useful to say that a hoop which is (cyclically) equal to its mirror-image is *symmetric*, and is otherwise *non-symmetric*. The two hoops above are each non-symmetric but form an *asymmetric pair*. Census com-

[2] See the sample Addition and Multiplication Tables below

[3] 'chirality' in physics, 'handed-ness' in geometry, 'enantiomorphism' in crystallography.

putations suggest that non-symmetric bit-hoops occur increasingly frequently as hoop size increases. For example, 16 of the 60 bit-hoops of size 9 are non-symmetric, and form 8 asymmetric pairs. It is not yet clear what rôles of 'non-symmetry' and 'asymmetry' might play. Perhaps the subset of all non-symmetric hoops of a given size might be especially significant, or can be used to reduce a hoop system to a smaller system by factoring with respect to the subset of non-symmetric hoops (*cf.* normal subgroups of a group). Nor is it clear whether we should introduce the idea of a *dual hoop system*, one in which the defining cyclical order is the 'opposite' of the original cyclical order.

No study of a new system with two main laws of composition is complete without testing for Commutativity, Associativity, and Distributivity (including the less familiar self-distributivity). In the present situation we also need to examine what happens when *e.g.* we add three or more copies of the same hoop. Many of these tasks are done and a census of the results given in table form. Many more remain to be carried throughb to completion.

Forty years ago the late Frederick Parker-Rhodes initiated the notions that gave rise to the Combinatoric Hierarchy of bit-strings in a discussion paper that was eventually published posthumously in an edited, corrected version[4] . The basis of his ideas was the creation (by a primitive 'observation & discrimination' process) of bit-strings (linear vectors over the field of two elements) and a way of breaking-up a bit-string to create a bit-matrix from its elements, and, conversely, to re-assemble the elements of a bit-matrix into a bit-string. In this way he could pass from one 'Level' in his system up to and down from another 'Level' according to whether an object at one level was regarded as a bit-string or a bit-matrix.

Of course there was a lot more to it than this, and subsequent developments of his ideas by Clive Kilmister, Ted Bastin and John Amson brought into play a vast paraphernalia of linear algebraic and group theoretic concepts.

---

[4] *Hierarchies of Descriptive Levels in Physical Theory*, by A.F.Parker-Rhodes & J.C.Amson, *Int.J.General Systems*, , 1998, Vol.27(1-3), pp.57-80. *Guest Editor:* Keith Bowden

In particular, objects in one (vector space) Level are identified with objects in a lower Level (a vector place of lower dimension) by regarding each higher level object as linear matrix operator identifiable with the set of its 'fixed-vectors'[5] in the lower level.

The motivation behind this present paper is our attempt to re-work Frederick's basic ideas in the context not of bit-strings and bit-matrices but of bit-hoops and bit-hoopoids.

In this paper we take the first steps on a long march.

There is a great deal of novel machinery to setup and to get working before we can properly address the main task.

[5] Frederick happened to use the notion of 'eigenvector' where we now use 'fixed-vector', but of course these are essentially one and the same thing when working with bit-strings and bit-matrices.

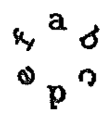
§ 1. HOOPS and HOOPOIDS

hoops & bit-hoops

HOOP By a *hoop, of hoop-circumference n,* we mean a family $(a_k)_{k \in \mathbf{K}}$ of *elements* a_k belonging to some non-empty set $\mathbf{A}$, indexed by an index-set $\mathbf{K} = \{1, \ldots, n\}$ of natural numbers equipped with a cyclical order $\prec$ such that $k \prec k+1$ for all $k = 1, \ldots, n-1$ and $n \prec 1$.

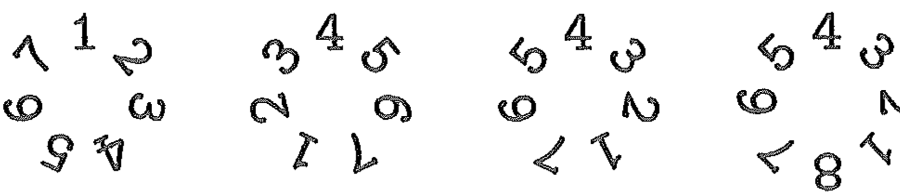
For brevity we use the alternative 'size' for 'hoop-circumference'.

We may use a single symbol to denote a hoop, *e.g.* $\mathbf{a} = (a_k)_{k \in \mathbf{K}}$, and we may use a special 'angle-bracket in-line notation' such as $\mathbf{a} = \langle\, a_1, a_2, a_3, a_4 \,\rangle$ to display the elements the elements explicitly (where practical), or occasionally John's previous 'petal-notation' such as



Whenever we write a hoop as an inline list of elements we shall not usually use separating commas or spaces between those elements, unless there is a risk of ambiguity or a lack of clarity; for example we write h = ⟨ xyz ⟩ for h = ⟨ x, y, z ⟩ or h = ⟨ x y z ⟩.

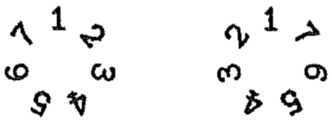
'*Hoop equality*' : Two hoops are *equal* if and only if they have the same size, same index set, same cyclical order, and same family of elements. For example, using the unambiguous 'petal notation' the first two of these hoops are 'equal', the last three are not...



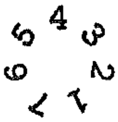
(why not?).

By a *constant* hoop we mean a hoop all of whose elements have one and the same value.

By the *mirror-image* of a hoop we mean the hoop whose elements are those of the original hoop but indexed in the 'opposite order'. In effect, the cyclical order $\prec$ on the index set $\mathbf{K}$ is replaced by the 'opposite cyclical order' $\succ$ such that if $k \succ k+1$ for all $k = 1, \ldots, n-1$ and $n \succ 1$. For example, the hoops $\langle\, 1\ 2\ 3\ 4\ 5\ 6\ 7\ \rangle$ and $\langle\, 1\ 7\ 6\ 5\ 4\ 3\ 2\ \rangle$ *i.e.*



are the mirror-images of each other. The latter could of course have been written (with an inconsequential cyclical shift) as $\langle\, 4\ 3\ 2\ 1\ 7\ 6\ 5\ \rangle$ *i.e.*



'Forming the mirror-image' is obviously a symmetric operation. A hoop that is equal to its mirror-image is said to be *symmetric*, and otherwise *non-symmetric.*

If a hoop is non-symmetric then that hoop and its (unequal) mirror-image are said to form an *asymmetric pair* (of hoops).

BIT-HOOP By a *bit-hoop* we mean a hoop whose elements belong to the two-set $\mathbf{A} = \{\mathrm{o}, 1\}$, thus a bit-hoop is a cyclically-ordered list of os and 1s.

Here are six examples of bit-hoops:

(1) $\langle$ ooo1ooo11oo1ooo1o $\rangle$, an example with 'hoop-circumference' 16 (*i.e.* with 16 elements), having *Eleven* os and *Five* 1s, *i.e.* 'Hamming-norm' 5;

(2) $\langle$ 11111o11 $\rangle$, with hoop-circumference 8 and Hamming-norm 7;

(3) $\langle$ o11o $\rangle$, with hoop-circumference 4 and Hamming-norm 2;

(4) $\langle$ o1o1 $\rangle$, with hoop-circumference 4 and Hamming-norm 2.

(5) $\langle$ oooo $\rangle$, constant, with hoop-circumference 4 and Hamming-norm 0.

(6) $\langle$ 1111 $\rangle$, constant, with hoop-circumference 4 and Hamming-norm 4.

In any such bit-hoop the starting position for writing down the elements is immaterial, the ordering is cyclical and the relationship of the 'last' element to the 'first' element in the list as written is the same as between any other intermediate element and its 'successor' to the 'right'. The example (3) could be written in any one of the four equivalent forms

$$\langle \text{o11o} \rangle \quad \langle \text{oo11} \rangle \quad \langle \text{1oo1} \rangle \quad \langle \text{11oo} \rangle$$

(using a 'shift-right with end-around-carry', in Signal Processing terms), and example (4) in either of the two equivalent forms

$$\langle \text{o1o1} \rangle \quad \langle \text{1o1o} \rangle$$

We also see that the Hamming-norm is not sufficient to distinguish between these two different bit-hoops of hoop-circumference 4 and Hamming-norm 2.

All cyclical-shifts of a constant hoop are indistinguishable.

hoopoids & bit-hoopoids

HOOPOID By a *hoopoid, of hoopoid-circumferences* n, m, [*i.e.* an $n{\times}m$ *hoopoid*] we mean a family $(a_{k,j})_{k,j \in \mathbf{K}\times\mathbf{J}}$ of $n \times m$ *elements* $a_{k,j}$ belonging to some non-empty set $\mathbf{A}$, indexed by an $n \times m$ index-set $\mathbf{K} \times \mathbf{J}$, $\mathbf{K} = \{1, \ldots, n\}$, $\mathbf{J} = \{1, \ldots, m\}$, of pairs of natural numbers equipped with a 2-dimensional cyclical order $\lll$ such that $(k, j) \lll (k', j')$ if and only if $k \prec k$ or $j \prec j'$ By a *column* resp. a *row* in a hoopoid $(a_{k,j})_{k,j \in \mathbf{K}\times\mathbf{J}}$ we mean a subfamily consisting of those elements for which the RIGHT resp. LEFT index is held fixed.

Plainly, each column is a hoop of size n (there are m columns), and each row is a hoop of size m (there are n rows). The cyclical order $\prec$ in each column-hoop or row-hoop is the restriction of the double-cyclical-order $\lll$ to the column or row subset. Again, for brevity we use the alternative 'sizes' for 'hoopoid-circumferences'.

Here is an example of a hoopoid of size 2×3 : $\left\langle \begin{matrix} \text{abc} \\ \text{def} \end{matrix} \right\rangle$

In any such hoopoid the starting position for writing down the elements is immaterial, the ordering $\lll$ is *toroidally-cyclical.* Indeed, the following 6 hoopoids are all equivalent under 'toroidal-cycling' :

$$\left\langle \begin{matrix} \text{abc} \\ \text{def} \end{matrix} \right\rangle \approx \left\langle \begin{matrix} \text{bca} \\ \text{efd} \end{matrix} \right\rangle \approx \left\langle \begin{matrix} \text{cab} \\ \text{fde} \end{matrix} \right\rangle \approx \left\langle \begin{matrix} \text{def} \\ \text{abc} \end{matrix} \right\rangle \approx \left\langle \begin{matrix} \text{efd} \\ \text{bca} \end{matrix} \right\rangle \approx \left\langle \begin{matrix} \text{fde} \\ \text{cab} \end{matrix} \right\rangle$$

In effect, each hoopoid here is formed from another either by cyclically 'rotating' all columns laterally or by cyclically 'rotating' all rows vertically. (Louis Gidney suggested visualizing an O-ring being rolled up/down a broomhandle and the broomhandle being rotated clock/anticlock-wise.)

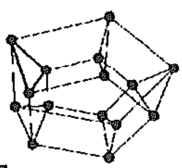
Just as we can decompose a matrix into a list of columns or a list of rows we can decompose a hoopoid into a list of 'column hoops' or a list of 'row-hoops', as in the example 2×3 hoopoid above :

$$\left\langle \begin{matrix} \mathsf{abc} \\ \mathsf{def} \end{matrix} \right\rangle \equiv \left\langle \begin{matrix} \langle\, \mathsf{abc} \,\rangle \\ \langle\, \mathsf{def} \,\rangle \end{matrix} \right\rangle \equiv \left\langle \left\langle \begin{matrix} \mathsf{a} \\ \mathsf{d} \end{matrix} \right\rangle \left\langle \begin{matrix} \mathsf{b} \\ \mathsf{e} \end{matrix} \right\rangle \left\langle \begin{matrix} \mathsf{c} \\ \mathsf{f} \end{matrix} \right\rangle \right\rangle$$

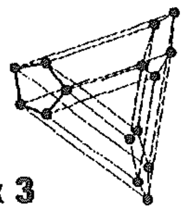
with an obvious but acceptable abuse of notation. But beware: this notation does not mean that you can cyclically rotate any individual 'column' independently of the other columns, or any individual 'row' independently of the other rows.

> **"For 'hoopoids' don't just think 'matrices'**
> **— think toroidal 'anchor-rings' instead."**

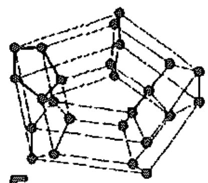
Here are diagrams of the toroidal frameworks for hoopoids of three different sizes:-



3 x 5



5 x 3



5 x 5

BIT-HOOPOID A *bit-hoopoid* is a hoopoid whose elements belong to the two-set $\mathbf{2} = \{\mathrm{o}, 1\}$, thus a bit-hoopoid is a toroidally-ordered double-list of os and 1s.

Here is an example of a bit-hoopoid: $\left\langle \begin{matrix} \mathrm{o11o} \\ \mathrm{1o1o} \end{matrix} \right\rangle$ with hoop-circumferences $m = 4$, $n = 2$ and Hamming-norm 4. For a (limited) census of some bit-hoopoids of small size, see the end of Appendix **A1**. We are indebted to Louis Gidney for his patient explorations of bit-hoopoids of larger size, the details of which will appear in a sequel to this present paper.

A census of all bit-hoops of size 7 or less

(The superscripts display the Hamming-norm of each bit-hoop.)
(Pairs under the duality $\mathrm{o} \leftrightarrow 1$ are kept vertically close.)

[1] there are 2 bit-hoops of size 1: $\langle \mathrm{o} \rangle^0$ and $\langle 1 \rangle^1$

[2] 3 bit-hoops of size 2: $\langle \mathrm{oo} \rangle^0$ $\langle \mathrm{o1} \rangle^1$
$\langle 11 \rangle^2$

[3] 4 bit-hoops of size 3: $\langle \mathrm{ooo} \rangle^0$ $\langle \mathrm{oo1} \rangle^1$
$\langle 111 \rangle^3$ $\langle \mathrm{11o} \rangle^2$

[4] 6 bit-hoops of size 4: $\langle \mathrm{oooo} \rangle^0$ $\langle \mathrm{ooo1} \rangle^1$ $\langle \mathrm{oo11} \rangle^2$ $\langle \mathrm{o1o1} \rangle^2$
$\langle 1111 \rangle^4$ $\langle \mathrm{111o} \rangle^3$

[5] 8 bit-hoops of size 5:
$\langle \mathrm{ooooo} \rangle^0$ $\langle \mathrm{oooo1} \rangle^1$ $\langle \mathrm{ooo11} \rangle^2$ $\langle \mathrm{oo1o1} \rangle^2$
$\langle 11111 \rangle^5$ $\langle \mathrm{1111o} \rangle^4$ $\langle \mathrm{111oo} \rangle^3$ $\langle \mathrm{11o1o} \rangle^3$

[6] 14 bit-hoops of size 6:
$\langle \mathrm{oooooo} \rangle^0$ $\langle \mathrm{ooooo1} \rangle^1$ $\langle \mathrm{oooo11} \rangle^2$ $\langle \mathrm{ooo1o1} \rangle^2$ $\langle \mathrm{oo1oo1} \rangle^2$
$\langle 111111 \rangle^6$ $\langle \mathrm{11111o} \rangle^5$ $\langle \mathrm{1111oo} \rangle^4$ $\langle \mathrm{111o1o} \rangle^4$ $\langle \mathrm{11o11o} \rangle^4$

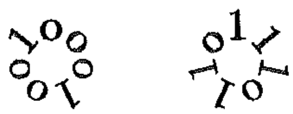
$\langle \mathrm{ooo111} \rangle^3$ $\langle \mathrm{oo1o11} \rangle^3$ $\langle \mathrm{o1o1o1} \rangle^3$
$\langle \mathrm{11o1oo} \rangle^3$

[7] 20 bit-hoops of size 7:
$\langle \mathrm{ooooooo} \rangle^0$ $\langle \mathrm{oooooo1} \rangle^1$ $\langle \mathrm{ooooo11} \rangle^2$ $\langle \mathrm{oooo1o1} \rangle^2$ $\langle \mathrm{ooo1oo1} \rangle^2$
$\langle 1111111 \rangle^7$ $\langle \mathrm{111111o} \rangle^6$ $\langle \mathrm{11111oo} \rangle^5$ $\langle \mathrm{1111o1o} \rangle^5$ $\langle \mathrm{111o11o} \rangle^5$

$\langle \mathrm{oooo111} \rangle^3$ $\langle \underline{\mathrm{ooo1o11}} \rangle^3$ $\langle \underline{\mathrm{ooo11o1}} \rangle^3$ $\langle \mathrm{oo1oo11} \rangle^3$ $\langle \mathrm{oo1o1o1} \rangle^3$
$\langle \mathrm{1111ooo} \rangle^4$ $\langle \mathrm{111o1oo} \rangle^4$ $\langle \mathrm{111oo1o} \rangle^4$ $\langle \mathrm{11o11oo} \rangle^4$ $\langle \mathrm{11o1o1o} \rangle^4$

The two non-mirror-equivalent hoops are underlined.)

Using the 'petal notation' of John Amson's earlier article, the 5-*th* and 10-*th* bit-hoops in the last list would appear as



The answer to the counting question : 'how many bit-hoops of size n are there?' is given in Appendix **A1**.

Bit-hoop Systems

By a *bit-hoop system (of size n)* we mean a set **H** of bit-hoops of size n, equipped with at least four laws of composition, *addition*, *multiplication*, and *mirror-imaging*. We may write $\mathbf{H}_{(n)}$ if the size n needs to be mentioned specifically.

It is natural to define the laws of composition 'element-wise'. Because of the cyclical ordering of the elements of a hoop, any element-wise definition of *e.g.* hoop-addition runs into the problem of deciding which element in one hoop is to be added to which element in the other hoop. This is in contrast to the element-wise addition of two vectors in which the linear ordering of the elements makes it natural first to add the first two elements, then the next two, and so on. There are different ways in which this problem might be tackled. The method we choose here is to introduce a composite procedure making use of an 'inner-product' of the elements of the two hoops, and repeating the process after cyclically shifting one bit-hoop against the other, thereby successively building up the cyclical elements of the sum of the two bit-hoops.

In order to have available some notion of what is to be meant by the addition and multiplication of individual 'elements', we choose to identify the element set $\mathbf{A} = \{\text{o}, \mathbf{1}\}$ with the field $\{0, 1\}$ of two elements **2** in which addition and multiplication can be taken to be addition, resp. multiplication, modulo 2 (or equivalently as Boolean logical XOR, resp. Boolean logical AND), denoted by $+$ resp. $\times$.

Two alternative (dual) inner-products of bit-hoops

By the *sumprod* ('SP') inner-product of two bit-hoops x and y both of size n we mean the sum of the pair-wise products of their elements, formed for example[6] , by

$$SP \stackrel{\text{def}}{=} (x_1 \times y_1) + \ldots + (x_n \times y_n)$$

[6] The choice of the 'starting index '1' for the element 'x_1' or 'y_1' is immaterial, because of the cyclical ordering.

By the *prodsum* ('*PS*') inner-product of two bit-hoops $\mathbf{x}$ and $\mathbf{y}$ both of size n we mean the product of the pair-wise sums of their elements, formed for example, by

$$PS \stackrel{\text{def}}{=} (x_1 + y_1) \times \ldots \times (x_n + y_n)$$

ADDITION of Bit-Hoops

The sum $\mathbf{s} = \mathbf{x} \oplus \mathbf{y}$ of two bit-hoops of size n has elements s_k successively formed by taking the *prodsum* of the elements of $\mathbf{y}$ cyclically-shifted against those of $\mathbf{x}$, *viz.*:

$$\begin{array}{lllll}
s_1 & = (x_1 + y_1) & \times (x_2 + y_2) \times \ldots \times (x_{n-1} + y_{n-1}) & \times (x_n + y_n) \\
s_2 & = (x_1 + y_2) & \times (x_2 + y_3) \times \ldots \times (x_{n-1} + y_n) & \times (x_n + y_1) \\
\ldots & & & \\
s_{n-1} & = (x_1 + y_{n-1}) & \times (x_2 + y_n) \times \ldots \times (x_{n-1} + y_{n-1}) & \times (x_n + y_n) \\
s_n & = (x_1 + y_n) & \times (x_2 + y_1) \times \ldots \times (x_{n-1} + y_{n-2}) & \times (x_n + y_{n-1})
\end{array}$$

Here is an illustration using bit-hoops of size 3, $\mathbf{x} = \langle\, a \;\; b \;\; c \,\rangle$, $\mathbf{y} = \langle\, d \;\; e \;\; f \,\rangle$, with sum $\mathbf{s} = \mathbf{x} \oplus \mathbf{y}$:

	$\langle a \;\; b \;\; c \rangle$	$\langle a \;\; b \;\; c \rangle$	$\langle a \;\; b \;\; c \rangle$
	$\langle d \;\; e \;\; f \rangle$	$\langle e \;\; f \;\; d \rangle$	$\langle f \;\; d \;\; e \rangle$
(*SP*)	$(a{+}d)(b{+}e)(f{+}c)$	$(a{+}e)(b{+}f)(c{+}d)$	$(a{+}f)(b{+}d)(c{+}e)$
(s)	s_1	s_2	s_3

Note that adding $\mathbf{x}$ to $\mathbf{y}$ to form the reversed sum $\mathbf{v} = \mathbf{y} \oplus \mathbf{x}$ gives :

	$\langle d \;\; e \;\; f \rangle$	$\langle d \;\; e \;\; f \rangle$	$\langle d \;\; e \;\; f \rangle$n
	$\langle a \;\; b \;\; c \rangle$	$\langle b \;\; c \;\; a \rangle$	$\langle c \;\; a \;\; b \rangle$n
(*SP*)	$(a{+}d)(b{+}e)(f{+}c)$	$(b{+}d)(c{+}e)(a{+}f)$	$(c{+}d)(a{+}e)(b{+}f)$n
(v)	v_1	v_2	v_3n

(We use *e.g.* '$(a{+}d)(b{+}e)$' to mean '$(a{+}d)\times(b{+}e)$', *etc.*, and we use the commutativity of $+$ and $\times$ in **2.**)

Inspection of the hoop Addition Tables for hoops of size $2, 3, 4, 5, 6$, below, shows that it is commutative in these cases. It is conjectured that it is commutative in every case, but a general proof is still to be found.

These Addition Tables also show some unfortunate features: for example, Hoop Addition is far from being injective; and . the sum of many non-zero bit-hoops is the Zero-hoop $\mathrm{o} = \langle\, 0\ 0, \ldots, 0\,\rangle$. There are many others unfamiliar aspects of hoop Addition to be explored.

....

MULTIPLICATION of Bit-Hoops

The product $\mathrm{p} = \mathbf{x} \otimes \mathrm{y}$ of two bit-hoops of size n has elements p_k successively formed by taking the *sumprod* of the elements of y cyclically-shifted against those of x, *viz.*:

$$
\begin{aligned}
p_1 &= (x_1 \times y_1) \quad + (x_2 \times y_2) + \ldots + (x_{n-1} \times y_{n-1}) + (x_n \times y_n) \\
p_2 &= (x_1 \times y_2) \quad + (x_2 \times y_3) + \ldots + (x_{n-1} \times y_n) \quad + (x_n \times y_1) \\
&\ldots \\
p_{n-1} &= (x_1 \times y_{n-1}) + (x_2 \times y_n) + \ldots + (x_{n-1} \times y_{n-1}) + (x_n \times y_n) \\
p_n &= (x_1 \times y_n) \quad + (x_2 \times y_1) + \ldots + (x_{n-1} \times y_{n-2}) + (x_n \times y_{n-1})
\end{aligned}
$$

Here is an illustration using bit-hoops of size 3, $\mathrm{x} = \langle\, a\ \ b\ \ c\,\rangle$, $\mathrm{y} = \langle\, d\ \ e\ \ f\,\rangle$, with product $\mathrm{p} = \mathrm{x} \otimes \mathrm{y}$:

	$\langle\, a\ \ b\ \ c\,\rangle$	$\langle\, a\ \ b\ \ c\,\rangle$	$\langle\, a\ \ b\ \ c\,\rangle$
	$\langle\, d\ \ e\ \ f\,\rangle$	$\langle\, e\ \ f\ \ d\,\rangle$	$\langle\, f\ \ d\ \ e\,\rangle$
(SP)	$ad + be + fc$	$ae + bf + cd$	$af + bd + ce$
(p)	p_1	p_2	p_3

Note that multiplying y on the right by x to form the reversed product $\mathrm{v} = \mathrm{y} \oplus \mathrm{x}$ gives :

	$\langle d \quad e \quad f \rangle$ $\langle a \quad b \quad c \rangle$	$\langle d \quad e \quad f \rangle$ $\langle b \quad c \quad a \rangle$	$\langle d \quad e \quad f \rangle$ $\langle c \quad a \quad b \rangle$
(SP)	$ad + be + fc$	$bd + ce + af$	$cd + ae + bf$
(v)	v_1 $(=p_1)$	v_2 $(=p_3)$	v_3 $(=p_2)$

(We use *e.g.* '*ad*' to mean '$a \times d$', and we use the commutativity of $+$ and $\times$ in **2**.)

Because in general we do not have $v_2 = p_2$ nor $v_3 = p_3$, despite having $v_2 = p_3$ and $v_3 = p_2$, we see that the two bit-hoops $\mathbf{x} \oplus \mathbf{y} = \mathbf{s} = \langle\, p_1 \;\; p_2 \;\; p_3 \,\rangle$ and $\mathbf{y} \oplus \mathbf{x} = \mathbf{v} = \langle\, v_1 \;\; v_2 \;\; v_3 \,\rangle = \langle\, p_1 \;\; p_3 \;\; p_2 \,\rangle$ and it would appear that their cyclic ordering is reversed and hence they may not be the same bit-hoops; but this does not follow; why? because it may (and can) happen that *e.g.* $p1 = p3$ and then $\mathbf{x} \oplus \mathbf{y} = \mathbf{s} = \langle\, p_1 \;\; p_2 \;\; p_1 \,\rangle$ and $\mathbf{y} \oplus \mathbf{x} = \mathbf{v} = \langle\, p_1 \;\; p_1 \;\; p_2 \,\rangle$ are cyclically equivalent. Hence we cannot use such an argument to assert that hoop multiplication is not commutative.

Indeed, inspection of the hoop multiplication tables for hoops of size $2, 3, 4, 5$, below, shows that it is commutative in these cases. Further inspection shows that muliplicative commutative breaks down for the first time with hoops of size 5 as seen in the table below.

Multiplication by a constant hoop

There are two (and only two) constant bit-hoops for each size n, namely the 'Zero-hoop' and the 'Unit-hoop'. Hoop products in which one of the multiplicands is a constant are special. In the case where the right-hand multiplicand is a constant bit-hoop, then the *sumprods* of every cyclical-shift of y is also constant and therefore the product $\mathbf{p} = \mathbf{x} \otimes \mathbf{y}$ is also a constant bit-hoop; the constant values of its elements is computed by a single *sumprod*. We see this immediately:

Let $\mathbf{o} = \langle\, 0 \ldots 0 \,\rangle$ be the 'Zero-hoop' of size n, and let $\mathbf{x} = \langle\, x_1 \ldots x_n \,\rangle$ be any bit-hoop of size n. Since $x_k \times 0 = 0 = 0 \times x_k$ for every element x_k, the *sumprod* of x and any cyclically-shifted copy of y is also 0. That is, $\mathbf{x} \otimes \mathbf{o} = \mathbf{o} = \mathbf{o} \otimes \mathbf{x}$. Hence (as expected) the Zero-hoop acts multiplicatively on the right as a 'multiplicative zero'. Small changes to this argument show that it also acts on the left as as 'multiplicative zero'. It is unique, for if W were another multiplicative zero then $W = W \otimes o = o$.

Let $1 = \langle\, 1 \ldots 1 \,\rangle$ be the 'Unit-hoop' of size n, and let $x = \langle\, x_1 \ldots x_n \,\rangle$ be any bit-hoop of size n. Since $x_k \times 1 = x_k = 1 \times x_k$ for every element x_k, the *sumprod* of x and any cyclically-shifted copy of y is either the Zero-hoop or the Unit-hoop — depending only on whether the Hamming-norm of x is even or odd. Hence (unexpectedly) the Unit-hoop acts multiplicatively on the right as if it were a 'parity operator' acting on the Hamming-norm of its multiplicand x. Small changes to this argument show that it also acts on the left in the same way.

For example, if $x = \langle\, 0\ 0\ 0\ 1 \,\rangle$ with Hamming-norm 1, then

(x)	⟨0 0 0 1⟩	⟨0 0 0 1⟩	⟨0 0 0 1⟩	⟨0 0 0 1⟩
(1)	⟨1 1 1 1⟩	⟨1 1 1 1⟩	⟨1 1 1 1⟩	⟨1 1 1 1⟩
(*SP*)	01+01+01+11	01+01+01+11	01+01+01+11	01+01+01+11
i.e.	0+0+0+1	0+0+0+1	0+0+0+1	0+0+0+1
i.e.	1	1	1	1
(p=1)	p_1	p_2	p_3	p_4

Whilst, if $x = \langle\, 0\ 0\ 1\ 1 \,\rangle$ with Hamming-norm 2, then

(x)	⟨0 0 1 1⟩	⟨0 0 1 1⟩	⟨0 0 1 1⟩	⟨0 0 1 1⟩
(1)	⟨1 1 1 1⟩	⟨1 1 1 1⟩	⟨1 1 1 1⟩	⟨1 1 1 1⟩
(*SP*)	01+01+11+11	01+01+11+11	01+01+11+11	01+01+11+11
i.e.	0+0+1+1	0+0+1+1	0+0+1+1	0+0+1+1
i.e.	0	0	0	0
(p=o)	p_1	p_2	p_3	p_4

However, the Unit-hoop 1 is not unique in acting as a parity operator on the Hamming-norm, since it is easy to modify the last two illustrations by replacing $1 = \langle\, 1\,1\,1\,1 \,\rangle$ with $1' = \langle\, 0\,0\,1\,1 \,\rangle$ — the same result arises!

Two Easy Exercises For The Reader: [1] Show that the bit-hoop 'A' of size n and Hamming-Norm 1 (*e.g.* ⟨ 0001 ⟩) acts as the Multiplicative Unit, *i.e.* $A \otimes x = x \otimes A = x$ for every bit-hoop x of same size, and is unique.

[2] What multiplicative rôle does the the bit-hoop 'Z' of size n and defect-Hamming-Norm 1 (*e.g.* ⟨ 0111 ⟩) play ?

MIRRORING of Bit-Hoops

We have already defined the notion of the 'mirror-image' of a hoop, the notion of a hoop being 'symmetric' if it is (cyclically) equal to itself, and of the idea of an 'asymmetric pair' of hoops.

By the *mirroring operation* we mean the operation of forming the mirror-image of a hoop. Thus for each hoop size n the mirroring operation is defined by the mapping

$$\mu_n \; : \; \mathbf{H}_{(n)} \longrightarrow \mathbf{H}_{(n)} \; : \; \mathtt{h} \mapsto \mu(\mathtt{h})$$

where $\mu_n(\mathtt{h})$ is the mirror-image of $\mathtt{h}$. For convenience we may ignore the subscripts n and write μ_n and $\mathbf{H}_{(n)}$ simply as μ and $\mathbf{H}$.

The mapping μ is an involution, that is $\mu \circ \mu = id$ (the identity mapping), *i.e.* $\mu(\mu(\mathtt{h})) = \mathtt{h}$.

The image of the mapping μ is a subset $\mu(\mathbf{H})$ of $\mathbf{H}$, the subset of symmetric hoops. By computerised inspection (see Appendix **B**) we see that $\mu(\mathbf{H}) = \mathbf{H}$ for $n = 1,2,3,4,5$ and that $\mu(\mathbf{H}) \subset \mathbf{H}$ for $n = 6,7,8,9$. It is conjectured that $\mu(\mathbf{H})$ is also a proper subset of $\mathbf{H}$ for all $n \geq 10$. (This should be easy to prove?)

The complement of the subset of symmetric hoops contains the non-symmetric hoops; we can denote this by $\alpha(\mathbf{H})$ where the 'asymmetric' mapping α is defined by the notation

$$\alpha(\mathbf{H}) \stackrel{\text{def}}{=} \complement(\mu(\mathbf{H}))$$

This subset $\alpha(\mathbf{H})$ consists of an even number of hoops which are pairwise mirror-images of each other — the pairs of asymmetric hoops.

Q: What is the combinatoric formula for calculating the number of asymmetric pairs of hoops of any given size n ?

Note added during the Revision of this paper, 2004-dec-31

"Loops" and "Hoops" as used in this paper are instances of combinatorial objects known in the literature under other names — "necklaces" and "periodic sequences", for example. The notion goes back at least to a paper by that reknowned combinatorialist, Major P.A.MacMahon, over a century ago[7]. A "necklace" is a "string of p beads each of which can take one of q colours". The "Hoops" in this paper correspond to necklaces in which the arrangement of beads is invariant under a cyclic permutation. The "Hoops which equal their "mirror-image" ('flip-hoops') belong to the class of necklaces in which the arrangement of beads is invariant under a dihedral permutation.

For instance, the rotations of a plane polygon (with n vertices) in its plane form the Cyclical subgroup (of order n) of the group of all permutations of its vertices (the symmetric group, of order $n!$), the rotations together with a reflection across a diagonal form the Dihedral subgroup of permutations (of order $2n$).

See Appendix **A1** resp. **A2** for "cyclical' bit-hoop" resp. "dihedral 'flip' bit-hoop" counting formulæ. Useful references are:

Gilbert, E. N. and Riordan, J. "Symmetry Types of Periodic Sequences." *Illinois J. Math.* **5**, 657-665, 1961.

Lint, J.H. & Wilson, R.M.CUP, *A Course in Combinatorics*, CUP, 1992, (reprinted 1996).

The web site: "http://mathworld.wolfram.com/Necklace.html" has a useful introduction and list of more references.

[End of Revision Note.]

ASSOCIATIVITY, DISTRIBUTIVITY, etc., of Bit-Hoops

Associativity As usual, if we have $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ identically, then we say that addition $\oplus$ is *additive-associative.* If the same holds with $\otimes$ in place of $\oplus$ then multiplication $\otimes$ is *multiplicative-associative.*

Distributivity As usual, if we have $x \oplus (y \times z) = (x \otimes y) \oplus (x \otimes z)$ identically, then we say that addition $\oplus$ is *left-distributive over* $\otimes$. If the same holds with $\otimes$ and $\oplus$ exchanged then we say multiplication $\otimes$ is *left-distributive over* $\oplus$.

[7] *Proc.Lond.Math.Soc.*, Vol.23, (1891-92), pp.305-313.

Plainly, *right distributivity* is defined in the obvious way. If addition is both left- and right- distributive over multiplication then we say that addition is *distributive.* Analogously for multiplication.

Self-Distributivity Less familiar: if we have $x \oplus (y \oplus z) = (x \oplus y) \oplus (x \oplus z)$ identically, then we say that addition $\oplus$ is *self-distributive on the left.* If the same holds with $\otimes$ in place of $\oplus$ then multiplication $\otimes$ is *self-distributive on the left.* Then self-distributivity *on the right*, and *self-distributivity* are also defined in the obvious way.

Here is a summary of these attributions of Addition and Multiplication for Bit-Hoops of sizes from 2 to 8, based on computer experiments:-

SIZE	2	3	4	5	6	7	8
$\oplus$							
COMM.	Y	Y	Y	Y	Y	Y	Y
ASSOC.	No	No	No	No	No	No	No
ZERO?	No	No	No	No	No	No	No
$\otimes$							
COMM.	Y	Y	Y	Y	No	Y	Y
ASSOC.	Y	Y	Y	Y	No	No	No
ZERO? †	Y	Y	Y	Y	Y	Y	Y
UNIT? ‡	Y	Y	Y	Y	Y	Y	Y
DISTRIB.							
$\oplus$ OVER $\oplus$	No	No	No	No	No	No	No
$\otimes$ OVER $\otimes$	No	Y	No	Y	No	No	No
$\oplus$ OVER $\otimes$	No	No	No	No	No	No	No
$\otimes$ OVER $\oplus$	No	No	No	No	No	No	No

† The multiplicative Zero is o for $n = 2, 3, 4, 5, 6, 7, 8$
‡ The multiplicative Unit is A for $n = 2, 3, 4, 5, 6, 7, 8$
(where $o \approx \langle 00 \ldots 00 \rangle$, $A \approx \langle 00 \ldots 01 \rangle$)

Note: Wherever an attribute failed, the computer program reported the elements involved in the failure (full details are available from the authors).

To illustrate Bit-Hoop Addition and Multiplication we can display the Addition Tables and Multiplication Tables for bit-hoops of sizes 2, 3, 4, 5, 6.

It is convenient to assign labels to the various bit-hoops in each system. (We use '$\approx$' to mean 'cyclically equivalent to'.)

When $n = 2$ we have three bit-hoops :

o : $\langle 00 \rangle$
A : $\langle 01 \rangle \approx \langle 10 \rangle$ (*'mirror-image'*)
1 : $\langle 11 \rangle$

When $n = 3$ we have four bit-hoops :

o : $\langle 000 \rangle$
A : $\langle 001 \rangle \approx \langle 100 \rangle \approx \langle 010 \rangle$
B : $\langle 011 \rangle \approx \langle 101 \rangle \approx \langle 110 \rangle$
1 : $\langle 111 \rangle$

When $n = 4$ we have six bit-hoops :

o : $\langle 0000 \rangle$
A : $\langle 0001 \rangle \approx \langle 1000 \rangle \approx \langle 0100 \rangle \approx \langle 0100 \rangle$
B : $\langle 0011 \rangle \approx \langle 1001 \rangle \approx \langle 1100 \rangle \approx \langle 0110 \rangle$
C : $\langle 0101 \rangle \approx \langle 1010 \rangle$ (*'mirror-image'*)
D : $\langle 0111 \rangle \approx \langle 1011 \rangle \approx \langle 1101 \rangle \approx \langle 1110 \rangle$
1 : $\langle 1111 \rangle$

...continued overleaf...

When $n = 5$ we have eight bit-hoops :

o : $\langle$ 00000 $\rangle$

A : $\langle$ 00001 $\rangle \approx \langle$ 10000 $\rangle \approx \langle$ 01000 $\rangle \approx \langle$ 00100 $\rangle \approx \langle$ 00010 $\rangle$

B : $\langle$ 00011 $\rangle \approx \langle$ 10001 $\rangle \approx \langle$ 11000 $\rangle \approx \langle$ 01100 $\rangle \approx \langle$ 00110 $\rangle$

C : $\langle$ 00101 $\rangle \approx \langle$ 10010 $\rangle \approx \langle$ 01001 $\rangle \approx \langle$ 10100 $\rangle \approx \langle$ 01010 $\rangle$

D : $\langle$ 00111 $\rangle \approx \langle$ 10011 $\rangle \approx \langle$ 11001 $\rangle \approx \langle$ 11100 $\rangle \approx \langle$ 01110 $\rangle$

E : $\langle$ 01011 $\rangle \approx \langle$ 10101 $\rangle \approx \langle$ 11010 $\rangle \approx \langle$ 01101 $\rangle \approx \langle$ 10110 $\rangle$

F : $\langle$ 01111 $\rangle \approx \langle$ 10111 $\rangle \approx \langle$ 11011 $\rangle \approx \langle$ 11101 $\rangle \approx \langle$ 11110 $\rangle$

1 : $\langle$ 11111 $\rangle$

When $n = 6$ we have fourteen bit-hoops :

o : $\langle$ 000000 $\rangle$

A : $\langle$ 000001 $\rangle\approx\langle$ 100000 $\rangle\approx\langle$ 010000 $\rangle\approx\langle$ 001000 $\rangle\approx\langle$ 000100 $\rangle\approx\langle$ 000010 $\rangle$

B : $\langle$ 000011 $\rangle\approx\langle$ 100001 $\rangle\approx\langle$ 110000 $\rangle\approx\langle$ 011000 $\rangle\approx\langle$ 001100 $\rangle\approx\langle$ 000110 $\rangle$

C : $\langle$ 000101 $\rangle\approx\langle$ 100010 $\rangle\approx\langle$ 010001 $\rangle\approx\langle$ 101000 $\rangle\approx\langle$ 010100 $\rangle\approx\langle$ 001010 $\rangle$

D : $\langle$ 000111 $\rangle\approx\langle$ 100011 $\rangle\approx\langle$ 110001 $\rangle\approx\langle$ 111000 $\rangle\approx\langle$ 011100 $\rangle\approx\langle$ 001110 $\rangle$

E : $\langle$ 001001 $\rangle\approx\langle$ 100100 $\rangle\approx\langle$ 010010 $\rangle$

F : $\langle$ 001011 $\rangle\approx\langle$ 100101 $\rangle\approx\langle$ 110010 $\rangle\approx\langle$ 011001 $\rangle\approx\langle$ 101100 $\rangle\approx\langle$ 010110 $\rangle$

G : $\langle$ 001101 $\rangle\approx\langle$ 100110 $\rangle\approx\langle$ 010011 $\rangle\approx\langle$ 101001 $\rangle\approx\langle$ 110100 $\rangle\approx\langle$ 011010 $\rangle$

H : $\langle$ 001111 $\rangle\approx\langle$ 100111 $\rangle\approx\langle$ 110011 $\rangle\approx\langle$ 111001 $\rangle\approx\langle$ 111100 $\rangle\approx\langle$ 011110 $\rangle$

I : $\langle$ 010101 $\rangle\approx\langle$ 101010 $\rangle$ (*'mirror-image'*)

J : $\langle$ 010111 $\rangle\approx\langle$ 101011 $\rangle\approx\langle$ 110101 $\rangle\approx\langle$ 111010 $\rangle\approx\langle$ 011101 $\rangle\approx\langle$ 101110 $\rangle$

K : $\langle$ 011011 $\rangle\approx\langle$ 101101 $\rangle\approx\langle$ 110110 $\rangle$

L : $\langle$ 011111 $\rangle\approx\langle$ 101111 $\rangle\approx\langle$ 110111 $\rangle\approx\langle$ 111011 $\rangle\approx\langle$ 111101 $\rangle\approx\langle$ 111110 $\rangle$

1 : $\langle$ 111111 $\rangle$

ADDITION TABLE
FOR BIT-HOOPS OF SIZE 2

⊕	o	A	1
o	o	o	1
A	o	A	o
1	1	o	o

symmetric

ADDITION TABLE
FOR BIT-HOOPS OF SIZE 3

⊕	o	A	B	1
o	o	o	o	1
A	o	o	A	o
B	o	A	o	o
1	1	o	o	o

symmetric

ADDITION TABLE
FOR BIT-HOOPS OF SIZE 4

⊕	o	A	B	C	D	1
o	o	o	o	o	o	1
A	o	o	o	o	A	o
B	o	o	A	o	o	o
C	o	o	o	C	o	o
D	o	A	o	o	o	o
1	1	o	o	o	o	o

symmetric

ADDITION TABLE
FOR BIT-HOOPS OF SIZE 5

⊕	o	A	B	C	D	E	F	1
o	o	o	o	o	o	o	o	1
A	o	o	o	o	o	o	A	o
B	o	o	o	o	A	o	o	o
C	o	o	o	o	o	A	o	o
D	o	o	A	o	o	o	o	o
E	o	o	o	A	o	o	o	o
F	o	A	o	o	o	o	o	o
1	1	o	o	o	o	o	o	o

symmetric

The Addition Table for bit-hoops of size 6 is given two pages below.

MULTIPLICATION T.
FOR BIT-HOOPS OF SIZE 2

⊕	o	A	1
o	o	o	o
A	o	A	1
1	o	1	o

symmetric

MULTIPLICATION TABLE
FOR BIT-HOOPS OF SIZE 3

⊗	o	A	B	1
o	o	o	o	o
A	o	A	B	1
B	o	B	B	o
1	o	1	o	1

symmetric

MULTIPLICATION TABLE
FOR BIT-HOOPS OF SIZE 4

⊗	o	A	B	C	D	1
o	o	o	o	o	o	o
A	o	A	B	C	D	1
B	o	B	C	1	B	o
C	o	C	1	o	C	o
D	o	D	B	C	A	1
1	o	1	o	o	1	o

symmetric

MULTIPLICATION TABLE
FOR BIT-HOOPS OF SIZE 5

⊗	o	A	B	C	D	E	F	1
o	o	o	o	o	o	o	o	o
A	o	A	B	C	D	E	F	1
B	o	B	C	F	C	F	B	o
C	o	C	F	B	F	B	C	o
D	o	D	C	F	E	A	B	1
E	o	E	F	B	A	D	C	1
F	o	F	B	C	B	C	F	o
1	o	1	o	o	1	1	o	1

symmetric

The Multiplication Table for bit-hoops of size 6 is given on two pages below.

ADDITION TABLE

FOR BIT-HOOPS OF SIZE 6

⊕	o	A	B	C	D	E	F	G	H	I	J	K	L	1
o	o	o	o	o	o	o	o	o	o	o	o	o	o	1
A	o	o	o	o	o	o	o	o	o	o	o	o	A	o
B	o	o	o	o	o	o	o	o	A	o	o	o	o	o
C	o	o	o	o	o	o	o	o	o	o	A	o	o	o
D	o	o	o	o	A	o	o	o	o	o	o	o	o	o
E	o	o	o	o	o	o	o	o	o	o	o	E	o	o
F	o	o	o	o	o	o	o	A	o	o	o	o	o	o
G	o	o	o	o	o	o	A	o	o	o	o	o	o	o
H	o	o	A	o	o	o	o	o	o	o	o	o	o	o
I	o	o	o	o	o	o	o	o	o	I	o	o	o	o
J	o	o	o	A	o	o	o	o	o	o	o	o	o	o
K	o	o	o	o	o	E	o	o	o	o	o	o	o	o
L	o	A	o	o	o	o	o	o	o	o	o	o	o	o
1	1	o	o	o	o	o	o	o	o	o	o	o	o	o

symmetric

...continued overleaf...

MULTIPLICATION TABLE

FOR BIT-HOOPS OF SIZE 6

⊗	o	A	B	C	D	E	F	G	H	I	J	K	L	1
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
A	o	A	B	C	D	E	F	G	H	I	J	K	L	1
B	o	B	C	H	E	K	J	J	C	1	H	K	B	o
C	o	C	H	C	K	K	H	H	H	o	C	K	C	o
D	o	D	E	K	I	1	D	D	K	I	E	o	D	1
E	o	E	K	K	1	o	E	E	K	1	K	o	E	o
F	o	G	J	H	D	E	L	A	C	I	B	K	F	1
G	o	F	J	H	D	E	A	L	C	I	B	K	G	1
H	o	H	C	H	K	K	C	C	C	o	H	K	H	o
I	o	I	1	o	I	1	I	I	o	I	1	o	I	1
J	o	J	H	C	E	K	B	B	H	1	C	K	J	o
K	o	K	K	K	o	o	K	K	K	o	K	o	K	o
L	o	L	B	C	D	E	G	F	H	I	J	K	A	1
1	o	1	o	o	1	o	1	1	o	1	o	o	1	o

non-symmetric

G⊗A = F ≠ A⊗G = G;   F⊗A = G ≠ A⊗F = F
L⊗G = F ≠ G⊗L = G;   L⊗F = G ≠ F⊗L = F

...continued overleaf...

MULTIPLICATION TABLE

FOR BIT-HOOPS OF SIZE 6

SHOWING ONLY HAMMINGNORM OF PRODUCTS

(ROWS & COLUMNS RE-ORDERED)

$\otimes$	o_0	A_1	B_2	C_2	E_2	D_3	F_3	G_3	I_3	H_4	J_4	K_4	L_5	1_6
o_0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A_1	0	1	2	2	2	3	3	3	3	4	4	4	5	6
B_2	0	2	2	4	4	2	4	4	6	2	4	4	2	0
C_2	0	2	4	2	4	4	4	4	0	4	2	4	2	0
E_2	0	2	4	4	0	6	2	2	6	4	4	0	2	0
D_3	0	3	2	4	6	3	3	3	3	4	2	0	3	6
F_3	0	3	4	4	2	3	5	1	3	2	2	4	3	6
G_3	0	3	4	4	2	3	1	5	3	2	2	4	3	6
I_3	0	3	6	0	6	3	3	3	3	0	6	0	3	6
H_4	0	4	2	4	4	4	2	2	0	2	4	4	4	0
J_4	0	4	4	2	4	2	2	2	6	4	2	4	4	0
K_4	0	4	4	4	0	0	4	4	0	4	4	0	4	0
L_5	0	5	2	2	2	3	3	3	3	4	4	4	1	6
1_6	0	6	0	0	0	6	6	6	6	0	0	0	6	0

Note that "the parity of [the HammingNorm of a product]" is the same as "the parity of [the product of the HammingNorms] of the factors",

e.g. $|B| = 2$, $|D| = 3$, $2 \times 3 = 6$, even; $B \otimes D = E$, $|E| = 2$, even.

e.g. $|D| = 3$, $3 \times 3 = 9$, odd; $D \otimes D = I$, $|I| = 3$, odd.

Q: Is this true in general ?

APPENDIX A1 THE NUMBER OF BIT-HOOPS OF SIZE n AND BIT-HOOPOIDS OF SIZE $n \times m$

Unlike the ease with which the number of bit-strings of length n can be counted (*viz.* 2^n) the problem of counting the number of bit-hoops of size n is a much more complicated combinatoric task. For example, when $n = 7$ the 20 distinct bit-hoops are counted as follows:

1 with no 1s	⟨ 0000000 ⟩
1 with one 1	⟨ 0000001 ⟩
3 with two 1s	⟨ 0000011 ⟩ ⟨ 0000101 ⟩ ⟨ 0001001 ⟩
5 with three 1s	⟨ 0000111 ⟩ ⟨ 0001011 ⟩ ⟨ 0001101 ⟩ ⟨ 0010011 ⟩ ⟨ 0010101 ⟩
5 with four 1s	⟨ 0001111 ⟩ ⟨ 0010111 ⟩ ⟨ 0100111 ⟩ ⟨ 0011011 ⟩ ⟨ 0101011 ⟩
3 with five 1s	⟨ 0011111 ⟩ ⟨ 0101111 ⟩ ⟨ 0110111 ⟩
1 with six 1s	⟨ 0111111 ⟩
1 with seven 1s	⟨ 1111111 ⟩

The same sort of step by step trial-and-error counting process can be used for most small bit-hoop sizes, but it rapidly becomes unmanageable and error-prone. As in the bit-string Combinatoric Hierarchy system, there is a potential need to count the numbers of bit-hoops of size n when n is an even power of a power of 2, eg 16, 128, 256 and so on. We need an analytic result. Here it is:-

Let $H(n)$ denote the number of bit-hoops of size n for any positive integer $n > 1$. Then

$$H(n) = \frac{1}{n} \sum_{d|n} \phi\left(\frac{n}{d}\right) 2^d$$

where the summation control $d|n$ for $H(n)$ means (as usual) that the summation is over just those 'integers d that divide integer n' (*i.e.* the divisors, or factors, of n), and where ϕ is the *Euler function* (*i.e.* $\phi(m)$ is the number of integers k with $1 \leq k \leq m$ having only 1 as a divisor in common with m) which is given in closed form by:

$$\phi(m) = m \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right)$$

for any positive integer m expressed as a product of prime numbers $m = p_1^{a_1} p_2^{a_2} \ldots p_r^{a_r}$. (See *Handbook of Mathematical Functions* by M.Abramovitch & I.A.Segun, (Dover, New York), 1968 (5th printing), page 826 and the tables on pages 840-843.) In particular, if m is a power of 2 ($m = 2^\alpha$), then $r = 1 = i$, $p_i = 2$ and $\phi(m) = \phi(2^\alpha) = \frac{m}{2} = 2^{\alpha-1}$.

Some useful illustrations are tabled in Appendix **A2** below.

The problem is not new: it is a special case of the problem of counting "the number of necklaces with beads of (two) different colours" or, equivalently, of counting "the number of circular sequences of 0s and 1s, where two sequences obtained by a rotation are considered the same". For a detailed discussion see for example the valuable text-book *A Course in Combinatorics* by J.H.Lint & R.M.Wilson, CUP, 1992, (reprinted 1996), where it occurs as EXAMPLE 10.5 on page 75. These authors point out that it is also given by *Burnside's lemma* (Cauchy, Frobenius) in Group Theory, which they show as their THEOREM 10.5:-

Let G be a permutation group acting on a set X. For $g \in G$ let $\psi(g)$ denote the number of points of X fixed by g. Then the number of orbits of G is equal to $\frac{1}{|G|} \sum_{g \in G} \psi(g)$.

As such, it potentially has a very particular rôle to play in view of the importance (*cf.* Discriminate Closure, *etc.*) of such similar ideas in the previous development and explorations of the 'Bit-String Combinatoric Hierarchy'.

....

For an illustration, here is the 'computation scheme' for $H(6)$:

The divisors d_i of 6 are $d_1 = 1$, $d_2 = 2$, $d_3 = 3$, $d_4 = 6$.

The quotients $q_i = 6/d_i$ are $q_1 = 6$, $q_2 = 3$, $q_3 = 2$, $q_4 = 1$.

The Euler number $\phi(6)$ can be got by inspection of which integers $1 \leq k \leq 6$ have only 1 as a common divisor with 6 (underlined): $\underline{1}$, 2, 3, 4, $\underline{5}$, 6, therefore $\phi(6) = 2$. (For small values of n the product formula for $\phi(n)$ given above is an overkill for hand calculations, though needed for a computer algorithm.) And $\phi(3) = 2$, $\phi(2) = 1$, $\phi(1) = 1$.

The summands $\phi\left(\frac{n}{d}\right) \times 2^d$ in the formula for $H(n)$ are

$$
\begin{aligned}
s_1 &= \phi(q_1) \times 2^{d_1} = 2 \times 2^1 = 4 \\
s_2 &= \phi(q_2) \times 2^{d_2} = 2 \times 2^2 = 8 \\
s_3 &= \phi(q_3) \times 2^{d_3} = 1 \times 2^3 = 8 \\
s_4 &= \phi(q_4) \times 2^{d_4} = 1 \times 2^6 = 64
\end{aligned}
$$

hence $H(6) = (s_1 + s_2 + s_3 + s_4) \div 6 = 84 \div 6 = 14$

When n is prime, the computation scheme is even simpler, *e.g.* when $n = 7$:

$n = 7$	$q_i = n/d_i$	$\phi(q_i)$	s_i	$\sum s_i$
$d_1 = 1$	$q_1 = 7$	6	$6 \times 2^1 = 12$	12
$d_2 = 7$	$q_2 = 1$	1	$1 \times 2^7 = 128$	140

hence $H(7) = 140 \div 7 = 20$.

When n is a power of 2, the computation scheme has many regularities, *e.g.* when $n = 2^3 = 8$:

$n = 8$	$q_i = n/d_i$	$\phi(q_i)$	s_i	$\sum s_i$
$d_1 = 1$	$q_1 = 8$	4	$4 \times 2^1 = 8$	8
$d_1 = 2$	$q_1 = 4$	2	$2 \times 2^2 = 8$	16
$d_2 = 4$	$q_2 = 2$	1	$1 \times 2^4 = 16$	32
$d_1 = 8$	$q_1 = 1$	1	$1 \times 2^8 = 256$	288

hence $H(8) = 288 \div 8 = 36$.

....

Counting bit-hoopoids

The first suggestion that the number $H(n, m)$ of bit-hoopoids of size $n \times m$ is just the product $H(n, m) = H(n) \times H(m)$ is rejected by the simplest example, a bit-hoopoid of size 2×2. Here is the census of this example :-

$$\begin{array}{ccc}
\langle 0\,0 \rangle & \langle 0\,0 \rangle & \langle 0\,0 \rangle \\
\langle 0\,0 \rangle & \langle 0\,1 \rangle & \langle 1\,1 \rangle \\
& & \\
& \langle 0\,1 \rangle & \langle 0\,1 \rangle \\
& \langle 0\,1 \rangle & \langle 1\,1 \rangle \\
& & \\
& & \langle 1\,1 \rangle \\
& & \langle 1\,1 \rangle
\end{array}$$

We count $H(2,2) = 6$ ($= H(4)$), whilst $H(2) \times H(2) = 4$.
And here is the census of bit-hoopoids of size 2×3 :-

$$\begin{array}{cccc}
\langle 0\,0\,0 \rangle & \langle 0\,0\,0 \rangle & \langle 0\,0\,0 \rangle & \langle 0\,0\,0 \rangle \\
\langle 0\,0\,0 \rangle & \langle 0\,0\,1 \rangle & \langle 0\,1\,1 \rangle & \langle 1\,1\,1 \rangle \\
& & & \\
& \langle 0\,0\,1 \rangle & \langle 0\,0\,1 \rangle & \langle 0\,0\,1 \rangle \\
& \langle 0\,0\,1 \rangle & \langle 0\,1\,1 \rangle & \langle 1\,1\,1 \rangle \\
& & & \\
& & \langle 0\,1\,1 \rangle & \langle 0\,1\,1 \rangle \\
& & \langle 0\,1\,1 \rangle & \langle 1\,1\,1 \rangle \\
& & & \\
& & & \langle 1\,1\,1 \rangle \\
& & & \langle 1\,1\,1 \rangle \\
& & & \\
& & \langle 0\,0\,1 \rangle & \langle 0\,0\,1 \rangle \\
& & \langle 0\,1\,0 \rangle & \langle 1\,1\,0 \rangle \\
& & & \\
& & & \langle 0\,1\,1 \rangle \\
& & & \langle 1\,0\,1 \rangle \\
& & & \\
& & & \langle 1\,0\,0 \rangle \\
& & & \langle 1\,1\,0 \rangle
\end{array}$$

We count $H(2,3) = 14$ ($= H(6)$), whilst $H(2) \times H(3) = 12$.
Does this suggest instead that $H(n, m) = H(n \times m)$? ?
And that for generalised bit-hoopoids of size ($n_1, n_2, \ldots, n_N$) do we have

$$H(n_1, n_2, \ldots, n_N) = H(n_1 \times n_2 \times \ldots \times n_N) \;?\;?\;?$$

These are open questions.

.... [End of Appendix **A1**] ....

APPENDIX A2 THE NUMBER OF DIHEDRAL BIT-HOOPS OF SIZE n

The number $DH(n)$ of dihedral ('flip symmetric') bit-Hoops of size n ($n \geq 4$) is given in terms of the number of cyclical bit-hoops $H(n)$ (see Appendix **A1** above) by

$$\text{(even } n\text{)} \qquad DH(n) = H(n)/2 + 3 \cdot 2^{((n-4)/2)}$$

$$\text{(odd } n\text{)} \qquad DH(n) = H(n)/2 + 2^{((n-1)/2)}$$

(That $DH(n) = H(n)$ for $n = 1, 2$ is plain by inspection.) Some illustrations are given below:

$H(n)$, $DH(n)$: THE NUMBER OF
BIT-HOOPS & DIHEDRAL BIT-HOOPS OF SIZE n

n	Hoops $H(n)$	DHoops $DH(n)$	Euler $\phi(n)$
2	3	3	1
3	4	4	2
[2^2] 4	6	6	2
5	8	8	4
6	14	13	2
7	20	18	6
[2^3] 8	36	20	4
[3^2] 9	60	46	6
10	108	78	4
11	188	126	10
12	352	224	4
13	632	380	12
14	1,182	687	6
15	2,192	1,224	8
[4^2] 16	4,116	2,200	8

continued

$H(n)$, $DH(n)$: THE NUMBER OF
BIT-HOOPS & DIHEDRAL BIT-HOOPS OF SIZE n
CONTINUED

n	Hoops $H(n)$	DHoops $DH(n)$	Euler $\phi(n)$
17	7,712	4,112	16
18	14,602	7,685	6
19	27,596	14,310	18
20	52,488	27,012	8
21	99,880	50,964	12
22	190,746	96,909	10
23	364,724	184,410	22
24	699,252	352,698	8
$[5^2]$ 25	1,342,184	675,188	20
26	2,581,428	1,296,858	12
$[3^3]$ 27	4,971,068	2,493,726	18
28	9,587,580	4,806,078	12

.... [End of Appendix **A2**] ....

APPENDIX B SOME NON-SYMMETRIC BIT-HOOPS

Computer experiments show us that:

bit-hoop size	Number of bit-hoops	Number of non-symmetric	%
2	3	0	0
3	4	0	0
4	6	0	0
5	8	0	0
6	14	2	14
7	20	4	20
8	36	12	33
9	60	28	47
10	105	?	?

Here are the non-symmetric bit-hoops of sizes $n = 6, 7, 8, 9$ (the superscript is the Hamming-norm), they are grouped in '[asymmetric pairs]' :

$n = 6$

$[\ \langle 001011 \rangle^3 \quad \langle 001101 \rangle^3\ ]$

$n = 7$

$[\ \langle 0001011 \rangle^3 \quad \langle 0001101 \rangle^3\ ] \quad [\ \langle 0010111 \rangle^4 \quad \langle 0011101 \rangle^4\ ]$

$n = 8$

$[\ \langle 00001011 \rangle^3 \quad \langle 00001101 \rangle^3\ ] \quad [\ \langle 00010011 \rangle^3 \quad \langle 00011001 \rangle^3\ ]$

$[\ \langle 00010111 \rangle^4 \quad \langle 00011101 \rangle^4\ ] \quad [\ \langle 00101011 \rangle^4 \quad \langle 00110101 \rangle^4\ ]$

$[\ \langle 00101111 \rangle^5 \quad \langle 00111101 \rangle^5\ ] \quad [\ \langle 00110111 \rangle^5 \quad \langle 00111011 \rangle^5\ ]$

$n = 9$

$[\ \langle 000001011 \rangle^3 \quad \langle 000001101 \rangle^3\ ] \quad [\ \langle 000010011 \rangle^3 \quad \langle 000011001 \rangle^3\ ]$

$[\ \langle 000100101 \rangle^3 \quad \langle 000101001 \rangle^3\ ] \quad [\ \langle 000010111 \rangle^4 \quad \langle 000011101 \rangle^4\ ]$

$[\ \langle 000100111 \rangle^4 \quad \langle 000111001 \rangle^4\ ] \quad [\ \langle 000101011 \rangle^4 \quad \langle 000110101 \rangle^4\ ]$

$[\ \langle 001001011 \rangle^4 \quad \langle 001001101 \rangle^4\ ] \quad [\ \langle 000101111 \rangle^5 \quad \langle 000111101 \rangle^5\ ]$

$[\ \langle 000110111 \rangle^5 \quad \langle 000111011 \rangle^5\ ] \quad [\ \langle 001010111 \rangle^5 \quad \langle 001110101 \rangle^5\ ]$

$[\ \langle 001011011 \rangle^5 \quad \langle 001101101 \rangle^5\ ] \quad [\ \langle 001011111 \rangle^6 \quad \langle 001111101 \rangle^6\ ]$

$[\ \langle 001101111 \rangle^6 \quad \langle 001111011 \rangle^6\ ] \quad [\ \langle 010110111 \rangle^6 \quad \langle 010111011 \rangle^6\ ]$

.... [End of Appendix **B**] ....

oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo

Goops $_{\text{v2.0}}$

Keith Bowden

Theoretical Physics Research Unit, Birkbeck College, University of London, Malet St, London WC1E 7HX

k.bowden@physics.bbk.ac.uk

This paper is based on joint work with John Amson and Louis Gidney. All errors, omissions and style are my fault. Note my erroneous usage of the phrase "equivalence class" in earlier versions of this paper (see below. I still prefer it, especially in this context!) Note particularly that my use of XXXXoids is different to John's. I use the expression to include *only* equivalence classes of circoids and *not*, for instance, loops of loops. This is not a formal paper; it is just meant to tell a story. ☺ I hope to include a formal paper in next year's volume.

Abstract

Goop is what stuff is made of. The American Philosopher David Louis talked of "Atomless Gunk" in his book "Parts of Classes". He then gave structure to the atoms. We do the same.

Definitions

A bit vector is an ordered set of binary bits, or bit-string. A circlor is an ordered set where the order is circular (ie, a loop). We usually refer to a bit circlor as a circ.

LOOPS is the set of equivalence classes of circs in (simply connected) two dimensional space. *(Note that in earlier versions of this paper I incorrectly said that LOOPS was "the equivalence class of circs", because I can never remember the rule that an equivalence class is a set which may not be a class, that is contained in a class which may not be a set!)*

n-LOOPS is the set of equivalence classes of n-circs (circs with n bits) in (simply connected) two dimensional space.

That is, if you take two circs and can rotate and translate one so that it becomes identical with the other then they are equivalent. The members of (n-)LOOPS are (n-)loops respectively. A loop is thus a set (because it is an equivalence class). It is one of the sets of free floating bit circs that are indistinguishable in two dimensional space without flipping them over. A loop is referred to by a label that is any one of its circs written clockwise as a bit string, by breaking the circ at any one point. Then 010011 = 100110, but 010011 and 110010 are different 6-loops. This is the smallest loop to show chirality (handedness), ie. that is different to its mirror image or flip.

In this paper we will generally think of loops of bits, but in general we can have loops of any objects, even loops of loops, which we will mention briefly below.

The number of loops L(n) of size n is given by the necklace formula[1], eg L(6)=14. This converges rapidly towards the value given by $(2^n-2)/n+2$ with n and is exact for n prime. There are 2^n bit strings of length n. Now a bit string is a circ which is cut at some arbitrary point. There are n ways of making the cut. So there must be (in the limit) $2^n/n$ loops of length n. However we can improve this formula by noticing that the 0-loop and the 1-loop can only be cut in one way to make a bit string. The new formula is then $(2^n-2)/n+2$. This converges slightly faster. We observe that for primes $L(p) = (2^p-2)/p+2$ exactly (by substituting n=p in the necklace formula). Then for primes $(2^n-2)/n = L(n)-2$ which is an integer. This remarkable corollary to Fermat's Little Theorem was rediscovered in this manner by Louis Gidney. It generates all the primes plus a very small number of non-primes (called pseudoprimes).

All the above is well known in the literature where loops are referred to as necklaces and hoops are referred to as bracelets (because they are smaller). John Amson brought loops to the attention of ANPA and has referred to them as hoops, but (for intuitive reasons):

(n-)HOOPS is the set of equivalence classes of (n-)circs (under translation and rotation) in (simply connected) three dimensional space. This is equivalent to the introduction of flipping into the definition of loops (in two dimensional space). That is, if you take two circs and can rotate, flip and translate one so that it becomes identical with the other then they are equivalent.

The members of HOOPS are hoops (we will drop the (n-) where its use is obvious). A hoop is thus a set. It is one of the sets of circs that are indistinguishable in three dimensional space. A hoop is referred to by a label that is any one of its circs written as a bit string, by breaking the circ at any one point. Then 010011 and 110010 are the same 6-hoop and we write 010011=110010. The number of hoops H(n) of size n is then given by the bracelet formula[1], eg H(6)=13.

The Inner Product

If it is unnecessary to distinguish between hoops and loops then we talk about goops and GOOPS. Louis says that the "g" stands for generic, but I suspect that it stands for Gidney. We now wish to define operations like

$*,+ : \text{n-GOOPS} \times \text{n-GOOPS} \rightarrow \text{n-GOOPS}$

Applying any one n-loop to any other with an arbitrary relative orientation we get an object. (We will now drop the n when the meaning is obvious and assume that we are talking about n-loops.) Applying any one loop to any other of the same size in all n possible relative orientations we get n objects. Keith Bowden first observed that these naturally form a loop with the obvious ordering and that if * is an operation on the loop labels, that * is then independent of the label that we use to represent each loop, as it should be. In other words

010011*010101 = 100110*010101

as it should.

Applying any one n-hoop to any other with an arbitrary relative orientation we get an object. Applying any one hoop to any other of the same size in all n possible relative orientations and flip we get 2n objects. These form a pair of loops with the natural orderings. That the pair of loops are equal is not difficult to see. If * is an operation on the hoop labels, the ability to flip one object relative to the other makes nonevermind and it is clear that * is then independent of the label that we use to represent each hoop, as it should be. In other words

010011*010101 = 110010*101010

as it should.

Consider * as an inner product on its operands, then it is related to the definitions of * and + on bits in the obvious way, and its commutativity, associativity and distributivity depends in turn upon their commutativity, associativity and distributivity, but not in any obvious way. There are 16 possible definitions of each of the bit operators, and so 256 possible inner products. Define + on loops as an inner product on its operands, then there are 256 possible inner sums. Consider the special one with the definitions of bit * and + swapped over relative to the inner product * on loops.

Most of these loop operators are neither commutative, associative, self distributive, nor distributive relative to its partner. Clearly, however, * is commutative for n<6. For n>5 the noncommutative pairs in the multiplication table are always chiral pairs (see the tables in John's paper in this volume for n=6 with bit operators AND and XOR). So for n=6 they are 110010 and 010011. Whether or not there are non-trivial operators such that the associated loops are associative and form categories is unknown. However the hoop operators are commutative by the equality of their "chiral pairs". It was originally thought that they are also associative and thus would form categories, but this turns out not to be true. The formal definition of HOOP(OID)S will be as the ***abelianisation*** of LOOP(OID)S obtained by the appending of the symmetrised elements from the LOOP(OIDS) to their ***centre*** by setting each non centre element equal to its chiral (or conjugate) pair (or, more generally, just identifying gh and hg whenever they are not equal). Thus

order(n-HOOPS)=(order(centre(n-LOOPS)) + order(n-LOOPS))/2

eg H(6) = (C(6)+L(6))/2 = (12+14)/2 = 13.

Categorically LOOP(OIDS) are groupoids (in the old fashioned sense of a set with an operator). In general they have neither units nor all inverses (but do have centres and abelianisations).

Goopoids

If a loop is an equivalence class of circs then a loopoid is an equivalence class of circoids under smooth homotopic relative motion of their surfaces ("rolling and rotation"). A circoid is made by joining both the sides and the top and bottom of a rectangular bit matrix together to make a torus. There are two ways of doing this and the transformation between the two is the circoid (or loopoid) transpose. A circoid formed from a rectangular matrix can either be short and fat or long and thin. We write (m,n)-loopoids for m×n loopoids. An (m,n)-loopoid is a 2-loopoid whereas a 1-loopoid is a loop.

We now wish to define operations like

$*,+ : \text{GOOPOIDS} \times \text{GOOPOIDS} \rightarrow \text{GOOPOIDS}$

Applying any one (m,n)-loopoid to any other with an arbitrary relative orientation (that is we can rotate or roll one with respect to the other) we get an object. (We will now drop the (m,n) when the meaning is obvious and assume that we are talking about (m,n)-loopoids.) Applying any one loopoid to any other of the same size in all mn possible relative orientations we get mn objects. These naturally form a loopoid with the obvious ordering and if * is an operation (eg, inner product) on the loop labels, then * should be independent of the label that we use to represent each loop. In other words eg,

010*010 = 100*010
011 101   110 101

as it should.

Hoopoids are formed by defining chiral pairs of loopoids in the multiplication table to be equal (or more formally by ***abelianisation*** as mentioned above). Applying any one (m,n)-hoopoid to any other with an arbitrary relative orientation we get an object. Applying any one hoopoid to any other of the same size in all mn possible relative orientations we get mn objects. These form a hoopoid with the obvious natural ordering. If * is an operation on the hoopoid labels, then it is clear that * is then independent of the label that we use to represent each hoop, as it should be. In other words

010*010 = 110*101
011 101   010 010

as it should.

Eventually we would like to construct appropriate products

$*:\text{GOOPOIDS} \times \text{GOOPS} \rightarrow \text{GOOPS}$

etc. These products are not obvious; care must be taken with contravariance and covariance (upper and lower indices).

For an m×n loopoid it is surprisingly not necessary for either m or n to be greater than 5 for chirality to appear. This is because in general the torus in 3D space is not the same as its mirror image. This initially made us believe that the chirality of the loopoid was different to that of the loop.

However it has recently been noticed that there is an uncanny resemblance between the 2×3 loopoids and the 6-loops. There are 14 of each. The smallest loopoid to show chirality is the (2,3)-loopoid

001
011

and its chiral partner

100
110

Reading the bit patterns in the two loopoids shown above rowwise we get

001011 and 100110

which is the 6-loop chiral pair. We can easily match the 0 and 1 loops with the 0 and 1 loopoids in this manner. There is only one loop and one loopoid with a Hamming norm (number of 1's) of either 1 or 5.

This leaves three of each with Hamming norms of 2 or 4 and two of each with Hamming norms of 3. There is a also natural way of pairing these which has implications about how one should cut and join loops.

However all this depends on reading the loopoids rowwise. Reading columnwise gives a (partially) different pairing. This gives each hoop a natural (TRANSPOSE) dual (which may be itself) to add to the chiral (MIRROR) dual and the inverse (NOT) dual. It is not clear if the three dual operators commute.

It was thought that the implied generalisation of the (m,n)-loopoid → (m×n)-loop mapping may eventually justify John Amson's (unproven) rule that L(m,n)=L(m×n) (and its generalisations), but the generalisation has now been disproved by the discovery that the number of (2,4) loopoids is 40 and not 36 as it "should be" if L(2,4)=L(8). It is now thought that this will not change the situation much. Instead of a collection of equivalent (isomorphic!) isomorphisms between LOOPOIDS(2,4) and LOOPS(8) it is now thought there will be a collection of isomorphic monomorphisms from LOOPS(8) into LOOPOIDS(2,4) under the appropriate "inner products".

(m,n)	L(m,n)		L(m×n)		H(m,n)	H(m×n)
2,2	6	=	L(4)=6	(Equal)	6	6
2,3	14	=	L(6)=14	(Equal)	13	13
2,4	40	>	L(8)=36		31	30
3,3	64	>	L(9)=60		?	46
3,4	352	=	L(12)=352	(Why the equality?)		
4,4	4154	>	L(16)=4116			

Finally, Louis Gidney has recently pointed out that the mirror dual and the flip dual are different from 3×3 loopoids upwards. Consider the chiral quadruplet

```
000  000  000  000
001  100  110  011
011  110  100  001
```

These four 3×3 loopoids are all different. You cannot get from one label to another by rotating and rolling. The question of how they will occur in chiral pairs in the chiral symmetry of the (noncommutative) multiplication table is still open. There are reasons for believing that only the mirror dual will appear with chiral symmetry but this is not known yet. All this will become clear after we have constructed multiplication tables for Loopoids (and hence Hoopóids) and can see how the chiral pairs are constructed.

n-Loops

An (n)-loop is a loop with n elements (grade 1, dimension n).

An (n)-loopoid is a grade n loopoid.

An (n,m,k)-loop(oid) is a (3)-loopoid with 3 dimensions of sizes n, m, and and nmk elements.

A 2-loop is a loop of loops.

A 3-loop is a loop of loops of loops.

A 3-loopoid is a loopoid of loopoids of loopoids.

An (n, m; (k, l)) 2-loopoid is a (n,m)-loopoid of (k,l)-loopoids.

A 2-loop is NOT a loopoid!

Proof for 2×3 loopoids:

There are 14 2×3 loopoids. There are 13 2×3 hoopoids.

But there are only 10 (2; (3)) 2-loops. And there are only 11 (3; (2)) 2-loops. QED.

Multiplication of 2-loops is just hierarchical overloaded loop multiplication.

ON THE OTHER HAND:

A circ of circs IS a circoid. Now any particular circoid, C is a label for an equivalence class of loopoids.

And any circ is a label for an equivalence class of loops.

Now replace C by its circ of circs rowwise. So C is a circ of labels of equivalence classes of loops. But this is NOT a circ of loops because it is a circ of FIXED labels. You cannot replace a label by another equivalent label at will!

References

"A Handbook of Integer Sequences" by N.J.A.Sloane, Acad.Press 1971.

"An Efficient Algorithm for Generating Necklaces with Fixed Density" by Frank Ruskey and Joe Sawada. SIAM J.COMP. 1999, V29, No 2, pp 671-685.

Three Routes to Quantization

G. N. Ord*
M.P.C.S.
Ryerson University
Toronto. Ont.

In this article we sketch three generic methods for inducing quantization in a simple one-dimensional model. The first method is formal, well-known, and routinely used. The second method originates in Nelson's stochastic mechanics, is less formal but less well-known. The third technique is new, constructive, but of unknown generality.

It is now over a hundred years since Max Planck made the suggestion that electromagnetic energy might be quantized. The suggestion, made in order to account for the blackbody radiation spectrum, was daring in the extreme. Taken at face value it meant that light could have both wave *and* particle properties. This would have made little physical sense at the time. Particles are abstractions of objects that are strongly localized in space, waves are abstractions of strongly de-localized objects. Suggesting there are objects which behave as if they simultaneously possess such mutually exclusive features was surely inspired madness. Even today, knowing as we do that Nature fully supports Planck's suggestion, there is still no physical theory that explains precisely *how* nature reconciles wave and particle.

So what is wave-particle duality? Apparently this is a feature of Nature that allows objects to behave either as particles or as waves, depending on circumstance. How Nature accomplishes this feat is not part of conventional quantum theory, however it is quantization that decrees that wave and particle be merged. It is this process that we examine here by sketching three distinct routes to wave equations from a particle paradigm. Of the three routes, only the first route, which we refer to as canonical quantization(CQ), is routinely used in mainstream physics. CQ effectively combines three features that old quantum theory struggled to patch into classical physics.

The first feature was the apparent 'duality' of particles and waves. Blackbody radiation and the photo-electric effect could only be understood by treating what were perceived as classical waves, as particles. The hydrogen atom could only be understood by treating particles as waves. Nature's confusion over what classical physicists considered mutually exclusive concepts came to a head with Heisenberg's

*corresponding author gord@acs.ryerson.ca

uncertainty principle in 1925. $\Delta x \Delta p \geq \hbar$ clearly signalled the end of the classical particle paradigm in the domain of the very small. CQ implemented wave-particle duality by replacing classical one-particle systems by related systems with an infinite number of degrees of freedom, changing ordinary differential equations to partial differential equations.

The second feature of CQ is the extraction of discrete states, the primary empirical motivation for QM. The discretization is accomplished through 'reasonable' requirements (like single-valuedness) placed on the solutions of the equations of quantum mechanics.

The third feature of CQ is the connection with classical physics. The elegant $p \to -i\hbar\nabla$, $E \to i\hbar\frac{\partial}{\partial t}$ or its equivalent for the Dirac equation extracts the empirically correct *rules* from classical physics as if by magic. Although the transition back to classical physics via measurement theory is less-than-transparent, there can be no doubt that CQ somehow captures a vast amount of physics simply by replacing classical dynamical variables by operators. However the formality of this replacement means that quantum mechanics really is an algorithm rather than a physical theory, and it is for this reason that the two other quantization methods sketched here deserve some attention. Both alternative methods focus on what formal quantization *does*, and how Nature might go about producing phenomena whose description involves quantum mechanics.

Now the physical motivation for CQ is wave-particle duality which is often associated with the uncertainty principle. In the two other quantization methods, the uncertainty principle arises through Fractal geometry, not wave-particle duality. This is because a particle moving on a fractal curve of dimension D obeys a 'classical' uncertainty principle of the form $\Delta x \Delta v \geq \text{const}(\Delta x)^{-2(D-2)}$[1]. This result ties quantum mechanics to stochastic processes because Brownian and Poisson paths are known to have fractal dimensions $D = 2$ on appropriate scales. Noticing this fact is important in understanding what really distinguishes quantum mechanics from classical physics. When we compare for example the diffusion equation.

$$\frac{\partial u}{\partial t} = D\,\nabla^2 u \tag{1}$$

with the Schrödinger equation

$$\frac{\partial \psi}{\partial t} = i\frac{\hbar}{2m}\,\nabla^2 \psi \tag{2}$$

we see that the most important aspect of CQ is that it involves a formal analytic continuation(FAC)! It is the FAC that underlies quantum interference and the resulting discretization. For this reason we will pay little attention to other aspects of CQ and simply call CQ a FAC. We then call the second method of quantization a " globally forced analytic continuation" or GFAC. In this method, first exploited by Nelson and Fenyes, you force a mathematical model into the domain of complex numbers, (or spinors) by global 'physical' requirements that cannot be met by simple real numbers. The advantage of procedures of this type is that you gain insight into possible physics behind the empirical rules.

The last and most recent quantization method which we call geometrically induced quantization (GIQ) induces the algebra of complex numbers specifically through the local space-time geometry of particle paths. The advantage of this approach is that it builds in a specific ontology for wave-particle duality. There is no need for a formal analytic continuation, or a global enforcement of a restriction on ensemble properties. The algebra of quantum mechanics arises directly from local space-time geometry, by construction. The wave properties of the 'quantum system' follow directly from the geometry of a single space-time path. This is a very exciting result. The drawback of the result is that it is currently only implemented in one dimension. It remains to be seen how much of CQ can be accounted for by a GIQ.

In this article we take a single basic stochastic model through the three routes to quantization just to compare and contrast the treatments. Each path to the PDE displays different features of the quantization procedure. The first path is quick, elegant, and well-known (Section 1). The second is less elegant but shows that we can in principle get the Dirac equation by forcing a diffusive system to be globally 'causal'(Section 2). The third method shows that it is possible by local geometry alone to force an equilibrium state that conforms to the Dirac algebra (Section 3).

I. THE KAC MODEL AND FORMAL ANALYTIC CONTINUATION

We start with a Stochastic model of the telegraph equations due to Marc Kac[2]. We imagine a particle moving on a discrete lattice in 1-D. The particles move with speed $v = \frac{\Delta x}{\Delta t}$ and change direction with probability $a\Delta t << 1$. If $P_\pm(x,t)$ represents the probability that the particle is at (x,t) moving in the $\pm$ direction then the master equation for $P_\pm(x,t)$ is:

$$P_\pm(x, t + \Delta t) = (1 - a\Delta t)P_\pm(x \mp \Delta x, t) + a\Delta t P_\mp(x \pm \Delta x, t) \qquad (3)$$

Equation (3) just expresses the fact that particles cannot disappear or be created. If we let Δx and Δt go to zero with fixed v we find:

$$\frac{\partial P}{\partial t} = -v\,\sigma_z \frac{\partial P}{\partial x} - aP + a\sigma_x P \tag{4}$$

where $P = [P_+, P_-]^T$ and the σ_k are the usual Pauli matrices. Equation (4) is just a two component form of the Telegraph equations. The second order form:

$$\frac{\partial^2 P_\pm}{\partial t^2} - v^2 \frac{\partial^2 P_\pm}{\partial x^2} = -2a\frac{\partial P_\pm}{\partial t} \tag{5}$$

is recovered by iteration and we see that if we let $v \to \infty$ in such a way that $\frac{v^2}{2a} \to D$, we get the diffusion equation with diffusion constant D. The solutions of (5) decay exponentially in time so if we write

$$P(x,t) = u(x,t)e^{-at} \tag{6}$$

to see the fluctuations about the exponential we find that u satisfies:

$$\frac{\partial u}{\partial t} = -\sigma_z \frac{\partial u}{\partial x} + a\sigma_x u. \tag{7}$$

As first shown by Gaveau et. al.[3], we can now connect this form of the Telegraph equations to the Dirac equation through a FAC. If we replace the real positive constant a by im and v by c (7) becomes:

$$\frac{\partial u}{\partial t} = -c\sigma_z \frac{\partial u}{\partial x} + im\sigma_x u. \tag{8}$$

The FAC that transforms (7) into (8) does not represent a physical operation. It changes a phenomenological equation(Telegraph) with a solid ontological basis (Poisson paths) to a 'Fundamental' equation (Dirac) with no known ontological basis. Note that the Feynman Chessboard paths[4–6] that underlie (8) are not 'real' in the sense that the Poisson paths underlying (7) are real. Diffusing particles have trajectories that are approximations of Poisson paths, but the same cannot be said for quantum 'particles' being described by chessboard paths.

The FAC involved in transforming (7) to (8) was by decree. It had nothing to do with the physical process underlying the Telegraph equations. The question then arises, "Can we force a stochastic process to give us (8) without explicitly dictating a FAC?"

II. FORCING AN ANALYTIC CONTINUATION

In the above 'derivation' of the Dirac equation, it is hard to see any physics behind the FAC. The model has a well-defined ontological basis up to equation (7), but the FAC removes the logical connection between equation (8) and the physical arguments underlying equation (7). What is needed is an argument that will in some sense relate the FAC to requirements on a physical system. The first example of such an argument can be found in the work of Nelson[7]. In that work, diffusion, which is macroscopically irreversible, was forced to be reversible. The reversibility requirement was enforced in just such a way that the result was Schrödinger's equation. In this section we shall force an analytic continuation on a generalization of the Kac model of section 1, following the work of McKeon[8]. To do this we will allow particles to move in *either* direction in time on small scales.

As before consider an ensemble of point particles moving on a space-time lattice with spacings Δx and Δt. We fix the ratio $\frac{\Delta x}{\Delta t}$ at $c = 1$ and construct the dynamics in the following way. Let $F_{\pm}(x,t)$ be proportional to the number of particles moving forward in time going in the direction specified by the subscript $\pm$. Thus for example F_- denotes a density of particles moving in the $(-x, +t)$ direction. Similarly let $B_{\pm}(x,t)$ be proportional to the number of particles moving backward in time. Suppose now that at each lattice site, any particles contributing to the forward-in-time densities turn left with probability $a_L\Delta t$ and turn right with probability $a_R\Delta t$. Conservation of particle number then gives us

$$\begin{aligned} F_{\pm}(x,t) &= (1 - a_L\Delta t - a_R\Delta t)\, F_{\pm}(x \mp \Delta x, t - \Delta t) \\ &\quad + a_{L,R}\Delta t B_{\pm}(x \mp \Delta x, t + \Delta t) \\ &\quad + a_{R,L}\Delta t F_{\mp}(x \pm \Delta x, t - \Delta t) \end{aligned} \tag{9}$$

To find the remaining two difference equations for the two time reversed densities we impose the 'causality condition' [8]

$$F_{\pm}(x,t) = B_{\mp}(x \pm \Delta x, t + \Delta t)\,. \tag{10}$$

on equation(9) to order Δt. This will force the difference equations to be 'causal' in the sense that given initial conditions at $t = 0$ can be iterated forward in t. It can also be seen as a detailed balance condition that strongly inhibits charge separation. Substituting (10) into (9) gives

$$\begin{aligned} B_{\pm}(x \mp \Delta x, t + \Delta t) &= (1 - (a_L + a_R)\Delta t)B_{\pm}(x,t) \\ &\quad + a_{L,R}\Delta t B_{\mp}(x,t) \\ &\quad + a_{R,L}\Delta t F_{\pm}(x,t) \end{aligned} \tag{11}$$

Shifting (11) one lattice spacing in x and t gives

$$\begin{aligned} B_{\pm}(x,t) &= (1-(a_L+a_R)\Delta t)B_{\pm}(x\pm\Delta x, t-\Delta t) \\ &\quad + a_{L,R}\Delta t B_{\mp}(x\pm\Delta x, t-\Delta t) \\ &\quad + a_{R,L}\Delta t F_{\pm}(x\pm\Delta x, t-\Delta t) \end{aligned} \tag{12}$$

Now writing

$$\begin{aligned} z_1 &= F_+ - B_- \\ z_2 &= F_- - B_+ \\ z_3 &= F_+ + B_- \\ z_4 &= F_- + B_+ \end{aligned} \tag{13}$$

we see from (9) and (13) that to first order in Δx and Δt

$$\begin{aligned} z_1(x,t) &= z_1(x-\Delta x, t-\Delta t) + \Delta t\ (\ (a_R - a_L)\ z_2 - (a_R + a_L)\ z_1\ ) \\ z_2(x,t) &= z_2(x+\Delta x, t-\Delta t) + \Delta t\ (\ (a_L - a_R)\ z_1 - (a_R + a_L)\ z_2\ ) \end{aligned} \tag{14}$$

and

$$\begin{aligned} z_3(x,t) &= z_3(x-\Delta x, t-\Delta t) + \Delta t\ (\ (a_R + a_L)\ z_4 - (a_R + a_L)\ z_3\ ) \\ z_4(x,t) &= z_4(x+\Delta x, t-\Delta t) + \Delta t\ (\ (a_L + a_R)\ z_3 - (a_R + a_L)\ z_4\ ). \end{aligned} \tag{15}$$

Here we note that the difference equations are decoupled to this order. Assuming that the solutions to the difference equations are sufficiently slowly varying on the scale of Δt that they may be approximated by smooth functions, we expand the right hand side of (14) about the point (x,t) to get

$$\begin{aligned} &z_1(x,t) + \left(-c\,\frac{\partial z_1}{\partial x} - \frac{\partial z_1}{\partial t} - \lambda\, z_2 - \mu\, z_1\right)\Delta t + O\left(\Delta t^2\right) \\ &z_2(x,t) + \left(c\,\frac{\partial z_2}{\partial x} - \frac{\partial z_2}{\partial t} + \lambda\, z_1 - \mu\, z_2\right)\Delta t + O\left(\Delta t^2\right) \end{aligned} \tag{16}$$

where we have written

$$\lambda = a_L - a_R \tag{17}$$

and

$$\mu = a_L + a_R. \tag{18}$$

Setting the coefficient of Δt to zero in this we get

$$\frac{\partial z_1}{\partial t} = -c\,\frac{\partial z_1}{\partial x} - \lambda\, z_2 - \mu\, z_1$$

$$\frac{\partial z_2}{\partial t} = c\,\frac{\partial z_2}{\partial x} + \lambda\, z_1 - \mu\, z_2 \tag{19}$$

Finally if we write $\Psi^T = (z_1, z_2)\, e^{\mu t}$ and $\lambda = mc^2/\hbar$, (19) may be written as

$$i\hbar\frac{\partial \Psi}{\partial t} = mc^2\sigma_y\Psi - ic\hbar\sigma_z\frac{\partial \Psi}{\partial x} \tag{20}$$

where σ_y and σ_z are Pauli matrices. This is a representation of the Dirac equation in 1+1 dimensions. Here note that in (19), z_1 and z_2 are real, since they are just approximations to the real solutions of the stochastic system (14). The wave aspects of the solutions arise from the fact that z_1 and z_2 represent a *net* flux of particles in the $+t$ direction.

Notice what has happened here. The first two relations (9) were generated by a simple physical requirement on the ensemble densities $F_\pm$. It is easy to see that these could be met by particles scattering with the appropriate probabilities. Had we just rotated these requirements in space-time the resulting equations would have described spiralling particles, but the ensemble averages would not have evolved causally in time. As a result we enforced a global causality condition which forced the system to evolve forward in time. The result was the Dirac equation. In the process we see that if we make the dynamics of individual particles such that (9) and (10) obtain, we then have a microscopic stochastic origin of the Dirac equation.

Unfortunately, the causal condition (10) is a restriction on ensemble averages, not individual trajectories. Thus we have not got, at this point, a fully ontological basis for (20).

There is however a strong hint in the calculation that in fact *there should be an ontological basis for (20)*. This hint arises from the fact that, in the continuum limit, the causality condition(10) forces the path densities F and B to diverge in comparison to z_1 and z_2. This means that in order to 'see' Ψ we must project *perfectly* onto the relevant eigenspace, meaning that our detectors must measure only the differences between F and B and must not attempt the subtraction themselves. Furthermore, the underlying stochastic process must be in perfect equilibrium. Statistical fluctuations that could leak components of z_3 or z_4 into the 'quantum' eigenspace would completely overwhelm the solutions of the Dirac equation. Thus our Dirac equation is completely ill-conditioned with respect to the continuum limit! This is a common feature of derivations of the Dirac equation which do not start in the continuum. Now the continuum itself is a convenient mathematical fiction which we routinely use to describe nature. We take real numbers for granted even though they would be impossible to measure. Part of the rationale for taking real numbers for granted is we assume that our mathematical models and the physical processes they describe are both transfinitely stable. If our calculations or observations are out in the 100'th decimal place we do not expect this to matter too much. The stability of Nature argues that this should be so. However the above calculation shows that we can construct sensible mathematical models in which the Dirac equation is spectacularly fragile. Either these models have no bearing on the same equation we use to describe Nature, or Nature has an intrinsic ontological element that insures stability, and our mathematical model has missed it.

The next section suggests that the latter is the case.

III. GEOMETRICALLY INDUCED ANALYTIC CONTINUATION AND ENTWINED PATHS

So far we have considered two routes to the Dirac equation. The first route was a simple FAC. By judicious replacement of 1 by i in a form of the Telegraph equations we invoked a subset of the Dirac algebra and obtained the required equation. In the second route, we forced the global requirement of causality on a physical system in such a way that the mathematical model responded with the appropriate algebra. This is like the familiar artifice of forcing all quadratics to have zeros, having discovered that some of them have real zeros. The result is an expansion of the number system with a 'new' algebra.

In the following derivation we avoid forcing an analytic continuation. Instead we pair paths in a way that induces the appropriate algebra through geometry. A more detailed treatment may be found in [9].

We shall be working in a two dimensional discrete space (z, t) with lattice spacings δ and ϵ respectively. Although we shall eventually think of t as time, it is convenient at this point to think of t as a spatial coordinate. Entwined paths can be generated by an ergodic process, and this is illustrated in Fig.1 A) and B). The particle proceeds forward in t (thick lines in figure), alternately reversing direction and dropping a marker according to a Poisson process. After some specified time the direction in t is reversed at a marker and the particle proceeds back to the origin

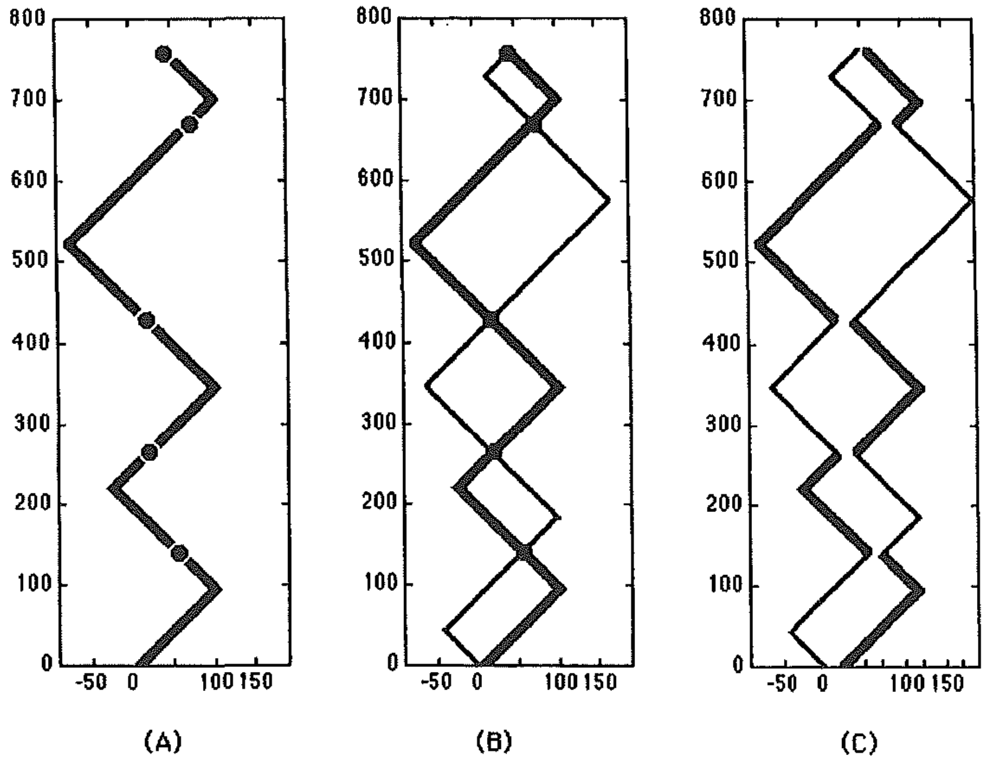


FIG. 1: Forming entwined paths in space-time. z is horizontal, t is vertical.
A) A particle starting at the origin travels at constant speed but occasionally reverses direction in response to a stochastic process. At every other indication from the stochastic process, a marker is dropped instead of a direction changed.
B) Past a specified time t_R, the particle reverses direction in t at the next marker. The particle then follows the 'light-cone' paths through the markers back to the origin (thin line in (B)).
C) The entwined path formed in (B) can be regarded as two osculating paths which we call envelopes. These are separated for clarity. We use the envelopes to count the oriented rectangles formed in (B).

along a light-cone path fixed by the markers (thin line in Fig. 1 B)). This process is repeated building up a background of oriented rectangles on spacetime. Here the 'orientation' comes from distinguishing the direction of traversal in the t direction. If the first rectangle in Fig. 1 B) has positive orientation, the second has a negative orientation and the orientation changes whenever the thick and thin lines in the path cross. Note that if we count oriented rectangles, +1 for positive and -1 for negative, we need only use an 'envelope' (Fig. 1 C)) for counting. Counting signed contributions here is appropriate since if 't' represents macroscopic time, the thick and thin lines represent 'particle' and 'antiparticle' trajectories respectively. As the particle trajectory repeats creation of these entwined pairs, the particle number will increase, however the 'net charge' will always remain zero since entwined paths always pair thick and thin contributions.

How do oriented rectangles behave in ensembles of entwined pairs? This is easy to calculate if we consider ensembles of paths in terms of their envelopes.

Consider the left envelope path in Fig.1C. Although the path itself is generated by the full path of Fig.1B, for the purpose of counting oriented rectangles we note that we could consider generating the path with a stochastic algorithm and a 'phase rule' to determine when the charge is +1 (thick) or -1 (thin). The rule is as follows. Starting at the origin, the particle proceeds in the $-z$ direction until a stochastic process indicates a direction change. At the first (*right*) turn, the particle just changes direction but not thickness. At the second (*left*) turn the path also changes thickness, and the charge changes sign. This process is repeated. Each right corner maintains the charge, each left turn changes the charge and the direction. The reader familiar with the Feynman chessboard model will recognize this rule as a version of Feynman's corner rule[6, 10]. In this context the rule is dictated by the geometry of entwined paths. Left turns in the left envelope are actually crossing points of the particle-antiparticle pair, and the origin of the sign change is physical.

If we now let $\phi_n^1(z)$ be the ensemble charge density from left envelope links parallel to the left light-cone at step n and $\phi_n^2(z)$ be the ensemble charge density from left envelope links parallel to the right light-cone, we can write:

$$\begin{aligned}\phi_n^1(z) &= (1-a\Delta t)\phi_{n-1}^1(z+c\Delta t) - a\Delta t\phi_{n-1}^2(z-c\Delta t) \\ \phi_n^2(z) &= (1-a\Delta t)\phi_{n-1}^2(z-c\Delta t) + a\Delta t\phi_{n-1}^1(z+c\Delta t)\end{aligned} \tag{21}$$

That is, most paths maintain their direction and colour as they pass through a lattice site. The proportion which do this is $(1 - a\Delta t)$. However a proportion $a\Delta t$ change direction at the site. When they scatter from the right light cone they change charge on scattering, so they *decrease* the net charge in the new direction in proportion to the density in the old direction. However when they scatter from the left light cone they maintain their charge on scattering, so they *increase* the net charge in the new direction in proportion to the density in the old direction. This rule is unmotivated if one considers an 'envelope' by itself. However, considering how an envelope is generated by entwined paths, it is a constraint imposed by geometry.

The right envelope is similar, except here it is the right turns which change charge. If we let $\phi_n^3(z)$ and $\phi_n^4(z)$ be the right envelope charges parallel to respectively left light cones and right light cones we have

$$\begin{aligned}\phi_n^3(z) &= (1-a\Delta t)\phi_{n-1}^3(z+c\Delta t) + a\Delta t\phi_{n-1}^4(z-c\Delta t) \\ \phi_n^4(z) &= (1-a\Delta t)\phi_{n-1}^4(z-c\Delta t) - a\Delta t\phi_{n-1}^3(z+c\Delta t).\end{aligned} \tag{22}$$

Equations (21-22) constitute a set of coupled difference equations in the four densities $\phi_1 - \phi_4$. Although their derivation is straightforward, their consequences as a description of a classical ergodic stochastic process are potentially far-reaching, since these are representations of the discrete Dirac equations.

Since the above arguments discuss only ensemble averages. It is possible that the above equations are correct for the ensemble average, but that the underlying stochastic process gives rise to such large fluctuations that the ensemble average survives only in the event of a uniformly covered ensemble. In this case, normal stochastic fluctuations would swamp the signal and the system would not exhibit the above equations except under the rare circumstance of an almost perfectly uniform coverage of the ensemble. In other words, if we watch the stochastic process assembling the propagator, will it converge or will the natural fluctuations swamp the signal?

In previous works[9, 11] we showed numerical evidence that if we fixed the origin of the stochastic process and let it run through many cycles, the resulting pattern converged to the Dirac propagator for a free particle. We have also shown that we can put the stochastic process in a periodic box and the resulting pattern of oriented rectangles recognize the box eigenfunctions. These studies verify that the artifice of pairing the forward and backwards paths is a sufficiently strong requirement on particle paths that it induces a form of quantization that is stable against 'vacuum fluctuations'.

It is tempting to speculate on the relevance of this result to quantum mechanics. Here we have been working on the level of the simplest possible system in QM, the one-dimensional free particle. It is thus *not* obvious that we can conclude that all quantization has an ontological basis in the geometry of space-time. However, on the positive side it is also no longer advisable to assert that such a scheme is impossible, since in one dimension the above demonstration provides a very straight-forward counter-example.

Also on the positive side we hope the result will help the renewed interest in Bohm's theory. In one dimension at least, the above result does support the general Bohm picture of a real particle trajectory and an associated wavefunction. In the above picture we have shown how a single particle can generate a wavefunction which appears to propagate unitarily through space-time. To use this as a model of a real particle, measurement of a particle property will have to fix the particle's macroscopic history in order to avoid violation of causality. This would prevent return paths backwards in time. Thus there will always be one unpaired path connecting successive observations. How close this is to the Bohm picture in detail is unknown, but we are optimistic that the study of EP's will help clarify what is really meant by 'quantization'.

Acknowledgments

This work was partly funded by NSERC. The author is grateful to the Perimeter Institute where this manuscript was completed.

---

[1] G. N. Ord, J.Phys.A **16**, 1869 (1983).
[2] M. Kac, Rocky Mountain Journal of Mathematics p. 4 (1974).
[3] B. Gaveau, T. Jacobson, M. Kac, and L. S. Schulman, Phys. Rev. Lett. **53**, 419 (1984).
[4] H. A. Gersch, Int. J. Theor. Physics **20**, 491 (1981).
[5] T. Jacobson and L. Schulman, J. Physics A **17**, 375 (1984).
[6] G. N. Ord, J. Stat. Phys. **66**, 647 (1992).
[7] E. Nelson, Physical Review **150**, 1079 (1966).
[8] D. G. C. McKeon and G. N. Ord, Phys. Rev. Lett. **69**, 3 (1992).
[9] G. N. Ord and R. B. Mann, Phys. Rev.A **67** (2003).
[10] L. H. Kauffman and H. P. Noyes, Phys. Lett. A **218**, 139 (1996).
[11] G. N. Ord and J. A. Gualtieri, Phys. Rev. Lett **89** (2002).

Beyond the Hilbert Space Formulation

Milan Glendza (ed KGB)

Theoretical Physics Research Unit, Birkbeck College,
Malet Street, London WC1E 7HX

M.Glendza@mail.cryst.bbk.ac.uk
http://www.bbk.ac.uk/tpru/

Abstract

We know that due to quantum mechanical fluctuations at the level of the Planck scale, the concept of space and time and the more general mathematical structure of a topological space is devoid of any meaning. Some authors, and among them especially J A Wheeler (1), have advocated the existence of a more fundamental level of reality that is essentially quantum mechanical in nature, from which our classical view of geometry would emerge. We advocate a similar physical basis built with the mathematical apparatus of Noncommutative geometry or Quantum Geometry coupled with von Neumann operator algebras, knots, Quantum Theory and non-Boolean (non-commutative) type logic.

Introduction to Quantum Spaces

Quantum Geometry brings in a totally new concept of space through the use of methods of classical geometry, non-commutative C^* algebras and functional analysis on a background of quantum theory. We know from mathematics that all the geometries have embedded in them the concept of space. In our particular case, we introduce the notion of *quantum spaces* where the classical concept of space becomes a limiting special case of the more general mathematical structure of Quantum Geometry.

We know that classical geometrical spaces are always assumed to be a *collection of points* that inhabit the relevant additional structure. We note that quantum spaces do not have this kind of character at all. In general, our quantum spaces do not have any points at all. Indeed, they are sometimes called pointless spaces!

The assumption that the underlying fabric of space-time is a smooth manifold is perhaps one of the main causes of different mathematical inconsistency problems that plague Quantum Field Theory. The same assumption is carried through into theories that attempt, without much success, to find a general mathematical structure which would contain both quantum theory and general relativity. We note that inadequacies with such classical concepts as the nature of space and time appear to emerge at exceedingly small distances, that is, at distances of the order of magnitude of the Planck length where we have quantum fluctuations.

The abstract concepts of a space-time point and that of space-time coordinates become indefinable and meaningless at the quantum level.

The Special Theory of Relativity treats the concepts of space and time as in classical theory but with an important difference since it places a significant role on light. The principle of the constancy of the velocity of light in all inertial systems together with the principle of relativity relativizes both space and time.

In Einstein's words, "...It is neither the point in space nor the instant in time at which something happens that has physical reality. There is no absolute (independent of the space of reference) relation in space and no absolute relation in time between two events but there is an absolute (independent of the space of reference) relation in space and time..." (2)

The mathematical formalism of Quantum Geometry, or as it is also known in the mathematics world by its other name of Non-Commutative Geometry, is a fusion, as we have noted, of global methods from classical differential geometry, non-commutative algebras and functional analysis.

When we generalize classical geometry to the non-commutative level, we need to perform two conceptual moves. First we transform it into a commutative algebraic structure and then secondly, we generalize this commutative structure into a non-commutative structure.

The first move involves us in re-expressing a geometrical structure existing on a classical space X in terms of the algebraic structure of the associated, commutative *algebra of the appropriate complex-valued functions on X. We note that the geometrical structure on X is always completely expressible in the language of the associated *algebra. The second move consists of the appropriate non-commutative generalization of such algebraically reformulated geometry. The basic idea is to be able to replace the function algebras by some more general non-commutative *algebras, and to extend in such a manner the concept of space through the introduction of quantum spaces.

We can interpret the theory of von Neumann algebras as a quantum generalization of classical measure theory. Hence, we note that commutative von Neumann algebras describe classical measurable spaces whilst the non-commutative von Neumann algebras represent *quantum measure spaces*. The building blocks and background of quantum geometry are present in the foundational papers by Murray and von Neumann (3, 4, 5)

We pause before we deal with some particular examples of quantum spaces to observe and take note of the fact that one and the same concept of classical geometry may have several very different generalizations in quantum geometry. We see that this happens because the procedure of generalizing objects of commutative algebra into the objects of non-commutative algebras is not at all unique.

We see also that one possible way to introduce symmetries in quantum geometry is to consider automorphisms of the corresponding non-commutative algebras. An essentially different procedure is to play with quantum groups, and define the action of quantum groups on quantum spaces, generalizing the classical concept of a group action.

Some Examples of Quantum Spaces

1) Quantum Groups

We have a very important class of examples of quantum spaces given by quantum groups. We note, these are, by definition, quantum spaces equipped with a 'group structure'. At this point, we shall outline how the concept of a compact group can be incorporated at the quantum level.

It is natural to assume that the group structure on a quantum space G is described by a *_homomorphism $\phi : A \to A \otimes A$ such that the diagram

$$\begin{array}{ccc} A & \xrightarrow{\phi} & A \otimes A \\ \phi \downarrow & & \downarrow id \otimes \phi \\ A \otimes A & \xrightarrow[\phi \otimes id]{} & A \otimes A \otimes A \end{array}$$

is commutative, and such that

$$A \otimes A = \overline{\left\{ \sum a\phi(b) \text{ s.th } a, b \in A \right\}}$$

$$A \otimes A = \overline{\left\{ \sum \phi(b) a \text{ s.th } a, b \in A \right\}}$$

In such a way we arrive at *compact quantum groups*.

As a very important special class of compact quantum groups, let us mention *matrix* groups. We note that these structures are specified by a C^*-algebra A, together with a *-homomorphism $\phi : A \to A \otimes A$ and a matrix $u \in M_n(A)$ such that the *-algebra A generated by the entries u_{ij} is everywhere dense in A and such that

$$\phi(u_{ij}) = \sum_k u_{ik} \otimes u_{kj}$$

It is also assumed that both u and its conjugate $\bar{u}$ are *invertible* matrices. It follows that

$$\phi(\mathrm{A}) \subseteq \mathrm{A} \otimes_{alg} \mathrm{A}$$

and that the above mentioned co-associativity and density properties are satisfied automatically. Matrix groups generalise the idea of a compact Lie group. The algebra A plays the role of polynomial functions over G. The matrix u correspond to the fundamental representation of the group G. The theory of compact quantum groups was systematically developed in (6, 7)

To see a basic example of a compact matrix quantum group, let us mention a quantum version of the $SU(2)$ group (8). By definition the corresponding C^* - algebra A is generated by elements α and γ, and relations

$$\alpha\alpha^* + \mu^2\gamma\gamma^* = 1 \qquad \alpha^*\alpha + \gamma^*\gamma = 1 \qquad \gamma\gamma^* = \gamma^*\gamma \qquad \alpha\gamma = \mu\gamma\alpha$$

$$\alpha\gamma^* = \mu\gamma^*\alpha \qquad \alpha^*\gamma = \frac{1}{\mu}\gamma\alpha^* \qquad \alpha^*\gamma^* = \frac{1}{\mu}\gamma^*\alpha^*$$

where $\mu \in [-1,1] \setminus \{0\}$. The co-product is specified by the above mentioned matrix rule

$$\phi(u_{ij}) = \sum_k u_{ik} \otimes u_{kj} \ ,$$

where the elements u_{ij} are the entries of a 2×2 A - matrix

$$u = \begin{pmatrix} \alpha & -\mu\gamma^* \\ \gamma & \alpha^* \end{pmatrix}$$

The defining relations for A are actually equivalent to the *unitarity* of the above matrix.

Quantum groups provide a conceptual framework for generalising the classical concept of symmetry. We see that indeed, in classical geometry, symmetries of the space X are interpretable as automorphisms of the associated algebra A. This is without much problem generalisable to the quantum level – we can define symmetries of a quantum space as appropriate automorphisms of the associated non-commutative algebra A. We see that symmetries always form a subgroup of the automorphism group $Aut(\mathrm{A})$. Another way of incorporating the idea of symmetry is to generalise the concept of the *group action* rather than the one of the individual symmetries. In such a way we arrive to the concept of *an action of a quantum group on a quantum space.*

A fundamental class of quantum spaces possessing a built-in quantum group symmetry is given by *quantum principal bundles*. We note that these objects are quantum counterparts of classical principal bundles. Quantum groups play the role of structure groups and general quantum spaces play the role of the base manifolds. The main geometrical idea is the same as in the classical theory – that of a fibered space on which the structure group acts freely on the right, so that the fibers are the orbits of this action. All basic areas of the classical theory such as differential calculus, the formalism of connections and other areas can be naturally generalised and incorporated in the non-commutative context.

2) Deformation Quantisation of Symplectic Manifolds

A very interesting class of examples of non-commutative *-algebras can be obtained by deforming the algebras of smooth functions over a symplectic manifold M, so that a *quantum correspondence principle* holds. We see that this requirement is actually a central problem in the deformation quantisation of symplectic manifolds. More precisely, let A be the commutative algebra of smooth functions on a symplectic manifold M. Let $\mathrm{A}(\nu)$ be the associated algebra of formal power series over a formal parameter ν. We say that a new associative product $*$ introduced in the space $A(\nu)$ satisfies the correspondence principle if and only if

$$f*g = fg+\frac{\nu}{2i}\{f,g\}+\nu^2 r\{f,g\}$$

where { , } are Poisson brackets associated with M. The motivating idea behind this definition is that M plays the role of the phase space of a classical mechanical system. We assume that this classical system has a quantum counterpart, described by a non-commutative algebra, which is in fact $A[\nu]$ equipped with the new product $*$. The parameter ν plays the role of the Planck constant. The correspondence principle tells us that the quantum commutator $\frac{i}{\nu}[*,*]$ coincides with the classical Poisson bracket, modulo terms of the order of ν.

We note that there exists an intrinsically geometrical construction (9) of a non-commutative product $*$ of the described type, for every symplectic manifold M. The following is a very brief sketch of this construction. We start from the Weyl algebra bundle $W[M]$ associated to $[M,\nu,\omega]$, where ω is the initial symplectic form on M. In other words, the fibers of $W[M]$ are the Weyl algebras associated to the tangent spaces $T_x(M)$, $x \in M$, equipped with the symplectic scalar product $\nu\omega_x$. Let W be the algebra of formal power series with coefficients in the smooth sections of $W[M]$. We see that it turns out that every symplectic torsion free connection ∇ on M induces an injective map $j_\nabla : A[\nu] \rightarrow W$ with the following properties,

(i) The image of j_∇ is a subalgebra of W. In fact this image coincides with the kernel of a naturally associated differential D, acting in the algebra $\Omega(M,W[M])$

(ii) A new product $*$ in $A[\nu]$ defined by

$$f*g = j_\nabla^{-1}\left[j_\nabla(f)\, j_\nabla(g)\right]$$

satisfies the above mentioned correspondence principle.

From the physical viewpoint, the final and crucial step in the quantisation of the considered system is to incorporate the construction in the conceptual framework of C^* -algebraic physics (10) We see that this is done by constructing a C^*-algebra $\hat{A}$, by completing the appropriate * subalgebra of $A[\nu]$ and considering irreducible representations and superselection sectors of the completed algebra $\hat{A}$.

(3) Penrose Tilings

The space of equivalence classes of certain tilings of the Euclidean plane, such as the Penrose tilings (11)

This space is defined as follows. Let us consider two triangles T_1 and T_2 of the Euclidean plane, defined by the lengths of edges - $(1,\tau,\tau)$ for the first triangle, and $(\tau,1,1)$ for the second. Here we have

$$\tau = \frac{1+\sqrt{5}}{2}$$

which is the golden ratio number. Both triangles are naturally coming from a regular pentagon.

Let us also assume that edges of both triangles are oriented, in the sense $(+,+,-)$ and $(-,-,+)$ respectively, say, relative to counterclockwise orientation. Let X be the set of all tilings of the Euclidean plane that can be obtained using the above two triangles, and the rules,

(1) We are allowed to perform reflections of triangles,

(2)The oriented edges are paired so that their orientation is the same.

Such tilings exist. The 3-parameter group $E(2)$ of isometric motions of the Euclidean plane is naturally acting on the space X. Let X be the corresponding orbit space. It can be shown that X possesses uncountably infinitely many points. However the points of X are effectively indistinguishable, because of the following remarkable property.

Let $T_1, T_2 \in \mathrm{X}$ be two arbitrary non-equivalent tilings. Then for every finite portion, consisting of finitely many triangles, of T_1 there exists the same modulo $E(2)$ portion of T_2.

Finally, it is important to mention that there exist two different interpretations of the relations between quantum spaces and the associated *-algebras. (1) We have already explained this interpretation

– it assumes that spaces determine algebras and algebras determine spaces, and, (2) the second interpretation, originally proposed by Connes, assumes that geometry is determined by the class of Morita-equivalent C^* -algebras. In other words, non-isomorphic but Morita equivalent C^* - algebras describe the same quantum space. By definition, two C^* - algebras A and B are Morita equivalent if

$$A \otimes K_\infty \cong B \otimes K_\infty,$$

where K_∞ is the ideal of compact operators of a separable Hilbert space. Morita-equivalent algebras have the same cyclic cohomology and K - groups.

The second approach is more suitable for constructions involving the factor spaces. However, it is not appropriate in cases involving quantum groups and quantum bundles, where passing to a Morita -equivalent algebras destroys the entire geometrical structure.

Von Neumann algebras

The space L^2 of all absolutely square integrable functions is a separable Hilbert space, H. An involutive symmetric subalgebra M of $L(\mathrm{H})$ containing the unit and closed by the weak topology is a von Neumann algebra (12). The distinction between von Neumann algebras and general *-algebras is essential: von Neumann algebras are closed under the weak topology while *-algebras are closed under the norm topology. Of course, von Neumann algebras are particular cases of *-algebras, but they are not, in general, separable by the norm topology.

If we rephrase all this, while introducing some more ideas. (13) A von Neumann algebra M is a nondegenerate self-adjoint algebra of operators on a Hilbert space H, which is closed under the weak operator topology. This is the locally convex topology in $L(\mathrm{H})$ induced by the family of semi-norms

$$x \in L(\mathrm{H}) \to \left|\langle x\zeta | \xi \rangle\right| \text{ for all } \xi, \zeta \in \mathrm{H}$$

The von Neumann *bicommutant theorem* says that $M = M'' : M$ is the commutant of its own commutant.

We notice that an important fact is that the set of all projectors on a von Neumann algebra M, with the identity included, generates M. We recall that a subset U of an algebra M is said to generate M if the set of all the polynomials obtained with all the members of U is dense in M. As the projectors are idempotents, polynomials in projectors are simply linear combinations.

The classification of von Neumann algebras is based on the properties of its projectors. The center of a von Neumann algebra is abelian. A *factor* is a von Neumann algebra with trivial center, $M \cap M' = C$. This means that the center is formed by the complex multiples of the identity.

One of the greatest qualities of von Neumann algebras is their receptivity to integration. In effect, it is possible to define on them a measure theory generalizing Lebesgue's, and that despite their noncommutativity. A factor, as said above, is a von Neumann algebra whose center is C. The space of factors contained in a von Neumann algebra M is itself Borel measurable. Let us call the measure μ. Each factor can be labelled by an index t belonging to a borelian set, and denoted by $M(t)$. Then the whole algebra M is given by the decomposition

$$M = \int M(t)\, d\mu(t)$$

This is a theorem by von Neumann. We see that this property justifies the name *factor* and reduces the problem of finding all the von Neumann algebras to that of classifying all the possible factors. The last grand steps in this measure algebraic program were given recently, mainly by A. Connes. They naturally opened the gates to noncommutative analysis and to noncommutative geometry. Another recent, astonishing finding by V Jones, is that von Neumann algebras are intimately related to knot invariants.

Projectors are ordered as follows. Let p and q be two projectors in the von Neumann algebra M. We say that p and q are equivalent, and write $p \approx q$, if there exists $u \in M$ such that $p = uu^*$ and $u^*u = q$. We say that $p \leq q$ if there exists $u \in M$ such that $p = uu^*$ and $u^*uq = u^*u$. This means that u^*u projects into a subspace of $q\mathrm{H}$, that is, that $u^*u \leq q$. We say further that $p \perp q$ if $pq = 0$.

Now, if M is a factor, then it is true that, given two projections p and q, the projector q is *finite* if the two conditions $p \leq q$ and $p \approx q$ together imply $p=q$. The projector q is *infinite* if the conditions $p \leq q, p \approx q$ and $p \neq q$ can hold simultaneously. The projector p is *minimal* if $p \neq 0$ and $q \leq p$ implies $q=0$. The factors are then classified in types, denoted I, II_1, II_∞, and III:

I : if there exists a minimal projector;
II_1: if there exists no minimal projector and all projectors are finite;
II_∞:if there exists no minimal projector and there are finite and infinite projectors;
III: if the only finite projection is 0.

On a factor there exists a dimension function $d\{projectionson\ M\} \rightarrow [0,\infty]$ with the suitable properties:

(i) $d(0)=0$;
(ii) $d\left(\sum_k p_k\right)=\sum_k d(p_k)$ if $p_i \perp p_j$ for $i \neq j$;
(iii) $d(p)=d(q)$ if $p \approx q$.

We have provided a short introduction to von Neumann algebras since they are important in our understanding of quantum spaces.

We see that it is difficult to interpret von Neumann's concept of quantum logic solely on the material of the seminal 1936 paper (14). During the period 1935-1936 whilst von Neumann was presenting the theory of quantum logic he was working at the same time on the theory of *rings of operators* better known in modern terminology as *von Neumann algebras.* In the year of the publication of the paper on quantum logic he also published a co-authored paper with J Murray, which established the classification theory of von Neumann algebras (3). The results of this classification theory of von Neumann algebras is very closely related to von Neumann's concept of quantum logic. For us to understand the special features of von Neumann's concept of quantum logic, we need to look into the earlier research of von Neumann. We see that in the second (16) of the famous three foundational papers (15,16,17) von Neumann showed how the quantum mechanical probability calculus can be derived from the frequency interpretation of probability. We note that this derivation had inconsistencies that von Neumann knew about. This resulted in him adopting a questioning position in regards to the Hilbert space formalism of quantum theory (18) and a search for a mathematical formalism of quantum logic that would serve as a mathematical structure of quantum theory rather than the use of the Hilbert space formalism.

We see that the paper (3) shows that there exists a non-distributive, modular lattice which is different from the Hilbert lattice of a finite dimensional Hilbert space. We note that the proof of the existence of such a structure is dealt in the dimension theory of von Neumann algebras. The key result in the research of dimension theory is the important result that there exist modular lattices of non-finite, linear dimensional projections on an infinite dimensional Hilbert space. We note also that on this lattice there exists a, unique up to normalization, dimension function d which can take every value in the interval $[0,1]$

We see that the von Neumann algebra generated by these projections goes under the label of type II_1 factor von Neumann algebra. (19, 20, 21, 22)

We see that von Neumann saw the modular lattice of a *type* II_1 *factor von Neumann algebra* as the proper quantum logic. This program generalises beyond the Hilbert space quantum theory. It uses a different mathematical structure based on the theory of von Neumann algebras, specifically on the theory of type II_1 algebras. The objections of von Neumann to the Hilbert space formalism at that time did not stem from physical reasons but from mathematical reasons. The main reasons for the objections can be summarized as following:

(i) The probabilities must be interpreted as relative frequencies as done in Mises' theory (23)

(ii) Quantum probability statements must be interpretable as conditional probability statements with the prior probability given by a trace.

We see that von Neumann saw these positions on quantum probabilities as flawed, that is if we are going to take the standard Hilbert lattice of an infinite dimensional Hilbert space as representing the random event structure. We can now see the reason which motivated him in suggesting the theory of type II_1 von Neumann algebras as a solution to a non-trivial, non-finite linear dimensional structure for quantum theory.

Today we have also objections to the Hilbert space formalism from physical grounds also. We shall not go into this here. Importantly, von Neumann was not totally happy with the theory of type II_1 factors and their lattices as quantum logic. (24, 25)

Conclusion

John von Neumann recognised that the future merging of the theory of logic, noncommutative probability theory, von Neumann algebras and quantum theory was an open question to be resolved. We see this today taking place in the related areas of noncommutative geometry, quantum spaces, knots, quantum logic and von Neumann algebras. We have gone beyond the initial formulation of quantum theory as a Hilbert space.

Acknowledgements

I would like to thank members of the Theoretical Physics Research Unit, Birkbeck College, University of London for the discussions at weekly seminars that have stimulated me in writing this paper.

Finally, I would like to thank Lilian Reascos Glendza and Dr. Keith Bowden for their patience.

References

(1) Wheeler J A, in *Quantum Theory and Gravitation*, edited by Marlow A C, Academic Press, (1980)

(2) Einstein A, *The Meaning of Relativity*, Chapman and Hall, (1962)

(3) Murray F J, von Neumann J, *On Rings of Operators*, Annals of Mathematics 37 116-229 (1936)

(4) Murray F J, von Neumann J, *On Rings of Operators 2*, Trans Amer Math Soc 41 208-248 (1937)

(5) Murray F J, von Neumann J, *On Rings of Operators 4*, Ann Math 44 716-808 (1943)

(6) Woronowicz S L, *Compact Matrix Pseudogroups*, Commun Math Phys III, 613 – 665 (1987)

(7) Woronowicz S L, *Tannaka – Krein Duality for Compact Matrix Pseudogroups ; Twisted SU(n) groups*, Invent Math 93, 35 – 76 (1988)

(8) Woronowicz S L, *Twisted SU(2) group : An example of a non-commutative differential calculus,* RIMS, Kyoto University 23, 117-181 (1987)

(9) Fedosov B, *Deformation Quantisation and Index Theory,* Akademie Verlag GmbH, Berlin (1996)

(10) Brateli O, Robinson D, *Operator Algebras and Quantum Statistical Mechanics,* Volumes 1, 2, Springer-Verlag (1979)

(11) Connes A, *Noncommutative Geometry,* Academic Press (1994)

(12) Dixmier J, *von Neumann Algebras,* North-Holland, Amsterdam (1981)

(13) Takesaki M, in Dell'Antonio G et al (eds.), Springer Lecture Notes in Physics 80, Springer, Berlin

(14) Birkhoff G, von Neumann J, *The logic of quantum mechanics,* Annals of Mathematics 37 (1936)

(15) Von Neumann J, *Mathematische Begründung der Quantenmechanik,* Göttinger Nachrichten, 1-57, (1927)

(16)Von Neumann J, *Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik,* Göttinger Nachrichten, 245-272, (1927)

(17) Von Neumann J, *Thermodynamik quantenmechanischer Gesamtheiten,* Göttinger Nachrichten, (1927)

(18) Redei M, *Why John von Neumann did not like the Hilbert space formalism of quantum mechanics (and what he liked instead),* Studies in the History and Philosophy of Modern Physics, 493-510, 27, (1996)

(19) Kalmbach G, *Orthomodular Lattices,* Academic Press, London, (1983)

(20) Takesaki M, *Theory of Operator Algebras, I.* Springer Verlag, New York, (1979)

(21) Bub J, *What does quantum logic explain?,* in Current Issues in Quantum Logic, eds. Beltrametti E, van Fraassen B C, Plenum Press, New York, (1981)

(22) Petz D and Redei M, *John von Neumann and the theory of operator algebras,* in 'The Neumann Compendium World Scientific Series of 20th Century Mathematics Vol I, EDS. Brody F., Vamos T, World Scientific, Singapore, pp 163-181 (1995)

(23) von Mises R, *Probability, Statistics and Truth,* Dover, New York, (1981)

(24) von Neumann J, *Letter to Dr. Silsbee, July 2, 1945,* John von Neumann Archive, Library of Congress, Washington, DC

(25) von Neumann J, *Unsolved problems in mathematics,* Unpublished address to the World Congress of Mathematics. Typescript John von Neumann Archive, Library of Congress, Washington, DC.

ANPA27

The next meeting of ANPA will be held in the New Common Room, Wesley House, Jesus Lane, Cambridge CB5 8BJ from Friday 5th to Tuesday 9th August 2005 (*five* days only). There will be a banquet on Monday 8th August at King's College. The Executive Council Meeting will be held Thursday 4th August in the Old Common Room, Wesley House.

The House Manager there is Iain Murton. His email is idm23@cam.ac.uk and the telephone number is 01223 741033, fax 01223 321177. Only contact Iain directly in an emergency, or if there is a need that *cannot* be satisfied by either Arleta Griffor or Keith Bowden. The Bursar is Jim Sturmey and the Principal is Philip Luscombe.

Could all members please adhere to the following rules

1. NO smoking anywhere indoors under pain of death.

2. The shower and kitchen facilities near the rooms in Wesley House are communal and you may use them. Please make sure that you leave the kitchen and all other facilities clean and tidy and that you do all your washing up promptly. Do NOT take other people's shampoo, conditioner, milk, bread etc without their permission!

3. Please read the Fire Regulations in the corridors and follow them accordingly.

4. Be aware of security issues. Please keep all downstairs windows and doors locked at night. Please keep the main doors to each corridor locked all the time.

5. *Nobody* is to be brought onto the premises without agreement with Keith Bowden specifically. The same goes for any expensive personal or professional equipment, with the exception of laptops. There are important insurance and security issues here.

6. The gates to Wesley House are locked between 10.30 and 11.30 pm and are not reopened until morning. You are provided with a key to the side entrance ten yards to the right of the main gates when facing Wesley House. You may use these out of hours but make sure that you return QUIETLY. ANPA has been criticised in the past for making a row in the

evenings. There will be other people staying at Wesley House. Please respect their needs.

7. Breakfast is provided as usual ONLY for those who have paid for it.

8. Please pay any money owing to either Tony Deakin or Keith Bowden. Spare copies of the Proceedings are £7 each. I have some spare copies of the Special ANPA Issue of IJGS. These are £20 each payable to me (Keith).

9. There may be someone recording talks in the meetings. If you do not wish to be recorded please make this clear before your talk.

10. Copies of the timetable and maps of Wesley House and Cambridge will be made available. Coffee and tea will be provided in the breaks. For those that want to lunch together the Bun Shop on the corner of Malcolm St and King St (five minutes from Wesley) is a popular venue.

11. You may use the chapel with the obvious provisos.

12. You may use the facilities in the New Common Room at any time that there are no talks. There is table tennis, pool, TV, video and tea and coffee making equipment. Please give yourself adequate time to put everything way before the next talk. Please try to keep the Common Room clean and tidy at all times and do any washing up promptly.

13. Any extra requirements that Keith or Arleta cannot satisfy, such as irons, mirrors etc. may be addressed to Iain Murton, but please remember that he has other responsibilities as well. In emergencies contact any one of us.

14. At the end of the meeting please RETURN ALL KEYS either to me, Arleta or Iain Murton.

Have a great meeting.

Keith

Alternative Natural Philosophy Association [26]

Statement of Purpose

1. The primary purpose of the Association is to consider coherent models based on a minimal number of assumptions, so as to bring together major areas of thought and experience within a Natural Philosophy alternative to the prevailing scientific attitude. The Combinatorial Hierarchy, as such a model, will form an initial focus of our discussions.

2. This purpose will be pursued by research, publications and any other appropriate means including the foundation of subsidiary organisations and the support of individuals and groups with the same objective.

3. The Association will remain open to new ideas and modes of action, however suggested, which might serve the primary purpose.

4. The Association will seek ways to use its knowledge and facilities for the benefit of humanity and will try to prevent such knowledge and facilities being used to the detriment of humanity.

Organisation (altered to reflect the current situation by KGB)

1. The Founder of the Association was Pierre Noyes. The Founder Members were Pierre, John Amson, Ted Bastin, Clive Kilmister and Frederick Parker-Rhodes. They will be known herein as the Founders. The Executive Council is the governing body of the Association. It consists of:

   (a) The Founders and all past Presidents of the Association, the Co-ordinator and the Treasurer,

   (b) The Executive Officers (the President and the Chairmen of the Executive Council and the Advisory Board),

   (c) Ordinary members nominated by classes (a) and (b), who serve for three years, with the possibility of re-nomination.

2. The Members of the Association are (a) the members of the Executive Council and (b) others nominated by the members and approved by the Executive Council.

3. The President is the official representative of the Association in external affairs, and has the responsibility for calling meetings of the Executive Council, at least annually, for the determination of overall policy.

4. The Treasurer is the responsible financial officer of the Association for the receipt and disbursement of funds and shall maintain and make available appropriate records, including annual accounts.

5. The President and the Co-ordinator may be paid an appropriate salary for their services, funds permitting. These services will include the organisation of meetings and the editing of the Proceedings of such meetings for publication, co-ordination of, and participation in, the research activities of the Association, preparation when appropriate of research reports and publication of such reports, and other such duties as may be assigned.

6. Members of the Executive Council may as appropriate receive funds for travel, expenses, etc.

7. The Executive Council has selected an independent Advisory Board. It may adopt its own rules for the operation and replacement of members. The Executive Council may nominate candidates to the Board. Any member of the Board, or the Board collectively, may make recommendations to the Executive Council, or directly to the Membership. Action taken on such recommendations must be promptly reported by the Executive Council to the Board in writing.

8. Dues are currently £35.00 per annum.

Executive Council: Mr. Anthony M. Deakin (Chairman), Dr. John Amson, Dr. Ted Bastin, Dr. Tom Etter, Prof. Louis Kauffman, Dr. Michael Manthey, Dr. David Roscoe, Dr. Fredric S. Young.

President: Dr. Keith Bowden, 139 Sandringham Road, Barking, Essex, IG11 9AH, UK.

[Tel: 0208 594 5064 Email: k.bowden@physics.bbk.ac.uk].

Co-ordinator: Arleta Dylus, 50A Grove Rd, North Finchley, London N12 9DY.

[Tel: 0208 369 5865 Email: a.griffor@physics.bbk.ac.uk]

Treasurer: David Roscoe, Department of Applied Mathematics, Sheffield University, Sheffield S3 7RH.

Advisory Board: Mike Horner (Chairman), Profs. G.F. Chew (Berkeley), C. Isham (Imperial College), M. Redhead (Cambridge and LSE), N. Cartwright (LSE), C. W. Kilmister (London), H. Pierre Noyes (Stanford).

ANPA Electronic Mailing Lists

It is **YOUR** responsibility to subscribe to these groups. If you do not do so you may miss important announcements about the annual conference or Proceedings contributions.

anpa-list

This group is for general announcements and is moderated (in varying degrees) by Keith Bowden

To subscribe to this group send **ANY** email to:
anpa-list-subscribe@yahoogroups.com

To post messages to this group email:
anpa-list@yahoogroups.com

To read messages on the web, go to:
http://groups.yahoo.com/group/anpa-list/

To unsubscribe from this group send **ANY** email to:
anpa-list-unsubscribe@yahoogroups.com

anpa-discussions

This is a group for general discussion and is largely unmoderated

To subscribe to this group, send **ANY** email to:
anpa-discussions-subscribe@yahoogroups.com

To post messages to this group email:
anpa-list@yahoogroups.com

To read messages on the web, go to:
http://groups.yahoo.com/group/anpa-discussions/
(this last not yet implemented I believe)

To unsubscribe from this group, send **ANY** email to:
anpa-discussions-unsubscribe@yahoogroups.com

Keith Bowden

ANPA26 Proceedings
Instructions to Printers

Please print 60 copies B&W reduced to A5 double sided and glue bound with wrap around glossy colour cover on white card. Print two up/two down, 30 off, reduce and cut to 60 off.

Enclosed front colour cover, back colour cover (it's a page from the proceedings) and spine. Leave inside of cover blank.

Number pages top centre starting from the first paper. Contents pages etc not to be numbered. Numbering must agree with Contents pages. Page numbers are written in pencil on the back of each page. Numbering must agree with these. Contents pages are numbered on the back of each page in pencil (i), (ii), ... for convenience only.

Keith Bowden,
Associate Research Fellow,
Theoretical Physics Research Unit,
Birkbeck College,
Malet St

Phone: 0208 594 5064
Email: k.bowden@physics.bbk.ac.uk

Invoice to:

ANPA c/o Keith Bowden,
139 Sandringham Rd,
Barking,
Essex IG11 9AH

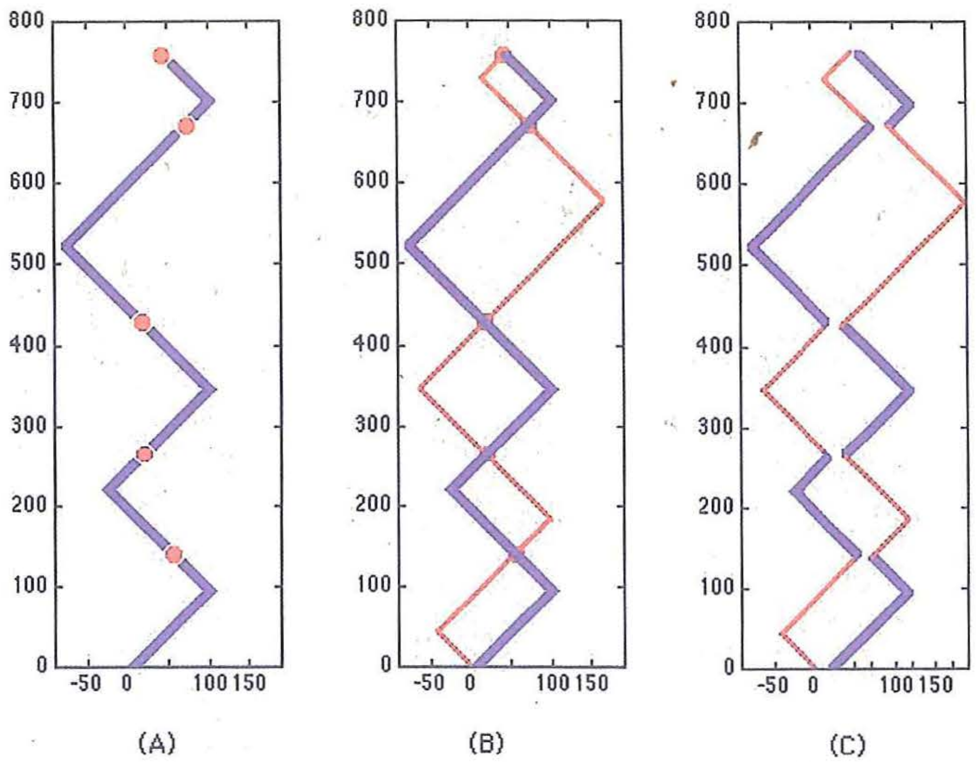


FIG. 1: Forming entwined paths in space-time. z is horizontal, t is vertical.
A) A particle starting at the origin travels at constant speed but occasionally reverses direction in response to a stochastic process. At every other indication from the stochastic process, a marker is dropped instead of a direction changed.
B) Past a specified time t_R, the particle reverses direction in t at the next marker. The particle then follows the 'light-cone' paths through the markers back to the origin (thin line in (B)).
C) The entwined path formed in (B) can be regarded as two osculating paths which we call envelopes. These are separated for clarity. We use the envelopes to count the oriented rectangles formed in (B).

along a light-cone path fixed by the markers (thin line in Fig. 1 B)). This process is repeated building up a background of oriented rectangles on spacetime. Here the 'orientation' comes from distinguishing the direction of traversal in the t direction. If the first rectangle in Fig. 1 B) has positive orientation, the second has a negative orientation and the orientation changes whenever the thick and thin lines in the path cross. Note that if we count oriented rectangles, +1 for positive and -1 for negative, we need only use an 'envelope' (Fig. 1 C)) for counting. Counting signed contributions here is appropriate since if 't' represents macroscopic time, the thick and thin lines represent 'particle' and 'antiparticle' trajectories respectively. As the particle trajectory repeats creation of these entwined pairs, the particle number will increase, however the 'net charge' will always remain zero since entwined paths always pair thick and thin contributions.