Mereologies

Proceedings of ANPA 18

Thomas L. Etter, *Editor*

Proceedings of the 18th Annual Meeting of the

Alternative Natural Philosophy Association

Wesley House, University of Cambridge
September, 1996

Published by **ANPA**
c/o Prof. C. Kilmister
Red Tiles Cottage, High St. Barcombe,
Lewes BN8 5DH, UK

July, 1997

These Proceedings bring together the collection of papers presented at the 18th annual meeting of the Alternative Natural Philosophy Association held in Cambridge in September, 1996.

Further copies of the proceedings may be obtained from
Prof. C. Kilmister, Red Tiles Cottage, High St. Barcombe,
Lewes BN8 5DH, UK. (Telephone 0273 400 922)

Mereologies

Proceedings of ANPA 18

Thomas L. Etter, *Editor*

The Alternative Natural Philosophy Association

Mereologies: Proceedings of ANPA 18/Thomas L. Etter, *Editor*

ISBN 0 9526215 18

published by ANPA, c/o Prof. C. Kilmister,
Red Tiles Cottage, High St. Barcombe,
Lewes BN8 5DH, UK.

Contents

C W Kilmister
The Hierarchy Reviewed 7

H Pierre Noyes
A Short Introduction to Bit String Physics 21

Viv Pope
The Gateway to Paradox 62

Keith G Bowden
Huygens' Principle Physics and Computers 67

F Selleri
The Rotating Platform Problem 86

V A Karmanov
On the Derivation of the Electron Propagator from a Random Walk 102

Emmanuel Ransford
From Naught to Aught: A Conceptual Inquiry 112

Viv Pope
The Cinematic Model of Quantum Touching 134

Michael Manthey
Distributed Computation, the Twisted Isomorphism and Auto-Poiesis 139

Michael Manthey
A Pure Process View of Anticipatory Systems 165

Peter J Marcer
From the Self Creation to the Creation of the Self 172

Ted Bastin
The Hierarchy and Physical Space 190

Thomas L Etter
A Very Brief Note on Link Theory 196

Geoffrey Constable
An Investigation, Arising from the Theory of Quantised Variables, into Inter-Particle Forces 197

D F Roscoe
An Analysis of 900 Rotation Curves of Southern Sky Spiral Galaxies: Are the Dynamics Constrained to Discrete States? 224

Adam Parker-Rhodes
A Representationalist Theory of Consciousness 237

Juan M Alvarez de Lorenzana
Some Issue in the Construction of Order in Self Organising Systems 265

Presitorial

As Tom is somewhat overloaded this Summer it has fallen to me to produce some words to precede the collection of papers contributed to this Proceedings. Tom is continuing to develop his work on Link Theory. He presented this to PHYSCOMP96 in Boston (continuing ANPA's presence there) and has been invited to submit a formal paper for the Physica D Special Issue edited by Tom Toffoli and arising from the PHYSCOMP meeting. This Summer he is working to tight deadlines on a project with Interval Research. My paper on Counterfactual Computation was too late for Toffoli's journal, but I hope to present a version of it to ANPA this coming year. The International Journal of General Systems Special Issue which I have been editing is now finally complete and should appear as a double issue, Vol 27, Nos 1 and 2, in Spring 1998. This will contain papers by, amongst others, Tom, myself, Ted Bastin and Clive Kilmister, Louis Kauffman and the remarkable and now famous "attic papers" of John Amson and Frederick Parker-Rhodes which have been lovingly reedited into a publishable form by John Amson. Another forthcoming project that will have a high ANPA presence will be Rainer Zimmerman's Natura Naturans conference at ZIF in Beilefeld immediately before ANPA 97. (I find it difficult to remember which was ANPA15 and which ANPA16. A labelling scheme which involved the year would improve this - what do people think? - but would fail after ANPA99, we could not really return to ANPA01, and ANPA00 sounds positively surreal). I guess ANPA2000, ANPA2001... sounds quite nice but would become cumbersome eventually. (Or am I just becoming sucked into ANPA's obsession with labelling schemes...?:-)

For the Proceedings Pierre Noyes has contributed an beautiful 40 page survey of his work over the last few years. This will be followed by his "From Bit Strings to Cosmology" at ZIF. I personally have enormous respect for the direction in which Pierre has steered his work recently; I take my hat off to him. Also resulting from the Kilmister-Noyes Combinatorial Hierarchy Survey material last year is the introductory paper "The Hierarchy Reviewed" by Clive Kilmister. Mike Manthey talked on "Computational Aspects of the Combinatorial Hierarchy". He unfortunately cannot attend this year (Daniel Dubois has seduced him away), so I persuaded him to contribute a couple of related papers. Tom Etter has contributed some notes on Link Theory. I have included my paper on things Mereological from the IJGS Special Issue.

This Proceedings is produced by the Print Unit of the **University of East London**. For their support and help I would particularly like to thank Al Freimanis and Anna Bass. 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. (Just stick to the style I have used in this issue.) At least, Times Roman, 12 point, double spaced, ***single sided, two copies,*** is preferred, although I will still accept typescripts in Courier. Main heading capitalised, centred, other headings capitalised to the left. 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 or address, email address and the **version number** of the draft. I often get sent more than one version of a paper and invariably mix them up! **Send copy to *Keith Bowden, 47 Park Avenue, Barking, Essex IG11 8QU*** (NOT to Tom). Please help me by conforming to all this as closely as you can.

Finally, as many people know I am joining Arleta Griffor by leaving UEL to become a doctoral student once more, this time for Basil Hiley at Birkbeck College. Working with the group there is most stimulating, despite the bad news recently that Birkbeck are to close the Physics Department. An interim name, the Theoretical Physics Research Unit has been given to the group that remains there under Basil Hiley; this status is now assured until 2001. I am planning to start a Quantum Information group within this Unit. Also over recent months Peter Marcer, along with Brian Oakley (former Chairman of the Alvey Commission) and various members of the British Computer Society (including myself) have formed a working party of the BCS with a view to set up a European Institute for Quantum Computation. This initiative is going well and after a recent visit to Brussels Brian is hoping to get funding for a first stage conference (a "Town Meeting") in London in the Autumn. Let me wish all who will be attending, an excellent ANPA19 meeting for 1997!

Keith Bowden,
Theoretical Physics Research Unit, Birkbeck College, London &
University of East London, August 1997

THE HIERARCHY REVIEWED

C W Kilmister

Red Tiles Cottage
High Street
Barcombe
Lewes
BN8 5DH

Four sessions arose at the conference because Pierre Noyes and I thought that it was time to try for definitive presentations of the positions we have now reached, to generate discussion of the relations between them. I have re-written mine as one paper, called:

AN ALGEBRAIC CONSTRUCTION AND ITS MEANING.

It is easiest for me to begin with fairly personal reminiscence. In the late forties both Ted Bastin and I were interested in Eddington's later speculations - this is an arena in which this may be declared openly without the lions being summoned. Now Eddington famously thought that in the relating of quantum mechanics and general relativity lurked numerical values for some natural constants. One analogy for this was the way in which Maxwell's unification of electricity and magnetism led to the constant $c^2\mu\kappa$ having the value 1. I say relating because Eddington held varying views of the exact relation; sometimes QM was to be seen as part of an extended GR, sometimes they were to be two different descriptions of a single problem and the numbers would arise by comparing solutions. In any case, Eddington had rediscovered what I call Planck's problem: that there are dimensionless physical constants which seem to have some great significance. What is this significance? Another person impressed by Planck's problem was P A M Dirac, who emphasised the importance for quantum electrodynamics of calculating the value of $e^2/\hbar \approx 1/137$, since the smallness of this fine structure constant was essential for the perturbation series to be asymptotic.

Everyone would now agree that Eddington's approach to Planck's problem was not coherent; how did this happen? His method was to derive numbers as the number of elements in certain algebras (in fact, in the Clifford algebras which had been brought to his attention by Dirac's electron equation) without attempting the basic reconstruction of physical theory which such a method needed. This lacuna showed up in a technical problem: even if you take $\hbar c/e^2$ as 137, a rough approximation, and you note from successive Clifford algebras that $4 = 1 + 3$, $1^2 + 3^2 = 10$, $2.1.3 = 6$, $6 + 10 = 16$, $10^2 + 6^2 = 136$, $136 + 1 = 137$, how can you hope to show that it is $\hbar c/e^2$? Eddington tried to do this by modifying the conventional equations by introducing the Clifford algebras and then chasing through the way in which the successive equations were changed to "spot the coefficient of the extra term". This is not obviously illogical but as a method it failed.

Now it is much easier to say that Eddington failed to attempt the necessary basic reconstruction than to carry it out. During the fifties Ted and I tried to formulate a new approach to scientific theories in our Concept of Order papers:

> 2. *The concept of order.* One would not wish to distinguish thoughts sharply from the spoken or written expression of them. For our purpose we wish to extend this idea and say that a physical theory consists of propositions which may be thoughts, sentences written or spoken, or manipulations with bits of the physical world. So we must not suppose that the calculus of a physical theory can exist, as it were, all in a lump in our minds, and distinct from bits of the physical world. Confusion like this is resolved if we substitute for the thought, in the theory, the experimental thing for which it stands. The difficulty, and the importance, of this substitution is that the experimental thing has meaning only as part of a theory. The theory may have different degrees of complexity, and there will be experimental procedures corresponding to each such degree. Thus the theory is a kind of language, but experiments in the theory are the same language. One cannot use experiments in a complex language to criticize a simpler theoretical language. Of course there are exploratory experiments (like the Michelson-Morley experiment) which we perform when we do not yet know the use of the language. The non-metrical theory we are to develop requires this idea of a series of theory-languages of increasing complexity, rather than that of the formal mathematical calculus and its physical interpretation.
>
> If one says one understands a theory, an important thing one means is that one can make the connexions between elements of the theory (thoughts, or physical manipulations) in a known and proper order. It is as though we can run through the elements in the correct order. We wish to show that physics is the study of certain specially simple kinds of such order, and therefore we call our method 'the concept of order'. Our awareness of these kinds of order is the empirical content—the appeal to sense—of the theory. The empirical appeal is in a different place from what is usual, but it is essential to see that for each of the 'theory-languages' it exists, and the 'order'

principle draws attention to it. It follows that the theory-languages do not differ in degree of physical reality, but only in complexity.

The method of this paper, and the view of the empirical content of a theory that has just been mentioned, cannot be maintained if we think of experimental facts about each of which a number of possible propositions may hold. We shall not think of facts in this way: when we speak of the understanding of a theory as consisting, in a way, in the ordering of its elements we do not at all mean that this ordering is an ordering of facts which could be done self-consciously in different ways.

Physics is interested in states, that is, sets of simultaneous events; simultaneous events are those which are not temporally ordered. Therefore a state—with this meaning—will be characterized by a certain ordering property; so that the first of the ordering properties we find, and the most important for physics, tells us whether or not our propositions serve to provide a unique time sequence. We define a state thus:

S is a state if

(i) *it is a sub-groupoid of V;*
(ii) *no relation in S serves to order its elements.*

Here (i) specifies the closed nature of a state, and (ii) its simultaneity.

It will be seen later that it is sufficient to fulfil this condition for triadic and tetradic relations. We may think of simultaneity in this way: we are given a set of n counters of various colours, and a certain arrangement of n boxes which is such that one counter may be put in each box to make, in all, a true proposition. Simultaneity is the requirement that any other arrangement of the same n counters in the boxes also gives a true proposition.

THEOREM 1. *A state is an Abelian group in which every element is of order* 2.

As you can see, this approach began with the undefined notion of "parts" or "elements" of a scientific theory and the two further ideas (i) that theorising is to put these parts in a certain order, (ii) that two parts can be combined to give a third. Then we looked for "states", that is, sets which are not in fact ordered by the only operation which can do the ordering i.e. the combining. We argued that this was captured by requiring that every equation which was true continued to hold when the symbols in it were permuted in any way. Such a requirement gives rise to abelian groups, for the assumptions are strong enough to prove the associativity,in which every element, except the identity, is of order two. By a well-known theorem, this means that any such group is a direct product of the cyclic groups of order 2. The simplest, after C_2 itself, is $C_2 \times C_2 = S$, Klein's Vierergruppe; this led us to toy with identifying the 3 + 1 nature of this with space and time but I now think this was mistaken.

In our second paper we moved on to emphasise the continual development of scientific theories. Here we saw the importance of a potentially infinite sequence of procedures as constituting a measurement and so the assigning of a number - the result - to the sequence.The promised treatment of the bounds to such measurements was then to come later. Looking back this seems to me misguided; the bounds need treating first and the other numbers later. But the paper has the virtue that the emphasis on continual development showed the need for a different mathematics from the conventional. This led to our flirtation with Brouwer's Intuitionism, using his Spread Theorem to characterise measurement. We would not now buy Brouwer lock, stock and barrel but the need to loosen the straitjacket of classical mathematics is important still. These papers raise as many questions as they answer: what are "parts" and what is combining? What sort of ordering is this and why, having defined it, seek its absence? And then, for the mathematician with one eye on the destination, why abelian groups instead of the hoped for Clifford ones?

These two 1954 papers and the later ones in the same series, were the context into which Frederick Parker-Rhodes' algebraic construction (the one in the title of this paper) burst. In the late fifties Ted was actively pursuing the possibility of modelling the notion of a theory with bounds on numerical quantities by means of a computer (as it was understood then). Does this seem an odd thing to try? It is the first appearance in my story of the realisation that process is important. Computer models in general, and therefore those for use in physics, must automatically have a process character, which is a great help if we are not yet sure exactly what constitutes such a character! Inevitably, if we use such a model, we are saying that, in essentials, the activity of the computer represents directly an aspect of the world. But not all models have physical significance, which is why the project was very tentative. And I should interject here that by now the project had widened out considerably. It was no longer Eddington's errors that were now the driving force but the idea that quantum mechanics needs understanding. Not many people at that time, except perhaps Richard Feynman, would stand up and say - as for example Roger Penrose did some ten years ago - "Quantum mechanics is a wonderful theory: the only trouble is that no-one understands it." Now it is a more widely held view. One consequence is that, although one can pray quantum mechanics in one's aid in making experimental checks on predictions, it is not available as an aid to understanding one's own theory. Returning to Ted's machine, then, Parker-Rhodes put forward a (highly modified) "realisation" of what was being attempted:

The Frederick Construction

Consider vector spaces, V_n, of dimension n over the field $\mathbb{Z}_2$ which has the two elements 0, 1 and in which therefore x + x = 0 whether x = 0 or 1. A basis for V_2 is $\begin{pmatrix}1\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\end{pmatrix}$ and the remaining vector in V_2 is the sum of these. Any two of them could be chosen as a basis. The sum, +, is called discrimination because it discriminates in the sense that a + b = 0 if and only if a = b. Define a discriminately closed subset (dcs) by the rule:

S is a dcs if and only if two DIFFERENT elements a,b in S have a + b in S.

Thus (but unhelpfully) dcss "are groups with the identity left out". For V_2 the chosen basis generates 3 dcss: $\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}$, $\left\{\begin{pmatrix}0\\1\end{pmatrix}\right\}$, $\left\{\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}\right\}$ and these can be characterised by the only non-singular 2 x 2 matrices having them as their sets of eigenvectors (with necessarily eigenvalues 1 since the non-singular requirement forbids the only other value):

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

Re-express these in some fixed way as vectors in V_4: $\begin{bmatrix}1\\0\\1\\1\end{bmatrix}$, $\begin{bmatrix}1\\1\\0\\1\end{bmatrix}$, $\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$. These three generate $2^3 - 1 = 7$ dcss in V_4.

* Characterise these in turn by choosing 7 non-singular 4 x 4 matrices which (new additional requirement) are linearly independent and have the sets as their sets of eignvectors.

Repeat * but writing 7 for 3, 16 for 4 and so $2^7 - 1 = 127$ for 7.

Repeat again, but at the next stage there are 127 vectors in V_{256} which generate $2^{127} - 1 = 10^{38}$ (approx) dcss. The non-singular 256 x 256 matrices to characterise these lie in a V_{65536} so evidently cannot be chosen linearly independent and the construction terminates. The cumulative totals of matrices chosen are 3, 10, 137, 10^{38} and what has been constructed is the Combinatorial Hierarchy. Frederick also showed shortly afterwards that there were essentially no other hierarchies (i.e. you have to use $\mathbb{Z}_2$ and you have to start with two elements, otherwise you can't get beyond one level change.)

You see how this construction triumphantly gave the sequence of numbers 3, 10, 137, 10^{38} and a stop to further numbers in that sequence. Here then were the bounds but again many questions were raised: why $\mathbb{Z}_2$? Why vectors? Why matrices, eigenvectors? (As well as some non-problems which seemed serious at the time such as the necessity to re-express the square matrices as vectors). More surprisingly, what was not noticed at the time was the isomorphism with the algebra of the Concept of Order papers. It is just that, what was written ab there is here written a + b (and so for I is written 0). So likewise, the overtones of this choice were not discussed: (i) if one binary operation is given, is there perhaps another and so a ring structure instead of a group? Then the x will distribute over the + and not vice versa, whichever was the original one. (ii) What of the automorphisms of the system, that is, the 1:1 transformations preserving the binary operation? In the + case these are provided by the very non-singular matrices in the Frederick construction. For the x case it is not so easy to specify them; though if the algebra had been a Clifford algebra, which it was not, it would have been easy because then the automorphisms are all inner and expressed by $q(.)q^{-1}$ operations.

Now Pierre comes into the story; the drive becomes one to interpret the algebra in a very specific way in high-energy physics. I must emphasise the great importance of that project, not only because it would provide a selling point - which is certainly not to be despised - but because the high-energy field is the one in which our ideas about the nature of theory can be seen most clearly. For there is a sense in which high-energy physics has a simplicity of structure, though not of detail. But having said that, I must also say that even if a wholly satisfactory interpretation of that kind were given, it would not help at all with the questions I have raised about the Frederick construction. This project gave rise to the PITCH paper - which incidentally was finalised at the very first ANPA meeting which took place at the address on the top of my paper in 1979. I leave it to Pierre to deal with that part of the story in his papers.

So far I have been using hindsight to give an impression of a uniform progress; now I will become more honest and deal with disparate topics as they come up. But first to give an overview: It is already clear from the way in which the Frederick construction replaces a dcs by an element at the next level, so that at each step it has to be decided whether we are taking a set of elements as individuals or as a single entity, that we cannot re-interpret it in a conventional mathematical framework. Developing this, the CH is comprehensible if and only if it is seen as a process. The so-called mathematics must mirror this. (I say "so-called" because it is unhelpful to use the term mathematics - it conjures up too much of the orthodox kind.)

We did not come to this view at once or easily. It is a story of many years of trial and error. You can read an account of the various blind alleys in our book Combinatorial Physics (CP), and indeed in the earlier Proceedings of ANPA. I must say a little about CP. The only review I have seen so far is in an abstracting journal and so is of four lines only. But I am so delighted with those four that I will read them to you: "The authors aim to reinstate a spirit of philosophical enquiry in physics. They abandon the intuitive continuum concepts and build up constructively a combinatorial mathematics of process. This radical change alone makes it possible to calculate the coupling constants of the fundamental fields which - via high energy scattering - are the bridge from the combinatorial world into dynamics." I could not have done so well myself. Back to ANPA: since $\mathbb{Z}_2 = \{0, 1\}$, the construction is in terms of bit-strings. Hence the question raised at these meetings: "What do the bits in the CH mean?". In this programme of a successful formalism searching for a rational interpretation, we are, I must emphasise, in no worse case than quantum mechanics. In each case the need for understanding is raised by the problem of extending the system to get to new results.

The question of what the bits mean can be paraphrased into those I asked earlier: why $\mathbb{Z}_2$, why vectors, why matrices and why eigenvectors? To these I should add a deeper problem which came up only more recently: in the stages of the construction after the first, one "chooses the matrices to be linearly independent". Who does the choosing?

A MODEL OF THE DEVELOPMENT OF A SCIENTIFIC INVESTIGATION:

1. "Entities" become known.
2. When is such an entity a repetition of an earlier one? (call such "equal")
3. Group such equal entities into a (continually developing) set.
4. Begin theoretical analysis by taking this set as an "element" and give it a symbol or "label", a say. Labels are chosen from some (ordered) set to be specified, W say.
5. Answer 2 by "testing" a new entity against a, or later against a set of such a's, The test must give a "signal" (not a label) if the putative new entity is old.
6. In CP the choice of signals is specialised without loss of generality to one, called 0. It is also the case (which limits the generality) that 0 is the only non-label in the analysis.

7. Following the lead of Conway, utilise the ordering of the set W by proving that the nature of an equivalence relation enables the choice of the testing function f(a, b) = c to be defined by

$$f(a, b) = \min_x [f(a', b) \neq x \ \& \ f(a, b') \neq x]$$

where min refers to the ordering of W, 0 is counted as less than any member of W and a', b' are any labels less than a, b respectively. Since f(a, b) then turns out to be a commutative and associative binary operation, write it as a + b and call it discrimination.

EXAMPLE (hurricane labelling). W = [a,b,c,d,...] ordered lexicographically.

The rule gives 0 + 0 = 0, a + 0 = 0 + a = a, a + b = c = b + a, a + c = c + a = b..

Closed sets with 2^r elements (including 0) i.e. $2^r - 1$ labels.

OTHER CHOICES OF LABELLING: CONWAY takes W = {1,2,3,..}with obvious ordering.

BASTIN & KILMISTER takeW = all words in the alphabet {1,2,3...} with the ordering 1,2,12,3,13,23,123,4,....

CONNEXION with PARKER-RHODES' VECTORS: Our 134 = $[1\ 0\ 1\ 1]^t$

This derives the "use of $\mathbb{Z}_2$","vectors over $\mathbb{Z}_2$" from the process idea. By a similar argument one can extend testing to testing against a set of elements and this derives the Parker-Rhodes matrices, and the eigenvectors.

That reminds you of the answer given to all these questions in CP and to some extent in previous ANPA meetings. Some comments:

(i) The ideas obviously hark back to Concept of Order II, the emphasis there on the continually developing quality of an investigation and the need for a different kind of mathematics to capture this. The "entities" here play the same role as the parts there. The introduction of entities in this way also builds in discreteness - the origin of quantum discreteness.

(ii) The theoretical analysis which begins by labelling is the "new kind of mathematics" but the emphasis is different from what we expected in CO2. We would now prefer not to refer to it as mathematics at all because it is only like that in the sense of putting marks on paper. "Label" suggests interference from outside. This must not be a human interference, so it requires a parallel and separately operable system (which will later turn out to be Parker-Rhodes' matrix mapping space).

(iii) The signal cannot be a label since, if it were, another test would be needed to see if it were the special label corresponding to equality, which would lead to an infinite regress.But if the entity is really new, the result of the test is a label (rather than another special signal) because this prevents the loss of information.

(iv) The example of hurricane labelling stays with the letters of the alphabet without enshrining them as girl's names for simplicity.

(v) There is an important way in which Frederick's vectors differ from the labels derived here by the notion of process; that is, the fact that they have a definite dimensionality imposed from outside. For example, the label 134 could also stand for any of the infinite set: $[1\ 0\ 1\ 1\ 0]^t$, $[1\ 0\ 1\ 1\ 0\ 0]^t$ and so on.I believe that Frederick's construction is limited and ours more general, but the limitation may well be important in confining attention to physics. This needs more investigation. It is only one indication of what is present, and important, in this new version of the construction. That is what I call level-ambiguity. In the new formulation, given one element, it is no longer possible, as it was in Frederick's version, to say to what level it belongs. Of course, changes of level are still there and so are comparisons of level of two elements. This level ambiguity can be important in developing the theory; so also is the means of limiting it.

I turn to the deeper question of who does the choosing and of course the answer I must make is "no-one". The system is an autonomous one and (in order to use the statistical arguments which we need to conjecture the constitution of the unknown part of the universe from the known) we postulate an ergodic principle, that everything that could happen in the process will eventually happen. There is no need to choose the operators to be linearly independent; they may be, in which case the next level of the construction proceeds in the way Frederick said. Or they may not be, in which case the dimensionality at the next stage will be less than he said. In fact, a detailed investigation of the available operators for the seven which will give rise to the 127 dcss show that linear independence is rather likely at that stage. But when the other cases are taken into account, the number 137 thrown up by the original construction is modified to a value 137·035. The system is a self-organising one; it may go up a level or it may not. The right way to understand the Frederick construction is to see it as a bounding construction, corresponding to the system having the greatest degree of self-organisation possible. What is more important here is the way in which the whole logical situtation has changed with regard to the dimensionless constants. This ties in with our view that one must take the scale-constants as primary. Here one can use the occurrence of 137 in the original construction as IDENTIFYING the particular scale-constant under discussion. There is then no need for an Eddington-type chase through the equations to show that this is $\hbar/e^2$; we know which scale constant it is, because the others have such vastly different values. Then the correction to the integer value is quite close to the observed value (137·036) and this is a worthwhile prediction. Two comments in answer to questions:

(i) The recognition of the change in logical status of the calculation is essential, even though it may well make this calculation not publishable in Phys.Rev. The logical change is all part and parcel of the process nature of the theory.

(ii) Could there be a possible world in which the scale-constants had different values? I find the notion of possible worlds a very difficult one, because I never know what is assumed to be held constant. But here I would say, since the number has come from nothing but an analysis of the sequence of an investigation, a possible world for which the value was different would have to be one which differs from ours in the whole way in which things become known. I cannot imagine such a world, which is just a statement about my imagination.

These ideas, which all come from the making sense of the construction, also lead to seeing the system as a graded algebra so that a complete element of the algebra will have four parts; because of the process starting-point this is evidently a prediction that all our experience will have to be seen as taking place in space and time. (More needs saying to establish the details including the distinction between space and time; but space precludes that here.) But notice what is at issue here: not only that the dimensionality of space is three, an old problem, but also that the notion of dimensionality is applicable to our experience at all - experience which is purely sequential. Again, it shows the 3 + 1 character of whatever will come up so that there is at last an explanation of why it is that the quantum "particle" (the conceptual carrier of a set of variates as Eddington puts it, or of a set of quantum numbers as Pierre rephrases it) looks so like a classical particle.

I conclude with some remarks about Clifford algebras. Eddington, as I said, saw them as important; it is a little surprising that they do not seem to come up in our theory. Moreover, our own Mike Manthey, who is evidently going on the same journey as us, has been using them and the Bohm-Hiley school has shown how much of genuine quantum mechanics can be expressed in Clifford algebra terms.Has the original algebraic argument of the previous hand-out excluded a useful possibility?

CLIFFORD ALGEBRAS

1. $xx = x^2 = e$ (for all labels x)

2. RULE $xy = \min_z [xy' \neq z \ \& \ x'y \neq z] \quad x \leq y$

$xy = \min_z [xy' \neq z \ \& \ x'y \neq z \ \& \ yx \neq z] \quad x > y$

(Note: As $a + b \rightarrow ab$, $0 \rightarrow e$ & is still counted as least)

3. $xe = x(xx) = (xx)x = ex$ assuming associativity, so the signal must commute with all labels; hence confine the rule to labels.

4. Even then, ab = c, ba = d and cd ≐ ab^2a = aea = ea^2 = e^2 = dc in the same way. That is, for all x, y xy and yx are conjugate. Write * for conjugacy. Conjugate labels commute so confine the rule to x, y NOT a conjugate pair.

5. Using these results and associativity alone:

↗	a	b	c	c*
a	e	c	eb	(eb)*
b	c*	e	(ea)*	ea
c	(eb)*	ea	e	e^2
c*	eb	(ea)*	e^2	e

But using the rule gives (eb)* = ac* = b, so eb = b*, and e^2b = b** = b. Similarly for a, c so e^2 acts like 1.

6. Complete the table:

↗	a	b	c	c*	b*	a*
a	e	c	b*	b	c*	1
b	c*	e	a	a*	1	c
c	b	a*	e	1	a	b*
c*	b*	a	1	e	a*	b
b*	c	1	a*	a	e	c*
c*	1	c*	b	b*	c	e

Writing x* = −x gives quaternions in a more familiar notation.

A few comments on this rather shortened version: one starts by changing notation to ab instead of a + b since non-commutativity is required and 0 is replaced by e, so that for any label x, x^2 = e. The rule might be modified to read as shown. Next associativity is assumed, because the consequences of not doing so are that the former "closing off" to give successive closed subsets does not happen. Perhaps a stronger argument is needed? Then since xe = x(xx) = (xx)x = ex, e commutes with all labels – so evidently complete non-commutativity is impossible. Let us confine attention to labels not commuting. What is mysterious about much of quantum mechanics and about Clifford algebras is that when objects do not commute they anti-commute; why should that be? We make no such assumption here. But next, note that even amongst labels the phenomenon of conjugate pairs arises; so the rule needs to be restricted to applying only to non-conjugate pairs. The rest of the argument is as on the sheet but the conclusion deserves emphasis; essentially the first closed set we have reached is quaternions the first Clifford algebra, as can be seen by writing e = -1, so a* = -a and so on. And if more elements enter, the higher Clifford algebras arise in the same way.

What is important here is not the algebraic result that the rule we use automatically brings in the anti-commutation which we were at pains not to introduce ad hoc but that the new algebras which arise here (the Clifford) are just as good as the ones we have used up to now in the CH.

REFERENCES

E W Bastin & C W Kilmister, "The Concept of Order I.The Space-Time Structure" Proc.Camb.Phil.Soc.50, 278 (1954).

E W Bastin & C W Kilmister, "Eddington's Theory in terms of the Concept of Order" Proc.Camb.Phil.Soc.50, 439 (1954).

E W Bastin & C W Kilmister, "The Concept of Order II. Measurement" Proc. Camb.Phil.Soc.51, 454 (1955)

E W Bastin & C W Kilmister, "The Concept of Order III. General Relativity as a Technique for Extrapolating over Great Distances" Proc.Camb.Phil. Soc.53, 462 (1957).

E W Bastin & C W Kilmister, "The Concept of Order IV. Quantum Mechanics" Proc.Camb.Phil.Soc.55, 66 (1959).

T Bastin "On the Scale Constants of Physics" Studia Philosophica Gandensia 4, 77 (1965)

T Bastin, H P Noyes, J Amson & C W Kilmister "On the Physical Interpretation & the Mathematical Structure of the Combinatorial Hierarchy" Int. Jour. Theor.Phys.18, 455 (1979)

T Bastin & C W Kilmister, Combinatorial Physics, World Scientific, 1995.

J H Conway, On Numbers and Games, Academic Press, 1976.

A Short Introduction to BIT-STRING PHYSICS*

H. Pierre Noyes

Stanford Linear Accelerator Center

Stanford University, Stanford, CA 94309

Email: noyes@slac.stanford.edu; FAX: (415) 926-2525

Abstract

This paper starts with a personal memoir of how some significant ideas arose and events took place during the period from 1972, when I first encountered Ted Bastin, to 1979, when I proposed the foundation of ANPA. I then discuss program universe, the fine structure paper and its rejection, the quantitative results up to ANPA 17 and take a new look at the handy-dandy formula. Following this historical material is a first pass at establishing new foundations for bit-string physics. An abstract model for a laboratory notebook and an historical record are developed, culminating in the bit-string representation. I set up a tic-toc laboratory with two synchronized clocks and show how this can be used to analyze arbitrary incoming data. This allows me to discuss (briefly) finite and discrete Lorentz transformations, commutation relations, and scattering theory. Earlier work on conservation laws in 3- and 4- events and the free space Dirac and Maxwell equations is cited. The paper concludes with a discussion of the quantum gravity problem from our point of view and speculations about how a bit-string theory of strong, electromagnetic, weak and gravitational unification could take shape.

1 Pre-ANPA IDEAS: A personal memoir

1.1 First Encounters

When I first met Ted Bastin in 1972 and heard of the Combinatorial Hierarchy (hereinafter CH), my immediate reaction was that it must be dangerous nonsense. Nonsense, because the two numbers computed to reasonable accuracy — $137 \approx \hbar c/e^2$ and $2^{127} + 136 \approx \hbar c/Gm_p^2$ — are *empirically determined*, according to conventional wisdom. Dangerous, because the idea that one can gain insight into the physical world by "pure thought" without empirical input struck me then (and still strikes me) as subversive of the fundamental Enlightenment rationality which was so hard won, and which is proving to be all too fragile in the "new age" environment that the approach to the end of the millennium seems to encourage [84, 86].

*Work supported by Department of Energy contract DE-AC03-76SF00515.

Consequently when Ted came back to Stanford the next year (1973)[8], I made sure to be at his seminar so as to raise the point about empirical input with as much force as I could. Despite my bias, I was struck from the start of his talk by his obvious sanity, and by a remark he made early on (but has since forgotten) to the effect that *the basic quantization is the quantization of mass.* When his presentation came around to the two "empirical" numbers, I was struck by the thought that some time ago Dyson[19] had proved that if one calculates perturbative QED up to the approximation in which 137 electron-positron pairs can be present, the perturbation series in powers of $\alpha = e^2/\hbar c \approx 1/137$ is no longer uniformly convergent. Hence, the number 137 *as a counting number* already had a respectable place in the paradigm for relativistic quantum field theory known as renormalized quantum electrodynamics (QED). The problem for me became *why* should the arguments leading to CH produce a number which *also* supports this particular physical interpretation.

As to the CH itself, I refer you to Clive Kilmister's introductory talk in these proceedings[39], where he discusses an early version of the bit-string construction of the sequence of discriminately closed subsets with cardinals $2^2 - 1 = 3 \rightarrow 2^3 - 1 = 7 \rightarrow 2^7 - 1 = 127 \rightarrow 2^{127} - 1 \approx 1.7 \times 10^{38}$ based on bit-strings of length 2,4,16,256 respectively.. The first three terms can be mapped by square matrices of dimension $2^2 = 4 \rightarrow 4^2 = 16 \rightarrow 16^2 = 256$. The 256^2 discriminately independent matrices made available by squaring the dimension needed to map the third level are many two few to map the $2^{127}-1$ discriminately closed subsets in the fourth level, terminating the construction. In the historical spirit of this memoir, I add that thanks to some archeological work John Amson and I did in John's attic in St. Andrews, the original paper on the hierarchy by Fredrick Parker-Rhodes, drafted late in 1961, is now available[79].

I now ask you to join with me here in my continuing investigation of how the CH can be connected to conventional physics. As you will see in due course, this research objective differs considerably from the aims of Ted Bastin and Clive Kilmister. They, in my view, are unnecessarily dismissive of the results obtained in particle physics and physical cosmology using the conventional (if mathematically inconsistent) relativistic quantum field theory, in particular quantum electrodynamics (QED), quantum chromodynamics (QCD) and weak-electromagnetic unification (WEU).

Before we embark on that journey, I think it useful to understand some of the physics background. Dyson's argument itself rests on one of the most profound and important papers in twentieth century physics. In 1937 Carl Anderson discovered in the cosmic radiation a charged particle he could show to be intermediate in mass between the proton and electron. This was the first to be discovered of the host of particles now called collectively "mesons" . *One* such particle had already been postulated by Yukawa, a sort of "heavy photon" which he showed, using a "massive QED", gave rise to an exponentially bounded force of finite range. If the mass of the Yukawa particle was taken to be a few hundred electron masses, this could be the "nuclear force quantum". Anderson's discovery prompted Gian Carlo Wick to try to see if the existence of such a particle could be accounted for simply by invoking the basic principles of quantum mechanics and special relativity. He succeeded brilliantly, using only one column in *Nature*[96]. We summarize his argument here.

Consider two massive particles which are within a distance R of each other during a time Δt. If they are to act coherently, we must require $R \leq c\Delta t$. [Note that this postulate, in the context of my *neo-operationalist* approach[68] based on measurement accuracy[76], opens the door to *supraluminal* effects at short distance, which I am now starting to explore[69]]. Because of the uncertainty principle this short-range coherence tells us that the energy is uncertain by an amount $\Delta E \approx \hbar/\Delta t$. But then mass-energy equivalence allows a particle of mass μ or rest-energy $\mu c^2 \geq \Delta E$ to be present in the space time-volume of linear dimension $\approx R\Delta t$. Putting this together, we have the *Wick-Yukawa Principle*:

$$R \leq c\Delta t \approx \frac{c\hbar}{\Delta E} \leq \hbar/\mu c \tag{1}$$

Put succinctly, if we try to localize two massive particles within a distance $R \leq \hbar/\mu c$, then the uncertainty principle allows a particle of mass μ to be present. If this meson has the proper quantum numbers to allow it to transfer momentum between the two massive particles we brought together in this region, they will experience a force, and will emerge moving in different directions than those with which they entered the *scattering region*. Using estimates of the range of nuclear forces obtained from deviations from Rutherford scattering in the 1930's one can then estimate the mass of the "Yukawa particle" to be $\approx 200 - 300$ electron masses.

We are now ready to try to follow Dyson's argument. By 1952, one was used to picturing the result of Wick-Yukawa uncertainty at short distance as due to "vacuum fluctuations" which would allow N_e electron-positron pairs to be present at distances $r \leq \hbar/2Nm_ec$. This corresponds to taking $\mu = 2Nm_e$ in Eq. 1. Although you will not find it in the reference[19], in a seminar Dyson gave on this paper he presented what he called a crude way to understand his calculation making use of the non-relativistic coulomb potential. I construct here my own version of the argument.

Consider the case where there are N_e positive charges in one clump and N_e negative charges in the other, the two clumps being a distance $r = \hbar/m_ec$ apart. Then a single charge from one clump will have an electrostatic energy $N_ee^2/r = N_e[e^2/\hbar c]m_ec^2$ due to the other clump and visa versa. I do recall that Dyson said that the system is dilute enough so that non-relativistic electrostatic estimates of this type are a reasonable approximation. If under the force of this attraction, these two charges we are considering come together and scatter producing a Dalitz pair ($e^+ + e^- \rightarrow 2e^+ + 2e^-$) the energy from the fluctuation will add another pair to the system. Of course this process doesn't happen physically because like charges repel and the clumps never form in this way. However, in a theory in which like charges attract [which is equivalent to renormalized QED with $\alpha_e = [e^2/\hbar c] \rightarrow -\alpha_e$ in the renormalized perturbation series], once one goes beyond 137 terms such a process will result in the system *gaining* energy by producing another pair and the system collapses to negatively infinite energy. Dyson concluded that the renormalized perturbation series cannot be uniformly convergent, and hence that QED cannot be a fundamental theory, as I have subsequently learned from Schweber's history of those heroic years[85].

Returning to 1973, once I had understood that, thanks to Dyson's argument, 137 can be interpreted as a *counting number*, I saw immediately that $2^{127} + 136 \approx 1.7 \times 10^{38} \approx \hbar c/Gm_p^2$ could *also* be interpreted as a counting number, namely the number of baryons of protonic mass which, if found within the Compton wavelength of any one of them, would form a black hole. These two observations removed my objection to the calculation of two pure numbers that, conventionally interpreted, depend on laboratory measurements using arbitrary units of mass, length and time. I could hardly restrain my enthusiasm long enough to allow Ted to finish his seminar before bursting out with this insight. If this cusp turns out to be the point at which a new fundamental theory takes off — as I had hoped to make plausible at ANPA 18 — then we can tie it firmly into the history of "normal science" as the point where a "paradigm shift", in Kuhn's sense of the word[40], became possible.

However, my problem with *why* the calculation made by Fredrick Parker-Rhodes[79, 80] lead to these numbers remained unresolved. Indeed, I do not find a satisfactory answer to that question even in Ted and Clive's book published last year[10]. [I had hoped to get further with that quest at the meeting (ANPA 18, Sept.5-8, 1996), but discussions during and subsequent to the meeting still leave many of my questions unanswered. I intend to review these discussions and draw my own conclusions at ANPA 19 (August 14-17,1997).]

1.2 From "NON-LOCALITY" to "PITCH": 1974-1979

My interest in how to resolve this puzzle has obviously continued to this day. I was already impressed by the quality of Fredrick's results in 1973, and made a point of keeping in contact with Ted Bastin. This led to my meeting with Fredrick Parker-Rhodes, Clive Kilmister and several of the other Epiphany Philosophers at a retreat in the windmill at Kings Lynn, followed by discussions in Cambridge. At that point the group were trying to put together a volume on *Revisionary Philosophy and Science.* I agreed to contribute a chapter, and finished about half of the first draft of what was to become "Non-Locality in Particle Physics" [55, 56] on the plane going back to Stanford. In the course of finishing that article, I noted for the first time that the Dyson route to the hierarchy number places an energy cutoff on the validity of QED at $E_{max} = 2 \times 137 m_e c^2$, which is approximately equal to the pion (Yukawa particle) mass. I have subsequently realized that this explains a puzzle I had been carrying with me since I was a graduate student.

This puzzle, which I have sometimes called the *Joe Weinberg memnonic*, came from quite another direction[91]. An easy way to remember the hierarchy of nuclear, QED, and atomic dimensions expressed in terms of fundamental constants is the fact that

$$1.4\ Fermi \approx \frac{e^2}{2m_e c^2} = \left[\frac{e^2}{\hbar c}\right] \frac{\hbar}{2m_e c} = \left[\frac{e^2}{\hbar c}\right]^2 \frac{\hbar^2}{2m_e e^2} \approx 0.265\ Angstrom \tag{2}$$

Why nuclear dimensions should be approximately half the "classical electron radius" (i.e. $\frac{e^2}{2m_e c^2} \approx 1.4 \times 10^{-15} meter$) and hence $[1/137]^2$ smaller than than the radius of the positronium atom (i.e. $\frac{\hbar^2}{2m_e e^2} \approx 2.65 \times 10^{-10} meter$) was almost completely mysterious in 1947. It was known that the mass of the electron attributed to it's electrostatic "self-energy" as due to its charge distributed over a spherical shell fixed at this radius would have the mass m_e, but the success of Einstein's relativity had shown that this electron model made no sense[78]. The square of this parameter was also known to be proportional to the cross section for scattering a low energy electromagnetic wave from this model electron (Thompson cross section $[8\pi/3](e^2/m_e c^2)^2$), but again why this should have anything to do with nuclear forces was completely mysterious.

As we have already seen, it *was* known that the Wick-Yukawa principle [96] accounted roughly for the range of nuclear forces if those forces were attributed to a strongly interacting particle intermediate in mass between proton and electron. However, the only known particle in that mass range (the muon) had been shown experimentally to interact with nuclei with an energy 10^{13} times smaller than the Yukawa theory of nuclear forces demanded[18]. The Yukawa particle (the pion) was indeed discovered later that year, but there was still no reason to connect it with the "classical electron radius". Joseph Weinberg left his students to ponder this puzzle.

The trail to the solution of this conundrum starts with a 1952 paper by Dyson[19], despite the fact that neither he nor I realized it at the time. Two decades later, when I first heard a detailed account of the *combinatorial hierarchy*[8], and was puzzled by the problem of how a counting number (i.e. 137) could approximate a combination of empirical constants (i.e. $\hbar c/e^2$), I realized that this number is both the number of terms in the perturbation series and the number of virtual electron-positron pairs where QED ceases to be self-contained. But, empirically, $m_\pi \approx 2 \times 137 m_e$. Of course, if neutral this system is highly unstable due to 2γ decay, but if we add an electron-antineutrino or a positron-neutrino pair to the system, and identify the system with π^- or π^+ respectively, the system *is* stable until we include weak decays in the model. This suggests that the QED theory of electrons, positrons and γ-rays breaks down at an energy of $[2(\hbar c/e^2) + 1]m_e c^2$ due to the formation of charged pions, finally providing me with a tentative explanation for the Joe Weinberg memnonic. As noted above, I first presented this speculative idea some time ago [55, 56].

By the time I wrote "NON-LOCALITY", I was obviously committed to engaging in serious research on the combinatorial hierarchy as part of my professional activity. Ted was able to get a research contract to spend a month with me at Stanford. I had hoped that this extended period of interaction would give me a better understanding of what was going on; in the event little progress was made on my side. By 1978 I had met Irving Stein, and was also struggling to understand how he could get both special relativity and the quantum mechanical uncertainty principle from an elementary random walk. His work, after much subsequent development, is now available in final form [88].

Meanwhile Ted had attended the 1976 Tutzing Conference organized by Carl Friedrick von Weizsacker and presented a paper on the combinatorial hierarchy by John Amson. I agreed to accompany Ted to the 1978 meeting and present a joint paper. I arrived in England to learn of the startlingly successful successful calculation of the proton-electron mass ratio, which Ted and I had to discuss and digest in order to present Fredrick's result [11, 81] at the Tutzing meeting, which followed almost immediately thereafter. This formula has been extensively discussed at ANPA meetings. It was originally arrived at by assuming that the electron's charge could come apart, as a statistical fluctuation, in three steps with three degrees of freedom corresponding to the three dimensions of space and that the electrostatic energy corresponding to these pieces could be computed by taking the appropriate statistical average cut off at the proton Compton radius $\hbar/m_pc$. The only additional physical input is the CH value for the electronic charge $e^2 = \hbar c/137$. Take $0 \leq x \leq 1$ to be the fractional charge in these units and $x(1-x)$ the charge factor in Coulomb's law. Take $0 \leq y \leq 1$ to be the inverse distance between the charge fractions in that law in units of the proton Compton radius. Then, averaging between these limits with the appropriate weighting factors of $x^2(1-x)^2$ and $1/y^3$ respectively, Fredrick's straightforward statistical calculation gives

$$\frac{m_p}{m_e} = \frac{137\pi}{< x(1-x) >< \frac{1}{y} >} = \frac{137\pi}{(\frac{3}{14})[1 + \frac{2}{7} + \frac{4}{49}](\frac{4}{5})} \tag{3}$$

At that time the result was within a tenth of a standard deviation of the accepted value. I knew this was much too good because, for example, the calculation does not include the effect of the weak interactions. I was therefore greatly relieved when a revision of the fit to the fundamental constants changed the empirical value by 20 standard deviations, giving us something to aim at when we know how to include additional effects.

I also learned from the group during those few days before Tutzing that up to that point no one had proved the *existence* of the combinatorial hierarchy in a mathematical sense! Subsequent to the

Tutzing meeting, thanks to the kind hospitality of K.V.Laurikainen in Finland, I was able to devote considerable time to an empirical attack on that problem and get a start on actually *constructing* specific representations of both the *level* 2 ⟶ *level* 3 and the *level* 3 ⟶ *level* 4 mappings.

It turned out that neither John Amson's nor our contributions to the Tutzing conferences, despite promises, appeared in the conference proceedings. Fortunately we had had an inkling at the meeting that this contingency might arise. In the event we were able to turn to David Finkelstein and write a more careful presentation of the developments up to that point for publication in the *International Journal of Theoretical Physics*[11]. The first version, called "Physical Interpretation of the Combinatorial Hierarchy" (or PICH for short) still lacked a formal existence proof, but Clive came up with one; further, he and John Amson (whose unpublished 1976 Tutzing contribution had been extended and completed to serve as an Appendix) were able to say precisely in what sense the CH is *unique*. The final title was therefore changed to "Physical Interpretation and mathematical structure of The Combinatorial Hierarchy" affectionately known as PITCH. The finishing touches on this paper were completed at the first meeting of ANPA. This brings my informal history to the point at which Clive ended his historical sketch in his first lecture.

1.3 ANPA 1: The foundation of the organization

Although I was obviously putting considerable time into trying to understand the CH, and the Parker-Rhodes formula for m_p/m_e showed that there might be more to the physics behind it than the basic coupling constants, I was by no means convinced that the whole effort might not turn out in the long run to be an unjustifiable "numerology". I therefore, privately, took the attitude that my efforts should go into trying to derive a clear *contradiction* with empirical results which would prove the CH approach to be wrong. Then I could drop the whole enterprise and get back to my (continuing) conventional research, where I felt more at home. I was not the only one with doubts at this time. Clive told us, years later, that he had been somewhat afraid to examine the foundational arguments too closely for fear that the whole scheme would dissolve!

In the spring of 1979 I happened to make the acquaintance of an investment counselor named Dugal Thomas who was advising a large fraction of the private charitable foundations in the US. He offered to help me with fundraising if I could put together a viable organization for supporting Ted Bastin's type of research. I threw together a proposal very quickly. Dugal located a few prospective donors; like all subsequent efforts to raise substantial funds for ANPA this initial effort came a cropper. Soon after that effort started I also learned that I had received a Humboldt U.S.Senior Scientist award, giving me the prospect of a year in Germany and some extra cash. Consequently I felt encouraged to approach Clive to see if he would serve as treasurer for the proposed organization.

Clive agreed to approach Fredrick to see if he would match the small amount of "seed money" I was prepared to invest in ANPA. [The name and original statement of purpose came from the proposal I had already written. I intended that the term "natural philosophy" in the name of the organization would hark back to the thinkers at the start of the scientific revolution who were trying to look at nature afresh and shake themselves loose from the endless debates of the "nominalist" and "realist" metaphysicists of the schools.] With Fredrick's promise in hand, Clive and I approached Ted Bastin with the invitation to be the Coordinator, and asked John Amson to join us as a founding member.

The result of all this was what can be properly called ANPA 1, which met in Clive's Red Tiles Cottage near Lewes in Sussex in the early fall of 1979. John Amson was unable to attend, but endorsed our statement of purpose (modified by Ted to include specific mention of the CH) and table of organization. Once these details were in hand we had a proper scientific meeting, including thrashing out an agreed manuscript for PITCH. I gave a paper on the quantum mechanical three and four body problem, which I was working on in Germany. I noted in particular that the three channel Faddeev equations go to the seven channel Faddeev-Yakubovsky equations when one goes from three to four independent particles, reminiscent of the CH $1 \longrightarrow 2$ level transition. It is taken a long time to see what the relationship is between these two facts, but now that I am developing a "bit-string scattering theory" with Ed Jones[75], this old insight is finding an appropriate home.

Selected Topics

All meetings subsequent to ANPA 1 have been held annually in Cambridge, England. Proceedings were prepared for ANPA 7[58], and some of the papers incorporated in a SLAC-PUB[59]. The ANPA 9 proceedings [60] are available from ANPA West. Proceedings ANPA's 10 to 17 are available from Clive Kilmister. This is obviously not the place to attempt the impossible task of summarizing 16 years of work by more than 20 dedicated people in a way that would do justice to their varied contributions. I have therefore chosen to pick a few topics where I still find continued discussion both interesting and important.

2 Program Universe

2.1 Origin of Program Universe

About a decade and a half ago, Clive attempted to improve the clarity of what Ted has called "the canonical approach" [9] by admitting into the scheme a second operation besides the *Discrimination* operation, which had been central to the project ever since John Amson introduced it [2] and related it to subsequent developments [4]. Clive called this second operation *Generation* because at that stage in his thinking he saw no way to get the construction off the ground without generating bit-strings as well as discriminating them. I think he had in mind at that time a random sequence of G and D operations, but did not quite know how to articulate it. Because Mike Manthey and I were unsure how to construct a specific theory from what we could understand of this new approach, we decided to make a simple-minded computer model of the process and see how far it would lead. The first version [42] turned out to be unnecessarily complicated, and was replaced [73] by the version described below in section 2.3.

One essential respect in which the construction Mike Manthey and I turned out differs from the canonical approach is that we explicitly introduced a random element into the generation of the bit-strings rather than leaving the background from which they arise vague. Some physicists, in particular Pauli, have seen in the random element that so far as proved to be inescapable in the discussion of quantum phenomena an entrance of the the "irrational" into physics. This seems to me to equate "rationality" with determinism. I think this is too narrow a view. Statistical theories of all sorts are used in many fields besides physics without such approaches having to suffer from being castigated as irrational. In particular, biology is now founded on the proposition that evolution is (mainly) explicable as the natural selection of heritable stability in the presence of a random background. The caveat "mainly" is inserted to allow for historical contingencies[33]. Even in physics, the idea of a random "least step" goes back at least to Epicurus, and of a least step to Aristotle. I would characterize Epicurus as an exemplary rationalist whose aim was to help mankind escape from the superstitious terrors generated by ancient religions. This random element enters program universe via the primitive function "flipbit" which Manthey uses to provide either a zero or a one by unsynchronized access to a closed circuit that flips these two bits back and forth between two memory locations. Before discussing how this routine is used, we need to know a bit more about bit-strings and the operations by which we combine them.

2.2 Bit-Strings

Define a bit-string $\mathbf{a}(a;W)$ with length W and Hamming measure a by its W ordered elements $(\mathbf{a})_w \equiv a_w \in 0,1$ where $w \in 1,2,...,W$. Define the Dirac inner product, which reduces two bit-strings to a single positive integer, by $\mathbf{a}\cdot\mathbf{b} \equiv \Sigma_{w=1}^{W} a_w b_w$. Hence $\mathbf{a}\cdot\mathbf{a} = a$ and $\mathbf{b}\cdot\mathbf{b} = b$. Define *discrimination* between bit-strings of the same length, which yields a third string of the same length, by $(\mathbf{a}\oplus\mathbf{b})_w = (a_w - b_w)^2$. Clive and I arrived at this way of representing discrimination during a session in his office after ANPA 2 or 3. From this representation the *basic bit-string theorem* follows immediately:

$$(\mathbf{a}\oplus\mathbf{b})\cdot(\mathbf{a}\oplus\mathbf{b}) = a + b - 2\mathbf{a}\cdot\mathbf{b} \tag{4}$$

This equation could provide the starting point for an alternative definition of "$\oplus$" which avoids invoking the explicit structure used above.

We also will need the *null string* $\mathbf{\Phi}(W)$ which is simply a string of W zeros. Note that $\mathbf{a} \oplus \mathbf{a} = \mathbf{\Phi}(W)$, that $(\mathbf{a} \oplus \mathbf{a}) \cdot (\mathbf{a} \oplus \mathbf{a}) = 0$ and that $\mathbf{a} \cdot \mathbf{\Phi} = 0$. The complement of the null string is the *anti-null string* $\mathbf{W}(W)$ which consists of W ones and has the property $\mathbf{W} \cdot \mathbf{W} = W$. Of course $\mathbf{W} \cdot \mathbf{\Phi} = 0$.

Define *concatenation*, symbolized by "$\|$", for two string $\mathbf{a}(a; S_a)$ and $\mathbf{b}(b; S_a)$ with Hamming measures a and b and respective lengths S_a and S_b and which produces a string of length $S_a + S_b$, by

$$\begin{aligned} (\mathbf{a}\|\mathbf{b})_s &\equiv a_s \;\; if \;\; s \in 1, 2, ..., S_a \\ &\equiv b_{S_a - s} \;\; if \;\; s \in S_a + 1, S_a + 2, ..., S_a + S_b \end{aligned} \tag{5}$$

For strings of equal length this doubles the length of the string and hence doubles the size of the bit-string space we are using. For strings of equal length it is sometimes useful to use the shorthand but somewhat ambiguous "product notation" **ab** for concatenation. Note that while "$\cdot$" and "$\oplus$" are, separately, both associative and commutative, in general concatenation is not commutative even for strings of equal length, although it is always, separately, associative.

2.3 Program Universe

To generate a growing universe of bit-strings which at each step contains $P(S)$ strings of length S, we use an algorithm known as *program universe* which was developed in collaboration with M.J.Manthey [42, 73]. Since no one knows how to construct a "perfect" random number generator, we cannot start from Manthey's "flipbit", and must content ourselves with a pseudo-random number generator that, to some approximation which we will be wise to reconsider from time to time, will give us either a "0" or a "1" with equal probability. Using any available approximation to "flipbit" and assigning an order parameter $i \in 1, 2, ..., P(S)$ to each string in our array, Manthey[42] has given the coding for constructing a routine "PICK" which picks out some arbitrary string $\mathbf{P}_i(S)$ with probability $1/P(S)$. Then program universe amounts to the following simple algorithm:

PICK any two strings $\mathbf{P}_i(S)$,$\mathbf{P}_j(S)$, $i, j \in 1, 2, ..., P$ and compare $\mathbf{P}_{ij} = \mathbf{P}_i \oplus \mathbf{P}_j$ with $\mathbf{\Phi}(S)$.

If $\mathbf{P}_{ij} \neq \mathbf{\Phi}$, adjoin $\mathbf{P}_{P+1} := \mathbf{P}_{ij}$ to the universe, set $P := P + 1$ and recurse to PICK. [This process is referred to as ADJOIN.]

Else, for each $i \in 1, 2, ..., P$ pick an arbitrary bit $\mathbf{a}_i \in 0, 1$, replace $\mathbf{P}_i(S + 1) := \mathbf{P}_i(S)\|\mathbf{a}_i$, set $S := S + 1$ and recurse to PICK. [This process is referred to as TICK.]

It is important to realize that if we take a snapshot of the universe of bit-strings so constructed at any time, with the $\mathbf{P}_i$ written as rows of 0's and 1's in a rectangular array containing S columns, there is nothing in the *process* that generated them which distinguishes this universe from any of the $S!$ other universes of 0's and 1's of this height and width which could be obtained by using any of the $S!$ possible permutations of the columns. In this sense any run of program universe up to this point could just as well have produced any of these other universes. The point here is that, since the rows are produced by discrimination, and the order of the bits is the same for each row, the result is independent of the order of the bits. Similarly, since the column of bits which is adjoined to this block representation just before $S \to S+1$ is some (hopefully good!) approximation to a Bernoulli sequence, the probability of it having k 1's and $P(S)-k$ 0's is simply $P(S)!/k!(P(S)-k)!$ independent of how the rows are ordered by the order parameter i. That is, even though we have introduced an order parameter for the rows in order to make it easy to code the program in a transparent way, this parameter in itself is not intended to play any role in the physical interpretation of the model. At this stage in our argument, this means that program universe can end up with any one of the $2^{P(S)}S!$ possible block rectangles containing only 0's and 1's of height $P(S)$ and width S with some probability which is presumably calculable. This probability is relevant when we come to discuss cosmology. Nevertheless, if we look at the *internal structure* of some fixed portion of *any one* of these universes, the way in which they are constructed will allow us to make some useful and general statements. Further, these rectangular blocks of "0" 's and "i" 's are *tables* and hence have *shapes* in the precise sense defined by Etter's *Link Theory* [22, 23, 24, 25, 26]. I hope to have time to discuss Etter's theory at ANPA 19.

I have called another symmetry of the universes so constructed *Amson invariance* in reference to his paper on the BI-OROBOUROS [5]. He notes that there is nothing in the discrimination operation which prevents us from using the alternative representation for discrimination given by

$$0 \oplus' 0 = 1;\ 0 \oplus' 1 = 1 = 1 \oplus' 0;\ 1 \oplus' 1 = 1 \tag{6}$$

This will produce a dual representation of the system in which the roles of the *bits* "0" and "1" (which obviously can no longer be thought of as integers in a normal notation) are interchanged. Then when the construction of the *combinatorial hierarchy* is completed at level 4, one will have the complete system and its dual. But then, one can answer the question which has been asked in these meetings: "Where do the bits in the CH come from?" in an interesting way. In John's construction the bits are simply the two dual representations of the CH! Consequently one has a nested sequence of CH's with no beginning and no end. The essential point for me here is not this nested sequence —- which will be difficult to put to empirical test —- but the emphasis it gives to the fact that the two symbols are *arbitrary* and hence that their interchange is a *symmetry operation.* This has helped me considerably in thinking about how the particle-antiparticle symmetry and CPT invariance come about in bit-string physics.

Note that in the version of program universe presented here the arbitrary bits are concatenated only at one growing end of the strings. Consequently, once the string length S passes any fixed length L the $P(L)$ strings present will consist of some number $n_L \leq L$ of strings which are discriminately independent. Further, once $S > L$, the portion of all string of length L changes only by discrimination between members of this collection. Consequently it can end up containing at most $2^{n_L} - 1$ types of distinct, non-null strings no matter how much longer program universe runs. Whether it ever even reaches this bound, and the value of n_L itself, are *historically contingent* on which run of program universe is considered. This observation provides a model for *context sensitivity.* One result of this feature of program universe is that at any later stage in the evolution of the universe we can always separate any string into two portions, a *label string* $\mathbf{N}_i(L)$ and a *content string* $\mathrm{C}_i(S-L)$ and write $P_i(S) = \mathbf{N}_i(L)\|\mathrm{C}_i(S-L)$ with $i \in 1, 2, ..., n_L$, making the context sensitivity explicit.. Once we separate labels from content, the permutation invariance we talked about above can only be applied to the columns in the label and/or to the columns in the content parts of the strings separately. Permutations which cross this divide will interfere with any physical interpretation of the formalism we have established up to that point.

In preparation for a more detailed discussion on the foundations of bit-string physics, we note here that the alternatives ADJOIN and TICK correspond *precisely* to the production of a virtual particle represented by a 3-leg "Feynman" diagram, or "3-event", and to the scattering process represented by a 4-leg "Feynman" diagram, or "4-event"respectively. We have to use quotes around Feynman, because our diagrams obey finite and discrete conservation laws consistent with measurement accuracy. This whole subject will be more fully developed elsewhere, for instance in discussing bit-string scattering theory [75].

Another aspect of program universe is worth mentioning. We note that TICK has a *global* character since a 4-event anywhere in the bit-string universe will necessarily produce a "simultaneous" increase of string length in our space of description. This means that it will be a candidate for representing a coherent cosmological time in an expanding universe model. The time interval to which TICK refers is the shortest meaningful (i.e. finite and discrete) distance that the radius of the universe can advance in our theory divided by the velocity of light. We will return to this idea on another occasion when we discuss cosmology.

3 Lessons from the rejection of the Fine Structure paper

3.1 Background

In preparation for ANPA 9, Christoffer Gefwert[32], David McGoveran[49] and I[61] prepared three papers intended to present a common philosophical and methodological approach to discrete physics. Unfortunately, in order to get the first two papers typed and processed by SLAC, I had to put my name on them, but I want it in the record that my share in Gefwert's and McGoveran's papers amounted mainly to criticism; I made no substantial contribution to their work. We started the report on ANPA 9[60] with these three papers, followed by a paper on Combinatorial Physics by Ted Bastin, John Amson's Parker Rhodes Memorial Lecture (the first in this series), a second paper by John, and a number of first rate contributed papers. Clive Kilmister's concluding remarks closed the volume. I went to considerable trouble to get the whole thing into camera ready format and tried to get the volume into the Springer-Verlag lecture note series, but they were unwilling to accept such a mixed bag. They were interested in the first three papers and were willing to discuss what else to include, but I was unwilling to abandon my comrades at ANPA by dropping any of their contributions to ANPA 9. We ended up publishing the proceedings ourselves, with some much needed help on the physical production from Herb Doughty, which we gratefully acknowledge.

David and I did considerably more work on my paper, and I tried to get it into the mainstream literature, but to no avail. Our joint version ended up in *Physics Essays*[73]. In the interim David had seen how to calculate the fine structure of hydrogen using the discrete and combinatorial approach, and presented a preliminary version at ANPA 10[46]. I was so impressed by this result (see below) that I tried to get it published in *Physical Review Letters.* It was rejected even after we rewrote it in a vain attempt to meet the referee's objections. In order for the reader to form his own opinion about this rejection, I review the paper[51] here and quote extensively from it.

The first three pages of the paper reviewed the arguments leading to CH and the essential results already achieved. These will already be familiar to the careful reader of the material given above. With this as background we turned to the critical argument:

> We consider a system composed of two masses, m_p and m_e — which we claim to have computed from first principles[62] in terms of $\hbar, c$ and $G_{[Newton]}$ — and identified by their labels using our quantum number mapping onto the combinatorial hierarchy [73]. In this framework, their mass ratio (to order α^3 and $(m_e/m_p)^2$) has also been computed using only $\hbar, c$ and 137. However, to put us in a situation more analagous to that of Bohr, we can take m_p and m_e from experiment, and treat 1/137 as a counting number representing the coulomb interaction; we recognize that corrections of the order of the

square of this number *may* become important one we have to include degrees of freedom involving electron-positron pairs. We attribute the binding of m_e to m_p in the hydrogen atom to coulomb events, i.e. only to those events which involve a specific one of the 137 labels at level 3 and hence occur with probability 1/137; the changes due to other events average out (are *indistinguishable* in the absence of additional information). We can have any periodicity of the form $137j$ where j is any positive integer. So long as this is the only periodicity, we can write this restriction as $137j$ *steps* $= 1$ *coulomb event.* Since the internal frequency $1/137j$ is generated independently from the *zitterbewegung* frequency which specifies the mass scale, the normalization condition combining the two must be in quadrature. We meet the bound state requirement that the energy E be less than the system rest energy $m_{ep}c^2$ (where $m_{ep} = m_e m_p/(m_e + m_p)$ is used to take account of 3-momentum conservation) by requiring that $(E/\mu c^2)^2[1 + (1/137N_B)^2] = 1$. If we take $e^2/\hbar c = 1/137$, this is just the relativistic Bohr formula[14] with N_B the principle quantum number.

[Here I inserted into McGoveran's argument a discussion of the Bohr formula and how it might be derived from dispersion theory. This insertion was motivated by the vain hope that any referee would see that our reasoning was in fact closely related to standard physics. We will look at this result, called the handy-dandy formula, in a new way in the section of this paper carrying that title.]

The Sommerfeld model for the hydrogen atom (and, for superficially different but profoundly similar reasons [12], the Dirac model as well) requires two *independent* periodicities. If we take our reference period j to be integer and the second period s to differ from an integer by some rational fraction Δ, there will be two minimum values $s_0^{\pm} = 1 \pm \Delta$, and other values of s will differ from one or the other of these values by integers: $s_n = n + s_0$. This means that we can relate ("synchronize") the fundamental period j to this second period in two different ways, namely to

$$137j \frac{steps}{(coulomb\ event)} + 137 s_0 \frac{steps}{(coulomb\ event)} = 1 + e = b_+ \tag{7}$$

or to

$$137j \frac{steps}{(coulomb\ event)} - 137 s_0 \frac{steps}{(coulomb\ event)} = 1 - e = b_- \tag{8}$$

where e is an event probability. Hence we can form

$$a^2 = j^2 - s_0^2 = (b_+/137)(b_-/137) = (1 - e^2)/137^2 \tag{9}$$

Note that if we want a finite numerical value for a, we cannot simply take a square root, but must determine from context which of the symmetric factors [i.e. $(1-e)$ or $(1+e)$] we should take (c.f. the discussion about factoring a quadratic above). With this understood, we write $s_n = n + \sqrt{j^2 - a^2}$.

We must now compute the probability e that j and s are mapped to the same label, using a single basis representation constructed within the combinatorial hierarchy. We can consider the quantity a as an event probability corresponding to an event **A** generated by a global ordering operator which ultimately generates the entire structure under consideration. Each of the two events j and s can be thought of as derived by sampling from the same population. That population consists of 127 strings defined at level three of the hierarchy. In order that j and s be independent, at least the last of the 127 strings generated in the construction of s (thus completing level three for s) must not coincide with any string generated in the construction of j. There are 127 ways in which this can happen.

There is an additional constraint. Prior to the completion of level three for s, we have available the $m_2 = 16$ possible strings constructed as a level two representation basis to map (i.e. represent) level three. One of these is the null string and cannot be used, so there are 15 possibilities from which the actual construction of the label for s that completes level 3 are drawn. The level can be completed just before or just after some j cycle is completed. So, employing the usual frequency theory of probability, the expectation e that j and s as constructed will be indistinguishable is $e = 1/(30 \times 127)$.

In accordance with the symmetric factors $(1-e)$ or $(1+e)$ the value e can either subtract from or add to the probability of a coulomb event. These two cases correspond to two different combinatorial paths by which the independently generated sequences of events may close (the "relative phase" may be either positive or negative). However we require only the probability that all s_0 events be generated within one period of j, which is $1-e$. Hence the difference between j^2 and s^2 is to be computed as the "square" of this "root", $j^2 - s_0^2 = (1-e)^2$. Thus, for a system dynamically bound by the coulomb interaction with two internal periodicities, as in the Sommerfeld or Dirac models for the hydrogen atom, we conclude that the value of the fine structure constant to be used should be

$$\frac{1}{a} = \frac{137}{1 - \frac{1}{30\times 127}} = 137.0359\,674... \tag{10}$$

in comparison to the accepted empirical value of[1]

$$\frac{1}{\alpha} \simeq 137.0359\,895(61) \tag{11}$$

Now that we have the relationship between s, j and a, we consider a quantity H' interpreted as the energy attribute expressed in dynamical variables at the $137j$ value of the system containing two periods. We represent H' in units of the invariant system energy μc^2. The independent additional energy due to the shift of s_n relative to j for a period can then be given as a fraction of this energy by $(a/s_n)H'$, and can be added or subtracted, giving us the two factors $(1 - (a/s_n)H')$ and $(1 + (a/s_n)H')$. These are to be multiplied just as we multiplied the factors of a above, giving the (elliptic) equation $(H')^2/(\mu"2c^4) + (a^2/s_n^2)(H')^2/\mu^2c^4 = 1$, Thanks to the previously derived expression of $s = n + s_0$ this can be rearranged to give us the Sommerfeld formula[87]

$$H'/\mu c^2 = [1 + \frac{a^2}{(n + \sqrt{j^2 - a^2})^2}]^{-1/2} \tag{12}$$

Several corrections to our calculated value for α can be anticipated,....

3.2 The rejection

It is obvious to any physicist that if an *understandable* theory can be constructed which allows one to calculate the fine structure constant and Newton's gravitational constant to high accuracy, it should be possible to create a major paradigm shift in theoretical physics. But even though David McGoveran[46, 51] had showed us how to add four more significant figures to the calculation of the inverse fine structure constant, we were unable to make the chain of thought understandable to most of the physics community. To quote an anonymous referee for *Physical Review Letters*:

> I recommend that this letter be rejected. How happy we should all be to publish a physical theory of the fine structure constant! Any such theory, right or wrong, would be worth publishing. But this letter does not contain a theory which might be proved right or wrong. The formula for the fine-structure constant comes out of a verbal discussion which seems to make up its own rules as it goes along. Somewhere underlying the discussion is a random process, but the process is never precisely defined, and its connection to the observed quantities is not explained. I see no way by which the argument of this letter could be proved wrong. Hence I conclude that the argument is not science.

It should be obvious already that, because of my professional background, I have some sympathy with this criticism. In fact, though I was careful not to discuss this with the people in ANPA, for a number of years my research into the meaning of the CH was, in a sense, aimed at giving the canonical ANPA arguments sufficient precision so that they *could* be proved wrong. Then, I could drop my involvement with these ideas and get back to doing more conventional physics. What convinced me that ANPA must be on the right track was, in fact, the McGoveran calculation and his later extension of the same ideas to yield several more mass ratios and coupling constants in better agreement with experiment[47]. Only this past year have we succeeded in getting two publications about discrete physics into leading mainstream physics journals[35, 36]. But the basic canonical calculations are, even today, not in the kind of shape to receive that blessing. This is not the place to review my disheartening attempts to get this and other ANPA calculations before a broader audience. As an illustration of this failed strategy of hoping that the quality of our results would do the job by itself I remind you in the next section what these results were.

4 Quantitative Results up to ANPA 17

We emphasize that the only experimental constants needed as input to obtain these results are $\hbar, c$ and m_p.

The bracketed rational fraction corrections given in bold face type are due to McGoveran[47]. The numerical digits given in bold face type emphasize remaining discrepancies between calculated and observed values.

Newton's gravitational constant (gravitation):

$$G_N^{-1}\frac{\hbar c}{m_p^2} = [2^{127} + 136] \times [\mathbf{1 - \frac{1}{3 \cdot 7 \cdot 10}}] = 1.693\ 31\ldots \times 10^{38}$$

$$\textit{experiment} = 1.693\ \mathbf{58}(21) \times 10^{38}$$

The fine structure constant (quantum electrodynamics):

$$\alpha^{-1}(m_e) = 137 \times [\mathbf{1 - \frac{1}{30 \times 127}}]^{-1} = 137.0359\ \mathbf{674}....$$

$$\textit{experiment} = 137.0359\ 895(61)$$

The Fermi constant (weak interactions — β-decay):

$$G_F m_p^2/\hbar c = [256^2\sqrt{2}]^{-1} \times [\mathbf{1 - \frac{1}{3 \cdot 7}}] = 1.02\ \mathbf{758}\ldots \times 10^{-5}$$

$$experiment = 1.02\ 682(2) \times 10^{-5}$$

The weak angle (gives weak electromagnetic unification, the Z_0 and $W^{\pm}$ masses).

$$sin^2\theta_{Weak} = 0.25[\mathbf{1} - \frac{\mathbf{1}}{\mathbf{3 \cdot 7}}]^{\mathbf{2}} = 0.2267\ldots$$
$$experiment = 0.2259(46)]$$

The proton-electron mass ratio (atomic physics):

$$\frac{m_p}{m_e} = \frac{137\pi}{< x(1-x) >< \frac{1}{y} >} = \frac{137\pi}{(\frac{3}{14})[1 + \frac{2}{7} + \frac{4}{49}](\frac{4}{5})} = 1836.15\ \mathbf{1497}\ldots \quad (13)$$
$$experiment = 1836.15\ 2701(37)$$

The standard model of quarks and leptons (quantum chromodynamics):

The pion-electron mass ratios

$$m^{\pm}_{\pi}/m_e = 275[1 - \frac{\mathbf{2}}{\mathbf{2 \cdot 3 \cdot 7 \cdot 7}}] = 273.12\ \mathbf{92}\ldots$$
$$experiment = 273.12\ 67(4)$$

$$m_{\pi^0}/m_e = 274[\mathbf{1} - \frac{\mathbf{3}}{\mathbf{2 \cdot 3 \cdot 7 \cdot 2}}] = 264.2\ \mathbf{143}\ldots$$
$$experiment = 264.1\ 373(6)]$$

The muon-electron mass ratio:

$$m_\mu/m_e = 3 \cdot 7 \cdot 10[1 - \frac{3}{3 \cdot 7 \cdot 10}] = 207$$
$$experiment = 206.768\ 26(13)$$

The pion-nucleon coupling constant:

$$G^2_{\pi N\bar{N}} = [(\frac{2M_N}{m_\pi})^2 - 1]^{\frac{1}{2}} = [195]^{\frac{1}{2}} = 13.96....$$
$$experiment = 13.3(3),\ or\ greater\ than\ 13.9$$

I eventually came to the conclusion that the only way to get the attention of the establishment would be to show, in detail, that these results can be derived from a finite and discrete version of relativistic quantum mechanics (it turns out, in a finite and discrete Foch space) which is compatible with most of the conventional approach. The rest of the paper is devoted to a sketch of what I think are constructive accomplishments in that direction. The next section is a draft of the start of a paper illustrating the new strategy.

5 The Handy-Dandy Formula

One essential ingredient missing from current elementary particle physics is a *non-perturbative* connection between masses and coupling constants. We believe that one reason that contemporary conventional approaches to relativistic quantum mechanics fail to produce a *simple* connection between these two basic classes of parameters is that they start by quantizing classical, manifestly covariant continuous field theories. Then the uncertainty principle necessarily produces infinite energy and momentum at each space-time point. While the renormalization program initiated by Tomonoga, Schwinger, Feynman and Dyson succeeded in taming these infinities, this was only at the cost of relying on an expansion in powers of the coupling constant. Dyson[19] showed in 1952 that this series cannot be uniformly convergent, killing his hope that renormalized QED might prove to be a fundamental theory [85]. Despite the technical and phenomenological successes of non-Abelian gauge theories, this difficulty remains unresolved at a fundamental level. What we propose here as a replacement is an expansion in *particle number* rather than in coupling constant. The first step in this direction already yields a simple formula with suggestive phenomenological applications, as we now show.

We consider a two-particle system with energies e_a, e_b and masses m_a, m_b which interact via the exchange of a composite state of mass μ. We assume that the exterior scattering state is in a coordinate system in which the particles have momenta of equal magnitude p but opposite direction. The conventional S-Matrix approach starts on energy-shell and on 3-momentum shell with the algebraic connections[6]

$$\begin{aligned} e_a^2 - m_a^2 &= p^2 = e_b^2 - m_b^2 \\ M^2 &= (e_a + e_b)^2 - |\vec{p}_a + \vec{p}_b|^2 \\ |\vec{p}|(M; m_a, m_b) &= \frac{[(M^2 - (m_a + m_b)^2)(M^2 - (m_a - m_b)^2)]^{\frac{1}{2}}}{2M} \end{aligned} \tag{14}$$

but then requires an analytic continuation in M^2 off mass shell. Although this keeps the problem finite in a sense, it leads to a non-linear self-consistency or *bootstrap* problem from which a systematic development of *dynamical* equations has yet to emerge.

We take our clue instead from non-relativistic multi-particle scattering theory [27, 92, 93, 94, 28, 3, 97] in which once a two-particle bound state vertex opens up, at least one of the constituents must interact with a third particle in the system before the bound state can re-form. This eliminates the singular "self energy diagrams" of relativistic quantum field theory from the start. Further, the algebraic structure of the Faddeev equations automatically guarantees the unitarity of the three particle amplitudes calculated from them [31]. The proof only requires the unitarity of the two-body input[54, 57]. This suggests that it might be possible to develop an "on-shell" or "zero range" multi-particle scattering theory starting from some two-particle scattering amplitude formula which guarantees s-wave on-shell unitarity.

In order to implement our idea, rather than use Eq.15 we define the parameter k^2, which on shell is the momentum of either particle in the zero 3-momentum frame, in terms of the variable s which in the physical region runs from $(m_a + m_b)^2$ (i.e. elastic scattering threshold) to the highest energy we consider by

$$k^2(s; m_a + m_b) = s - (m_a + m_b)^2 \tag{15}$$

Then we can insure on-shell unitarity for the scattering amplitude $T(s)$ with the normalization $Im\ T(s) = \sqrt{s - (m_a + m_b)^2}|T|^2$ in the physical region by

$$\begin{aligned} T(s) &= \frac{e^{i\delta(s)} sin\ \delta(s)}{\sqrt{s - (m_a + m_b)^2}} = \frac{1}{k\ ctn\ \delta(s) - i\sqrt{s - (m_a + m_b)^2}} \\ &= \frac{1}{\pi} \int_{(m_a+m_b)^2}^{\infty} ds' \frac{\sqrt{s' - (m_a + m_b)^2}|T(s')|^2}{s' - s - i\epsilon} = \frac{2}{\pi} \int_0^{\infty} dk' \frac{sin^2\delta(k')}{k^2 - (k')^2 - i\epsilon} \end{aligned} \tag{16}$$

We arrived at this way of formulating the two-body input for multi-particle dynamical equations in a rather circuitous way. It turns out that this representation does indeed lead to well defined and soluble *zero range* three and four particle equations of the Faddeev-Yakubovsky type[57, 75], and that *primary singularities* corresponding to bound states and CDD poles [17] can be introduced and fitted to low energy two particle parameters without destroying the unitarity of the three and four particle equations. However, if we adopt the S-Matrix point of view which suggests that elementary particle exchanges should appear in this non-relativistic model as "left hand cuts" starting at $k^2 = -\mu_x^2/4$, where μ_x is the mass of the exchanged quantum [72], then we discovered[57] that the unitarity of the 3-body equations can no longer be maintained; our attempt to use this model as a starting point for doing elementary particle physics was frustrated.

We concluded that a more fundamental approach was required, in the pursuit of which[11, 73] the non-perturbative formula which is the subject of this paper was discovered[51]. However the reasoning was considered so bizarre as, according to one referee, not even to qualify as science. This paper aims to rectify that deficiency by carrying through the derivation in the context of a relativistic scattering theory, which we will call *T-Matrix theory* in order to keep it distinct from the more familiar S-Matrix theory from which it evolved. Thanks to a comment by Castillejo[16] in the context of our treatment of the fine structure of the spectrum of hydrogen[51], we finally realized that the success of our new approach required us from the start to view our *T-Matrix* as embedded in a multi-particle space. This can be accomplished using the relativistic kinematics of Eq.16 rather than of Eq.15 for the off-shell extension which leads to dynamical equations.

As we know from earlier work on partial wave dispersion relations[72], if we know that the scattering amplitude has a pole at $s_\mu = \mu^2$, or equivalently at $k^2+\gamma^2 = 0$ where $\gamma = +\sqrt{(m_a+m_b)^2 - \mu^2}$ then a subtraction in the partial wave dispersion relation given by Eq.17 easily accommodates the constraint while preserving on-shell unitarity in the physical region. This allows us to define the dimensionless *coupling constant* g^2 as the "residue at the bound state pole" with appropriate normalization. We choose to do this by the alternative definition of $T(s)$ given below:

$$\begin{aligned} T(s; g^2, \mu^2) &= \frac{g^2\mu}{s-\mu^2} = \frac{g^2\mu}{k^2(s)+\gamma^2} \\ &= \frac{1}{k\ ctn\ \delta(s) + ik(s)} \end{aligned} \tag{17}$$

Consistency with the dispersion relation, assuming a constant value for g^2, then requires that at $k^2 = 0$

$$\begin{aligned} T((m_a+m_b)^2); g^2, \mu^2) &= \frac{1}{\gamma} = \frac{g^2\mu}{\gamma^2} \\ k\ ctn\ \delta((m_a+m_b)^2) &= \gamma \end{aligned} \tag{18}$$

Consequently $g^2\mu = \gamma$ and by taking γ^2 also from Eq. 190 we obtain our desired result, the *handy-dandy formula* connecting masses and coupling constants:

$$(g^2\mu)^2 = (m_a+m_b)^2 - \mu^2 \tag{19}$$

In the non-relativistic context where $\gamma_{NR}^2 = 2m_{ab}\epsilon_{ab}$, $m_{ab} = m_a m_b/(m_a+m_b)$, $\epsilon_{ab} = m_a+m_b-\mu$, this evaluation of the value of $k\ ctn\ \delta$ at low energy is equivalent to assuming that the phase shift is given by the *mixed effective range expansion*[53]:

$$k\ ctn\ \delta = \gamma + k^2/\gamma = -\gamma + (k^2+\gamma^2)/\gamma \tag{20}$$

corresponding to the *zero range* bound state wave function $r\psi(r) = e^{-\gamma r}$ which assumes its asymptotic form very close to point where the positions of the two particles coincide. As Weinberg discusses in considerable detail in his papers on the quasi-particle approach [92, 93, 94], this constraint requires the bound state to be purely composite — i.e. to contain precisely two particles with no admixture of effects due to other degrees of freedom. We believe that his analysis supports our contention that we can claim the same interpretation for our relativistic model of a bound state, and hence that we have derived the proper two-particle input for relativistic dynamical n-particle equations of the Faddeev-Yakubovsky type. These equations, which are readily solved for three and four particle systems, will be presented on another occasion[75].

What follows next is an unsystematic presentation of results, some of which were initially obtained using the *combinatorial hierarchy*[11, 73], but which we now claim to have placed firmly within at least the phenomenology of standard elementary particle physics.....

THE TIC-TOC LABORATORY: A Paradigm for Bit-String Physics

Just prior to ANPA 19 I will be attending a conference organized by Professor Zimmermann entitled *NATURA NATURANS: Topoi of Emergence.* The following notes are intended to serve as raw material for my presentation there. Some of these ideas came out of extensive correspondence I have had with Ted and Clive following ANPA 18, and owe much to their comments. In particular the section 6 should be compared to Clive's discussion of a scientific investigation in his paper in these proceedings[39].

I would also like to remind you before we start of Eddington's parable that if we set out to measure the length of the fish in the sea, and we find that they are all greater than one inch long we have the option of concluding (a) that all the fish in the sea are greater than one inch long or (b) that we are using a net with a one inch mesh. Thinking of my approach in this way, I seem to be finding out that *because* I insist on finite and discrete measurement accuracy together with standard methodological principles. I am bound to end up with something that looks like a finite and discrete relativistic quantum mechanics that has the "universal constants" we observe in the laboratory. Whether the cosmology we observe is also constrained to the same extent is an interesting question. My guess is that we will find that historical contingency plays a significant role.

6 A Model for Scientific Investigation

I restrict our formalism so that it can serve as an abstract model for physical measurement in the following way.

We assume that we encounter entities one at a time, save an entity so encountered, compare it with the next entity encountered, decide whether they are the same or different, and record the result. If they are the same, we record a "0" and if they are different we record a "1". The first of the two entities encountered is then discarded and the second saved, ready to be compared to the next entity encountered. The recursive pursuit of this investigation will clearly produce an ordered string of "0" 's and "1" 's, which we can treat as a bit-string. We further assume that this record — which is our abstract version of a *laboratory notebook* — can be duplicated, communicated other investigators, treated as the input tape for a Turing machine, cut into segments which can be duplicated, combined and compared using our bit-string operations, the results recorded, and so on.

Our second assumption is that if we cut this tape into segments of length N and determine how many such segments have the Hamming measure a, the probability we will find the integer a, given N will approach $2^{-N}\frac{N!}{a!(N-a)!}$ in the sense of the *law of large numbers*. Without further tests all such strings characterized by the two integers $a \leq N$ will be called *indistinguishable*. It should be obvious that I make this postulate in order to be able to, eventually, derive the Planck black body spectrum from my theory. Remember that Planck's formula has stood up to all experimental tests for 97 years, a remarkable achievement in twentieth century physics! We have recently learned that in fact it also represents to remarkable precision the cosmic background radiation at 2.73^oK. For those who want to know how and why the quantum revolution started with the discovery of Planck's formula, rather than just myths about what happened, I strongly recommend Kuhn's last major work[41].

Any further structure coming out of our investigation is to be found using the familiar operation of *discrimination* $\mathbf{a} \oplus \mathbf{b}$ between two strings $\mathbf{a}$, $\mathbf{b}$ of equal length, by *concatenation* $\mathbf{a}\|\mathbf{b}$ (which doubles the string length for equal length strings), and by taking the *Dirac inner product* $\mathbf{a} \cdot \mathbf{b}$ which takes two strings out of the category of *bit-strings* and replaces them by a positive integer. This third operation is *also* how we determine the Hamming measure of a single string: $\mathbf{a} \cdot \mathbf{a} \equiv a$. It will become our abstract version of *quantum measurement*, which we interpret as the determination of a *cardinal*.

Clearly the category change between "bit-string" and "integer" is needed if we are to have a theory of *quantitative measurement*. I take this to be the hallmark of *physics* as a science.

The category change produced by taking the inner product allows us to relate two strings which combine by discrimination to the integer equation:

$$2\mathbf{a} \cdot \mathbf{b} = a + b - (\mathbf{a} \oplus \mathbf{b}) \cdot (\mathbf{a} \oplus \mathbf{b}) \tag{21}$$

If it is taken as axiomatic that (a) we can know the Hamming measure of a bit string and (b) that this implies that we can know the Hamming measure of the discriminant between two bit-strings, then this *basic bit-string theorem* seems very natural.

Once we start combining bit-strings and recording their Hamming measures, and in particular writing down sequential records of these integers, the analysis clearly becomes *context sensitive*. It is our abstract model for a *historical record*.

The underlying philosophy is the assumption that in appropriate units *any* physical measurement can be abstractly represented by a positive integer with an uncertainty of $\pm\frac{1}{2}$. If we were using real numbers, this would be expressed by saying that the value of the physical quantity represented by $n \pm \frac{1}{2}$ has a 50% chance of lying in the interval between $n - \frac{1}{2}$ and $n + \frac{1}{2}$. But in discrete physics, such a statement is *meaningless* in the sense used by operationalists. Clearly, part of our conceptual problem is to develop a language describing the uncertainty in the measurement of integers which does *not* require us to construct the real numbers.

7 Remark on Integers

It is clear from our comment on measurement accuracy that we will find it useful to talk about *half-integers* as well as integers. This will also be useful when we come to talk about angular momenta and other "non-commuting observables" in our language. But how much farther must we go beyond the positive integers? It was Kronecker who said "God gave us the integers. All else is the work of man." One of our objectives is to keep this extra work to a minimum.

I am certain that the largest string length segment we will need to construct the quantum numbers needed to analyze currently available data about the observable universe of physical cosmology and particle physics is 256, and that all we need do with such segments is to combine or compare or reduce them by the operations listed above, i.e. discrimination, concatenation and inner product. Using as a basis bit-strings of length $16W$, I also see how to represent negative integers, positive and negative imaginary integers, and complex integer quaternions. Discussion of how far we need go in that direction, or into using rational fractions other than $\frac{1}{2}$ is, in my opinion, best left until we find a crying need to do so. In any case, we have to lay considerable groundwork before we do.

For the moment we assume all we need know about the integers is that

$$1 + 0 = 1 = 0 + 1; \quad 1 \times 0 = 0 = 0 \times 1; \quad 1 + 1 = 2 \tag{22}$$

that we can iterate the third equation to obtain the counting numbers up to some largest integer N that we pick in advance as adequate for the purpose at hand, and that given any integer n so generated other than 0 or N, any second integer n' will be greater than, equal to, or less than the first. That is, we assume that the three cases

$$n' > n \quad or \quad n' = n \quad or \quad n' < n \tag{23}$$

are *disjoint*. This already implies that we can talk about larger integers[34].

I have found that McGoveran's phrase "naming a largest integer in advance", used above, needs be give more structure in my theory. I assume that all the quantum numbers I need consider can be obtained using strings of length 256 or less. If we have 2^{256} such strings, we have more than enough to *count* — in the sense of Archimedes *Sand Reckoner* — the number electrons and nuclei ("visible matter") in our universe. The mass of the number of nucleons of protonic mass needed to form these nuclei is considerably less than current estimates of the "closure mass" of our universe, leaving plenty of room for the observed "dark matter". I also believe that considerably less than the 256! orders in which we could combine the 2^{256} distinct strings of length 256 will suffice to provide the raw material for a reasonable model of a historical record of both cosmic evolution and terrestrial biological, social and cultural evolution. Such a model can be correct without being complete.

8 From the Tic-Toc Lab to the Digital Lab

Our model for an observatory is a number of *input* devices which, relying on our general model, produce bit-strings of arbitrary length which we can segment, compare, duplicate and operate on using the three operations $\mathbf{a} \oplus \mathbf{b}$, $\mathbf{a} \| \mathbf{b}$, and $\mathbf{a} \cdot \mathbf{b}$, and record the results. Our model for a laboratory adds to these observatory facilities, *output* devices which convert bit-strings into signals we can *calibrate* in the laboratory by disconnecting our input devices from the (unknown) signals coming from outside the laboratory and connecting them to our locally constructed output devices. We require that the results correspond to the predictions of the theory which led to their construction. We then take the critical step from being observers to being *participant observers* by connecting our output devices to the outside of the laboratory and seeing if the input signals from the outside into the laboratory change in a correlated way. Just how we construct input and output devices is a matter of *experimental protocol*, which we must test by having other observer-participants construct similar devices and assuring ourselves that they achieve comparable results to our own. *All* of this "background" is presupposed in what I mean by the phrase "the practice of physics" which Gefwert, McGoveran and I employed in our discussion of methodology at ANPA 9[32, 49, 61]. It will be seen that from the point of view of *theoretical physics* I am claiming that all of our operations can, in principle, be reduced to bit-string operations looking at input tapes to the laboratory, preparing output tapes connected to the outside world, and comparing the new inputs which result from these outputs. It will sometimes be convenient to refer to these tapes by more familiar names, such as "clocks", "accumulating counters", etc., without going through the detailed translation into laboratory protocol that is required for the actual practice of physics.

The most important device with which to start either an astronomical observatory or a physical laboratory is a reliable clock. For us a *standard clock* will simply consist of an input device which produces a tape with an alternating sequence of "1" 's and "0" 's, which we will also respectively refer to as *tic*'s and *toc*'s. It may be a device we construct ourselves or something that occurs without our intervention other than what is involved in producing the tape. Note that the fact that we have constructed it still leaves the physical clock *outside* our (theoretical) bit-string world; it remains essentially just as mysterious as, for example, the pulsations of a pulsar, as recorded in our observatory.

We have two types of clock with period $2W$: a *tic-toc* clock in which the sequence of $2W$ alternating symbols starts with a "1", and a *toc-tic* clock in which the sequence starts with a "0". We will represent the first by a bit-string which we call $\mathbf{L}(W; 2W)$ and the second by a bit-string $\mathbf{R}(W; 2W)$. The arbitrariness of the designation R or L corresponds the fact that our choice of the symbols on the bit-strings is also arbitrary, reflecting what we called Amson invariance in our discussion of program universe above. Independent of the specific symbolization, these two bit-strings have the following properties in comparison with each other and with the (unique) anti-null string $\mathbf{I}(2W)$ of length $2W$ (which could be called a "tic-tic clock"):

$$\begin{aligned} \mathbf{R}\cdot\mathbf{R} = &\ W &= \mathbf{L}\cdot\mathbf{L} \\ \mathbf{R}\cdot\mathbf{I} = &\ W &= \mathbf{L}\cdot\mathbf{I} \\ \mathbf{R}\cdot\mathbf{L} = &\ 0 & \\ \mathbf{I}\cdot\mathbf{I} = &\ 2W & \\ \mathbf{R}\oplus\mathbf{L} = &\ \mathbf{I} & \end{aligned} \tag{24}$$

We now have two calibrated clocks one of which we can use to make measurements, and the second to obtain the *redundant* data which is so useful in checking for experimental error — a matter of laboratory protocol which I could expound on at some length, but will refrain from so doing. We now consider an *arbitrary* signal $\mathbf{a}(a; 2W)$, and compare it with the anti-null string. We must have either that (1) $\mathbf{a}\cdot\mathbf{I} < W$ or that (2) $\mathbf{a}\cdot\mathbf{I} = W$ or that (3) $\mathbf{a}\cdot\mathbf{I} > W$. Actually, we need consider only the first two cases, because if we define $\bar{\mathbf{a}}$ by the equation $\bar{\mathbf{a}} \equiv \mathbf{a}\oplus\mathbf{I}$, we can reduce the third case to the first simply by replacing $\mathbf{a}$ by $\bar{\mathbf{a}}$. If we know the Hamming measure a, for instance by running the string through an accumulating counter which simply records the number of "1" 's in the string, we do not have to make this test because, independent of the order of the bits in the string the definitions of the inner product, the anti-null string $\mathbf{I}$ and the conjugate string $\bar{\mathbf{a}}$ guarantee that

$$\begin{aligned} \mathbf{a}\cdot\mathbf{a} &= a = \mathbf{a}\cdot\mathbf{I} \\ \bar{\mathbf{a}}\cdot\bar{\mathbf{a}} &= 2W - a = \bar{\mathbf{a}}\cdot\mathbf{I} \\ \mathbf{a}\cdot\bar{\mathbf{a}} &= 0 \end{aligned} \tag{25}$$

Hence an accumulating counter, which throws away most of the information contained in the bit-string $\mathbf{a}$, still gives us useful structural information if we know the context in which it is employed to produce its single integer result a.

If we compare the bit-string $\mathbf{a}$ with our standard clock *before* throwing away the string, we get two additional integers, only one of which is independent. Define these by

$$a_R \equiv \mathbf{a}\cdot\mathbf{R}; \quad a_L \equiv \mathbf{a}\cdot\mathbf{L} \tag{26}$$

Without actually constructing them, we now know that there exist in our space of length $2W$ two bit-strings $\mathbf{a}_R$ $\mathbf{a}_L$ with the following properties

$$\mathbf{a}_R \oplus \mathbf{a}_L = \mathbf{a}$$

$$\begin{aligned}
\mathbf{a}_R \cdot \mathbf{a}_R &= a_R = \mathbf{a} \cdot \mathbf{R} \\
\mathbf{a}_L \cdot \mathbf{a}_L &= a_L = \mathbf{a} \cdot \mathbf{L} \\
\mathbf{a}_R \cdot \mathbf{L} = 0 = \mathbf{a}_R \cdot \mathbf{a}_L &= 0 = \mathbf{a}_L \cdot \mathbf{R} \\
a_R + a_L &= a
\end{aligned} \tag{27}$$

Suppose we have a second arbitrary string $\mathbf{b}$ coming from some independent input device. Clearly we can get some structural information in the same way as before, succinctly summarized by the three integers b, b_R and b_L and the constraint $b = b_L + b_R$. If the two sources are *uncorrelated*, these amount to a pair of *classical measurements*, which we can, given enough data of the same type, analyze statistically by the methods developed in classical statistical mechanics. But if the two sources are correlated *and* we construct the string $\mathbf{a} \oplus \mathbf{b}$ and take its inner product both with itself and with our standard clock before throwing it away, we will have the starting point for a model of *quantum measurement*. This is a deep subject, on which much light has been shed by Etter's recent papers on Link Theory [22, 23, 24, 25, 26]. We have started to investigate the connection to bit-string physics[70], but have only scratched the surface. We trust that the more systematic analysis started in this paper will, eventually, help in bringing the two together. Here, we will, instead, show how our tic-toc laboratory can give us useful information about the world in which it is embedded.

9 Finite and Discrete Lorentz Transformations

We now consider a situation in which our laboratory is receiving two independent input signals one of which, for segments of length $2W$, repeatedly gives Hamming measure a and the other $2W - a$. Because of our experience with the Doppler shift, we leap to the conclusion that our laboratory is situated between two standard clocks similar to our own which are sending output signals to us. We assume that they are at relative rest but that our own lab is moving toward the one for which the recorded Hamming measure is larger than W and away from the second one for which the recorded Hamming measure is smaller than W. We calculate our velocity relative to these two stationary, signalling tic-toc clocks as $v_{lab} = (a - W)/W$ measured relative to the velocity of light. If our lab is in fact a rocket ship and we have any fuel left, we can immediately test this hypothesis by turning on the motors and seeing if, after they have been on long enough to give us a known velocity increment Δv, our velocity measured relative to these external clocks changes to

$$v' = \frac{v + \Delta v}{1 + v\Delta v} \tag{28}$$

If so, we have established our motion relative to a given, external framework. Rather than go on to develop the bit-string version of finite and discrete Lorentz boosts, which is obviously already implicitly available, I defer that development until we have discussed the more general bit-string transformations developed in the section below on commutation relations. For an earlier approach, see[64].

This situation is not so far fetched as might seem at first glance. Basically, this is how the motion of the earth, and of the solar system as a whole, have been determined relative to the 2.73^oK cosmic background radiation in calibrating the COBE satellite measurements that give us such interesting information about the early universe.

To extend this "calibration" of our laboratory relative to the universe to three dimensions, we need only find much simpler pairs of signals than those corresponding to the background radiation, namely pairs, which for the moment we will call U and D which have the properties

$$\begin{aligned} \mathbf{U}\cdot\mathbf{U} &= W = \mathbf{D}\cdot\mathbf{D} \\ \mathbf{U}\cdot\mathbf{I} &= W = \mathbf{D}\cdot\mathbf{I} \\ \mathbf{U}\cdot\mathbf{D} &= 0 \\ \mathbf{U}\oplus\mathbf{D} &= \mathbf{I} \end{aligned} \tag{29}$$

These look just like our standard clock, but compared to it we find that

$$\begin{aligned} \mathbf{U}\cdot\mathbf{R} &= W+\Delta = \mathbf{D}\cdot\mathbf{L} \\ \mathbf{U}\cdot\mathbf{L} &= W-\Delta = \mathbf{D}\cdot\mathbf{R} \end{aligned} \tag{30}$$

where (for $\mathbf{U},\mathbf{D}$ distinct from $\mathbf{R},\mathbf{L}$) we have that $\Delta \in 1,2,...,W-1$. These form the starting point for defining directions and finite and discrete rotations. As has been proved by McGoveran[45], using a statistical result obtained by Feller[29], at most three independent sequences which repeatedly have the same number of tic's in synchrony can be expected to produce such recurrences often enough to serve as a "homogeneous and isotropic" basis for describing independent dimensions, showing that our tic-toc lab *necessarily* resides in a three-dimensional space.

It is well known that finite and discrete rotations of any macroscopic object of sufficient complexity, such as our laboratory, *do not commute*. It is therefore useful to develop the non-commutative bit-string transformations before we construct the formalism for finite and discrete Lorentz boosts *and* rotations as a unified theory. Once we have done so, we expect to understand better why finite and discrete commutation relations imply the finite and discrete version of the free space Maxwell[35] and Dirac [36]equations, which we developed in order to answer some of the conceptual questions raised by Dyson's report[20] and analysis[21] of Feynman's 1948 derivation[30] of the

Maxwell equations from Newton's second law and the non-relativistic quantum mechanical commutation relations; we intend to extend our analysis to gravitation because we feel that Tanimura's extension of the Feynman derivation in this direction[89] raises more questions than it answers.

10 Commutation Relations

If we consider three bit-strings which discriminate to the null string

$$\mathbf{a} \oplus \mathbf{b} \oplus \mathbf{h}_{ab} = \mathbf{\Phi} \tag{31}$$

they can always be represented by three *orthogonal* (and therefore *discriminately independent*) strings[66, 70]

$$\begin{aligned}
(\mathbf{n}_a \oplus \mathbf{n}_b \oplus \mathbf{n}_{ab}) \cdot (\mathbf{n}_a \oplus \mathbf{n}_b \oplus \mathbf{n}_{ab}) &= n_a + n_b + n_{ab} \\
\mathbf{n}_a \cdot \mathbf{n}_b &= 0 \\
\mathbf{n}_a \cdot \mathbf{n}_{ab} &= 0 \\
\mathbf{n}_b \cdot \mathbf{n}_{ab} &= 0
\end{aligned} \tag{32}$$

as follows

$$\begin{aligned}
\mathbf{a} &= \mathbf{n}_a \oplus \mathbf{n}_{ab} \Rightarrow a = n_a + n_{ab} \\
\mathbf{b} &= \mathbf{n}_b \oplus \mathbf{n}_{ab} \Rightarrow b = n_b + n_{ab} \\
\mathbf{h}_{ab} &= \mathbf{n}_a \oplus \mathbf{n}_b \Rightarrow h_{ab} = n_a + n_b
\end{aligned} \tag{33}$$

It is then easy to see that the Hamming measures a, b, h_{ab} satisfy the triangle inequalities, and hence that this configuration of bit-strings can be interpreted as representing and integer-sided triangle. However, if we are given only the three Hamming measures, and invert Eq.35 to obtain the three numbers n_a, n_b, n_{ab}, we find that

$$\begin{aligned}
n_{ab} &= \frac{1}{2}[+a + b - h_{ab}] \\
n_a &= \frac{1}{2}[+a - b + h_{ab}] \\
n_b &= \frac{1}{2}[-a + b + h_{ab}]
\end{aligned} \tag{34}$$

Hence, if either one (or three) of the integers a, b, h_{ab} is (are) *odd*, then n_a, n_b, n_{ab} are *half-integers* rather than integers, and we *cannot* represent them by bit-strings. In order to interpret the angles in the triangle as *rotations*, it is important to start with orthogonal bit-strings rather than strings with arbitrary Hamming measures.

In the argument above, we relied on the theorem that

if $\mathbf{n}_i \cdot \mathbf{n}_j = n_i \delta_{ij}$ when $i, j \in 1, 2, ..., N$,

then

$$(\Sigma^N_{\oplus, i=1} \mathbf{n}_i) \cdot (\Sigma^N_{\oplus, i=1} \mathbf{n}_i) = \Sigma^N_{i=1} n_i \tag{35}$$

which is easily proved [70]. Thus in the case of two discriminately independent strings, under the even-odd constraint derived above, we can always construct a representation of them simply by concatenating three strings with Hamming measures n_a, n_b, n_{ab}. This is clear from a second easily proved theorem:

$$(\mathbf{a}||\mathbf{b}||\mathbf{c}||....) \cdot (\mathbf{a}||\mathbf{b}||\mathbf{c}....) = a + b + c + ... \tag{36}$$

Note that because we are relying on concatenation, in order to represent two discriminately independent strings **a**, **b** in this way we must go to strings of length $W \geq a + b + h_{ab}$ rather than simply $W \geq a + b$, as one might have guessed simply from knowing the Hamming measures and the Dirac inner product.

If we go to *three* discriminately independent strings, the situation is considerably more complicated. We now need to know the *seven* integers $a, b, c, h_{ab}, h_{bc}, h_{ca}, h_{abc}$, invert a 7×7 matrix, and put further restrictions on the initial choice in order to avoid quarter-integers as well as half-integers if we wish to construct an orthogonal representation with strings of minimum length $W \geq a + b + c + h_{ab} + h_{bc} + h_{ca} + h_{abc}$. We have explored this situation to some extent in the references cited, but a systematic treatment using the reference system provided by tic-toc clocks remains to be worked out in detail.

The problem with non-commutation now arises if we try to get away with the scalars a, a_R, Δ, W arrived at in the last section when we ask for a transformation either of the basis $\mathbf{R}, \mathbf{L} \rightarrow \mathbf{U}, \mathbf{D}$ or the rotation of the string **a** under the constraint $a_R + a_L = a = a_U + a_D$ while keeping the two sets of basis reference strings fixed. This changes $a_R - a_L$ to a different number $a_U - a_D$, or visa versa. If one examines this situation in detail, this is exactly analagous to raising or lowering j_z while keeping j fixed in the ordinary quantum mechanical theory of angular momentum. Consequently, if one wants to discuss a system in which both j and j_z are conserved, one has to make a *second* rotation restoring j_z to its initial value. It turns out that, representing rotations by $\oplus$ and bit-strings then gives different results depending on whether j_z is first raised and then lowered or visa versa; finite and discrete commutation relations of the standard form result. We will present the details of this analysis on another occasion. In effect what it accomplishes is a mapping of conventional quantum mechanics onto bit-strings in such a way as to get rid of the need for continuum representations (eg. Lie groups) while retaining *finite and discrete* commutation relations. Then a new look at our recent results on the Maxwell[35] and Dirac[36] equations should become fruitful.

11 Scattering

We ask the reader at this point to refer back to our section on the Handy Dandy Formula when needed. There we saw (Eq. 17) that the unitary scattering amplitude $T(s)$ for systems of angular momentum zero can be computed at a single energy if we know the phase shift $\delta(s)$ at that energy. For the following analysis, it is more convenient to work with the dimensionless amplitude $a(s) \equiv \sqrt{s-(m_a+m_b)^2}T(s)$,which is related to the tangent of the phase shift by

$$a(s) = e^{i\delta(s)} tan\ \delta(s) = \frac{\tan\delta(s)}{1+i\ tan\ \delta(s)} \equiv \frac{t(s)}{1+it(s)} \tag{37}$$

In a conventional treatment, given a real interaction potential $V(s,s') = V(s',s)$, $T(s)$ can be obtained by solving the Lippmann Schwinger equation $T(s,s';z) = V(s,s') + \int ds'' T(s,s'';z)R(s'',s';z)T(s'',s';z)$ with a singular resolvent R and taking the limit $T(s) = \lim_{z\to s+i0^+, s'\to s} T(s,s';z)$. Here we replace this integral equation by an *algebraic* equation for $t(s)$:

$$t(s) = g(s) + g(s)t(s) = \frac{g(s)}{1-g(s)} \tag{38}$$

One can think of this equation as a sequence of scatterings each with probability $g(s)$ which is summed by solving the equation. Here $g(s)$ will be our model of a *running coupling constant*, which we assume known as a function of energy. We see that if $g(s_0) = 0$ there is no scattering at the energy corresponding to s_0, while if $g(s_0) = +1$, the phase shift is $\frac{\pi}{2}$ at the corresponding energy and $a(s_0) = -i$; otherwise the scattering is finite.

The above remarks apply in the physical region $s > (m_a+m_b)^2$, where in the singular case a phase shift of $\frac{\pi}{2}$ causes the cross section $4\pi sin^2\delta/k^2$ to reach the unitarity limit $4\pi\lambda_0^2$ where $\lambda_0 = \hbar/p_0$ is the de Broglie wavelength at that energy; this is called a resonance and the cross section goes through a maximum value at that energy. If, as in S-matrix theory, we analyticly continue our equation below elastic scattering threshold, the scattering amplitude is real and the singular case corresponds to a bound state pole in which the two particles are replaced by a single coherent particle of mass μ, within which the particles keep on scattering until some third interaction supplies the energy and momentum needed to liberate them. There can also be a singularity corresponding to a repulsion rather than attraction, which is called a "CDD pole" in S-matrix dispersion theory[17]. The corresponding situation in the physical region is a cross section which never reaches the unitarity limit. To cut a long story short, these four cases correspond to the four roots of the quartic equation (Eq. 19) called the handy-dandy formula, which we repeat here, replacing the running coupling constant by its value at the singularity which we call $g_0 = g(s_0)$

$$(g_0)^4\mu^2 = (m_a+m_b)^2 - \mu^2 \tag{39}$$

Again to cut a long story short, the model for a running coupling constant which Ed Jones and I are exploring[75] is simply

$$g_{m_a,m_b;\mu}(s) = \sqrt{\frac{\pm[(m_a+m_b)^2-\mu^2](m_a+m_b)^2}{[k^2(s)-(m_a+m_b)^2]s}} g_{m_a m_b;\mu}(0) \tag{40}$$

The singularity at $s = 0$ is included only when m_a and m_b have a bound state of zero mass, usually called a quantum.

We have seen that when $m_a = m_e$, $m_b = m_p$ and $\mu = m_H$ the handy-dandy formula gives the relativistic Bohr formula for the hydrogen spectrum. Replacing m_p by m_e in the formula gives the corresponding formula for positronium (i.e the bound state of an electron-positron pair). But for that system, one can think of the photons produced in electron-positron annihilation as bound states of the pair with zero rest mass. This interaction is important in high energy electron-positron scattering, where it is called "Bhabha scattering". Introducing the $s^{-\frac{1}{2}}$ in this way is supposed to insure that our theory gives the correct Feynman diagram (and hence cross section) for this effect, but until we have checked the detailed derivation and predictions I warn the reader to treat this formula (Eq.42) as a guess rather than as a result actually derived from the theory.

In my paper at ANPA WEST 13[70], I started to explore the connections of this type of scattering theory to bit-strings *and* to Etter's Link Theory by making the hypothesis that

$$tan\delta_{ab} = \frac{\pm(\mathbf{a}\oplus\mathbf{b})\cdot(\mathbf{a}\oplus\mathbf{b})}{\mathbf{a}\cdot\mathbf{b}} \tag{41}$$

Unfortunately the details are about a sketchy as presented here, but at least should provide insight into where I am headed.

12 Quantum Gravity

The initial intent of Eddington, and following him of Bastin and Kilmister, was to achieve the reconciliation of quantum mechanics with general relativity. I emphasize here that they were aware of the problem, and thought they had a research program which might solve it, long before the buzz-words "Grand Unified Theory", "Theory of Everything", "Final Theory" or "Ultimate Dynamical Theory" [90, 67] became popular. In a sense they achieved the first major step with the publication of Ted Bastin's paper in 1966[7]. According to John Amson, that paper came about after many attempts to understand Fredrick's breakthrough[79] had led John to the discovery of discriminate closure[2] which gave more mathematical coherence to the scheme, and also convinced Clive and Fredrick that it was time to publish. In the event, the four authors could not agree on a text in time to meet a deadline, and authorized Bastin to go ahead with his version as sole author. This paper really does unify electromagnetism with gravitation (which was also Einstein's long sought and unachieved goal) in the sense that both coupling constants are derived from a common theory. Ted also correctly identified $(256)^4$ with the weak interactions, but missed the $\sqrt{2}$ needed to connect it numerically with the Fermi constant because he was unfamiliar with the difference between 3-vertices (Yukawa-type couplings) and 4-vertices (Fermi-type couplings) in the quantum theory of fields. As we all know, this paper was met by resounding silence. I am optimistic, in spite of my past failures, that the bit-string theory now has enough points of contact with more conventional approaches to fundamental physics to get us all into court.

Since the critical problem for many physicists is how we deal with "quantum gravity", I start there. For weak gravitational fields it makes sense to start in a flat space-time[95]. Then it can be shown that spin 2 gravitons of zero mass lead to the Einstein field equations, but as Meisner, Thorne and Wheeler note[52], the "Resulting theory eradicates original flat geometry from all equations, showing it to be unobservable." Consequently they feel that this approach says nothing about "... the greatest single crisis of physics to emerge from these equations: complete gravitational collapse." My qualitative answer is that this crisis arises from using a continuum theory at short distance where only a quantum theory makes sense. I now try to make the alternative presented here plausible.

My first step is to establish the existence of quantum gravitational effects for *neutral* particles, namely neutrons. That neutrons are gravitating objects in the classical sense was proved at Brookhaven soon after the physicists there learned how to extract epithermal neutrons from their high flux reactor and send them down an evacuated pipe a quarter of a mile long. The neutrons fell (within experimental error) by just the amount that Galileo would have predicted. That they are quantum mechanical objects was proved by Overhauser[77] by cutting a single silicon crystal 10 centimeters long into three connected planes and using critical reflection, both *calculable* from measured $n - Si$ cross sections and demonstrable *experimentally*, to form in effect a two-slit apparatus for neutrons with the positions known to atomic precision over a distance of ten centimeters. Then the shift in the interference pattern between the case when two beams were both horizontal to the case when they were in the vertical plane with one higher than the other for part of its path was proved to be precisely that predicted by non-relativistic quantum mechanics using the Newtonian gravitational potential in the Schroedinger equation. It was this brilliant experiment which convinced me that quantum mechanics is a general theory and not just a peculiarity in the behavior of electrically charged particles at short distance. In my opinion, Overhauser deserves the Nobel prize for this work, which opened up the study of the foundations of quantum mechanics to high precision experimental investigation. In the hands of Rausch[82] and others this technique has led to many tests of the model of the neutron as a quantum mechanical particle acting coherently with a precisely known mass and magnetic dipole.

Having established that neutral particles react gravitationally to the Newtonian gravitational potential $V_N(m_1, m_2; r) = G_N \frac{m_1 m_2}{r}$ as expected, it makes sense to extend our relativistic bit-string model for the Coulomb potential $V_C(m_1, m_2; r) = \frac{Z_1 Z_2 e^2}{r}$ to the gravitational case. Here Z_1, Z_2 are the electric charges expressed in units of the electronic charge e. If we are guided by Bastin's remark quoted above to the effect that the basic quantization is the quantization of mass, the analogy suggests that there is a (currently unknown) unit of mass, which we will call Δm. Then, to complete the analogy with the Coulomb case, we can replace $\alpha_C = e^2/\hbar c \approx 1/137$ with a much smaller constant $\alpha_N = G_N \Delta m^2/\hbar c$. If we also define $N_i = m_i/\Delta m$ for any particle i with *gravitational mass* m_i, the quantized version of the Coulomb and Newtonian interactions become formally equivalent, differing only by two dimensionless constants and two *quantum numbers*, independent of the units of charge or mass:

$$V_C(Z_1, Z_2; r) = Z_1 Z_2 \alpha_C \frac{\hbar c}{r}; \quad V_N(N_1, N_2; r) = N_1 N_2 \alpha_N \frac{\hbar c}{r} \tag{42}$$

We can now apply the Dyson argument to gravitation with more precision. We have seen that renormalized QED extended to enough precision to generate $N_e = 137$ electron-positron pairs, becomes unstable because of (statistically rare) clumping of clusters with enough electrostatic energy to form another pair. We interpret the fact that this disaster does not occur to the formation of a pion with mass $m_\pi \approx 2 \times 137 m_e$. In the neutral particle case, if one assumes CPT invariance, one can still distinguish fermions from anti-fermions by their spin even if they have no other quantum numbers. Hence, independent of whether or how neutral fermions and anti-fermions interact, we can expect gravitational clumping to occur for each type separately. Recall Dyson's remark that the system is dilute enough so that the non-relativistic potential can be used reliably to estimate the interaction energy of the clump. We know experimentally that Δm is much smaller than the electron mass so that the critical radius of the clump is $\hbar/\Delta m c >> \hbar/m_e c$, so Dyson's comment still applies..

In contrast to the electromagnetic case where the cutoff mass-energy of the pion requires us to go outside QED for the physics, in the gravitational case we have a cutoff energy ready to hand, namely the Planck mass $M_{Pk} \equiv [\hbar c/G_N]^{\frac{1}{2}}$. If we assemble a Planck's mass worth of neutral particles of mass Δm at rest within their own Compton wavelength, i.e. $N_G \Delta m = M_{Pk}$, and nothing else intervenes, they will fall together until they are all within a distance of $\hbar/M_{Pk}c$. At that point they will have a gravitostatic energy $N_G G_N \Delta m M_{Pk}/[\hbar/c\Delta M] = M_{Pk}c^2$, which is just sufficient to contain the kinetic energy they acquired in reaching this concentration. Inserting the definition of the Planck Mass into this gravitational energy equation we find that it is algebraically equivalent to the boundary condition with which we started: $N_G \Delta m = M_{Pk}$. They will form a black hole with the Planck radius. We conclude that $N_G = M_{Pk}/\Delta m$ neutral, gravitating objects of mass Δm at rest within their own Compton wavelength will collapse to a black hole with the Planck radius, a quantum version of the disaster that Wheeler is concerned about.

If there are, in fact, neutral fermions with no other properties than their mass, they would form such such black holes and might serve as a model for the dark matter which we know to be at least ten times as prevalent in the universe as ordinary matter. It then becomes a question in big-bang cosmology whether or not they contribute the needed effects to correlate additional observations. We defer that question to another occasion.

If we assemble enough particles of the types we know about to add up to a Planck mass, they can start collapsing and will radiate much of their energy on the way down to higher concentrations. If attractions are balanced by repulsions they could end up close enough together to form a black hole. However, as Hawking showed, small black holes interact with the "vacuum" outside the event horizon and radiate electromagnetically with the consequence that they are "white hot" and soon evaporate; the calculation was extended to rotating, charged black holes by Zurek and Thorne[98]. At the quantum scale a new possibility enters, namely that a quantum number may be possessed by the system which *cannot* be radiated away by emitting a single particle with that quantum number while conserving energy, momentum and spin. The obvious candidates for such conserved quantum numbers are baryon number, charge and lepton number, suggesting that the lightest baryon (the proton), the lightest charged lepton (the electron) and the lightest neutral lepton (the electron-type neutrino) are *gravitationally stabilized* black holes with spin $\frac{1}{2}\hbar$. This idea did not make it into the mainstream literature[63]. We use it freely in what follows. The problem then is to explain why $(M_{Pk}/m_p)^2 \approx 2^{127}$, why $m_p/m_e \approx 1836$ and why $m_{\nu_e}/m_e \leq 5 \times 10^{-5} m_e$.

Before we leave gravitation, however, we need to show within bit-string physics that the graviton has spin 2. We know from our discussion of the handy-dandy formula that we can account for spin $\frac{1}{2}$ electrons, positrons and protons and their interactions with the appropriate spin 1 photons. We have shown elsewhere[65] that we can construct the quantum numbers of the standard model of quarks and leptons. In particular, this will include the electron, muon and tau neutrinos and their anti-particles. Extending this approach to a string of length 10 we can have 6 spin $\frac{1}{2}$ fermions and 2 spin 1 bosons. On another occasion I will show how these construct 5 gravitons and 5 anti-gravitons represented in terms of strings of length 10, and go on from that to make a model for dark matter that can be expected to be approximately 12.7 times as prevalent in the universe as electrons and nucleons.

It remains to show that we can meet the three classical tests of general relativity, a problem met on another occasion [71]. Briefly, any relativistic theory gives the solar red shift (Test 1), the factor of 2 compared to special relativity in the bending of light by the sun comes from the spin 1 of the photon (Test 2), and the factor 6 compared to special relativity for the precession of the perihelion of Mercury[13] from the spin 2 of the graviton (Test 3). The calculation by Sommerfeld on which the third argument partly depends comes from simply replacing the factors $N_i = m_i/\Delta m$ in the Coulomb potential by $E_i/\Delta m$, where E_i includes the changing velocity of the orbiting particle in elliptical orbits, and hence is natural in our theory.

13 STRONG, ELECTROMAGNETIC, WEAK, GRAVITATIONAL UNIFICATION (SEWGU): A look ahead

I now show, very briefly, how SEWGU *might* take shape, if all goes well, giving the zeroth approximation to some old results such as $[\alpha_e^{-1}(0)]_0 = 137$, some new results I believe such as $[m_\tau/m_\mu]_0 = 16$, and some guesses such as $[m_t/m_Z]_0 = 2$ which I have little confidence in. The notation makes no distinction between fact and speculation, so *CAVEAT LECTOR*!

Start with strings of length 8 to label the 6 L,R *bare* states of the three types of neutrinos ν_e, ν_μ, ν_τ and their anti-particles, together with two slots for three generations, $g_1 = (10), g_2 = (11), g_3 = (01)$. To get masses and energies we have to add content strings and do a detailed analysis using bit-string scattering theory. We expect the following results to emerge

$$[\frac{m_\mu}{m_e}]_0 = 210 = [\frac{m_{\nu_\mu}}{m_{\nu_e}}]_0 \tag{43}$$

$$[\frac{m_\tau}{m_\mu}]_0 = 16 = [\frac{m_{\nu_\tau}}{m_{\nu_\mu}}]_0 \tag{44}$$

together with the massless bosons $\gamma_L, \gamma_R, \gamma_C$ (two states of spin 1 photons with the coulomb interaction as a third state) and $g_{2L}, g_{1L}, g_0, g_{1R}, g_{2R}, g_N$ (five states of the graviton plus the Newtonian interaction).

The electromagnetic coupling at the mass of the Z_0, $[\alpha_e^{-1}(m_Z)]_0 = 128$.

The masses of the charged and neutral pion compared to the electron:

$$[\frac{m_{\pi^\pm}}{m_e}]_0 = 275 = 2[\alpha_e^{-1}(0)]_0 + 1 \tag{45}$$

$$[\frac{m_{\pi^0}}{m_e}]_0 = 274 = 2[\alpha_e^{-1}(0)]_0 \tag{46}$$

We can also use label strings of length eight to get (bare) quarks with eight colors (red,orange,yellow,green,blue,purple,black,white) which can form the colorless pion triplet (π^+, π^0, π^-) and the nucleon-antinucleon doublet $(n, p, \bar{n}, \bar{p})$. To identify these as first generation hadrons, we neead to extend the string length from 8 to 10. This allows us to go on to strings of length 16 and include the neutrinos with the two generation slots accounting for the shared coupling. After a few years of effort, we expect parameters of the full Cabbibo-Kobayashi-Maskawa coupling scheme to emerge. If this fails, we may have to abandon bit-string physics!

Calling the strong coupling constant α_π rather than α_s to emphasize the conceptual difference, we are confident that at low energy

$$[\alpha_\pi(m_\pi^2)]_0 = 1 \tag{47}$$

$$[\alpha_\pi^{-1}(4m_p^2)]_0 = 7 \tag{48}$$

At high energy we expect to show that

$$[\frac{m_Z}{m_p}]_0 = 2 \times 7^2 \tag{49}$$

$$[\frac{m_t}{m_Z}]_0 = 2 \tag{50}$$

where m_t is the top quark mass. If this works out we may be able to predict a new coupling between quarks and leptons that goes beyond the standard model which *might* explain the anomalous results recently obtained at ZEUS and HERA.

I expect to be able to derive the m_p/m_e formula in a way consistent with McGoveran's last paper on that problem[48], but now directly using bit-string dynamics.

I expect to be able to understand the mapping between Foch space labels and bit-string geometry in terms of the Eulerean rectangular block Kilmister told me about with edges of length 44, 117, 240 using strings of length 256 divided into a label of length 16 and content string of length 240.

Finally, I expect to recover the old result

$$[\frac{M_{Pk}^2}{m_p^2}]_0 = 2^{127} \tag{51}$$

in terms of a running coupling constant for gravitation normalized by $\alpha_G(M_{Pk}^2) = 1$ using bit-string scattering theory.

14 Epilogue

I realize all too well how sketchy these notes are and apologize for that. I hope to get a systematic and reasonably complete outline of the full theory hammered out in a year or two. "But always at my back I hear time's winged chariot hurrying near. And yonder all before us lie deserts of vast eternity.... The grave's a fine and private place, but none I think do there embrace..." theoretical physics! So I decided to rough out this paper to insure that it is part of the Fixed Past before some singular event in the Uncertain Future terminates my activities.

References

[1] M.Aguilar-Benitz, et. al., *Particle Properties Data Booklet*, North Holland, Amsterdam, 1988, pp. 2, 4.

[2] J.Amson, "A Discrimination System", unpublished workpaper dated 31 August 1965; presented with comments in *International Journal of General Systems*, (1996)[15].

[3] E.O.Alt,P.Grassberger and W.Sandhas, *Nucl.Phys.*, **B 2**, 167 (1967).

[4] J.Amson, Appendix to [11].

[5] J.Amson, " 'BI-OROBOUROS' — A Recursive Hierarchy Construction", in [58, 59] (unpublished).

[6] R.M.Barnett, et. al. *Phys.Rev.* **D**, No. 1, Part I, p.175 (1996).

[7] T.Bastin, "On the Scale Constants of Physics", *Studia Philosophica Gandensia*, **4**, 77 (1966).

[8] Seminar on [7] given at Stanford in 1973.

[9] T.Bastin, remark in private conversation with Clive Kilmister and HPN, May 9, 1996.

[10] T.Bastin and C.W.Kilmister, *Combinatorial Physics*, World Scientific, Singapore, 1995.

[11] T.Bastin, H.P.Noyes, J. Amson and C.W.Kilmister, *Int'l J. Theor. Phys.*, **18**, 445-488 (1979).

[12] L.C.Biedenharn, *Found. of Phys.*, **13**, 13 (1983).

[13] P.G.Bergmann, *Introduction to the Theory of Relativity*, Prentice-Hall, New York, 1942, pp 212-218 provides a clear summary of Sommerfeld's calculation.

[14] N.Bohr, *Phil. Mag.*, 332-335 (February 1915).

[15] K.Bowden, ed. " General Systems and the Emergence of Physical Structure from Information Theory", special issue of the *International Journal of General Systems* (in press).

[16] L.Castillejo, private comment to HPN c.1990.

[17] L.Castillejo, R.H.Dalitz and F.J.Dyson, *Phys. Rev.* **101**, 453 (1956).

[18] E.Conversi, E.Pancini and O Piccioni, *Phys. Rev.* **71**, 209 (1947).

[19] F.J.Dyson, *Phys Rev*, **85**, 631 (1952).

[20] F.J.Dyson,*Physics Today*, No.2, 32 (1989).

[21] F.J.Dyson, *Am. J. Phys*, **58**, 209 (1990).

[22] T.Etter, "Boolean Geometry", *ANPA West Journal* **6**, No. 1, 15-33 (1996).

[23] T.Etter, "How to Take Apart a Wire", *ANPA West Journal* **6**, No. 2, 16-28 (1996).

[24] T.Etter, "Quantum Mechanics as a Branch of Merology, *Proc. PhysComp96.*

[25] T.Etter, "Process, System, Causality, and Quantum Mechanics: A Psychoanalysis of Animal Faith", in [15].

[26] T.Etter, "Quantum Mechanics as Meta-Law", submitted to *Physica D.*

[27] L.D.Faddeev, *JETP*, **39**, 1459 (1960).

[28] L.F.Faddeev, *Mathematical Aspects of the Quantum Mechanical Three-body Problem*, Davey, New York, 1965.

[29] W.Feller, *An Introduction to Probability Theory and its Applications*, Vol.I, 3"*rd* Edition, p. 316; For more detail see the first edition, p. 247 et. seq, Wiley, New York, 1950.

[30] R.P.Feynman,c.1948. Conversation with F.J.Dyson, discussed by Dyson [20].

[31] D.Z.Freedman, C.Lovelace and J.M.Namyslowski, *Nuovo Cimento* **43 A**, 258 (1966).

[32] C.Gefwert, D.O.McGoveran and H.P.Noyes, "Prephysics", in [60] pp 1-36.

[33] S.J.Gould, "Darwinian Fundamentalism", *N.Y.Review of Books*, June 12, 1997, pp 34-37.

[34] L.Kauffman, phone conversations with HPN, spring, 1997.

[35] L.H.Kauffman and H.P.Noyes, *Proc. Roy. Soc. (London)*, **A 452**, 81-95 (1996).

[36] L.H.Kauffman and H.P.Noyes, "Discrete Physics and the Dirac Equation", *Physics Letters* **A 218**, 139-146 (1996).

[37] C.W.Kilmister, ed. *Discrete Physics and Beyond: Proc.ANPA 10*, ANPA, c/o C.W.Kilmister, Red Tiles Cottage, High Street, Barcombe, Lewes, BN8 5DH, UK.

[38] C.W.Kilmister, *Eddington's search for a Fundamental Theory: A key to the universe*, Cambridge, 1994.

[39] C.W.Kilmister, "The Hierarchy Reviewed", in *Proc. ANPA 18*, T. Etter, ed., ANPA, c/o C.W.Kilmister, Red Tiles Cottage, High Street, Barcombe, Lewes, BN8 5DH, UK.

[40] T.S.Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, 1962.

[41] T.S.Kuhn, *Black Body Theory and the Quantum Discontinuity, 1894-1912*, Oxford, 1978; afterword in University of Chicago reprint, 1989.

[42] M.J.Manthey, "Program Universe" in [58, 59], pp 101-110.

[43] M.J.Manthey, ed., *Objects in Discrete Physics: Proc. ANPA 11 (1989)*; ANPA, c/o C.W.Kilmister, Red Tiles Cottage, High Street, Barcombe, Lewes, BN8 5DH, UK.

[44] M.J.Manthey, ed., *Alternatives in Physics and Biology:Proc. ANPA 12 (1990)*; ANPA, c/o C.W.Kilmister, Red Tiles Cottage, High Street, Barcombe, Lewes, BN8 5DH, UK.

[45] Discussed as Theorem 13, [49] pp 59-60.

[46] D.O.McGoveran, "The Fine Structure of Hydrogen", in [37], pp. 28-52.

[47] D.O.McGoveran, "Advances in the Foundations",.in[43], pp 82-100.

[48] D.O.McGoveran, "Motivations for using the Combinatorial Hierarchy", in[44], pp 7-17.

[49] D.O.McGoveran and H.P.Noyes, "Foundations of a Discrete Physics" in [60], pp 37-104.

[50] D.O.McGoveran and H.P.Noyes, "Foundations of a Discrete Physics" SLAC-PUB-4526, June, 1989.

[51] D.O.McGoveran and H.P.Noyes, *Physics Essays*, **4**, 115-120 (1991).

[52] C.W.Misner, K.S.Thorne, and J.A.Wheeler, *Gravitation*, Freeman, New York, 1973, Box 18.1, p. 437..

[53] H.P.Noyes, *Ann.Rev.Nuc.Sci.*, **22**, 465-484 (1972).

[54] H.P.Noyes, *Czech.J.Phys.*, **B24**, 1205-1214 (1974).

[55] H.P.Noyes, "Non-locality in Particle Physics", SLAC-PUB-1405 (1974).

[56] 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).

[57] H.P.Noyes, *Phys.Rev* **C 26**, 1858-1877 (1982).

[58] H.P.Noyes, ed. *Proc. ANPA 7*, 1985; I have what may be the only copy.

[59] H.P.Noyes, "ON THE CONSTRUCTION OF RELATIVISTIC QUANTUM MECHANICS: A Progress Report, SLAC-PUB-4008, June 1986 (Part of Proc. ANPA 7).

[60] H..P.Noyes, ed. *'DISCRETE AND COMBINATORIAL PHYSICS: Proc. ANPA 9 (1987)*, ANPA West, 112 Blackburn Ave., Menlo Park, CA 94025.

[61] H.P.Noyes, "DISCRETE PHYSICS: Practice, Representation and rules of correspondence" in [60], pp 105-137.

[62] H.P.Noyes, in *Physical Interpretations of Relativity Theory.II.*, M.C.Duffy, ed. (Conf. at Imperial College, London, 1990), p. 196 and SLAC-PUB-5218, July, 1990.

[63] H.P.Noyes, "Comment on 'Statistical Mechanical Origin of the Entropy of a Rotating, Charged Black Hole", SLAC-PUB-5693, November 1991 (unpublished).

[64] H.P.Noyes, "The RQM Triangle", in *Reconstructing the Object* (Proc. ANPA WEST 8), F.Young, ed. 1992; ANPA West, 112 Blackburn Ave, Menlo Park, CA 94025.

[65] H.P.Noyes, "Bit-String Physics, a Novel 'Theory of Everything' ", in *Proc. Workshop on Physics and Computation (PhysComp '94)*, D. Matzke, ed.,94, Los Amitos, CA: IEEE Computer Society Press, 1994, pp. 88-94.

[66] H.P.Noyes, "On the Lorentz Invariance of Bit-String Geometry", SLAC-PUB-95-6760 (September 1995), a slightly revised version of the presentation at ANPA WEST 8, Feb. 18-20, 1995.

[67] H.P.Noyes, "Some Remarks on Discrete Physics as an Ultimate Dynamical Theory", in *Philosophies: Proc. ANPA 17 (1995)*, K.G.Bowden, ed., pp 26-37.

[68] H.P.Noyes, "Operationalism Revisited", *Science Philosophy Interface*, **1**, 54-79 (1996).

[69] H.P.Noyes, "Are Partons Confined Tachyons?", in *Proc. ANPA WEST 12 (1996)*, F.Young, ed., ANPA West, 112 Blackburn Ave., Menlo Park, CA 94025.

[70] H.P.Noyes, "On the Distinction Between Classical and Quantum Probabilities", in *Proc. ANPA WEST 13 (1997)*, T.Etter ed., ANPA West, 112 Blackburn Ave., Menlo Park, CA 94025.

[71] H.P.Noyes and D.O,McGoveran, "Observable Gravitational and Electromagnetic Orbits and Trajectories in Discrete Physics", in *Physical Interpretations of Relativity Theory.I*, (Proc. II), M.C.Duffy, ed, British Society for the Philosophy of Science, 1988, p.42.

[72] H.P.Noyes and D.Y.Wong, *Phys.Rev.Letters*, **3**, 191(1959).

[73] H.P.Noyes and D.O.McGoveran, *Physics Essays*, **2**, 76 (1989).

[74] H.P.Noyes, C.Gefwert and M.J.Manthey, "Toward a Constructive Physics", SLAC-PUB-3116 Rev., September, 1983.

[75] H.P.Noyes and E.D.Jones, "Zero Range Scattering Theory.II.A minimal model for relativistic quantum mechanical 2, 3, and 4 particle scattering" (in preparation).

[76] H.P.Noyes, *DISCRETE PHYSICS: A New Fundamental Theory*, J.C. van den Berg, ed. (in preparation).

[77] R.Collela, A.W.Overhauser and S.A.Werner, *Phys.Rev.Lett.*, **34**, 1472-4 (1975).

[78] A.Pais, *Subtle is the Lord*, Oxford, 1982, pp 155-159.

[79] A.F.Parker-Rhodes, "Hierarchies of Descriptive Levels in Physical Theory", Cambridge Language Research Unit, internal document I.S.U.7, Paper I, 15 January 1962. Included in [15].

[80] A.F.Parker-Rhodes, *The Theory of Indistinguishables*, *Synthese*, **150**, Reidel, Dordrecht, Holland (1981).

[81] In [80], p 184, Fredrick gives a different version of the m_p/m_e calculation.

[82] H.Rausch, "Quantum Measurements in Neutron Interferometry", in *Proc. 2nd Int. Symp. Foundations of Quantum Mechanics*, Tokyo, 1986, pp 3-17.

[83] "Review of Particle Properties", *Physical Review* **D**, Part I, 1 Aug. 1994, p. 1233.

[84] C.Sagan *The Demon-Haunted World*, Random House, New York, 1995.

[85] S.Schweber, *QED and the Men Who Made It*, Princeton, 1994, p.565.

[86] A.Sokal, "Transgressing the Boundaries: Toward a transformative hermeneutics of Quantum Gravity", *Social Text* 46/47, 1996; — , "A Physicist Experiments with Cultural Studies", *Lingua Franca*, May/June 1996.

[87] A.Sommerfeld, 1916 *Ann. d. Physik*, **51**, 1.

[88] I.Stein, *The Concept of Object as the Foundation of Physics*, Fritz Lang, New York, 1996.

[89] S.Tanimura, *Annals of Physics*, **220**, 229-247 (1992)

[90] G.'t Hooft, "Questioning the answers, or stumbling on good and bad Theories of Everything", in *Physics and our View of the World*, J. Hilgewoord, ed., Cambridge University Press, 1994.

[91] J.Weinberg, in graduate course on classical electrodynamics, UCBerkeley 1947.

[92] S.Weinberg, *Phys.Rev.*, **130**, 776 (1963).

[93] S.Weinberg, *Phys.Rev.*, **131**, 440 (1963).

[94] S.Weinberg, *Phys.Rev.*, **133 B**, 232 (1964).

[95] S.Weinberg, *Phys. Rev.* **B 138**, 988-1002 (1965).

[96] G.C.Wick, *Nature*, **142**, 993 (1938).

[97] O.A.Yakubovsky, *Sov.J.Nucl.Phys.*, **5**, 937 (1967).

[98] W.H.Zurek and K.S.Thorne, *Phys. Rev. Lett.*, **54**, 2171-2175 (1985).

The Gateway to Paradox

by

Viv Pope

Verbatim:

ANPA: the Alternative Natural Philosophy Association. For an outsider, what does this mean? Does it mean that there are many alternative natural philosophies to be considered at the Association? If so, then ANPA is very liberal – like PIRT, held at Imperial College, London. Or does it mean that there is just one of these possible alternatives presented, namely, the theory of the Combinatorial Hierarchies? If so, then ANPA is illiberal – like the other conferences of the sort which virtually exclude any contribution from members of what is, to them. the 'opposition'.

Where, then, on this scale of liberality/illiberality does ANPA stand? Let me speak as I find. What the majority of members of ANPA are doing, it seems to me, is similar to what people at other conferences are doing; that is, in all their various supporting and competing ways, producing variations on a single theme. At ANPA, this is to construct a programme of numbers and hierarchical combinations of numbers which, hopefully, will match the mathematical regularities of nature. That is essentially a metaphysical agenda which seems to need no particular philosophical underpinning. Basically, it is a code-cracking exercise. The code is that of nature – whatever, in the end, that might turn out to be – and the code-cracker is the theoretician, a Leibnitz, an Ayer or a Wittgenstein, an Eddington or a Parker-Rhodes. What matters, however, in this sort of venture, is not the philosophy, or the philosopher whose name is applied to this approach. What really matters, in the end, is simply the success or failure of this purely numerological exercise. Once the code is cracked, by whatever means – call it rational, empirical or whatever – no-one will give a hoot about what philosophy, or whose philosophy it was based on.

Now, whether or not such a programme of numbers is possible is anyone's guess. If discovered it will definitely be a world-renowned *fait accompli*. If not, then the possibility of its eventual discovery remains a matter of belief, and it is difficult to see how such a belief or the possibility of its fulfilment can ever be dispelled.

Well, that is one agenda; let's call it the *constructivist* agenda. There is, however, another agenda, which is by no means antagonistic to this metaphysical one and which *should*, by rights, be regarded as complementary to it. That is to say, if the one agenda is to take what we currently think of as 'the universe' and seek to explain it with systems of numbers and so on, then the other agenda has to be that of analysing our current concepts of this 'universe' to ascertain to what extent that 'universe' is nature or a product of our conventional imaginings. These two agendas, the constructivist and the analytic are, I repeat, complementary and should not be taken as excluding each other. We simply *have* to know, before embarking on methods of explanation, how much of what we seek to explain is true physical fact and how much is sheer illusion.

Now most outsiders I have spoken to also see what ANPA is doing as metaphysics rather than physics. As an ANPA member I have no difficulty with that. However, I find that in ANPA there are three sorts of metaphysical approaches more or less mixed – not to say confused. Sometimes the CH people speak as though what they are doing is to produce a programme of mathematics to match the universe as we currently conceive it. At other times they speak as though their mathematical programme is to be taken as a system, on purely æsthetic grounds, and our customary physical concepts adapted towards it. Or they speak as though the whole aim of the ANPA method is simply to rationalise, by eclectic mathematical means, all the conceptual accretion of our scientific history, without even pausing to analyse it – far less at the level of analytic philosophy.

In contrast to all these various approaches – but, as I stress, not necessarily antagonistic to them – is that of the analytic philosopher who, as Socrates did in days of yore, takes all current claims to knowledge and examines them, by a process of critical discussion (or dialectic) to see whether they truly are bits of knowledge or mere fancy. This sort of examination can cause distress – which, of course, is why Socrates suffered the fate he did at the hands of those whose authority he justifiably questioned. I have a friend, back home, who is a very personable and utterly competent engineer. This is because, in his youth, he diligently studied all the proper textbooks on his subject, which, of course, makes him so very well qualified and competent. What this very good friend *cannot* do, it seems, is accept that the principles of science that he studied at College, connected with the names of Newton, Faraday, Maxwell, *et al.*, were subscribed by ordinary mortals of the sort he rubs shoulders with in the street or at, say, the 'Blue Anchor', where my friends and I meet for philosophical discussions on Tuesday evenings. In his mind, these ideas that he studied are crystallised, fixed for all time, so that no argument which puts them in question, no matter how logical, can even be considered, let alone prevail.

Now that, I find, is the way it tends to be with the scientific community in general. The more highly qualified and competent a scientist is, the less likely he may be to accept that the principles he studied, and in which he was examined in qualifying for his profession, are – and never were – more than *tentative* philosophical proposals as to how we should interpret our scientific experience; and that far from being fixed for all time these ideas are held on lease until they may be clearly demonstrated to have outlived their conceptual usefulness.

It is a truism of history that such well-established and well-intentioned members of the scientific community, in their typical antagonism towards 'upstart' new ideas, as often as not turn out to be the enemies of Science rather than its friends. Just how deep their antagonism can become has been manifest, recently, in an attack on myself, via the INTERNET, for having the temerity even to *suggest* that on logical grounds we may apply Occam's 'razor' to our current conceptions of how light travels, and shave-off all the hoary accretion of paradoxes created by the venerated myth of the clairvoyant, space-travelling 'photon'. In response to that attack I have encapsulated, on INTERNET, a homepage, which has been offered to all ANPA members.*

Those who have read my Homepage have given it a very good press. They can see that for well-considered and well-discussed reasons connected with what I have identified as the necessary 'second agenda', the *analytic* one, I am absolutely serious in my advocacy of getting rid of the concept of the photon. My detailed arguments for this are as explained in the Homepage and in my publications to which it refers. Anyway, in this short space, I have to collapse the account; and here is how I explained it to my philosophical friends at the 'Blue Anchor'.

Another friend of mine, who renovates vintage cars, acquired, secondhand, what in the auto-industry is called a 'guillotine', a very large machine for cutting steel sheets. The parts of this had lain around in his yard for some time until eventually he got me to help him put the two main sections of the thing, the cutter and the bed, together – trying not to lose our fingers, which I nearly did. These sections were so large and heavy that we had to use pinch bars to lever them into place.

* (http://ourworld.compuserve.com/homepages/Viv_Pope)

The cutter and the bed, when well-greased, should have 'clunked' easily together. However, no matter what we did, they simply wouldn't go. After much cursing and theorising we eventually found what was wrong. Someone, for whatever misguided reason, had jammed onto one of the locating pins a very thin nut. That nut was foreign to the system but was jammed on so hard that we thought it was part of it. Eventually, however, we decided that the nut was no part of the design and that it would have to go. So, with some misgiving, we cut the thing off. As soon as we did that, the cutter and the bed slid neatly together.

The point to this story, as I related it to members of the Blue Anchor – one of whom, at the time, was Ted (Bastin), here – is that, metaphorically speaking, the '*photon*' is the 'nut' which prevents the two main sections of modern physics, Relativity and Quantum theory, from 'clunking' naturally together. If that nut – which another Blue Anchor friend aptly dubbed the 'Gordion Nut' – had not been cut away, those two bits of the guillotine would have remained apart until now. As my friend remarked, whatever else the machine might be used for, with the nut attached it could never be for its originally intended purpose.

And so it is, I suggested, with Relativity and Quantum physics. While the analogous 'Gordion Nut' of modern Physics, the 'photon', remains in place, whatever uses we may put our Physics to, it can never be for any truly natural purpose.

It is plain, then, surely, how a superstitious dread of new and radical solutions can work against our best interests. I am also reminded of the tale told by my father-in-law who was a civil engineer. On a building site that he was supervising, an earth-moving machine was out of commission because the mechanic had no spanner short enough to allow it to turn against a bulk-head. 'Have you got a spanner to fit the bolt?' my father-in-law asked. 'Yes,' came the reply, 'but they're all too long.' 'Then cut one of them,' ordered my father-in law. '*What*?' said the mechanic, 'Have you any idea what this set of tools cost?' 'Have you any idea' came my father-in-law's instant rejoinder, 'of the cost of this machine standing here idle?' The mechanic pondered the wisdom of my father-in law's decision, then duly sacrificed the spanner.

The principle is the same, of course, in science as in any other walk of life, except that in science, because of its exalted character, we are inclined to forget how perfectly obvious is the need, sometimes, for truly radical and adventurous solutions. But of course, there are always vested interests in preserving the status quo. For instance, the use to which my friend's wife had previously put that guillotine – as a shelf for displaying her potted plants – was 'ruined' by our removal of the 'Gordion nut' and the restoring of the machine to its proper use. In the same way, for the purposes of some present scientists, the uses – not to say misuses – to which they have personally and variously put Science may be 'ruined' by getting rid of the 'photon'. But it all depends what we want out of Science, a true understanding of nature, or just 'usefulness'. This 'usefulness' of retaining the 'photon' concept has sometimes, to my astonishment, been stressed almost to the exclusion of the need for an overall logical and philosophical understanding. In opposition to this I have argued that ANPA is – or at least purports to be – a whole philosophical 'building-site', so to speak, not a mere shed, or shrine, on that site dedicated to the protection and preservation of nuts and spanners.

Now Ted, here, doesn't like these homilies of mine – which he refers to as Viv's 'parables'. Nor is he alone in this. Other mathematicians and physicists have deemed this 'commonsense' way of arguing, which does not involve the usual mathematical formulæ, as 'unfair', 'unprofessional' ... you name it! As I said, I have even been castigated on Internet for arguing 'more like a 'Dylan Thomas than a natural philosopher.' But no amount of mathematical *formulæ*, however precise, can validate an argument which, at the level of commonsense logic, is clearly nonsensical. Anyone wishing to study my precise logical and mathematical backing for these 'homilies' of mine has only to read my homepage and the literature to which it refers. For now, let me finish with another homily which reiterates something I said in my first paper at PIRT, six years ago (My apologies to those of you who, like Pierre, here, may have already seen this.)

Shows OHP 1
(picture of the RAF homing station)

This is a true story from the days of National Service. In a very large field, on the approach to the main runway of a remote airfield was a building called the 'homer', which was used to give radio directional bearings to aircraft. Now the homer operators had a problem. They found it impossible to stay awake all night, alone on duty, listening to what was, for the most part, 'white noise' in their headphones. The authorities had no sympathy with this 'psychological' problem. They simply and stolidly considered it the 'duty' of these airmen to stay awake all night, no matter what. However, having been punished on too many occasions for being caught sleeping on duty, the airmen contrived a highly efficient loudspeaker alarm-system to wake them when duty called, and another device to detect the approach of the RAF police.

In this conflict of interests, the problem for the RAF special police – the SPs – was, of course, that of catching these operators sleeping on duty. But somehow, whenever they approached the homer they knew that the operators were alerted because they could see the lights in the building switch on as they approached. Suspecting that the operators had a trip-switch hidden either in or under the turf, the S.P.s decided to tread very carefully and start digging as soon as they saw the lights come on in the homer. With no success, and far from being unintelligent, they assumed that they had set off some delayed-action mechanism. So, as soon as they saw the lights come on they stopped and exactly retraced their steps, digging in all the likely places. But no matter where they looked they never found the switch.

But the homer operators *did* have a switch. However, in such a very large field, the S.P.s never realised just how far away from the homer it was. Their customary entrance to the field was by way of an ordinary five-bar gate, near the ruin of a hut which had originally housed relief operators during the second world-war.

shows OHP 2
(picture of ruined hut)

Between this hut and the homer was an abandoned deep lying, (i.e., 'bomb-proof') conduit:

shows OHP 3
(picture of conduit).

In that conduit were the original wires whose ends the operators had found in the fuse-box at the homer. To these they connected their 'system' and to the gate at the other end they attached their switch. The secret switch would be activated by the slightest movement of the gate in being opened or climbed.

The gate, as I say, was so far away from the homer that it never occurred to the S.P.s that *this* was the source of their trouble. So no matter where they dug in the field, or what theories or experiments they devised, they never did nor ever would , in that way, solve their problem.

This problem of catching the sleeping operators is, by analogy, let's say, that of understanding the physical world. The 'field' is that of modern physics and the 'gateway' is the customary entrance to that field by way of our scientific traditions. The 'SPs' are those conventional physicists who can contemplate no entry into the field other by the customary route. Thus, even when it becomes plain that they are digging in all the wrong places, they persist in doing so, no matter how many problems and paradoxes are generated in approaching the field by that conventional gateway.

Now I am not suggesting that ANPA or the Combinatorial Hierarchies people are the ones taking that abortive, conventionalist route. You have seen, for instance, my cartoon on the cover of the Newsletter, which shows the little buggy called ANPA going down a road very different from that taken by the Rolls-Royce of conventional physics. However, that route taken by ANPA concerns only the first agenda, the constructivist, or metaphysical one which, for physicists in general, is definitely non-standard. As I have said, I have no quarrel with ANPA on that adventurous first agenda. My criticisms concern only what I have described as the necessary – and, to me, very much neglected – 'second, or analytic agenda'. Imagine how it would be if all our scientific traditions were venerated to the extent that none was ever revised or turfed out! Imagine, for instance, CH having to explain not only the world of scale-constants and the Big Bang but also the 'luminiferous æther', the 'caloric' and 'phlogiston', etcetera, of 17th-18th century science; the angels and cherubim of the Canon Law Synthesis; the Greek Pantheon, the elephants and turtles of Hindu cosmology ... and so on and so on! And when does that process of revision and turfing-out officially stop? With 'wave-particles', 'electric charges', 'field-forces', 'quarks', the 'God particle', 'strings and superstrings', 'virtual particles', 'superluminality', clairvoyant 'photons', 'strangeness', 'upness and downness', particles which exist in two places at once, particles which endure backwards in time ... need I go on?

What I am saying is that this *analytic* agenda cannot be ignored and that we cannot afford to be any less creative and adventurous in this second agenda than in our first. How can our attempts to explain the 'universe' possibly succeed if our current ideas of that 'universe' remain unexamined? What price the Combinatorial Hierarchies if it succeeds in explaining something that may turn out, in the end, to be no more than a contemporary fantasy?

So you see, I am by no means antagonistic towards, or in competition with, the Combinatorial Hierarchies, as seems to have been assumed. Clive (Kilmister), here, has convinced me that the numerical regularities in nature do definitely require explanation; and there is little doubt in my mind that the explanation, when achieved, will be of the sort explored by CH.

However, so far as the *second* agenda is concerned, I have to say that my efforts, here at ANPA, to get these things discussed at the level of *analytic philosophy* have so far fallen on stoney ground – doubtless because my criticisms may have seemed to constitute 'an attack on CH'. So now I have reconsidered my position. This time I intend to show, by means of a very simple conceptual model, and without dragging in the dreaded 'philosophy', how to cut away that 'Gordion Nut' which prevents relativity and quantum theory from 'clunking' neatly and naturally together. The model destroys the myth that there is any 'Einstein separation' between the world of the object and the world of the observer – between quantum theory and relativity, in effect. It thus finally disposes of the bogus dualism which has bedevilled physics since the time of Descartes. The model is called the Cinematic Model of Quantum Touching and, in my view, it completely solves the problem of action-at-a-distance – as you shall see in my next talk, here, tomorrow, on **'The Cinematic Model of Quantum Touching'**.

5

Huygens' Principle, Physics and Computers Draft 4.2w
(*or* On General Physical Systems Theory II)

KEITH BOWDEN,
University of East London, Longbridge Rd,
Dagenham, Essex RM8 2AS, UK.
(keith@uel.ac.uk)

We discuss the emergence of topology from a consideration of set extensions in General Systems Theory. Boundaries arise in a natural way, separating independent elements or regions of the system. Our aim is a unification of Etter theory, Kron's method of Tearing and Jessel's formulation of Huygens' Principle. This should make explicit the equivalence between the objective, structural, holographic and the subjective, relative definitions of information, sought in [Bowden, 1994b], reprinted in this Special Issue. It connects the abstract generalisations of Schrodinger's equation and Born's rule derived in probabilistic Etter theory with the real world of electrical and other physical phenomena in General Physical Systems Theory. This paper can be considered as a continuation of [Bowden, 1990] and [Bowden, 1994a] or [Etter, 1996] and as a response to [Bowden, 1994b], reprinted in this issue.

We review the ideas behind Kron's Method of Tearing and Jessel's Principle of Secondary sources (both special cases of the above theory) and their equivalence. We follow Hiley's argument in [Hiley, 1996] to show how Schrodinger's equation can be thought of as specifying the evolution of (a series of) tearings in continuous space. These can be shown on a commutative diagram as a series of similarity transforms. We compare this with Etter's derivation in [Etter 1996]. We describe briefly a recently published derivation of Maxwell's equations from a noncommutative algebra and show how they fit onto a related commutative diagram. Finally we make some comments on applications of the general theory to computer systems. This paper is a series of vignettes of work in progress. It is designed to point the direction of work to come in Constructive Physics.

KEYWORDS: Kron, Tearing, Diakoptics, Networks, Jessel, secondary sources, Huygens' Principle, holography, information, boundaries, Homology, Physics, Schrodinger's equation

1. INTRODUCTION

In the 1950's Gabriel Kron [1963] introduced his Method of Tearing or Diakoptics. The technique enables the solution of Engineering problems, particularly boundary value problems, by tearing structures topologically into subsystems, solving the subproblem on each subsystem and then recombining into an exact overall solution. It has been recently rediscovered by the parallel processing community as "Domain Decomposition". The problem always took the form of an (electrical) network analogy or discretisation, thus Kron was one of the pioneers of fundamental theory of the Finite Element Method. (Indeed Roth (or Tonti) diagrams are proving to be of basic importance in modern Finite Element theory, eg. see [Bossavit, 1995]. This will be discussed in a forthcoming paper.) Kron's claims for the generality of the Method of Tearing caused some controversy until its proof by the mathematician J. Paul Roth [1955] in the context of Homology Theory or Algebraic Topology. In a recent paper [Bowden, 1990] in this Journal it was shown how Kron's Method is equivalent to Jessel's Principle of Secondary Sources, a reevaluation of Huygens' Principle prompted by the physicist Louis de Broglie. Jessel's Principle states that any wave source distribution (eg charge) can be replaced by a set of secondary sources on any boundary surrounding the original sources, such that from the outside the resulting waveform is indistinguishable from the original. This is the basis of Dennis Gabor's holography.

Both Kron and Jessel saw their work as natural philosophies. They are both information based theories of Engineering. Both are rich, yet different analyses, one in discrete space and one continuous. They are compatible with the hierarchical, structural, systemic definitions of information within physical structures due to Scarrot

"A system is an interdependent assembly of elements and/or systems. Information is that which is exchanged by the components of a system" [Scarrot, 1989]

"The information contained by a system is a function of the linkages binding simpler into more complex units. The universe is organised into a hierarchy of information levels" [Stonier, 1990]

and Stonier. These considerations, in the light of the information based physics of David Bohm and Basil Hiley [1994], and also of new tentative theories of discrete informational physics, such as that of Ted Bastin and Clive Kilmister [1995] (who derive physical structure from the process of observation or innovation) led the author to propose that it would be a major achievement [Bowden, 1994, reprinted in this issue] to relate these hierarchical, holographic, topological ideas of information to the relative, subjective, statistical, and often discrete, traditional concepts of information such as that of Shannon or Parker-Rhodes

"The information content of a statement is the number of YES/NO questions to which the answers are needed in order to identify the statement" [Parker-Rhodes, 198?]

(in true Twenty Questions style!)

In this issue Etter [1996] appears to have achieved this goal as a side effect of his search for a consistent, statistical interpretation of Quantum Mechanics. He shows how topological information can be reconstructed from data about statistical physical variables. In the process he shows how what are usually thought to be two empirical statements in orthodox physics, Schrodinger's equation and Born's rule, are really general theorems about links and tearing.

We have described three informational theories of the physical world, one discrete, one continuous and one statistical. The main aim of this paper is to provide a unified theory, which we call General Physical Systems Theory, as its structure is based on the work of Mesarovic and Takahara, although clearly the consideration of the emergence of topology puts a very different light on its applicability. We find that we can make very general statements about independence, regions, boundaries, decomposition and tearing without any reference to the kind of system to which we are referring.

The second part of the paper follows an argument by Basil Hiley in Hiley [1996] which starts from

Grassman's idea of a simplicial complex as modelling thought processes and ends by deriving Schrodinger's equation (which, of course, is about *observation*) from changes in this structure. We reproduce the results but emphasising our approach, in particular with respect to changes in the structure being due to the evolution of a succession of tearing operations. We use the opportunity to remind the reader of the mathematical structure of Kron's network theory (which indeed is isomorphic to the structure of simplicial complexes), the Method of Tearing, and of Jessel's Principle (which, of course, are two of the important special cases of our general theory).

Finally we note that the lack of commutativity in the variables implied by Schrodinger's equation leads directly, according to Feynman and Dyson [1990] in the continuous case and Kaufinan and Noyes [1996] in the discrete case, to Maxwell's equations. We suggest a new approach to their derivation using general noncommutative definitions of grad, div and curl. These should make clearer the structure of the derivation. The noncommutative field equations can be displayed on an extended Roth's diagram identical to the one in [Bowden, 1990]. These ideas will be investigated in more detail in a forthcoming paper.

2. HUYGENS' PRINCIPLE

A physical (or topological) system $\Sigma=(S,X)$ is a set of variables $X=\{x_1,x_2,...,x_n\}$, where each variable x_i takes its value(s) from a set X_i=span(x_i), where span(x_i) is defined as the set of all possible values of x_i, for all possible values of the other variables, and an "extension" or set of conditions on X which can be written S(X)=0. (Often S(X) will be nonzero valued in our examples below; but there is no inconsistency here as the value of S(X) can always be subtracted from the left hand side and S redefined appropriately.) The x_i may be real valued or discrete (eg, Boolean). They can be static or dynamic (indexed on continuous or discrete time and realised as polynomials in s, the Laplace transform or z^{-1}, the backward shift operator). (In [Bowden 1990] we noted that this could be extended to indexing on (continuous or discrete) space.) And they can be deterministic or stochastic (probabilistic), as in the following examples:

Example Mesarovic and Takahara [1975] consider the general deterministic system $\Sigma_d=(S, X)$ with $S\subset X_1\times X_2\times...\times X_n$, where equality would give the trivial system with all the variables independent. However, in their analysis, they do not consider the emergence of topology.

Example Dempster and Schafer [Sudkamp, 1992] consider the general probabilistic system Σ_p=(S, X), where the variables are propositions with a probability of being true, $p(x_i)\in[0,1]$, the closed unit interval, and $S:X_1\times X_2\times...\times X_n\rightarrow[0,1]$, giving the joint probability $p(x_1,x_2,...,x_n)$ = $p(x_1\&x_2\&...\&x_n)$ $\in$ [0,1]. They define a *basic probability assignment* S over X as a function $p:2^X\rightarrow[0,1]$ that satisfies (i) p(0)=0 and (ii) $\Sigma_{A\subset X}p(A)=1$ (where Σ here is summation). Etter [1996] considers the emergence of topology from a joint probability function; it is on this work that our ideas are based.

Σ_d can be considered as either a special case of Σ or of Σ_p by putting $S:X_1\times X_2\times...\times X_n\rightarrow\{0,1\}\subset[0,1]$ with 1 for $\{x_1, x_2, ..., x_n\}\in S\subset X_1\times X_2\times...\times X_n$ and 0 otherwise. Conversely Σ_p can be considered as a special case of Σ_d by considering the joint probability to be another variable, that is, appending I=[0,1] to the span of the system and putting $S\subset X_1\times X_2\times...\times X_n\times I$.

It is this extension, or set of conditions, that gives rise to the topology of the system as we will show. It should be clear that our general definitions are all categorical statements in that they deal with generic (polymorphic) data types. The crucial concepts are

Independence Two variables are said to be independent if directly changing the value of one has no effect on the value of the other. Another way of saying this is that information from x_i is never received (or, at least, is always ignored) by x_j and vice versa. (Note that this is a stricter definition than that of Etter who defines two variables to be dependent even if they are only both affected by a third, (that is we define all variables to be inputs, see below). This is related to the difference between $\partial x_i/\partial x_j=0$ and $dx_i/dx_j=0$.)

Example Two variables x_1 and x_2 in Σ_p are said to be independent iff $p(x_1\&x_2)=p(x_1)p(x_2)$. A set of variables $x_i, x_2, ..., x_n$ is said to be independent iff $p(x_1\&x_2\&...\&x_n)=p(x_1)p(x_2)...p(x_n)$.

Example Two variables x_1 and x_2 in Σ_d are said to be independent iff $S = span(x_1,x_2) = span(x_1)\times span(x_2) = X_1\times X_2$. A set of variables $x_1, x_2, ..., x_n$ is said to be independent iff $span(x_1,x_2,...,x_n)=span(x_1)\times span(x_2)\times...\times span(x_n)$, where $span(x_i)$ is defined as the set of all possible values of x_i, for all possible values of the other variables. A special case of Σ is the system with $S = X_i \times T$ where $T\subset X_1\times...\times X_i'\times...\times X_n$ and the prime indicates a missing term. In this case the variable x_i is said to be independent of the rest of the system.

Now consider Σ_d as a either a special case of Σ_p by putting $S:X_1\times X_2\times...\times X_n\rightarrow\{0,1\}\subset[0,1]$ with 1 for $\{x_1, x_2, ..., x_n\}\in S\subset X_1\times X_2\times...\times X_n$ and 0 otherwise as described above. Similarly $S:X_i\rightarrow\{0,1\}$ with 1 for $\{x_i\}\in X_i$ and 0 otherwise, thus $span(x_i)=\{x_i, S(x_i)=1\}$. The values 0 and 1 can be considered to be the *unnormalised* probabilities that the variables can take the associated values. Then two variables x_1 and x_2 in Σ_d are independent iff $S(x_1,x_2) = S(x_1)S(x_2) = S(x_1)\&S(x_2)$. We may write $p(x)=S(x)$. A set of variables $x_1, x_2, ..., x_n$ is said to be independent iff $span(x_1,x_2,...,x_n)=span(x_1)\times span(x_2)\times...\times span(x_n)$ or equivalently $S(x_1,x_2,...,x_n) = S(x_1)S(x_2)...S(x_n)$.

Independence of Sets If a pair of subsets, R_1 and R_2, of X are independent we write $\{R_1|R_2\}$. Note that it is not possible to define the independence of such subsets in terms of the independence of their elements as can be seen from the following. Consider the two sets of Boolean variables $A=\{a_1, a_2\}$ and $C=\{c\}$ with $c = a_1 XOR a_2$ (XOR is exclusive-or), then any pair $\{a_1, a_2\}$, $\{a_1, c\}$ or $\{a_2, c\}$ consists of two independent variables and we can write $\{a_1|C\}$ and $\{a_2|C\}$ but *not* $\{A|C\}$.

Example Two subsets $R=\{r_1, ..., r_n\}$ and $S=\{s_1, ..., s_m\}$ of X in Σ_p are said to be independent iff $p(r_1\&...\&r_n\&s_1\&...\&s_m)=p(r_1\&...\&r_n)p(s_1\&...\&s_m)$.

Example Two subsets $R=\{r_1, ..., r_n\}$ and $S=\{s_1, ..., s_m\}$ of X in Σ_d are said to be independent iff $S = span(r_1, ..., r_n, s_1, ..., s_m) = span(r_1, ..., r_n)\times span(s_1, ..., s_m)$.

Separability A variable y (or set of variables B) is said to separate two variables x and z iff fixing the value of y (or B) makes the values of x and z independent. This is Etter separability [Etter 1996].

Regions and boundaries A set of variables, or boundary, $B\subset X$ is said to separate two sets of variables $R_i \subset X$ called regions, iff

$R_1 \cap R_2 = 0$, $R_i \cap B = 0$ and
$R_1 \cup R_2 \cup B = X$ (although we might want to relax this condition)

and B separates every pair *of subsets* $r_1 \subset R_1$ and $r_2 \subset R_2$. If B separates A from C then we write $\{A|B|C\}$. The extension to $\{R_1, R_2, ..., R_m\}$ is obvious; there may be one or many internal boundaries. (An alternative definition, which Kron would have preferred, is

$R_1 \cap R_2 = B$, (Kron called B the *intersection*) and
$R_1 \cup R_2 = X$ (although we might want to relax this condition)

and B separates every pair *of subsets* $r_1 \subset R_1 - B$ and $r_2 \subset R_2 - B$.)

Etter has pointed out that **every B_i forms a basis for the associated R_i** (and maybe the complement of R_i in the Universe) so that given a network or a field equation we can interpret it as a system for which every boundary is a basis at least for its interior. Two questions then arise both for specific systems and in the general systems approach. 1. Given a set of time series data for the values of the variables in such a system can we reconstruct the associated sets of boundaries and regions from the data alone? 2. Given such a collection of boundaries and regions, under what circumstances can we recover the entire connectivity of the system. That is, under what circumstances can we reconstruct the topology from the data? These ideas will be the basis of a future paper.

Subsystems A subsystem, Σ' of Σ is a subset X' of X and an extension S' on X' such that S'=S on X'. Two important subsets of X will be referred to as the input set U and the output set Y. U is the set of variables that we can control and Y is the set of variables that we can observe. U and Y will often be the same as B, the boundary of the system, otherwise they will usually be larger. The Universe is the distinguished system with no system boundary. A physical system is a region of the Universe, ie a subsystem with a boundary B (or, more generally, an input set U and an output set Y) which separates it from the rest of the Universe. If U=Y=B for all subsystems, the system is said to be local.

Decomposition (Tearing) A system is said to be decomposable into a set of regions $\{R_1, R_2, ..., R_m\}$ iff the values of the variables $\{r_1, r_2, ..., r_{n(i)}\}$, for each R_i, can be calculated from the local boundary conditions $B_i \subset B$ and the local extension. R_i is said to be disconnected, or torn, from the rest of the system. Such a system is a **Physical System**. This is a generalisation of Kron's Method of Tearing (or "diakoptics") [Kron, 1963].

Fundamental Theorem of Disconnection x can be disconnected (torn) from z at y iff x and z are separable at y; ditto for regions. That is if fixing the boundary splits the problem then knowing the boundary splits the problem and vice versa. This is Etter's Fundamental Theorem of Disconnection [Etter, 1996]. The hypothesis that it is possible to tear apart a physical system is known as Kron's Principle [Bowden, 1990] or the "Diakoptical Proposition" [Dweck].

Suppose we have a disconnection of (S,X), ie two regions $R_1=(S,\{X_1,B_1\})$, and $R_2=(S,\{X_2,B_2\})$, such that when they are reconnected by setting the values of $B_1=B_2$, they make the values of $X=X_1 \cup X_2$ (as well as the set equality). Then regardless of the values of B_1 and B_2, the values of X_1 and X_2 must be independent, thus fixing $B_1=B_2=B$, a constant, leaves the values of X_1 and X_2 independent, so they are separable by B.

Conversely, suppose X_1 and X_2 are separable by B. Then fixing $B_1=B_2=B$ makes the values of X_1 and X_2 independent. Thus the values of X_1 and X_2 can be calculated from the local boundary conditions only, therefore R_1 can be disconnected from R_2 at B. QED

Etter has pointed out the further theorem that X_i can be torn out at B iff B forms a basis for X_i.

Example The Dean wants to reorganise the Computer Centre into two interacting departments. Most employees lie naturally in one department or the other, but for some the situation is not so clear. How does she simplify her problem? Our theorem says that the problem splits if we know the job specification and conditions of service of the people in the intersection of the two sets. How do we identify these people? They are just the set of people for whom the problem splits if we fix their contracts.

A corollary of Kron's Principle is the

Generalised Huygen's Principle (Jessel's Principle of Secondary Sources) A set of sources (input variables) within {outside} a region of a *Physical* System can be replaced by a new set of sources on the boundary of that region and from the outside {inside} of the region it will not be possible to tell the difference. This is a generalisation of Jessel's formulation of Huygens' Principle [Jessel 1962].

Suppose we have a disconnection or tear, T of $\Sigma = (S,X)$, into two regions $R_1=(S,\{X_1,B_1\})$, and $R_2=(S,\{X_2,B_2\})$, then fixing $B_1=B_2=B$ makes the values of X_1 and X_2 independent. Thus, assuming that R_1 is the inner region and R_2 the outer region with no sources, then X_2 can be calculated from B alone and, furthermore, it must be possible to choose a B such that the values of $X=X_1 \cup X_2$, as well as the set equality. Then we refer to B as the "secondary sources" and, if the original sources are denoted $Y=Y_1 \cup Y_2$, we can calculate B from the commutative diagram

$$\begin{array}{ccc} \{X_1,B_1\}\times\{X_2,B_2\} & \xrightarrow{S} & Y_1\times Y_2 \\ T\uparrow & & \uparrow T' = T+[S,T]S^{-1} \\ X & \xrightarrow[S]{} & Y \end{array}$$

A (special case of) another corollary is

The Generalised Principle of Optimality For a system in which the variables are indexed by time (or any other independent variable), and on which there is a criterion of optimality, then **subregions of an optimal trajectory are optimal**. That is, when calculating the optimal trajectory of a subregion it can be assumed that the rest of the trajectory is optimal; optimal trajectories need be calculated only in terms of the local boundary conditions. This is a generalisation of Bellman's Principle of Optimality.

Black boxes Mesarovic and Takahara consider the system $S \subset U \times Y$ or $S: U \to Y$ where U is the input set and Y the output set. For the sort of treatment we are developing we often won't distinguish between U and Y. If a variable has an impressed value it is an input, if it is left floating it is an output. This kind of system has never been treated in the literature before. We will call it a black box. We will generally use the notation above except that we will introduce a new symbol into the input set $X_i = \{\text{span}(x_i), ?\}$ where the ? means that this variable is left floating, ie it's an output. So a typical input to a system would be the string $\{x_1, x_2, ?, x_4, ?, x_6\}$; the corresponding output would be $\{y_1, y_2, y_3, y_4, y_5, y_6\}$ where $y_i = x_i$ iff $x_i \neq ?$.

The advantage of this notation is that it gives more structure to the system. The question of what is an input and what is an output is now part of the structure of the system itself. Kron always claimed that writing down the equations of a system threw away some of the structure of the system. It is this that we wish to avoid. We can permute the example above in order to partition the input and output strings into input and output sets (confusing jargon!) thus $\{x_1, x_2, x_4, x_6, ?, ?\}$

Topological systems Let $T \subset 2^X$ be the class of subsets of X consisting of all possible disconnected regions (and their boundaries) of X. Then T is a topology on X iff T satisfies the following axioms:

(1) X and {0} belong to T.
(2) The union of any number of sets in T belongs to T.
(3) The intersection of any two sets in T belongs to T.

The members of T are called open sets. (X,T) is a topological space. It would now be natural to look at the Eilenberg Steenrod axioms for a homology theory.

3. PHYSICS

Using Tom Etter's ideas we have shown how topological structure emerges from (in)dependence in the form of conditional probabilities on information streams in a parallel way to that which can be inferred from dependency properties of a deterministic classical black box, and in a language such that the theorems are general enough to apply to both paradigms, thus establishing a strong connection between two different approaches to "information". We are now going to follow through a train of thought given by Basil Hiley in [Hiley, 1996] but with some different emphases. Hiley starts from the "Algebra of *Process*" (my italics) with Grassman's original but forgotten idea (which led him on to define the Grassman algebra) of a process as an entity $[P_1, P_2]$. Grassman argued that mathematics was about thought, not about material reality. Hiley asks "How [does] one thought become another? Is the new thought *independent* of the old or is there some essential *dependence*?" (my italics). Such a process was generalised by Grassman by considering $[P_1, P_2]$ as a 1-simplex and and building simplicial complexes with lines, planes and volumes etc. connecting thoughts and their dependencies. Hiley notes that we can thus construct a complex of "relations of thought, process, activity or movement. The sum total of all such relations constitutes the holomovement." He emphasises *process* (as do Etter and Bastin and Kilmister) but does not mention dependence again. We consider *independence* to be fundamental, but the two approaches have many similarities. Hiley also shows, in a very similar context to [Manthey 1996] in this issue and elsewhere, that Grassman's structure forms a Clifford algebra.

Assume that we have constructed a topological space from dependencies among information streams. Then Etter argues that changes in this space are induced by cutting or tearing the links. Let us start from Kron's idea of tearing and then see if we can reproduce Hiley's results. Into Etter's work we introduce the mathematics of Kron and simplicial complexes; into Hiley's we introduce the concept of tearing or unlinking. We will use the equivalence beween Kron's Method of Tearing and Jessel's Principle of Secondary sources to make the jump from a discrete world to a continuous one, rather glossed over in [Hiley 1996].

Kron always used an electrical analogy when building models. An (electrical) network consists of a one dimensional network (and associated 2 dimensional "meshes" and 0-dimensional "nodes") through whose branches flow currents and across whose branches exist potentials (or voltages) responsible for the flows (or vice versa). The flows and potentials are subject to conservation laws (known as Kirchoff's laws) at nodes and around meshes. The branches consist of impedances governing the dynamic relationships (Ohm's law) between the local branch voltages and currents. Typically these impedances are either linear (resistance), first order integrators (capacitance) or first order differentiators (inductance). Also distributed around the network are a set of voltage and current generators or sources responsible for initiating and/or sustaining the dynamic evolution of the system variables, voltages and currents. So a branch will typically consist of an impedance and a voltage source in series. The system equations can be written and solved in two dual forms

$$
\begin{array}{llll}
V_1=ZJ^1 \text{ or} & & J^1=YV_1 \text{ or} & \\
E_1+e_1=Z(I^1+i^1) & \text{(Ohm)} & I^1+i^1=Y(E_1+e_2) & (1) \\
C'e_1=0 & \text{(Kirchoff)} & A'i^1=0 & (2) \\
i^1=Ci^2 & & e_1=Ae_0 & (3) \\
E_2=C'E_1 & & I^0=A'I^1 &
\end{array}
$$

where i^1 is the vector of *b* branch currents,
and I^1 of corresponding generators,
i^2 of *m* mesh currents,
e_1 of *b* branch voltages,
E_1 of corresponding generators,
e_0 of *n* node-pair voltages
and Z is the (diagonal) matrix of branch impedances,
Y is its reciprocal of branch admittances
and C is a $b \times m$ matrix of incidence numbers from the set {1,-1,0} depending whether the *i*th oriented branch is {positively, negatively, not} incident to the *j*th oriented mesh. In a similar way the A matrix relates mesh and node variables.

The subscripts and superscripts give the dimension of the simplex and define whether it is *covariant* or *contravariant* respectively. The mathematician J. Paul Roth showed how Kron's formulation is isomorphic to a simplicial complex with complex coefficients. In the context in which we are dealing, if covariant vectors represent the dependence between two thoughts then contravariant vectors would represent mappings beween those dependencies. Using the elements of homology theory, or algebraic topology, Roth and Sun Ichi Amari showed the general validity of Kron's Method of Tearing described below.

It can easily be seen that $A'C=0$ and that the following diagram, due to Roth [1955], commutes (gives the same answer whichever way you go round a loop)

$$\begin{array}{ccccccccc} & & & C & & A' & & & \\ 0 & \rightarrow & i^2 & \rightarrow & J^1 & \rightarrow & I^0 & \rightarrow & 0 \\ & & C'ZC\downarrow & & Z\downarrow\uparrow Y & & \uparrow A'YA & & \\ 0 & \leftarrow & E_2 & \leftarrow & V_1 & \leftarrow & e_0 & \leftarrow & 0 \\ & & & C' & & A & & & \end{array}$$

meshes branches nodes

(Actually the commutation is "weak" as $Y \neq C(C'ZC)^{-1}C'$. The mapping Z is Roth's famous "twisted isomorphism". These ideas will be discussed in a forthcoming paper.)

The rows of the diagram are (co)chain complex sequences and the arrows are boundary operators with $A'C=0$ and $C'A=0$ (the boundary of a boundary is zero). The upper sequence relates contravariant variables, ie., currents of successively lower dimension, via boundary operators. The lower sequence relates covariant variables, ie., voltages of increasing dimension via coboundary operators. The vertical mappings (impedances) are isomorphisms. Multiplying (1) through by (2) and substituting (3) gives the branch currents and voltages in terms of the system sources

$$i^1=C(C'ZC)^{-1}C'(E-ZI) \qquad e_1=A(A'YA)^{-1}A'(I-YE) \qquad (4)$$

Application of the method of tearing to an electrical network decomposes (by reordering) the matrix inversions in equations (4) into the form shown below. This is possibly only for sparsely connected networks. For example, for a linear 1-dimensional physical system with four subsystems, the system matrix looks like this (note the new notation!)

$$\begin{bmatrix} Z_1 & & & & C_1 \\ & Z_2 & & & C_2 \\ & & Z_3 & & C_3 \\ & & & Z_4 & C_4 \\ C_1' & C_2' & C_3' & C_4' & Y \end{bmatrix}$$

where the Z_i are the subsystem matrices, Y is the "intersection network" (subsystem boundary) matrix and the C_i are the connection matrices giving the topology of the interconnections between the subsystems and the intersection. As none of the subsystems have any common boundary the rest of the matrix is null. This can be considered to be no more than a particular permutation, or reordering, of the system matrix induced by the tearing operation.

So we can write now write the system equations

$$Zx + Cy = b \qquad (1)$$

$$C'x + Yy = c$$

where Z is the block diagonal system matrix,
Y is the intersection network system matrix,
C' is the partitioned connection matrix $[C_1' \; C_2' \; C_3' \; C_4']$,
x is the partitioned vector of subsystem solutions,
y is the intersection network solution vector,
b is the partitioned vector of subsystem boundary conditions (voltage and current sources)
and
c is the vector of intersection network boundary conditions

The last equation is the system equation of the intersection network which itself is adjacent to all the subsystems, hence the appearance of C_i' in every term. Note that for a two dimensional physical system the subsystem matrices Z_i are themselves of the form

$$\begin{vmatrix} Z & C \\ C' & Y \end{vmatrix}$$

where Z is block diagonal (again we have reused the symbols for different entities in this matrix). This recursive structure extends naturally to multidimensional systems.

Rearranging equations (1) gives the basic equations of "diakoptics" (network tearing) for a linear system

$$y = (Y - C'Z^{-1}C)^{-1}(c - C'Z^{-1}b) \quad (3)$$

which gives the intersection vector and

$$x = Z^{-1}(b - Cy) \qquad (4)$$

which gives the subsystem solutions. Note that the equation for y involves only the inversion of a matrix of the order of Y, which can usually be made quite small, and the inversion of Z, which is block diagonal and thus involves only the inversion of (all) the Z_i. Equation (3) can be thought of as a projection of the boundary conditions onto the intersection network. The equation for the *n* subsystem solutions naturally splits into *n* subsystem equations

$$x_i = Z_i^{-1}(b_i - C_i y) \qquad (5)$$

which as stated depend only on the intersection vector and the local boundary conditions. The intersection vector is given by

$$y = (Y - \Sigma_{j=1}^{n} C_j' Z_j^{-1} C_j)^{-1}(c - \Sigma_{k=1}^{n} C_k' Z_k^{-1} b_k) \quad (6)$$

Note that although all of the vectors and matrices given above are assumed to be real, the theory works over any suitable field eg., complex numbers or polynomials. In particular solution of the equations where the field is polynomials in the Laplace transform s, allows a dynamic analysis, a wave theory rather than a field theory.

In [Bowden, 1990] we showed the equivalence of Kron's Method of Tearing and Jessel's Principle of Secondary Sources. The starting point for Jessel was Huygens' Principle which, after researching Huygens' original work (in the original old Dutch), Jessel reformulated as follows, "The perturbation that goes out (or in) through a closed surface S that contains (excludes) a wave or field source is identical to the perturbation that can be obtained by cutting off the source and replacing it by appropriate sources distributed on the surface S". Consider the wave equation

$$OP\ F = S_{or}$$

where OP is a differential operator such as the Laplacian, but possibly with a time derivative, F is a field distribution in space and S_{or} is a set of (original) sources, (eg., charges). It should be noted that this is the continuous (distributed) version of the basic system equation that diakoptics tries to solve.

Kron's aim was a fast solution of the equation

$$F = OP^{-1}S_{or}$$

Jessel was more interested in changing the form of the field F by modifying the sources. He considered the identity

$$OP(sF) = sOPF + [OP,s]F$$

where the last term is the Lie bracket (OPs-sOP)F. The scalar space function s is a field modifier. For instance it can be used to define a Huygens' surface, S of secondary sources by setting its value to 1 outside the surface and 0 inside the surface. Its value on the surface is undefined but in general may be complex (ie., complicated). As s is a constant scalar everywhere else the Lie bracket vanishes everywhere but on the surface. The first term $sOPF = sS_{or}$ simply defines any sources that are not inside the surface. Thus the Lie bracket gives the secondary sources required to reproduce that element of the field not produced by any sources external to the Huygens' surface and we write

$$S^s_{or} = [OP,s]F$$

and $\quad OP(sF) = sS_{or} + S^s_{or}.$

If the sources are all inside the surface then the first term vanishes because either s or S_{or} is always zero. This is Huygens' Principle.

Jessel and Resconi claim that a change in the field F due to the modifier s induces a corresponding "anticausal" change s' in the sources. They show this on the commutative diagram

$$\begin{array}{ccc} & OP & \\ sF & \rightarrow & sS_{or}+S^{s}_{or} \\ s\uparrow & & \uparrow s'=s+[OP,s]OP^{-1} \\ F & \rightarrow & S_{or} \\ & OP & \end{array}$$

and $s' = OPsOP^{-1} = s+[OP,s]OP^{-1}$.

They call a diagram of this form an ELS or Elementary Logical System. s and s' are said to be similar when there exists a 1:1 relation OP such that the diagram commutes. For instance if OP is a matrix it must not be singular. The domain and codomain of s (and s') must be of the same type. Bertrand Russell [1919] saw similarity as a very basic concept and noted that "when two relations are similar, they share all properties that do not depend on the actual terms in their fields... Even statements involving the actual terms of the field of a relation, though they may not be true as they stand when applied to a similar relation, will always be capable of translation into statements that are analogous." Thus similarity is not unlike isomorphism which in turn is a special case of exactness.

Example Consider the electrostatic equation

$$\mathrm{div}(D) = \rho$$

where the charges, ρ are contained within a surface defined by s then the secondary sources to reproduce the field outside the surface are given by

$$\begin{aligned} S^{s}_{or} &= [\mathrm{div},s]D \\ &= \mathrm{div}(sD) - s\,\mathrm{div}D \\ &= D.\mathrm{div}(s) \end{aligned}$$

and $sD = \mathrm{div}^{-1}(\mathrm{div}^{-1}\rho).\mathrm{div}(s)$.

or

$$\begin{array}{ccc} & \mathrm{div} & \\ sD & \rightarrow & D.\mathrm{div}(s) \\ s\uparrow & & \uparrow \mathrm{div}^{-1}(.).\mathrm{div}(s) \\ D & \rightarrow & \rho \\ & \mathrm{div} & \end{array}$$

which gives the field in the region outside the surface in terms of secondary sources on the surface which in their turn are calculated from the original sources inside the surface.

Note how similar this is to Kron's procedure. We claim in [Bowden 1990] that the intersection network of Kron's Method of Tearing is in fact a Huygens' surface. The ELS for Kron's Method looks like this

$$x \xrightarrow{Z} b-Cy$$

$$[I\ 0]\uparrow \qquad \uparrow [I\ 0]-C(C'Z^{-1}C-Y)^{-1}[C'Z^{-1}-I]$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \xrightarrow[\begin{bmatrix} Z & C \\ C' & Y \end{bmatrix}]{} \begin{bmatrix} b \\ c \end{bmatrix}$$

If it is required only to solve the *i*th system then s will take the special form

[0 ... 0 I 0 ... 0]

where the I is in the *i*th position. The intersection vector is a holographic field containing all the information from the boundary conditions. The holographic form is easier to visualise if a dynamic system is considered.

Sometimes we may be interested in applying a succession of tearing operations. For instance, the "staircase method" tears out a subsystem at a time, (similarly hierarchical tearing [Bowden, 1994] tears the system into subsystems, then subsubsystems, etc.) Considering diakoptics to be a reordering of the system matrix induced by the tearing operation, where P is a reordering operator, usually a permutation matrix, leads us to see more clearly the relationship to Jessel's Principle of Secondary sources. Such a succession of transformations was referred to by David Bohm as an "ordering or enfolding", and by Jessel and Resconi as an Logical System and was represented by them on a commutative diagram. A Logical System is a sequence of ELS as shown below

$$F_0 \xrightarrow{OP_1} F_1 \xrightarrow{OP_2} F_2 \ldots F_n$$

$$s_0\uparrow \quad s_1\uparrow \quad s_2\uparrow \quad s_n\uparrow$$

$$F_0 \xrightarrow[OP_1]{} F_1 \xrightarrow[OP_2]{} F_2 \ldots F_n$$

Thus s_0 is similar to s_1 and hence to s_2 and so on. In the example above due to Jessel s is a scalar with value either 0 or 1. If these are thought of as probabilities then s can be considered to be a (probability) density matrix. For convenience we will now switch to the notation in [Etter, 1996]. Assuming $OP_i = T$ is constant and writing s_0=S, s_1 = S', s_2 = S'', etc. then

$$\begin{array}{ccccc} & T & & T & \\ F_0 & \rightarrow & F_1 & \rightarrow & F_2 \ldots \end{array}$$

$$S\uparrow \quad S'\uparrow \quad S''\uparrow$$

$$\begin{array}{ccccc} F_0 & \rightarrow & F_1 & \rightarrow & F_2 \ldots \\ & T & & T & \end{array}$$

commutes and the diagram represents a sequence of "anticausal" changes and we have

$$S' = TST^{-1} = S + [T,S]T^{-1}$$

etc. where, as mentioned above, S can be thought of as a density operator. Now assuming T=exp(iHt/h) where t is an "enfolding parameter" representing successive steps, h is a scaling constant and (remember T can be complex) H is an operator then for small t we can write

$$S' = (1 - iHt/h)\ S\ (1 + iHt/h)$$

so that

$$ih(S' - S)/t = HS - SH = [H, S].$$

So that in the limit as $t \rightarrow 0$ we have

$$ihdS/dt = [H, S]$$

which Hiley notes has the same form as the Heisenberg equation of motion. As S is a density matrix we can assume that it is factorisable in the form $S = \Psi\Psi^*$ and then we have

$$ihd\Psi\Psi^*/dt = ih(d\Psi/dt)\Psi^* + ih\Psi d\Psi^*/dt = (H\Psi)\Psi^* + \Psi(\Psi^*H)$$

which splits naturally into two equivalent formulae thus

$$ihd\Psi/dt = H\Psi \qquad \text{and} \qquad ihd\Psi^*/dt = -\Psi^*H$$

(postmultiply the first equation by Ψ^* and premultiply the second by Ψ and add. This ends Hiley's argument. Putting the Hamiltonian $H=V-\nabla^2h^2/2m$ gives the usual form of the Schrodinger equation, $ihd\Psi/dt = (V-\nabla^2h^2/2m)\Psi$.)

In detail, the commutative diagram for this Logical System is

$$\begin{array}{ccccc} & e^{iHt/h} & & e^{iHt/h} & \\ F_1 & \rightarrow & F_2 & \rightarrow & F_3 \ldots \end{array}$$

$\Psi\Psi^*\uparrow \quad \Psi'\Psi'^*\uparrow \quad \Psi''\Psi''^*\uparrow$

$$\begin{array}{ccccc} F_1 & \rightarrow & F_2 & \rightarrow & F_3 \ldots \\ & e^{iHt/h} & & e^{iHt/h} & \end{array}$$

Finally we look briefly at Kauffman and Noyes derivation of Maxwell's equations. In quantum mechanics it is usual to define position and momentum operators R_i and P_j such that for any location and momentum vectors **r** and **p** of a particle

$$\langle \mathbf{r} \mid R_i \mid \Psi\rangle = r_i\langle \mathbf{r} \mid \Psi\rangle \quad \text{and} \quad \langle \mathbf{p} \mid P_i \mid \Psi\rangle = p_i\langle \mathbf{p} \mid \Psi\rangle \quad i=1,2,3$$

where R_1, R_2, R_3 and P_1, P_2, P_3 designate respectively position and momentum operators X, Y, Z and P_x, P_y, P_z respectively. Then it is easy to show (eg., [Cohen-Tannoudji, 1977] p.150) the canonical commutation relations

$$[R_i, R_j] = 0,$$
$$[P_i, P_j] = 0,$$
and $[R_i, P_j] = ih\delta_{ij}$

for i,j = 1, 2, 3. In particular $R_iP_i-P_iR_i = ih$; all other variables commute. (For example in one dimension R and P are defined by the equations

$$R\Psi=x\Psi \quad \text{and} \quad P\Psi = (h/i)\partial\Psi/\partial t.$$

Thus

$$(RP-PR)\Psi = x(h/i)\partial\Psi/\partial x - (h/i)(\partial/\partial x)(x\Psi) = (ih)\Psi.$$

Hence RP - PR = ih.)

Following Dyson [1990], after an idea by Feynman, Kauffman and Noyes [1995] consider a point in a discrete space whose position and velocity are governed by the analogous equations

$$[X_i, X_j] = 0,$$
and $[X_i, X_j^t] = \kappa\delta_{ij}$ for i,j = 1, 2, 3

where X^t=dX/dt is defined as J(X'-X) where the operator J, whose sole purpose is to keep track of the time shifting, is defined thus

$$J' = J,$$
$$AJ = JA' \text{ for all } A$$

and XdX/dt becomes X'(X'-X) as we have to take a time step to make the differential. The commutator $[X, X^t]$ is clearly nonzero. They then develop the Discrete Ordered Calculus (DOC) and show that

$$d(XY)/dt = (dX/dt)Y + XdY/dt$$

as is appropriate. They define

$$E_j = d^2X_j/dt^2 - \varepsilon_{ijl}(dX_k/dt)H_l$$

(by analogy with $E = F - v \times H$) where

$$H_l = (1/2\kappa)\,\varepsilon_{ijl}[dX_j/dt, dX_k/dt]$$

and prove after a lengthy calculation that

$$\text{div}\,H = 0 \quad \text{and} \quad \partial H/\partial t + \nabla \times E = 0.$$

Following Dyson they define

$$\text{div}\,E = \rho \quad \text{and} \quad \partial E/\partial t - \nabla \times H = j$$

to complete Maxwell's equations. They note that from (1) and $\partial X_i/\partial X_j = \delta_{ij}$ that we can write

$$\partial X_i/\partial X_j = [X_i, X_j^t]/\kappa$$

which allows us to define

$$\partial G/\partial X_j = [G, X_j^t]/\kappa$$

To simplify and clarify their derivation, introduce div, grad and curl with noncummutative coefficients

$$\nabla G = [\ldots \partial G/\partial X_i \ldots] = [\ldots [G, X_i^t]/\kappa \ldots] = (GX^t - X^tG)/\kappa \qquad \text{(G scalar)}$$

$$\nabla.G = \partial G_i/\partial X_i = [G_i, X_i^t]/\kappa = (G.X^t - X^t.G)/\kappa$$

$$\nabla \times G = [\ldots \varepsilon_{ijk}\partial G_i/\partial X_j \ldots] = [\ldots \varepsilon_{ijk}[G_i, X_j^t]/\kappa \ldots] = -(G \times X^t + X^t \times G)/\kappa$$

with $A \times B = \varepsilon_{ijk}A_iB_j$. Note of course that $A \times B \neq -B \times A$ and thus that $A\times A \neq 0$! Thus we can write Kauffman and Noyes' definitions E and H,

$$E = X^{tt} + X^t \times H \text{ and } H = \nabla\times X^t/2\kappa$$

and $\partial H/\partial t = dH/dt + (\partial X_j/\partial t)\,\partial H/\partial X_j$

$$= d(X^t \times X^t)/\kappa dt - (X^t.\nabla)H$$

$$= (X^{tt} \times X^{t} + X^{t} \times X^{tt})/\kappa - (X^{t}.\nabla)H$$

$$= -\nabla \times X^{tt} - (X^{t}.\nabla)H$$

$$= -\nabla \times (E + X^{t} \times H) - (X^{t}.\nabla)H$$

$$= -\nabla \times E - X^{t}(\nabla.H) + H(\nabla.X^{t}) - (H.\nabla)X^{t} + (X^{t}.\nabla)H - (X^{t}.\nabla)H$$

$$= -\nabla \times E$$

where we have omitted the detailed calculations. These will be discused in a forthcoming paper. Mike Peskin has pointed out that if the J operator is considered as a time step operator of the form exp(iHt) then from AJ=JA' we recover ELS like structures as shown above.

The structure of Maxwell's equations can be shown on an extended Roth diagram [Roth, 1955 or Bowden, 1990] or Tonti diagram [Tonti, 1976 or Bossavit, 1992] of the form

$$\begin{array}{ccccccc}
 & \text{grad} & & \text{div} & & \text{curl} & \\
\psi & \rightarrow & M+H & \rightarrow & J+sD & \rightarrow & . \\
\downarrow & & \mu s \downarrow\uparrow & & \downarrow\uparrow \sigma + \varepsilon s & & \uparrow \\
. & \leftarrow & s(\mu M+B) & \leftarrow & E+sA & \leftarrow & -\phi \\
 & \text{div} & & \text{curl} & & \text{grad} &
\end{array}$$

on which the internal structure of the twisted isomorphism can be shown as in [Bowden, 1990]. This is a special case of the general form

$$\begin{array}{cccccccccccc}
 & & \partial & & \partial & & \partial & & \partial & & & \\
\ldots & 4 & \rightarrow & 3 & \rightarrow & 2 & \rightarrow & 1 & \rightarrow & 0 & & \text{(dimension)} \\
 & OP\downarrow\uparrow & & OP\downarrow\uparrow & & OP\downarrow\uparrow & & OP\downarrow\uparrow & & OP\downarrow\uparrow & & \\
\ldots & 4 & \leftarrow & 3 & \leftarrow & 2 & \leftarrow & 1 & \leftarrow & 0 & & \text{(dimension)} \\
 & & \delta & & \delta & & \delta & & \delta & & &
\end{array}$$

showing the dimensional structure of the Logical Systems described above. This structure is becoming more important in practical electrical field analysis [Bossavit, 1992 and 1996]. Indeed Bossavit asks "When can we discretise a Tonti diagram?" [Bossavit 1992].

4. COMPUTERS AND CONCLUSIONS

The title of this paper is that given by Jessel to the book he and I were to write together, before his untimely death in 1992. It was hoped to include some application of our general theory to computing in this paper, but this is still proving notoriously elusive. It is hoped that we will be forgiven for the title out of respect for Jessel. The problem essentially involves applying the theory over the group $Z_2=\{0,1\}$, rather than the reals or the integers. We have given some examples in the text, in particular the one relating to exclusive OR, showing that independence does not just depend on pairwise application, but in general we have found it hard to get useful results. It is believed that this new theory should help. Dubois [1996] describes the problem in his paper in this issue. It may be that Pawel Siwak's [1996] work also points the way. Recent work with Mike Manthey and Clive Kilmister on Manthey's Topsy, which is naturally defined over $Z_3=\{-1,0,1\}$ is looking more promising. It is hoped to publish some results in the near future.

The frequency of new publications on derivations of physical laws is remarkable. One physicist remarked that "Derivations of Schrodinger's equation are ten a penny!" This in no way detracts from the value of Etter's work which can be thought of as having an underlying (statistical) ontological picture contributing to the physical understanding of our world (such as why we add amplitudes instead of probabilities in Quantum Theory). The degree to which Etter's picture corresponds to that which we have developed in the past, relating the work of Jessel and Kron is remarkable. This has allowed the beginnings of the creation of a "covering theory" which is general enough to allow description of both the classical and Quantum worlds. This has been one of the major problems with the orthodox theory and is known as the Measurement Paradox. The value of these ideas for General Systems Theory itself should also not be underestimated. A general algorithm for recovering topological structure from unstructured data is implied. We hope to develop all these ideas much further in future publications.

ACKNOWLEDGEMENTS

Thanks are due to John Amson, Ted Bastin, Tom Etter, Basil Hiley, Clive Kilmister and, of course, to the late Maurice Jessel.

REFERENCES

Bastin, T. and C. W. Kilmister [1995], *Combinatorial Physics*. World Scientific, London.

Bohm, D. and B. J. Hiley [1993], *The Undivided Universe*. Routledge, London.

Bossavit A. [1992], "A New Viewpoint on Mixed Elements", *Meccanica*, 27, pp. 3-11.

Bossavit A. [1996], "Edge Elements for Magnetostatics", *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, **9**, pp. 19-34.

Bowden K. [1990], "On General Physical Systems Theories", *International Journal of General Systems*, **18**, pp. 61-79.

Bowden K. [1994], "Hierarchical Tearing: An Efficient Holographic Algorithm for System Decomposition", *International Journal of General Systems*, **23**(1), pp. 23-37.

Bowden K. [1996], "What is to be Done", *Proceedings of the 17th International ANPA Conference*, University of Cambridge, September 1995.

Cohen-Tannoudji C. et al [1977], *Quantum Mechanics Vol. 1*. Wiley, International and Hermann, Paris.

Dweck E., Editorials, *Matrix and Tensor Quarterly*.

Etter T. [1996], "Process, System, Causality and Quantum Mechanics", *International Journal of General Systems, this issue.*

Hiley B. J. [1996], "The Algebra of Process", *Conciousness at the Crossroads, Proceedings of the Tempus Project 'Phenomenology and Cognitive Science', Maribor, Slovenia, 23-7 August 1994,* pp. 52-67.

Jessel M. [1962], *Contribution aux Theories du Principe Huygens et de la Diffraction*. Thesis for Doctorate of Physical Sciences, University of Paris.

Kauffman L. H. and H. Pierre Noyes [1996], "Discrete Physics and the Derivation of Electromagnetism from the Formalism of Quantum Mechanics", *Proceedings of the Royal Society of London,* 452, pp 81-95.

Kron G. [1963], *Diakoptics: The Piecewise Solution of Large Scale Systems*. MacDonald, London.

Mesarovic M. D. and Yasuhiko Takahara [1975], *General Systems Theory: Mathematical Foundations*. Academic Press, New York.

Parker-Rhodes F. [198?], *The Inevitable Universe*, unpublished.

Roth J. P. [1955], "The Validity of Kron's Method of Tearing", *Proceedings of the National Academy of Sciences*, **41**.

Russel B. [1919], *Introduction to Mathematical Philosophy*. Allen and Unwin, London.

Sudkamp T. [1992], "The Semantics of Plausibility and Possibility." *International Journal of General Systems*, **21**, pp. 273-289.

Tonti E. [1976], "The Reason for Analogies between Physical Theories", *Appl. Math. Modelling*, **1**, pp. 37-50.

18th Annual International Meeting of the Alternative Natural Philosophy Association

5-8 September 1996, Cambridge, U.K.

THE ROTATING PLATFORM PROBLEM

F. Selleri

Università di Bari - Dipart. di Fisica
INFN - Sezione di Bari
Via Amendola 173
I-70126 Bari

Abstract. The large lifetime increase of muons observed in a storage ring experiment does not depend on muon acceleration but only on muon speed on the circular orbit. We assume this to be a manifestation of a general principle: only velocities matter for length contraction and time dilation phenomena. Therefore local measurements of space, time and velocity in an accelerated reference frame must give the same result as those performed in a "locally equivalent" inertial frame having the same instantaneous velocity. This principle is used for deducing the velocity of light both along the rim of a rotating platform and relative to inertial reference frames, and leads to noninvariant values. A set of transformations of space and time different from the Lorentz transformations is deduced and shown to have satisfactory composition properties. They do not form a group but multiply in a selective way, similar to that of time evolution operators.

1. TIME TRANSFORMATION TO A ROTATING DISK

Lifetimes of positive and negative muons were measured in the CERN Storage-Ring experiment [1] for muon speed $0.9994\,c$, corresponding to a γ factor of 29.33. Muons circulated in a 14 m diameter ring with an acceleration of 10^{18}g. Excellent agreement was found with the formula

$$\tau_0 = \frac{\tau_{rest}}{\sqrt{1-\beta^2}} \tag{1}$$

where τ_0 is the observed muon lifetime, τ_{rest} is the lifetime of muons at rest, and $\beta = v/c$, v being the laboratory speed of the muon on its circular orbit.

Consider an ideal platform rotating with the same angular velocity as the muon in the e.m. field (with respect to such a platform the muon is at rest). Consider also four different observers:

1. O_0 is the observer at rest in the laboratory reference frame S_0, assumed to be an inertial frame. Thus O_0 could be the CERN storage ring experimenter.

2. $\tilde{O}_a$ is an observer standing on the platform. He knows everything about the platform (the accelerated frame S_a) through which he can move freely;

3. O_a is a second observer on the platform S_a. He stands very near the muon which, to him, appears to be at rest; O_a has a local knowledge of the platform and of its physical properties extending only to the immediate surroundings of his position, including of course the muon;

4. O_T observes from an inertial frame S_T, with respect to which O_a is instantaneously at rest. S_T will be called the "tangent" inertial frame.

We next consider the muon lifetime according to each of these four observers.

According to O_0 the muon lifetime τ_0 is greatly enhanced with respect to that (τ_{rest}) of muons at rest in S_0. His observations are expressed by Eq. (1). According to the accelerated observer $\tilde{O}_a$ the clocks on the platform have a pace dependent on their distance from the centre, the fastest one being at the centre; in agreement with the equivalence principle he attributes this phenomenon to the presence of a position-dependent radial gravitational field of cosmic origin. He can check that the lifetime of muons placed near the rim of the platform is either τ_0 or τ_{rest} depending on the clock chosen (in the centre or near the rim, respectively) for measuring it. According to O_a, who knows only the time shown by his local clock, the muon lifetime is τ_{rest}. Of course O_a is under the action of a large acceleration (10^{18}g), which he detects as a radial gravitational field, but nevertheless his lifetime measurements give just τ_{rest}, as for muons at rest in S_0 observed by O_0. The value τ_{rest} found by O_a is explained by $\tilde{O}_a$ to be a consequence of a cosmic gravitational field which delays in the same way muon decay and the clock used by O_a for lifetime measurements. Finally, according to the observer O_T belonging to the tangent inertial frame S_T the lifetime is τ_{rest} (measured of course on muons at rest in his frame). Obviously the agreement between O_a and O_T also means that the paces of their clocks are observed to be identical during the very short time interval in which they are at rest with respect to one another.

The first conclusion to be drawn concerns the transformation of time given by Eq. (1) which can be extended to all kinds of processes. The laboratory time interval Δt_0 between any two events taking place at a given spot on the rotating disk (muon injection and decay, in the last example) is seen dilated by the well-known factor, compared with the corresponding time interval Δt measured by the accelerated observer O_a:

$$\Delta t_0 = \frac{\Delta t}{\sqrt{1 - \beta^2}} \tag{2}$$

For example, the pace of a clock C_a placed on the rim of the disk would be slowed down according to Eq. (2), if compared with clocks at rest in S_0. And a light signal propagating along the rim (e.g., on the internal surface of a cylindrical mirror) from a point P near C_a and back to P on a circular path, would require a time interval measurable both by the clock C_a itself and by the laboratory clocks, thus again satisfying Eq. (2). Notice that Eq. (2) holds unmodified, irrespective of muon position on the circular orbit: in other words time is transformed from the laboratory to the rim of the disk without a space-dependent time-lag term.

The second conclusion is that the observers O_a and O_T agree on the results of measurements of space and time, for example on the decay rate of muons at rest, even though O_a feels the presence of a radial gravitational field (of cosmic origin) while O_T does not. Of course this conclusion is not new, even though it was never used systematically. Already Einstein in his 1911 article [2] assumed that the acceleration of a clock C_a relative to an inertial system has no influence on the rate of C_a, and that the increase in proper time of C_a at any time is the same as that of the standard clocks of the inertial system in which C_a is momentarily at rest. A similar conclusion is assumed to hold for length: a measuring rod placed parallel to the circumference of the disk undergoes a contraction dependent only on the rotational velocity and indifferent to the centrifugal acceleration. Therefore its length is the same as that of another rod at rest in S_T set parallel to the velocity of S_T with respect to S_0. The identical conclusions of O_a and O_T concerning their measurements of time and space imply that the one-way speed of light found locally along the rim of the rotating disk should be the same as that observed in the "tangent" inertial frame. But in Einstein's theory of special relativity the latter speed is always c and we are so brought to conclude that the speed of light along the rim of the rotating disk should also be c. However this conclusion is not acceptable, as shown in the third section.

2. IS CLOCK SYNCHRONISATION CONVENTIONAL?

The one-way velocity of light in moving inertial systems (e.g., on Earth) has never been measured. Often people stress that in order to measure it one needs synchronised clocks, but that in order to synchronise clocks one must know the one-way velocity of light, so that the logical situation becomes circular. All the laboratory experiments (from Fizeau, Foucault, Michelson, to the recent ones) measured instead the two-way velocity of light. Since such measurements are obviously possible with just one clock, the synchronisation problem did not arise. When Einstein [3] formulated his theory of relativity he postulated that the velocity of light has always the same value c.

There seems to be agreement among several physicists on the conclusion that the constancy of the one-way velocity of light is a useful convention and that it must not be considered as something dictated by objective properties of the natural world. Mansouri and Sexl stated [4]: "When clocks are synchronised according to the Einstein procedure the equality of the velocity of light in two opposite directions is trivial and cannot be the subject of an experiment." On the conventionality of the velocity of light agreed, among others, Grünbaum [5], Jammer [6], Sjödin [7], Cavalleri [8], Ungar [9], and Vetharaniam and Stedman [10]. The idea is however rather old.

Already in 1898 Poincaré [11], discussing the independence of the speed of light of its direction of propagation stated: "This is a postulate without which it would be impossible to start any measurement of this velocity. It will always be impossible to verify experimentally the said postulate." Similarly in 1916 Einstein [12] wrote: "that light requires the same time to traverse the path AM ... as the path BM [M being the midpoint of the line AB] is in reality *neither a supposition nor a hypothesis* about the physical nature of light, but a *stipulation* which I can make of my own free will." In 1925 Reichenbach [13] considered the following situation: in an inertial system S a flash of light leaves point A at time t_1, is reflected back in point B at time t_2, and arrives again in A at time t_3. The problem is how to synchronise the clock near point B with the clock near point A. In the theory of relativity it is assumed that the one-way velocity of light has the same value from A to B as from B to A, so that $t_3 - t_2 = t_2 - t_1$, whence the clock-B time t_2, can be written in terms of the two clock-A times t_1 and t_3 as follows:

$$t_2 = t_1 + \tfrac{1}{2}(t_3 - t_1) \tag{3}$$

Reichenbach commented [13]: *"This definition is essential for the special theory of relativity, but it is not epistemologically necessary. If we were to follow an arbitrary rule restricted only to the form*

$$t_2 = t_1 + \varepsilon(t_3 - t_1)\,, \qquad 0 < \varepsilon < 1 \tag{4}$$

it would likewise be adequate and could not be called false. If the special theory of relativity prefers the first definition, i.e., sets ε equal to $1/2$, it does so on the ground that this definition leads to simpler relations."

Clearly, different values of ε correspond to different values of the one-way speed of light. In fact, one can write

$$t_2 - t_1 = \frac{L}{\tilde{c}(\theta)} \qquad \text{and} \qquad t_3 - t_2 = \frac{L}{\tilde{c}(\theta+\pi)} \tag{5}$$

where L is the $A-B$ distance, $\tilde{c}(\theta)$ is the one-way velocity of light from A to B, θ being the angle between the light propagation direction and the absolute velocity of the considered inertial frame. Similarly $\tilde{c}(\theta+\pi)$ is the one-way velocity from B to A. By adding the previous relations one gets

$$t_3 - t_1 = \frac{L}{\tilde{c}(\theta)} + \frac{L}{\tilde{c}(\theta+\pi)} = \frac{2L}{c} \tag{6}$$

the last step being a necessary assumption, because the two-way velocity of light has been measured with great precision and always found to be c. From (4), (5) and (6) one easily gets

$$\varepsilon = \frac{t_2 - t_1}{t_3 - t_1} = \frac{c}{2\tilde{c}(\theta)} \tag{7}$$

Therefore freedom of choice of ε would mean freedom of choice of the one-way velocity of light.

In the following section we will show that the one-way speed of light is not really a conventional feature of relativistic physics because there is only one way of choosing $\tilde{c}(\theta)$ which leads to a consistent theory allowing one to pass smoothly from an accelerated reference frame to an inertial reference frame, when they are istantaneously at rest with respect to one another. This choice is necessarily different from the standard assumption $\tilde{c}(\theta) = c$ because the latter leads to unsurmontable difficulties, as will be shown next.

3. TIME MEASUREMENTS ON THE ROTATING PLATFORM

The laboratory is assumed to be an inertial frame in which clocks have been synchronised with the standard relativistic method, that is by assuming that the one-way velocity of light is c in all directions.

We consider a clock C_Σ that is on the rim of a uniformly rotating platform having radius R and angular velocity ω. We assume this clock to be set as follows: When a clock of the laboratory momentarily very near to C_Σ shows time $t_0 = 0$ then also C_Σ is set at time $t = 0$. Given that C_Σ keeps a constant position on the platform we can apply Eq. (2) with $\Delta t_0 = t_0 - 0 = t_0$ and $\Delta t = t - 0 = t$ and conclude that any observer at rest in the laboratory near the rim of the platform whose clock marks the time t_0 will see the clock C_Σ passing by in that very moment marking the time

$$t_0 = \frac{t}{\sqrt{1-\beta^2}} \tag{8}$$

with $\beta = \omega R / c$.

On the rim of the platform besides the clock C_Σ there is also a light source Σ placed in a fixed position very near C_Σ. Two light flashes leave Σ at time t_1 of C_Σ.

The description of these light flashes circulating along the rim of the disk by the laboratory observers will be the following: Two light flashes leave Σ at time t_{01}. The first one propagates circularly in the sense opposite to the platform rotation and comes back to Σ after a 2π rotation at time t_{02}. The second one propagates circularly in the same rotational sense of the platform and comes back to Σ after a 2π rotation at time t_{03}. These laboratory times, all relative to events taking place in a fixed point of the platform very near C_Σ, are related to the corresponding platform times by

$$t_{0i} = \frac{t_i}{\sqrt{1-\beta^2}} \qquad (i = 1,2,3) \tag{9}$$

as a consequence of (8).

If L_0 is the disk circumference length measured in the laboratory, light propagating in the direction opposite to the disk rotation must cover a distance smaller than L_0 by a quantity $x = \omega R(t_{02} - t_{01})$ equalling the shift of Σ during the time $t_{02} - t_{01}$ taken by light to reach Σ. Therefore

$$L_0 - x = c(t_{02} - t_{01}) \quad ; \quad x = \omega R(t_{02} - t_{01}) \tag{10}$$

From these equations one gets:

$$t_{02} - t_{01} = \frac{L_0}{c(1+\beta)} \tag{11}$$

Light propagating in the rotational direction of the disk, instead, must cover a distance larger than the disk circumference length L_0 by a quantity $y = \omega R(t_{03} - t_{01})$ equalling the shift of Σ during the time $t_{03} - t_{01}$ taken by light to reach Σ. Therefore

$$L_0 + y = c(t_{03} - t_{01}) \quad ; \quad y = \omega R(t_{03} - t_{01}) \tag{12}$$

One now gets

$$t_{03} - t_{01} = \frac{L_0}{c(1-\beta)} \tag{13}$$

By taking the difference between (13) and (11) we see that the time delay between the arrivals of the two light flashes back in Σ is observed in the laboratory to be

$$t_{03} - t_{02} = \frac{2L_0\beta}{c(1-\beta^2)} \tag{14}$$

By the way, (14) is the well known time delay for the Sagnac effect calculated in the laboratory, but this is uninportant for our present purposes. We show next that these relations fix the time intervals on the disk. In fact (9) applied to (11) and (13) gives

$$t_2 - t_1 = \frac{L_0\sqrt{1-\beta^2}}{c(1+\beta)} \quad ; \quad t_3 - t_1 = \frac{L_0\sqrt{1-\beta^2}}{c(1-\beta)} \tag{15}$$

Notice that nothing was assumed, concerning length, in deducing (15). Therefore the result

$$\frac{t_3 - t_1}{t_2 - t_1} = \frac{1+\beta}{1-\beta}$$

depends only on the transformation of time and can safely be extended to the tangent frame. Of course one expects that

$$L_0 = L\sqrt{1-\beta^2} \tag{16}$$

because the rotating disk circumference length should appear to be contracted in the laboratory. Therefore the velocities of the two light flashes relative to the disk are measured to be

$$\tilde{c}(\pi) = \frac{L}{t_2 - t_1} = \frac{c}{1-\beta} \tag{17}$$

for the propagation in the direction of the disk rotation, and

$$\tilde{c}(0) = \frac{L}{t_3 - t_1} = \frac{c}{1+\beta} \tag{18}$$

for the propagation in the opposite direction. Eq.s (17)-(18) are particular cases of the formula $\tilde{c}(\theta) = c/(1+\beta\cos\theta)$ discussed at length in Ref. [14] and shown to be compatible with the experimental evidence at the special relativistic level (no accelerations). These results will be applied in the next section to inertial frames by using the continuity principle between the rim of the platform and the "tangent" inertial frame.

Starting from (17) and (18) it is a simple matter to show that

$$\frac{L}{\tilde{c}(0)} + \frac{L}{\tilde{c}(\pi)} = \frac{2L}{c} \tag{19}$$

which can be read as follows: the time required to a light signal for propagating along the whole disk circumference plus the time needed to go back (e.g., after reflection on a mirror) equals twice the circumference length divided by c. In other words the two-way velocity of light along the rim is just c. A comment about Eq. (19) is in order. Every line element of the circumference of the rotating disk is instantaneously at rest in some inertial reference frame (the "tangent" frame). Since in inertial frames the two-way velocity of light is experimentally well established to be c the local equivalence with the accelerated frame requires the validity of (19). Nothing different would be acceptable. It is therefore satisfactory that (19) is a consequence of our previous considerations.

Notice also that Eq. (9) applied to (14) gives

$$t_3 - t_2 = \frac{2L\beta}{c} \tag{20}$$

where also (16) was used. This is again the Sagnac formula, this time calculated on the rim of the disk. Thus L is the circumference length as obtained by measurements made with rods placed along the rim of the disk, and $t_3 - t_2$ is the time delay measured with the clock fixed on the rim of the disk.

If instead of (17)-(18) one had $\tilde{c}(0) = \tilde{c}(\pi) = c$ it would obviously follow that the two light pulses require the same time to go around the disk in opposite directions, or that $t_2 = t_3$. The prediction for the Sagnac effect would become $t_3 - t_2 = 0$, totally unacceptable, not only because it is incompatible with the experimentally certain existence of the Sagnac effect, but especially because it contradicts the theoretical prediction in the laboratory (14), which implies that the effect should be nonzero. A velocity of light along the rim of the circular disk equal to c is thus seen to lead to a logical contradiction within the theory. But such a value of the velocity of light is required by the principle of local equivalence between the accelerated frame and the "tangent" inertial frames if in the latter frames clocks are synchronised with Einstein's method, as it would follow from the assumed validity of the Lorentz transformations in all inertial frames. This is the reason why something fundamental has to be modified.

By accepting (17)-(18) we can overcome the long-standing "mystery" of the rotating platform and find a perfectly rational description of the Sagnac effect. The latter effect [15] is essentially the observation of a phase shift between two coherent beams travelling on opposite paths in an interferometer placed on a rotating disk. In the typical experiment of this type a monochromatic light source placed on the disk emits two coherent beams of light in opposite directions along the disk circumference until they reunite in a small region and interfere, after a 2π propagation. The positioning of the interference figure depends on the disk rotational velocity. Textbooks deduce the Sagnac formula in the laboratory (essentially our Eq. (14) above), but say nothing about the description of the phenomenon on the rotating platform. Exceptions to this trend are Langevin [16], Anandan [17], Dieks and Nienhuis [18], and Post [19], but dissatisfaction remains, because none of these treatments is free of ambiguities. For example Langevin's approach leads to difficulties at the general relativistic level [however in his 1937 paper he recognised the possibility of a non-standard velocity of light on the rotating platform and gave formulae which agree to first order with our results (17)-(18)]. As a second example, Post's relativistic formula is not generally valid, but limited to the arbitrary case where the origin of the "tangent" inertial frame coincides with the centre of the rotating disk.

4. VELOCITY OF LIGHT RELATIVE TO INERTIAL FRAMES

One can always choose Cartesian co-ordinate systems in two inertial reference frames S and S_0 and assume:

(1) that space is homogeneous and isotropic, and that time is also homogeneous;
(2) that relative to S_0 the velocity of light is the same in all directions, so that Einstein's synchronisation can be applied in this frame and the velocity v of S relative to S_0 can be measured;
(3) that the origins of S and S_0 coincide at $t = t_0 = 0$;
(4) that planes (x_0, y_0) and (x, y) coincide at all times t_0; that also planes (x_0, z_0) and (x, z) coincide at all times t_0; but that planes (y_0, z_0) and (y, z) coincide at time $t_0 = 0$ only.

It then follows [14] that the transformation law from S_0 to S has necessarily the form

$$\begin{cases} x = f_1(x_0 - vt_0) \\ y = g_2 y_0 \\ z = g_2 z_0 \\ t = e_1 x_0 + e_4 t_0 \end{cases} \tag{21}$$

where the four coefficients f_1, g_2, e_4, and e_1 can depend on v. A fifth coefficient, still present in Ref.[14], can be eliminated by invoking rotational invariance around the x axis. If at this point one assumes the validity of the relativity principle (including invariance of light velocity) the previous transformations reduce <u>necessarily</u> to the Lorentz ones.

It was shown in [14] that the most general transformation laws of the general type (21) between two inertial frames S_0 and S satisfying the conditions of constant two-way velocity of light and of time dilation according to the usual relativistic factor are a particular case of (21) with

$$f_1 = \frac{1}{R(\beta)} \quad ; \quad g_2 = 1 \quad ; \quad e_4 = R(\beta) - e_1 \beta c$$

where $\beta = v / c$, and

$$R(\beta) = \sqrt{1-\beta^2} \tag{22}$$

so that

$$\begin{cases} x = \dfrac{x_0 - \beta c t_0}{R(\beta)} \\ y = y_0 \\ z = z_0 \\ t = R(\beta) t_0 + e_1 (x_0 - \beta c t_0) \end{cases} \tag{23}$$

Thus e_1 is the only remaining unknown function of velocity v. Length contraction by the usual factor is also a consequence of (23). The velocity of light compatible with (23) was shown in [14] to be:

$$\frac{1}{\tilde{c}(\theta)} = \frac{1}{c} + \left[\frac{\beta}{c} + e_1 R(\beta)\right] \cos\theta \tag{24}$$

where θ is the angle between the direction of propagation of light in S and the absolute velocity $\vec{v}$ of S. The transformations (23) represent the complete set of theories "equivalent" to SRT: if e_1 is varied, different elements of this set are obtained, which, according to the conventionality thesis of Reichenbach, should be all equivalent as far as the explanation of experimental results is concerned (unless accelerations are considered). The Lorentz transformation is recovered as a particular case with $e_1 = -\beta / c\, R(\beta)$ whence it also follows $\tilde{c}(\theta) = c$. Different values of e_1 are obtained from different clock-synchronisation conventions. In all cases but that of SRT such values exclude the validity of the strong relativity principle, and imply the existence of a privileged frame [14].

In the previous sections we found values of the one-way velocity of light along the rim of the disk (relative to the disk itself) different from c and given by Eq.s (17) and (18). Our principle of local equivalence between the rim of the disk and the "tangent" inertial frame requires (17) and (18) to apply in the latter frame as well. Eq. (24) applied to the cases $\theta = 0$ and $\theta = \pi$ gives

$$\frac{1}{\tilde{c}(0)} = \frac{1}{c} + \left[\frac{\beta}{c} + e_1 R(\beta)\right] \quad ; \quad \frac{1}{\tilde{c}(\pi)} = \frac{1}{c} - \left[\frac{\beta}{c} + e_1 R(\beta)\right] . \tag{25}$$

This agrees with (17) and (18) if and only if

$$e_1 = 0 \tag{26}$$

The space dependent term in the transformation of time is thus seen to disappear from (23).

5. THE INERTIAL TRANSFORMATIONS

In the previous section we showed that the principle of continuity between the rim of the disk and the tangent frame implies that the condition $e_1 = 0$ has necessarily to be used. Therefore from all positions in S_0 the time in S will be seen to be the same, and no position dependent time-lag factor will be present in the transformation of time. This gives rise to a transformation different from the Lorentz one, but nevertheless particularly simple [14]:

$$\Omega(0, \beta): \begin{cases} x = \dfrac{x_0 - \beta c t_0}{R(\beta)} \\ y = y_0 \\ z = z_0 \\ t = R(\beta) t_0 \end{cases} \tag{27}$$

where we introduced the abstract symbol $\Omega(0, \beta)$ to indicate the transformation from the privileged frame S_0 (having absolute velocity 0) where Einstein's synchronisation can be applied, and the frame S moving with absolute velocity v ($\beta = v / c$). The velocity of light relevant to a theory based on (27) can easily be found by putting $e_1 = 0$ in (24):

$$\frac{1}{\tilde{c}(\theta)} = \frac{1 + \beta \cos\theta}{c} \tag{28}$$

The transformation (27) can be inverted and gives:

$$\Omega(\beta, 0): \begin{cases} x_0 = R(\beta)\left[x + \dfrac{\beta c}{R^2(\beta)} t\right] \\ y_0 = y \\ z_0 = z \\ t_0 = \dfrac{1}{R(\beta)} t \end{cases} \tag{29}$$

Note that there is a formal difference between (27) and (29). The latter implies, for example, that the origin of S_0 (satisfying $x_0 = y_0 = z_0 = 0$) is described in S by $y = z = 0$ and by

$$x = -\frac{\beta c}{1 - \beta^2} t$$

This origin is thus seen to move with speed $\beta c/(1-\beta^2)$, which can exceed c, but cannot be superluminal. In fact a light pulse seen from S to propagate in the same direction as S_0 has $\theta=\pi$, and thus [using (27)] has speed $\tilde{c}(\pi)=c/(1-\beta)$, which can be checked to satisfy

$$\frac{c}{1-\beta} \geq \frac{c\beta}{1-\beta^2}$$

One of the typical features of these transformations is of course the presence of velocities which can grow without limit *when they are relative to moving systems* having absolute velocities βc near to c. Absolute velocities can instead never exceed c [14]. In STR one is used to relative velocities that are always equal and opposite, but this symmetry is a consequence of the particular synchronisation used and cannot be expected to hold more generally [14].

Consider now a third inertial system S' moving with velocity $\beta' c$ and the transformation ti its space and time coordinates from those of S_0, which of course is

$$\Omega(0,\beta'): \begin{cases} x' = \dfrac{x_0 - \beta' c t_0}{R(\beta')} \\ y' = y_0 \\ z' = z_0 \\ t' = R(\beta') t_0 \end{cases} \tag{30}$$

where $R(\beta')$ is given by (22) with β' replacing β. By eliminating the S_0 variables from (30) and (29) one obtains the transformation between the two moving systems S and S':

$$\Omega(\beta,\beta'): \begin{cases} x' = \dfrac{R(\beta)}{R(\beta')}\left[x - \dfrac{\beta' - \beta}{R^2(\beta)} c t\right] \\ y' = y \\ z' = z \\ t' = \dfrac{R(\beta')}{R(\beta)} t \end{cases} \tag{31}$$

A transformation having the form (27) has once been written down by Tangherlini, while (29) and (31) do not seem to exist in the scientific literature: a possible name for (27)-(29)-(31) is "inertial transformations". In its most general form (31) the inertial transformation depends on two velocities (v and v'). When one of them is zero, either S or S' coincide with the privileged system S_0 and the transformation (31) becomes either (27) or (29).

A feature characterising the transformations (27)-(29)-(31) is the existence of absolute simultaneity: two events taking place in different points of S but at the same t are judged to be simultaneous also in S' (and vice versa), this property being consequence of the absence of space variables in the transformation of time. Of course the existence of absolute simultaneity does not imply that time is absolute: on the contrary, the β-dependent factor in the transformation of time gives rise to time-dilation phenomena similar to those of STR. *Time dilation* in another sense is however also absolute: a clock at rest in S is seen from S_0 to run slower, but a clock at rest in S_0 is seen from S to run faster so that both observers will agree that motion relative to S_0 slows down the pace of clocks. Quantitatively one has for both situations:

$$\Delta t = \sqrt{1-\beta^2}\Delta t_0 \tag{32}$$

where Δt and Δt_0 are the time intervals between any two given events measured with clocks at rest in S and in S_0, respectively. The difference with respect to STR is however more apparent than real: a meaningful comparison of rates implies that a clock T_0 at rest in S_0 must be compared with clocks at rest in different points of S, and the result is therefore dependent on the convention adopted for synchronising the latter clocks.

Absolute length contraction can also be deduced from (27)-(29). A rod at rest on the x axis of S between the points with co-ordinates x_2 and x_1 is seen in S_0 to have end points x_{02} and x_{01} at a common time t_0, where from (27):

$$x_2 = \frac{x_{02} - vt_0}{\sqrt{1-\beta^2}} \quad ; \quad x_1 = \frac{x_{01} - vt_0}{\sqrt{1-\beta^2}} \tag{33}$$

From this one obtains

$$x_2 - x_1 = \frac{1}{\sqrt{1-\beta^2}}(x_{02} - x_{01}) \tag{34}$$

The reasoning can be inverted by considering the rod at rest in S and observed from S_0, and using the transformation (29). One gets then, after a few simple steps:

$$x_{02} - x_{01} = \sqrt{1-\beta^2}(x_2 - x_1) \tag{35}$$

which could also be obtained by inverting (34). The two results are thus mathematically equivalent and lead to the conclusion (with which both observers agree) that motion relative to S_0 leads to contraction. This is obviously an absolute effect, but again the discrepancy with the STR is due to the different conventions concerning clock synchronisation: the length of a moving rod can only be obtained by marking the simultaneous positions of its end points, and is therefore dependent on the very definition of simultaneity of distant events.

6. COMPOSITIONS OF INERTIAL TRANSFORMATIONS

In the present section we will study the multiplication properties of the inertial transformations and show that they do not form a group. In fact if $\Omega(\beta, \beta')$ is the transformation (31), dependent on the two dimensionless absolute velocities β and β', and if $I = \{\Omega(\beta, \beta')\}$ is the set of all such transformations, two elements of I differ from one another only for the value of one or both velocities β and β'. The Tangherlini transformation (27) is $\Omega(0, \beta)$; its inverse (29) is $\Omega(\beta, 0)$, so that they both belong to I. It follows that:

[1] The **identical transformation** is an element of I because for arbitrary $\beta = \beta'$ (31) becomes

$$\begin{cases} x' = x \\ y' = y \\ z' = z \\ t' = t \end{cases}$$

which can be written as $\Omega(\beta, \beta) \in I$.

[2] The **inverse transformation** of $\Omega(\beta, \beta')$ is obtained by inverting (31):

$$\begin{cases} x = \dfrac{R(\beta')}{R(\beta)}\left[x' - \dfrac{\beta - \beta'}{R^2(\beta')}ct'\right] \\ y = y' \\ z = z' \\ t = \dfrac{R(\beta)}{R(\beta')}t' \end{cases} \tag{36}$$

Obviously this inverse of $\Omega(\beta, \beta')$ is $\Omega(\beta', \beta) \in I$. Therefore the inverse of a transformation is obtained by interchanging the two absolute velocities β and β'.

[3] The **product of two inertial transformations** is obtained as follows: consider the inertial transformation $\Omega(\beta', \beta'')$ from S' to S'':

$$\Omega(\beta', \beta''): \begin{cases} x'' = \dfrac{R(\beta')}{R(\beta'')}\left[x' - \dfrac{\beta'' - \beta'}{R^2(\beta')}ct'\right] \\ y'' = y' \\ z'' = z' \\ t'' = \dfrac{R(\beta'')}{R(\beta')}t' \end{cases} \tag{37}$$

By inserting (31) in it one obtains

$$\begin{cases} x'' = \dfrac{R(\beta)}{R(\beta'')}\left[x - \dfrac{\beta'' - \beta}{R^2(\beta)}ct\right] \\ y'' = y \\ z'' = z \\ t'' = \dfrac{R(\beta'')}{R(\beta)}t \end{cases} \tag{38}$$

which is $\Omega(\beta, \beta'') \in I$. The previous result can also be written

$$\Omega(\beta, \beta')\Omega(\beta', \beta'') = \Omega(\beta, \beta'') \tag{39}$$

This is the multiplication law of inertial transformations: as one can see the common velocity disappears from the product.

[4] The **associative law** of the multiplication of inertial transformations can now be established. Consider four inertial frames S, S', S'' and S''' and the following transformations

$$\Omega(\beta, \beta') : S \Rightarrow S'$$
$$\Omega(\beta', \beta'') : S' \Rightarrow S''$$
$$\Omega(\beta'', \beta'''): S'' \Rightarrow S'''$$

By applying (39) one easily gets:

$$[\Omega(\beta, \beta')\Omega(\beta', \beta'')]\Omega(\beta'', \beta''') = \Omega(\beta, \beta'')\Omega(\beta'', \beta''') = \Omega(\beta, \beta''')$$

and

$$\Omega(\beta, \beta')[\Omega(\beta', \beta'')\Omega(\beta'', \beta''')] = \Omega(\beta, \beta')\Omega(\beta', \beta''') = \Omega(\beta, \beta''')$$

so that the associative law is satisfied.

It should however be noticed that it is possible to multiply only elements of I having a common velocity. Thus $\Omega(\beta, \beta')\Omega(\beta', \beta'')$ makes sense, but $\Omega(\beta, \beta')\Omega(\beta'', \beta''')$ does not, if $\beta' \neq \beta''$. Because of this restriction the set I is not a group. This is no serious problem, however, because one only needs to insure the validity of (39), while a product like $\Omega(\beta, \beta')\Omega(\beta'', \beta''')$ with $\beta' \neq \beta''$ does not have any physical relevance.

The situation is analogous to that of time evolution operators in quantum mechanics. As it is well known the Schrödinger equation has the formal solution:

$$\psi(t) = T(t,t_0)\,\psi(t_0) \tag{40}$$

where

$$T(t,t_0) = \exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')\right\} \tag{41}$$

is the unitary time evolution operator given as a functional of the generally time-dependent Hamiltonian $H(t)$ and satisfying

$$T^{+}(t,t_0)T(t,t_0) = T(t,t_0)T^{+}(t,t_0) = I \tag{42}$$

where I is the unity operator. Through (40) $T(t,t_0)$ gives rise to the time evolution of wave-functions ("Schrödinger description") and, symmetrically, to the time evolution of operators ("Heisenberg description") through

$$A(t) = T^{+}(t,t_0)\,A\,T(t,t_0) \tag{43}$$

The integral in (41) is made over the unbroken interval (t_0, t). If one multiplies as follows two time evolution operators one gets

$$T(t_2,t_1)T(t_1,t_0) = \exp\left\{-\frac{i}{\hbar}\int_{t_1}^{t_2} dt' H(t')\right\}\exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t_1} dt' H(t')\right\} =$$

$$= \exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t_2} dt' H(t')\right\} = T(t_2,t_0)$$

where the interval defining $T(t_2,t_0)$ is again unbroken. No time evolution operator is instead obtained by considering the product

$$T(t_3,\, t_2)T(t_1,\, t_0) \tag{44}$$

with $t_2 \neq t_1$. Therefore the time evolution operators do not form a group under multiplication. This is no serious problem, however, because products of the type (44) are physically uninteresting. The situation is obviously analogous to the one we considered above for the inertial transformations [20].

8. CONCLUSIONS

Our choice of synchronisation ("absolute", according to Mansouri and Sexl [4]) has been made by considering accelerations. It is worth stressing that the normally accepted relationship between SRT and accelerated systems is far from negligible. Accelerations are essential in the so-called twin paradox, which is a prediction of SRT made before the general theory was conceived. Very large accelerations enter in the CERN Muon Storage Ring experiment, which is considered the most accurate quantitative test of time dilation. Earth acceleration is essential for perceiving the retardation/anticipation of the eclipses of Jupiter's satellites. The same can be said about the detection stellar aberration. And so on. Accelerations are instead avoided when they seem to generate difficulties within the existing theory. This is perhaps an understandable initial reaction, but it is of course not acceptable in the long run. We have tried to do better.

There remain at least two fundamental problems to solve, before a theory based on the inertial transformations can be considered reasonably complete:

(1) Maxwell's equations must be reformulated. They will take a more general form, dependent on the absolute velocity of the inertial frame with respect to which they are considered, and will assume the usual form only in the privileged frame. This work is underway and will probably generate a good agreement with experiments, given the results already obtained [14] concerning experiments on optical interference.

(2) Also the general theory needs modifications because the ds^2 cannot be considered invariant anymore, given that the inertial transformations predict a frame-dependent one-way velocity of light. It will be necessary to show that the modified theory makes the same successful predictions as the general theory of relativity. We can say even before starting that the new theory will have an important advantage over the old one in the validity of the continuity principle between accelerated frames and locally tangent inertial frames. The work on this problem is in progress.

The kinematics of high energy particle interactions must also be reconsidered, but here the situation is already settled because this problem was solved in Ref. [14], where it was shown that a complete equivalence exists between our predictions and those of SRT. In fact energy and momentum are defined in such a way as to coincide numerically (not analytically) with those of the SRT for all particles and in all inertial frames, once they coincide in the fundamental frame. Therefore the kinematics of high energy processes, the determination of particle masses, and so on, do not require a different analysis from the one successfully carried out by particle physicists up to the present time.

Acknowledgements: I wish to thank Dr. Jan Czerniawski, Prof. Jerome A. Goldstein, Clive Kilmister and Tom Phipps for kind suggestions and criticisms.

REFERENCES

[1] J. Bailey et al., Nature **268**, 301 (1977).

[2] A. Einstein, Ann. Phys. **35**, 898 (1911).

[3] A. Einstein, Ann. der Physik, 17, 891 (1905).

[4] R. Mansouri e R. Sexl, General Relat. and Gravitation, **8**, 497 (1977).

[5] A. Grünbaum, *Philosophical Problems of Space and Time* (Reidel, Dordrecht, 1973).

[6] M. Jammer, Some Fundamental Problems in the Special Theory of Relativity, in: *Problems in the Foundations of Physics* (G. Toraldo di Francia ed., North Holland, Amsterdam, 1979).

[7] T. Sjödin, Nuovo Cim. **51B**, 229 (1979).

[8] G. Cavalleri, Nuovo Cim. **104B**, 545 (1989).

[9] A. A. Ungar, Found. Phys. **21**, 691 (1991).

[10] I. Vetharaniam and G.E. Stedman, Phys. Lett., **A183**, 349 (1993).

[11] Quoted by A. Tyapkin, *Relatività speciale* (Jaca Book, Milano, 1993), p. 67.

[12] A.Einstein, *Relativity, The Special, The General Theory,* (Chicago, 1951), p. 18

[13] H. Reichenbach, *The Philosophy of Space &Time* (Dover Publ., New York, 1958).

[14] F. Selleri, Found. Phys. **26**, 641 (1996); Found. Phys. Lett. **9**, 43 (1996).

[15] G. Sagnac, Compt. Rend. **157** (1913) 708; *ibidem* 1410.

[16] P. Langevin, Compt. Rend. **173** (1921) 831; Compt. Rend. **205** (1937) 304.

[17] J. Anandan, Phys. Rev. **D 24** (1981) 338.

[18] D. Dieks and G. Nienhuis, Am. J. Phys. **58** (1990) 650.

[19] E. J. Post, Rev. Mod. Phys. **39** (1967) 475.

[20] J.A. Goldstein, *Semigroups of Linear Operators and Applications*, Oxford University Press (1985).

On the derivation of the electron propagator from a random walk

V.A. Karmanov
Lebedev Physical Institute, Leninsky Prospekt 53, 117924 Moscow, Russia
e-mail: karmanov@sci.lpi.msk.su

February 27, 1997

Abstract

A representation for the electron propagator is found in the form of the sum of the amplitudes corresponding to the following process. An electron propagates at the speed of light as a massless particles. At a random Poisson distributed time moments its direction of motion and projection of spin on this direction are changed. The average distance covered between these time moments is exactly equal to the electron Compton length. The connection of this representation with the structure of space-time and with nonconservation of P- and CP-parity is discussed.

1 Introduction

In the papers [1, 2] a representation for the propagator of a Dirac particle (say, electron) was found which corresponds in the one-dimensional case to the following process. The electron passes with the speed of light a distance ϵ, and can then change its direction of motion to the opposite one or continue to move in the initial direction. The amplitude of the rectilinear motion is equal to one, every change of the direction of motion contributes to the amplitude a factor $i\epsilon m$, where m is the electron mass. The sum of the amplitudes over all such trajectories in the limit $\epsilon \to 0$ gives the electron propagator. Such a representation was investigated in detail in refs. [3, 4, 5]. In ref. [1] the problem for the three-dimensional case has been solved as well. Related problems were considered in ref. [6].

A similar process was studied in ref. [7]. Namely, in a three-dimensional space the electron propagates at the speed of light (like as a massless neutrino). At random Poisson distributed moments of time its helicity and direction of motion are simultaneously changed. An operator constructed in [7] averaged with the Poisson distribution also results in the electron propagator.

In the present paper, based on ref. [8], we give another representation for the electron propagator, which exactly corresponds to this physical process. Our representation differs

102

from the one given in ref. [7] by its derivation and by the physically transparent form. Instead of formal averaging the operator introduced in [7] we explicitly calculate the sums of the amplitudes corresponding to the neutrino (or antineutrino) propagation and the change of its spin projection on the momentum to the opposite one. The moments of time when the changes of helicity and direction of motion occur are Poisson distributed, like in ref. [7]. The average path covered between these two moments of time turns out to be equal to the Compton length of the electron. We admit that such a representation is not a mathematical curiosity but may reflect a real physical processes taking place in the nature. In its turn, these processes have to be explained by a more fundamental theory. Therefore the representation derived below in which these processes manifest themselves explicitly have a heuristic value and can help to guess this theory. This is the main motivation of the present paper.

Examples of the search for this fundamental theory are given in many talks at the ANPA meetings and, in particular, in approaches developed in refs. [9]-[17]. These approaches emphasize the fact that the concept of the space-time continuum implies the existence of a physical process allowing one to measure the spatial coordinates and time in principle with arbitrary high precision. The mathematical structure of classical mechanics is perfectly adjusted for the description of such processes. In refs. [9]-[17] in different ways the point of view is developed according to which in the realm of quantum mechanics, restricting the localization of an elementary particle by the Compton length, the foundations should be revised. They must be constructed by means of concepts which naturally consistent with the primary notion of quanta. For example, one can start with the idea of a quantum event (or process), and the space-time could appear as a secondary notion together with the transition to the limit of classical mechanics. The hypothesis of underlying space-time continuum is considered in ref. [10] "as naive as the caloric fluid, the homulculus, and other important historical reifications". Such an approach must, of course, reproduce the usual quantum theory, i.e., admit a reformulation in terms of the wave function, defined on the continium. The conceptual difficulties inherent to quantum mechanics (such as "the collapse of the wave function", EPR paradox, infinities and their renormalizations, etc.) would turn out to be a consequence of embedding the initial concept of the quantum event in the space-time continium. This idea is expressed and developed in a series of papers (see refs. [11]-[14]), and references therin). In the paper [14] the Dirac equation has been derived in a discrete approach, using the calculus of finite differences.

The papers [15] are devoted to the derivation of both quantum mechanics and the special theory of relativity from a random walk with a step length exactly equal to the Compton length (in refs. [15] the fixed step length is considered, but not the average one, in contrast to refs. [1]-[5]). A model for discrete quantum mechanics is developed in [16]. Another approach developed in ref. [17] begins with the idea of indistinguishable objects (twins), a set of which cannot be ordered even for a finite cardinal number. Its development has intersections with the results of refs. [11, 12] obtained from a different starting point. We believe that very interesting ideas of ref. [17] deserve further careful investigation.

We begin the derivation of a propagator for an especially transparent one-dimensional case. Let a particle starts its motion with the speed of light from the origin of a coordinate system in the positive direction of the x axis. Then it changes its direction of motion to the opposite one at an odd number $k = 2p + 1$ of points and comes to the point x moving in the negative direction (see fig. 1). We assume that the amplitude of the process after

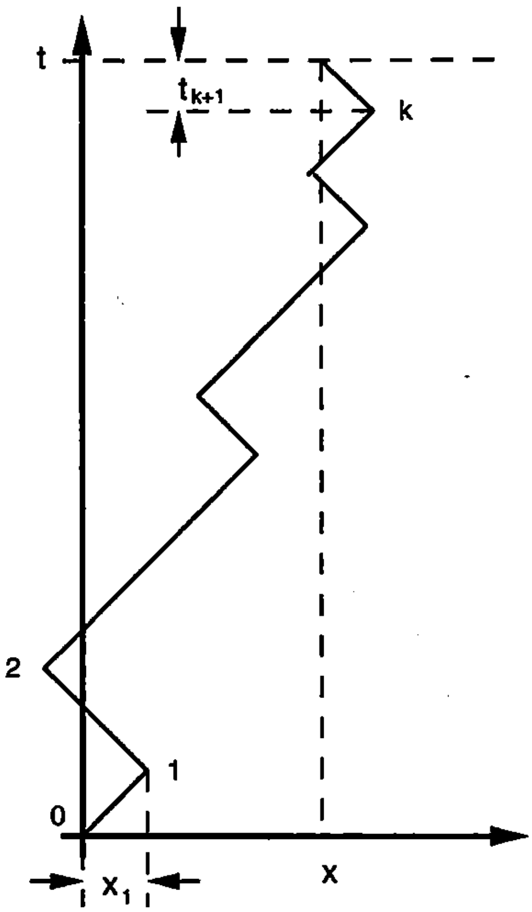


Fig.1. Random walk of a particle in one-dimensional space.

one turn is the imaginary unit i, the turns are distributed homogeneously in the time t and in the interval x, the number of turns in the interval dx is dx/l_0, where $l_0 = m/\hbar c$ is the Compton length. Then the amplitude for the path with k turns constrained by the initial and final conditions is the product $i^k I_k$, where I_k is the number of corresponding paths determined by the integral over these paths. We emphasize that this path integral differs from the usual one by the domain of integration and by the integrand. The integration is carried out only over the paths corresponding to the speed of light (fig. 1). The paths with k turns differ from each other by the positions of turns. Hence, the path integral is simply the integral over positions of the turns. This integral has the form:

$$I_k = \int_D \frac{dx_1}{l_0} \cdots \frac{dx_k}{l_0}. \qquad (1)$$

104

The integrand in eq. (1) is the constant $1/l_0^k$. The integration domain D includes all the paths with k turns, i.e., with $k+1$ time and space intervals. It is determined by the conditions

$$x_1, x_2, \ldots, x_k \geq 0, \tag{2}$$

$$t_1 + t_2 + \ldots + t_k + t_{k+1} = t, \tag{3}$$

$$x_1 - x_2 + x_3 - \ldots + x_k - x_{k+1} = x, \tag{4}$$

where $x_1 = ct_1$ is the path covered between the initial position of the particle and the first turning point during the time t_1, x_i is the path covered between the turns number $i-1$ and i during the time t_i, x_{k+1} is the path covered between the last turning point k and the final position x during the last time interval t_{k+1}. According to eq (2), x_i is positive. The condition (3) constrains the sum of the time intervals, the condition (4) fixes the final position of the particle at the time t in the point x.

The total amplitude K_{+-} is the sum of the amplitudes $i^k I_k$ corresponding to different numbers of the turning points:

$$K_{+-} = \sum_k i^k I_k. \tag{5}$$

Introducing the θ- and δ-functions automatically providing the integration over the domain D, we represent I_k in the form:

$$\begin{aligned} I_k &= \int_{-\infty}^{\infty} \theta(x_1)\ldots\theta(x_{k+1})\delta(x_1 + \ldots + x_{k+1} - ct) \\ &\times \delta(x_1 - x_2 + \ldots + x_k - x_{k+1} - x)\frac{dx_1}{l_0}\cdots\frac{dx_k}{l_0}dx_{k+1}. \end{aligned} \tag{6}$$

The integration over all the variables is carried out in the infinite limits. After integration over dx_{k+1} there is still one constraint on the position of the turning points.

Derivation of the transition amplitude K_{+-} is reduced now to calculation of the integral (6) and then of the sum (5).

Representing the θ- and δ-functions in eq. (6) in the integral form:

$$\theta(x_i) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\exp(i\alpha_i x_i)}{\alpha_i + i0}d\alpha_i,$$

$$\delta(x_1 + \ldots x_{k+1} - ct) = \int_{-\infty}^{\infty}\exp[-i\alpha(x_1 + \ldots + x_{k+1} - ct)]\frac{d\alpha}{2\pi},$$

$$\delta(x_1 - x_2 + \ldots + x_k - x_{k+1} - x) = \int_{-\infty}^{\infty}\exp[i\beta(x_1 - x_2 + \ldots + x_k - x_{k+1} - x)]\frac{d\beta}{2\pi},$$

and then integrating over $dx_i d\alpha_i$ $(i = 1, \ldots, k+1)$, we find

$$I_k = \frac{(-i)^{k+1}}{(2\pi)^2 l_0^k}\int_{-\infty}^{\infty}\frac{\exp(i\alpha t)\exp(-i\beta x)d\alpha d\beta}{(\alpha - \beta - i0)^{p+1}(\alpha + \beta - i0)^{p+1}}, \tag{7}$$

where $k = 2p + 1$. Introducing the variables $\alpha' = \alpha - \beta$, $\beta' = \alpha + \beta$ and calculating the two residues, we obtain:

$$I_k = \frac{m(m\tau)^{2p}}{2(p!)^2 2^{2p}} \tag{8}$$

where $\tau^2 = t^2 - x^2 > 0$ is the space-time interval. The calculation of the integral (5) at $\tau^2 < 0$ gives $I_k = 0$. In eq.(8) and almost everywhere below we put $\hbar = c = 1$.

The substitution of eq.(8) in the sum (5) results in the Bessel function J_0:

$$K_{+-} = \sum_k i^k I_k = \frac{1}{2} im \sum_p \frac{(-1)^p (m\tau)^{2p}}{2^{2p}(p!)^2} = \frac{1}{2} im J_0(m\tau). \tag{9}$$

Expression (9) coincides with a non-diagonal matrix element of the two-dimensional Green function for the electron in coordinate space. For K_{-+} we similarly get: $K_{-+} = K_{+-}$. An analogous calculation of the amplitude of a process which starts and ends with the particle moving in the same (right or left) direction gives the diagonal matrix elements:

$$K_{++} = \frac{m(-t-x)}{2\tau} J_1(m\tau), \quad K_{--} = \frac{m(-t+x)}{2\tau} J_1(m\tau), \tag{10}$$

where J_1 is the Bessel function of the first order. In this way we obtain the 2×2 matrix coinciding with the electron Green function for the two-dimensional case (cf. ref. [3]).

There is some freedom in the rules associating any given path with the corresponding amplitude, which influences the explicit expression for K. This corresponds to the different unitary equivalent representations of the Dirac matrices. In particular, it is possible to remove the factor i from the rules and from the final result, that corresponds to the Majorana representation.

If the change of the direction of motion of the particle is considered as a random process, then this process is a Poisson one. This immediately follows from our assumptions and can be easily seen in a explicit form after removing from the integral (6) the last δ-function providing the boundary condition (this is equivalent to integrating the function I_k over dx). Since the integral (6) is the number of trajectories with given boundary conditions, then after removing this δ-function we obtain the total number of all the trajectories with k turns, i.e., simply the number $N_k(t)$ of turns for the time t:

$$\begin{aligned} N_k(t) &= \int_{-ct}^{ct} I_k(x,t)dx \\ &= \int_{-\infty}^{\infty} \theta(x_1)\ldots\theta(x_{k+1})\delta(x_1 + \ldots + x_k + x_{k+1} - ct)\frac{dx_1}{l_0}\cdots\frac{dx_k}{l_0}dx_{k+1}. \end{aligned} \tag{11}$$

The integral (11) is calculated analogously to the integral (6). The result has the form

$$N_k(t) = \frac{(t/t_0)^k}{k!},$$

where $t_0 = l_0/c = \hbar/mc^2$ ($= l_0$ at $c = 1$). Calculating the normalization $N = \sum_k N_k = \exp(t/t_0)$, we find the Poisson probability:

$$w_k(t) = \frac{N_k(t)}{N} = \frac{(t/t_0)^k}{k!}\exp(-t/t_0). \tag{12}$$

The average number of events for the time t is, evidently, $\bar{N} = \sum_k k w_k(t) = t/t_0$, the average time between two turns is $\bar{t} = t/\bar{N} = t_0$ and the average path between them is the Compton length: $\bar{x} = c\bar{t} = ct_0 = l_0$.

The integral I_k in eq.(6) can be represented in the form:

$$\begin{aligned} I_k = & \int D_+(x_1,t_1)D_-(x_2,t_2)\ldots D_-(x_{k+1},t_{k+1})\delta(x_1+x_2+\ldots+x_{k+1}-x) \\ & \times\delta(t_1+t_2+\ldots+t_{k+1}-t)dx_1\frac{dt_1}{t_0}\ldots dx_k\frac{dt_k}{t_0}dx_{k+1}dt_{k+1}, \end{aligned} \tag{13}$$

where $D_\pm(x,t) = \theta(t)\delta(x \mp t)$ is the propagator of a particle moving to the right (left) with the speed of light.

The three-dimensional case is given by evident generalization of formula (13):

$$\begin{aligned} I_k^{(3)} = & \int D_+(\vec{x}_1,t_1)D_-(\vec{x}_2,t_2)\ldots D_-(\vec{x}_{k+1},t_{k+1})\delta^{(3)}(\vec{x}_1+\vec{x}_2+\ldots+\vec{x}_{k+1}-\vec{x}) \\ & \times\delta(t_1+t_2+\ldots+t_{k+1}-t)d^3x_1\frac{dt_1}{t_0}\ldots d^3x_k\frac{dt_k}{t_0}d^3x_{k+1}dt_{k+1}, \end{aligned} \tag{14}$$

where $D_\pm(\vec{x},t)$ is the propagator of the right (left) polarized antineutrino (neutrino):

$$D_\pm(\vec{x},t) = \left(-\frac{\partial}{\partial t} \pm \vec{\sigma}\cdot\frac{\partial}{\partial\vec{x}}\right)\frac{\theta(t)\delta(\lambda^2)}{2\pi} = -i\int\frac{q_0 \pm \vec{\sigma}\cdot\vec{q}}{q^2 + i0\,\mathrm{sgn}(q_0)}\exp(-iqx)\frac{d^4q}{(2\pi)^4}, \tag{15}$$

$\vec{\sigma}$ are the Pauli matrices, $\lambda^2 = t^2 - \vec{x}^{\,2}$, $\mathrm{sgn}(q_0) = 1$ at $q_0 > 0$ and $\mathrm{sgn}(q_0) = -1$ at $q_0 < 0$. After substitution of the integral representation of (15) for $D_\pm(\vec{x},t)$ in eq.(14) the integral for $I_k^{(3)}$ is calculated analogously to the integral in eq.(6). For $k = 2p+1$ and $p = 0$ we obtain $I_k^{(3)} = m\delta(\lambda^2)/2\pi$. For $p \geq 1$ the result has the form:

$$I_k^{(3)} = -\frac{m}{4\pi r}\frac{\partial}{\partial r}\frac{(\lambda m)^{2p}}{2^{2p}(p!)^2}, \tag{16}$$

where $r = (\vec{x}^{\,2})^{1/2}$. Summing over k we find:

$$K_{+-} = \sum_k i^k I_k^{(3)} = \frac{im}{2\pi}\delta(\lambda^2) - \frac{im^2}{4\pi\lambda}J_1(m\lambda). \tag{17}$$

Similarly we get $K_{-+} = K_{+-}$ and for the sum over even values of k:

$$K_{\pm\pm}(\vec{x},t) = \left(\frac{\partial}{\partial t} \mp \vec{\sigma}\cdot\frac{\partial}{\partial\vec{x}}\right)\frac{m}{4\pi\lambda}J_1(m\lambda). \tag{18}$$

Each of them is a 2×2-matrix. Four of these matrices form a 4×4 matrix, its Fourier transform being connected with the electron propagator $1/(\hat{p} - m)$:

$$K = iS\gamma_0 = -i \int \frac{(\hat{p} + m)\gamma_0}{p^2 - m^2 + i0\,\mathrm{sgn}(p_0)} \exp(-ipx) \frac{d^4p}{(2\pi)^4}, \tag{19}$$

where $\hat{p} = p_0\gamma_0 - \vec{p}\cdot\vec{\gamma}$, γ_μ are the Dirac matrices. The factor $i\gamma_0$ in eq.(19) appears because of the usual initial condition for the Green function S: $S(\vec{x}, t = 0) = i\gamma_0\delta^{(3)}(\vec{x})$, whereas by construction at $t = 0$ we have a unit matrix (generally speaking, multiplied by a phase factor): $K(\vec{x}, t = 0) = -\delta^{(3)}(\vec{x})$. Apparently, we have calculated the so-called retarded Green function which is equal to zero at $t < 0$.

From a formal point of view the sum in eqs.(9) and (17) over the turning points are series in the electron mass. Therefore the representations (1) and (14) can be apparently derived by expanding the electron propagator (multiplied at the right by $i\gamma_0$):

$$K = iS\gamma_0 = -[i\gamma_0(\hat{p} - m)]^{-1} \equiv (\Pi - i\beta m)^{-1} \tag{20}$$

in degrees of m. The operator Π in the momentum space has the form: $\Pi = i(p_0 - \vec{\alpha}\cdot\vec{p})$, $\vec{\alpha} = \gamma_0\vec{\gamma}$, $\beta = \gamma_0$. In the coordinate space: $\Pi = -(\partial/\partial t + \vec{\alpha}\cdot\partial/\partial\vec{x})$. In the spinor representation with

$$\vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

we have

$$\Pi = \begin{pmatrix} \Pi_+ & 0 \\ 0 & \Pi_- \end{pmatrix},$$

where

$$\Pi_\pm = -\left(\frac{\partial}{\partial t} \pm \vec{\sigma}\cdot\frac{\partial}{\partial\vec{x}}\right)$$

and $\Pi_\pm D_\pm = \delta^{(4)}(x)$, that is $\Pi_\pm^{-1} = D_\pm$, $D_\pm$ is defined by eq.(15). Expanding the propagator represented in the form of the last term in eq.(20) in a series over powers of m, we find:

$$K = \Pi^{-1} + \Pi^{-1}(i\beta m)\Pi^{-1} + \Pi^{-1}(i\beta m)\Pi^{-1}(i\beta m)\Pi^{-1} + \dots \,. \tag{21}$$

Taking into account that

$$\Pi^{-1} = \begin{pmatrix} \Pi_+^{-1} & 0 \\ 0 & \Pi_-^{-1} \end{pmatrix} = \begin{pmatrix} D_+ & 0 \\ 0 & D_- \end{pmatrix}$$

we obtain

$$K = \begin{pmatrix} D_+ & 0 \\ 0 & D_- \end{pmatrix} + (im)\begin{pmatrix} 0 & D_+D_- \\ D_-D_+ & 0 \end{pmatrix} + (im)^2\begin{pmatrix} D_+D_-D_+ & 0 \\ 0 & D_-D_+D_- \end{pmatrix} + \dots \,. \tag{22}$$

The series (22) just corresponds to the representation found above. For example, in kth order for odd k the decomposition (22) contains two nondiagonal terms only. One of them (after extracting the factor i^k) is the product $m^k D_+ D_- D_+ \ldots D_-$, containing $k+1$ $D_\pm$ factors, that in the coordinate space is exactly the integral (14). The introduction in refs. [1]- [5] of the step ϵ tending to zero corresponds to the approximation of the integrals by the Riemann sums.

Note that the representation in which the electron manifests itself as a particle moving with the speed of light is consistent with the well known fact that eigenvalues of the electron velocity operator $c\alpha_i$ are equal to the speed of light $\pm c$ (zitterbewegung).

In this way, the direct calculation of the amplitude for the physical process in which an electron propagates as a massless neutrino and at the random Poisson distributed moments of time changes simultaneously its direction of motion and helicity indeed reproduces the electron propogator.

If we suppose that the left and right polarized massless particles propagate through space in different ways (e.g., if we replace D_- by $(1+\delta)D_-$ and D_+ by $(1-\delta)D_+$ at $\delta \ll 0$, assuming different "transparency" of the space for these states, then we obtain nonconservation of P-parity. Indeed, since

$$\gamma_5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

this is equivalent to the multiplication of the matrix

$$\begin{pmatrix} D_+ & 0 \\ 0 & D_- \end{pmatrix}$$

by $1+\gamma_5\delta$. In its turn this means that the matrix Π and, hence, $\hat{p}$ are multiplied at the right by $1-\gamma_5\delta$. Hence, the propagator becomes:

$$[\hat{p}(1-\gamma_5\delta)-m]^{-1} \approx \frac{\hat{p}(1-\gamma_5\delta)+m}{p^2-m^2}. \tag{23}$$

The amplitude of any process containing such a propagator has no definite parity. For example, in the case of interaction with the vector particle such an amplitude can be transcribed in the form with the usual parity conserving propagator but with the parity nonconserving part included in the interaction vertices:

$$\gamma_\mu[\hat{p}(1-\gamma_5\delta)+m]\gamma_\nu \approx [\gamma_\mu(1+\frac{1}{2}\gamma_5\delta)][\hat{p}+m]\gamma_\nu(1+\frac{1}{2}\gamma_5\delta)]. \tag{24}$$

At $\delta \ll 1$ this resembles the usual situation with weak parity nonconservation at the parity conserving background of electromagnetic or strong interaction.

Analogously, the assumption that the amplitudes of the transitions in the turning points between the states with helicities $1/2 \to -1/2$ and opposite are not equal to each other (i.e., these amplitudes are proportional to $1-i\eta$ and $1+i\eta$ correspondingly at $\eta \ll 1$) leades to the propagator containing the numerator

$$\hat{p}+(1+i\gamma_5\eta)m \approx (1+\frac{1}{2}i\gamma_5\eta)(\hat{p}+m)(1+\frac{1}{2}i\gamma_5\eta). \tag{25}$$

The corresponding amplitude has no definite CP-parity[1] (as well as both P- and T-parities), but conserves CPT. It is important that the origin of the P- and CP-parity nonconservations in this model is connected with the properties of the space (or "prespace"), but not with the interaction.

It is of interest to derive the representation for the electron propagator obtained above starting from an approach in the spirit of refs. [9]-[17]. In this connection we note that the combinatorial formulas in refs. [11]-[14] are very close to the formulas (8), (16). It could be also useful to introduce explicitly the objects (vacuum fluctuations -?) scattering from which results in change by electron its direction of motion and helicity.

I am sincerely grateful to B. Desplanques, T. Etter, D. McGoveran, I.S. Shapiro, I. Stein, P. Suppes and especially to H.P. Noyes for their interest, stimulating discussions and useful remarks. I am also indebted to H.P. Noyes for the support allowing to attend the meeting ANPA-18.

References

[1] G.V. Ryazanov, Zh. Eksp. Teor. Fiz. **33** (1957) 1437 [Sov. Phys. - JETP **6** (1958) 1107].

[2] R.P. Feynman and A.R. Hibbs, Quantum mechanics and path integrals (New York, 1964), p. 34.

[3] T. Jacobson and L.E. Schulman, J. Phys. **A17** (1984) 375.

[4] T. Jacobson and L.E. Schulman, J. Phys. **A17** (1984) 2433; J. Math. Phys. **26** (1985) 1600.

[5] G. Rosen, Formulation of classical and quantum dynamical theory (Academic Press, New York, 1969), p.118-122.

[6] O.A. Barut and N. Zanghi, Phys. Rev. Lett. **52** (1984) 2009; **53** (1984) 2355.

[7] B. Gaveau, T. Jacobson, M. Kac and L.S.Schulman, Phys. Rev. Lett., **53** (1984) 419.

[8] V.A. Karmanov, Phys. Lett. **A174** (1993) 371.

[9] D. Finkelstein, Phys. Rev. **184** (1969) 1261; **D5** (1972) 320, 2422; D. Finkelstein, G. Frye and L. Suskind, Phys. Rev. **D9** (1974) 2231.

[10] D. Finkelstein, Phys. Rev. **D9** (1974) 2219.

[11] D. McGoveran and P. Noyes, Foundation of discrete physics, preprint SLAC-PUB-4526 (1989); Phys. Essays, 4 (1991) 115.

[12] P. Noyes and D. McGoveran, Phys. Essays **2** (1989) 76.

[1] My attention to this fact was drawn by B. Desplanques

[13] H.P. Noyes et al., preprint SLAC-PUB-4347 (1987);
H.P. Noyes, preprint SLAC-PUB-5752 (1992); Physics Essays, **8** (1995) 434.

[14] L.H. Kauffman and H.P. Noyes, Phys. Lett. **A218** (1996) 139.

[15] I. Stein, Phys.Essay **1** (1988) 155; **3** (1990) 66;
The concept of Object as the Foundation of Physics, New York, 1995.

[16] S.P. Gudder, J. Math. Phys. **27** (1986) 1722; Found. of Phys. **18** (1988) 751.

[17] A.F. Parker-Rhodes, The theory of indistinguishables, Synthese Library, Vol. 150 (Reidel, Dordrecht, 1981).

III

FROM NAUGHT TO AUGHT: A CONCEPTUAL INQUIRY

Emmanuel RANSFORD

DRET/DCE, 16 bis, ave Prieur de la Côte d'Or, 94114 Arcueil Cedex, France

Somewhere there must be an unexplained and inexplicable being: either the world we live in, or a being that created it, or another being that created the creator, and so on.
T. Penelhum

What constitutes reality ?
- Infinite possibility constitutes reality.

ABSTRACT

This paper sets out to explore some aspects of the baffling mystery of ontogenesis, with a view to put forward a logically sound, if tentative, approach. (I surmise that the world as such is ultimately intelligible, at least partially - otherwise my undertaking would be pointless.) Ontogenesis *deals with the coming-into-being (or genesis) and sustained existence of anything-that-be. It is a mind-boggling puzzle, which seemingly lies excruciatingly beyond the farthest reaches of human ken.*

It is hard to imagine, let alone grasp, that our surrounding world could have suddenly and spontaneously popped into existence, from flat-out nothing. So there must *be a reason for its being out there. Can we ever get the hang of it?*

My bet is that a rational, if partial, approach to the enigma of ontogenesis is possible - on the understanding that ***existence comes at a price****, and that the price tag involved can (up to a point) be unearthed on sheer logical grounds.*

Put differently, the idea is that whatever exists concretely complies with a number of rules and constraints, which are absolute requisites for actual existence. *If so, I believe that a better understanding of these requirements will bring us closer to sorting out the enigma at hand. My overriding goal is to show, or at least suggest, precisely that.*

In a nutshell, the approach outlined hereafter hinges on the threefold assumption that: a) there is a rational *explanation for ontogenesis; b) this explanation is* causal; *and c) it does not allow for 'pure chance' to exist. On that basis, the message I shall try to get across is that:* a) a rational approach to the riddle of existence is an option worth considering; *and:* b) such an undertaking may prompt a better in-depth understanding of quantum physics. *(I shall not dwell at much length on the latter aspect. It will only be broached, very skimpily, near the end of this short essay.)*

[*Most of the speculations to come call for a formal framework; but, given the limited scope of this introductory paper, I shall supply none. However, when deemed necessary, I shall briefly sketch the definitions of some of the notions brought into play.*]

FOREWORD

Virtually everybody has, one day or another, wondered about the deep mystery of existence. Poets, philosophers, and the man in the street - all of them felt, in the bosom of their heart, an overwhelming sense of awe when pondering over it.

Why is there something ("aught") *rather than nothing* ("naught")? This nagging question has been around since time immemorial, and yet hasn't gotten the faintest glimmer of an answer to this day. Is it an utterly unanswerable one? When considering it, are we facing an impossible task - like the blind man in a dark room who looks for a black hat which isn't there? I do not think so. I dare believe that it is possible to make a few strides forward, beyond the current level of accepted understanding.

The odds to crack the mystery of ontogenesis are dauntingly long, and perhaps infinitely long. Yet there may be room to enhance our intellectual grasp of it, and altogether win a better grasp of our tangible world. The interesting bit is that this would in turn lead to a sharper insight into quantum physics. (Ontology is a gateway to epistemology.)

In this paper, my stance is to gainsay the 'cheap' and hollow explanation afforded by sheer indeterminism: I posit that there is no consequence without cause, no causeless effect. (Otherwise the 'enigma' of ontogenesis vanishes outright.) As for "a *rational* explanation of ontogenesis", it implies that I shall try and stick to a (logically) sound line of reasoning - and hope that I am not wrestling with an insoluble conundrum!

Nothingness, naught, 'nowt': what the heck is this? Is it a sink, is it a source - or is it entirely different still? Admittedly, it is not much of a source. We cannot *rationally* fancy our world popping out of utter nothingness - just like that, for the mere thrill of it! But then, we may hit on the idea that something could unwittingly skulk in-between nothingness (*viz. naught*) and full-blown existence (*viz. aught*). This "something" (accordingly christened later on a *be-able*, or *beable*) would pull off just the right trick: it would serve as a go-between of sorts, bridging the abysmal chasm lying between aught and naught.

Its nature will be explored in due course (but not much beyond a slight scratching of the surface, with a lot left to be done). It will turn out to be ethereal rather than substantial - not quite a full-blooded being. If "full-blown beingness" were white and naught black, then the *beable* would be an ever-shifting shade of grey - frisking across the whole spectrum of greyish hues.

Images are more telling than words, and one may fancy this outlandish *beable* springing back and forth from naught, like the lingering whiff of a ghost that would never really be 'out there' nor the opposite....

The main drift of my approach is to show that such a "half-aught, half-naught" entity enters the overall landscape of existence almost naturally; and that it manages to hyphenate the two extremes of aught and naught. *(As we shall see, a downside to full-blown existence is the availability of a limited range of possibilities only. Contrariwise, a* beable *is both shorter on beingness and immeasurably longer on potentialities. Here is the central trade-off of existence!)*

How can we tackle the huge mystery of existence? The tale of the two rods hereafter may help us to find out what basic conceptual inputs could prove fruitful to that effect.

OPENING TALE

Once upon a time (and a long time ago at that) the universe was simple, much simpler than it is today. It consisted of no more than a bunch of scattered rods, which roamed free - seemingly at random - through space.

Apart from them, there was nothing, flat nothing!

Yet, on closer inspection, it turned out that these rods displayed a few beguiling features. Not only could they be of any length, but they exhibited differing properties. Some were stark rigid ("r-rods"), with a fixed but arbitrary length ("r-length"); others were elastic ("e-rods"), with a variable length (or "e-length", e) that could be anything (with the proviso: $0 \leq e < \infty$).

(Incidentally, in this universe naught was tantamount to a 0-length rod.)

Indeed, elastic rods came in two kinds. Some, already mentioned, were wholly elastic. Others - and actually most of them - were mixed. These "m-rods" were part rigid, part elastic: their "m-length", m, was characterized by either $\{0 < \mu \leq m < \infty\}$ or $\{0 < \mu \leq m < \mu' < \infty\}$ (the lower and upper bounds μ and μ' were arbitrary and fixed for each m-rod considered).

The amazing fact was, the e-rods *could* and *did* vanish, every now and again! More baffling still was the fact that, equally easily, some of the e-rods happened to pop into existence - without warning, just out of the blue. *How and why would they pull off such an uncanny trick?* That is the question.

Now, upon a little reflection, we may wake to the idea that these astounding events were a straight consequence of the *full* elasticity of the e-rods - which by their very nature allowed for them (as a mere, if far-reaching, matter of consistency).

This sheds a crucial new light on the very problem of existence: the hint here is that (full) 'elasticity' reaches farther into the deep recesses of reality. Unlike 'rigidity', it connects with naught; it seamlessly "touches" it both ways (*i.e.* to and fro). This in turn points to the fact that 'elasticity' somehow comes first, ontogenesis-wise. 'Rigidity', by contrast, would arise later, as a derivative construct.

This raises a number of issues, such as:

- What underpins elasticity, what "pulls its strings"?
- The utter fickleness of elasticity cannot beget, no matter what, the steadiness of rigidity. How then can rigidity be arrived at, from such an unseemly background? (How to firmly build a house on quicksands? Not exactly the easiest thing to do in the world...)

In anticipation to what will follow, here are my proposed answers to the above:

- 'Elasticity' is brought about by *endo-causation;* or else, by *endo-determinacy.* (This property may hold the key to the puzzle of ontogenesis. Recall e-rods: unlike m-rods and r-rods, they are empowered to pop in and out of existence - thus bridging the forbidding gap which sets aught and naught apart...)
- 'Rigidity' likewise harks back to *exo-causation* or *exo-determinacy.* This property is tantamount to *determinism.* It does not come first: instead, it stems (somewhat wryly) from endo-causation. There is a definite process of "stiffening" at work here. (The challenge is to figure out what this process amounts to: How do we get m-rods and r-rods from e-rods?)

WORDS ARE MISLEADING

Beware of words! As Rumi aptly noticed a few centuries ago, "language is a tailor's shop where nothing fits". This is especially true where ontogenesis is concerned, because we are trying to think beyond the limits of thought. This is bound to entail stretching ordinary language beyond what it would normally bear - and bungle every attempt at thinking fair and square.

Ontogenesis deals with any world-that-be (or rather: with any world-that-*could*-be). With it, we must forget about our world and its contingent features. The hardest part still is to forget about our space-time. We must *unlearn* the very notions of space and time. Any remotely relevant ontological speculation must be explicitly untimed and 'unspaced' (or rather, it must be *space-time-less*). Otherwise it is flawed, in a deep way, at the outset.

Well, how can we do without our timed or tensed language? How for instance can we talk about causation without a timed framework, as it is understood that a cause must always *precede* its effect? (Part of the answer is that, when it comes to endo-causation, the cause and effect are one and the same thing.) No doubt, a brand new "space-time-less" language is wanted. But I shall make no attempt at it: there is no room here for this far too confusing and taxing task!

In order to steer clear of common pitfalls, I lay down a few caveats. The first one bears on 'reality'. Throughout this paper, I am concerned with *actual* (or tangible) reality, as opposed to conceptual (or abstract) reality. When talking about 'being' or 'existence', I refer to *actual* existence exclusively. (A strawberry exists in that sense, but neither an equation nor a sociological theory.)

The second caveat has to do with logic and consistency, a.k.a. non-contradiction. Admittedly, consistency is an incontrovertible requisite for any rational undertaking worth its salt. Yet it does not go without problem, since there is not *one* logic, but a batch of competing strands of them. Thus, taken at face-value, it is no longer crystal-clear what consistency amounts to. (I shan't dwell on this huge and all-important issue, would that be only because - being no logician - I am not qualified to handle it properly.)

I shall get away with one remark: some of these strands (*e.g.* the so-called 'deontic' logic) have no currency here; and in the end, there ought to be a way to pin down a notion of "logical kernel" - relevant to ontology - whereby an 'intrinsic' notion of consistency finally comes forth. (In fact, it seems that only the logic of intuitionism currently deserves an in-depth scrutiny.)

A MATTER OF CHOICE

The world *is* magical. In the face of it, a wide variety of ways to account for its existence have been contemplated.[(3)] [(7)] [(8)] Here is a glimpse of some of them:

- A flight into wild metaphysics (!), that makes no attempt to use the lights of reason. Being a matter of "leap of faith", it can be very fulfilling subjectively but is hardly enlightening, let alone compelling, on the conceptual level considered here.
- Conventional metaphysics: in most western traditions, it revolves around a personal God, who plays the alleged part of an "uncreated creator" or "uncaused cause". Thus naught is shrouded behind God's creative willpower, and the problem vanishes altogether. ("Either there is a God or the existence of the world is ultimately unintelligible.") Conceptually speaking, this all but shifts the unknowns from the scientific to the religious categories.
- The *brute fact* argument. According to it, existence as such requires no explanation. It is just a fact of experience - and that's it. (A version thereof assumes that the universe is beginningless: it has always been around, which leaves the question of its coming-into-being pointless and meaningless.) The problem is dissolved by fiat.
- The stark denial. This is one of the most efficient, if daftest, of all. It rests on the claim that "nothing does exist". Why not? "No thing, no problem"!
- The infinite regress escape route. The premiss is that anything that exists must have a cause other than itself. In seeking the 'ultimate' cause, we therefore run through an endless chain of external causes. The upshot is that the ascending causal chain grows

into an infinite regress, and no full explanation is ever at hand. We'd better give up from the start, since whatever our pig-headed efforts we'll flunk it.

- The fluctuating quantum vacuum. This vacuum comes in handy, for it is a-throbbing. It fluctuates. Taken at face value, it gives something for nothing. *Something for nothing:* in this dreamworld of mock emptiness, one stands a real chance to snap up a free lunch. (And what a whopping free lunch: a whole universe in a row!)

Let me briefly comment on the last two proposals. The infinite regress, to start with: it has a hollow ring to it; for infinity is not intrinsic. It lies in the eye of the beholder. How's that? Reason is, we can always break a *finite* causal chain down into an *infinite* yet equivalent chain of sub-causes. Conversely, it is always possible to construe or coalesce an infinite chain into a finite one.

The quantum vacuum, now. It fluctuates: fine. This makes it stand out as something really special. But the trouble here is that this vacuum is definitely not naught (in the acknowledged sense of being 'flat-out nothing'). On the contrary, it is fraught with physical laws and teems with energy. Not quite the ontological naught we have in mind here!

True to it, this vacuum is amazing. It keeps converting part of its (huge!) energy into particle-antiparticle pair creation. *(That its energy is staggeringly huge is underscored by this reckoning by Richard Feynman: "A coffee mug of vacuum energy would be enough to boil the oceans.")* When we think of it, all this is but a consequence of Einstein's discovery: *ergo*, it is less of a feat than it is trumpeted. (Remember, matter equals energy: $E = mc^2$...)

Still, quantum vacuum strikes a deep ontological chord, for it traces our intricate material world to a near-emptiness of pristine simplicity. Not quite naught, but definitely 'something' closer to nothing than our surrounding universe.

I am personally not fully satisfied with these offers, even though some of them are impressive in their own right. Part (and part only) of my dissatisfaction stems from my choice: I decided to stick to the straight and narrow path of rationality. It is certainly a painful one, but I believe that it is also one of the potentially most rewarding.

To be more specific, my approach arises from a simple question, namely: *What ideas of causation can be made ontologically relevant?* Our wonted notion of 'cause' is perhaps oversimplistic, and would therefore stand in the way of our success to get something of a conceptual toehold on the mystery of existence. If so, can we amend it? I'll take up this issue in the next Section. Meanwhile, the first order of my business is to spell out some key assumptions. Here are two of them:

(H1) There is no causeless effect in actual reality.

(H2) There is no warlock at work in actual reality.

Remarks

- **(H1)** rules out pure chance or sheer indeterminacy, taken as a genuinely 'hollow' (*i.e.* content-free) non-determinism. Were it not for **(H1)**, ontology would raise no difficulty, since reality - the effect - could be traced to just nothing (*i.e.* no cause). Too easy to be true!

- **(H2)** rules out any "warlock" or "wizard", that would explain it all without being itself amenable to a causal explanation. (By "warlock" I mean any "witchsome" agency which: a) would have the might to create a world; and: b) would lie, *allegedly or by definition*, outside the sphere of rational explanations.)

In this paper, **(H1)** and **(H2)** will hold throughout. Any chunk of reality complying with them both will be labelled *'ontologically fair'* or *fair*. (This *fairness* subsumes *consistency*:

wooden square circles, for instance, are neither consistent nor fair. No actual causal chain - which has to be consistent - can lead to their coming into being.)

THE HIDDEN KEY

With **(H1)** and **(H2)** taken on board, there isn't much leeway. If we are to explain anything at all, there seems to be just one option left, which is:

(A1) Whatever there is exists by bootstrapping itself into existence.

So, in the real world everything is there at the price of harbouring the inner resource to bring and sustain itself into being. This is no mean feat! **(A1)** throws a new, if tentative, light on ontogenesis. It narrows the range of what a being or 'thing' can be: no matter what, it *must* bootstrap or "self-breed" itself. And this could prove a harsh restriction indeed.

My hope is to reap some worthwhile information from this putative price tag of existence. *What profile this bootstrapping "self-breeder" must have to be up to the job?* It may be helpful, in addressing this core question, to reframe **(A1)** as:

(A2) Anything that be is essentially a self-breeding process.

The message is: *No self-breeding process, no actual existence.* That simple! In short, a being is an ongoing course of bootstrapping action. It is therefore dynamical: no ontologically fair being is inherently static. We must forget about static *being-things,* and embrace dynamical *being-events.*

(If something is intrinsically static, then it is 'unfair': it cannot *hence* does not *exist; or else, it is naught. This illustrates that ontological speculations, for all their far-fetchedness, are not without bearing on the practical nature of the world out there...)*

Definitions

A self-breeding or bootstrapping process can *a priori* be either fickle or steady; whence the following definitions:

- A *fickle* self-breeding process is said to be ***endo-causal*** or ***endo-determined.***
- A *steady* self-breeding process is said to be ***exo-causal*** or ***exo-determined.***

Properties

In the tangible world of real facts, only the outward features of endo- and exo-causation can be observed (not their inner content). Therefore:

- *Exo-determinacy* corresponds to a *deterministic* behaviour.
- *Endo-determinacy* goes with a *nondeterministic,* or *indeterministic,* behaviour.

In the first case, the being holds no sway over its own bootstrapping processes, it cannot change it. With nondeterminism on the other hand, the being is fully in charge. It can bring about different, varying processes; thus becoming truly unpredictable. The idea here is that causation comes in two guises, namely:

- *Exo-causation,* whereby the causal agency of existence lies outside the being at stake
- *Endo-causation,* whereby the causal agency of existence is internal to the being at stake

Besides, a determinism is an instantiation of exo-causation - little wonder, a determinism is never self-supportive - whereas endo-causation manifests itself as nondeterminism. (We usually - and mistakenly - conflate exo-causation and causation, without further ado.)

Exo-determinacy amounts to a stiffened sub-class in the broader basket of endo-determinacy. (A fickle process can always be "frozen" into a steady one.) The notion that determinism is a special case of nondeterminism thus dawns upon us. An endo-causal being has the knack of prompting and modifying whatever self-breeding process it is subject to, owing to its "decision-making" ability. In order not to sound too anthropocentric, I shall say that this "decision-making" ability arises from the "shedding of a *deed*".

A *deed* is shed, released or cast off, whenever a being tampers with its own bootstrapping processes. A ***deed*** is a sort of creative spurt that springs from within. *No deed, no change:* wholly exo-causal (or deterministic) self-breeders are 'deedless'. (Recall that **(H1)** warrants this, by ruling out any causeless effect - or, for that matter, any causeless change.)

By the same token, when an endo-determined being stops - *reversibly* - shedding deeds, its goes through a stage of mock or sham exo-determinacy. I call ***sham-determinism*** this type of endo-causation; which *reversibly* puts on the trappings of exo-causation. This notion will later prove helpful, when coming to grips with the question of the *onset of determinism.* (It addresses the issue of how *genuine* exo-determinacy can be achieved from endo-determinacy. In short, and going back to the opening tale: How can we get r-rods and m-rods from e-rods? *How can we build an unshakable house on quicksands?*)

At this juncture, a few definitions are in order *(Again, a satisfactory treatment would require a fitting formal setting, which is beyond the scope of this paper).* They are, if B tags any actual *(hence 'fair')* being:

- The deed ∂ is a *B-deed* iff *(i.e. if and only if)* it is a deed that can be released by B. (One then writes: $\partial \in \delta_B$, where δ_B is the set of all the deeds that are available to B.)
- Given a being B which is not wholly deterministic, B can change itself from one *contingent state* b (b $\in$ B) to another state b' (b' $\in$ B) (b and b' are also labelled *B-states*). It does so by throwing off a deed ∂. (This deed, available to B, is thus a *B-deed:* $\partial = \partial_B \in \delta_B$.) One has: $b' = \partial(b) = \partial_B(b)$, which amounts to say that there is a transition T - or more accurately: an *onto-transition*, meant to be both spontaneous and direct - from b to b': $\{b' = \partial_B(b) \Leftrightarrow \exists\, T,\ T = T^{(B)}{}_b^{b'}\}$. (Depending on the adopted definition, this transition - either "single-deeded" or "multi-deeded" - is either reversible or *skew-reversible*. More on these matters soon.)
- Let T_B^B be the set of all the onto-transitions which link up any couple of B-states. One defines, using easy-to-understand notations: $T_B^B = \{T^{(B)}{}_b^{b'} \mid (b, b') \in B^2\}$. (In a proper framework, $T^{(B)}{}_b^{b'}$ would be construed as an equivalence class.)
- Define Π_B as the *field of B-potentialities:* it contains whatever contingent (*i.e.* "deed-prompted") state b B can be in. In the main, it is a compound of self-breeding (or bootstrapping) processes, or sub-processes. One has:

$$\Pi_B = \{\Pi_b \mid b \in B\}; \quad \text{where in turn: } \Pi_b = \{b' \mid b' = \partial(b), \partial \in \delta_b\}.$$

- For any state b, b $\in$ B, there is a *local* neighbourhood of b, noted Π_b as above, which contains all the B-states which can be *directly* (or "single-deededly") arrived at from b. This *local* b-neighbourhood is included in the *(global)* field Π_B, which similarly contains all the B-states that can ever come into being: $\{\Pi_b \subset \Pi_B\}$. Any self-breeder b belonging to Π_B (b $\in$ B) also belongs to at least one local sub-field $\Pi_{b'}$ (where as usual: $b' \in B$, $\{\exists\, \partial = \partial_{b'}: b = \partial(b')\}$). Now Π_b proper ($\Pi_b \neq \Pi_{b'}$) comprises all the B-states that can be reached forthwith from b, and these only: $\Pi_b = \{b' \mid b' \in B, b' = \partial_b(b)\}$.
- Π_B could formally be defined as the (indeed, *topologically loaded)* space of all the B-states. Any contingent state of B, call it b, is such that: $\{b \in \Pi_b \subset \Pi_B\}$. One could similarly define δ_B as the (dual) space of all the B-deeds, whereby any deed - say, ∂ - that can be shed by B, whilst in the state b, verifies: $\{\partial \in \delta_b \subset \delta_B\}$. *(B-states and B-deeds are somehow mutually* dual, *as a full presentation would show. One can define a*

correspondence law leading to equivalences such as: $\{\partial \equiv b;\ \delta_b \equiv \Pi_b$ *(local);* $\delta_B \equiv \Pi_B$ *(global)*$\}$.)

- In like manner, define Π as the (reality-wide) *field of potentialities,* meant to contain whatever 'fair' or actual self-breeder there can ever be. One has: $\{\forall B,\ \Pi_B \subset \Pi\}$ ($D_B = n$ or $D_B \neq n$). Define Π_W as well in the restricted sense of the *field of the purely endo-causal potentialities* (even though W is larger than that - and in a definite way encompasses the *whole* of Π, as we shall make out).

To sum up this Section, my suggested 'hidden key' to ontogenesis is the notion of *endo-causation.* It ties in with the kindred notion of *deed,* which refers to some inward, "self-willed action". ("Self-willed" is to be understood in an ontological way!) A *deed* is a creative act *par excellence.* Most naturally, it jumped on board here....

THE FOUR FACES OF WHAT-THERE-IS

In view of the foregoing, a being can either be endo-determined or exo-determined - or both. Accordingly, one can write the generic being, say, B, down as: B = (S, D); where S and D are respectively the endo- and exo-causal parts. (S is self-determined whereas D is deterministic.) I shall occasionally write $B = (S_B, D_B)$ instead of B = (S, D).

One draws from this simple formula four differing kinds of beings or self-deeders, depending on whether S and D are really there or not. *(If B is bereft of either S or D, I shall write "n" instead of S or D. So, unless otherwise specified, B = (S, D) normally implies:* $\{S \neq n,\ D \neq n\}$*.)* These four optional kinds - or "faces" - of self-deeders are:

B = (S, D) :	B is an (ontological) ***stuff***	[with $\{S \neq n, D \neq n\}$]
B = (n, D) :	B is a *degenerate stuff*	[with $D \neq n$]
B = (S, n) = W :	W is an *ur-stuff,* or ***whiff***	[with $S = S_W \neq n$]
B = (n, n) = N :	N is (the ontological) ***naught***	

Remarks

- N, *naught* or utter nothingness, is by definition the absolute lack of any self-breeding process - be it fickle or steady. (It is the state of "deedlessness"; it is therefore a completely static, or inherently motionless, state.) A key, if eerie-sounding, question will later turn out to be: *Is naught reversibly or irreversibly deedless - and why?*
- The idea of - generally, nondegenerate - *stuff* has a pointed relevance to *our* universe. I shall go back to this later on.
- The ontological *whiff,* W, stands out as very special indeed. Its properties and features - which, I bet, can be partly unravelled on logical grounds - are likely to be odd and counter-intuitive. (Human intuition is honed on everyday objects, it is out of its depth when dealing with W....)
- As a starter, let us take note that W is virtually restless and ever fluctuates, has no abiding core and thereby cannot stash any bit of information whatsoever - and this includes the one regarding its own identity! (This flies in the face of the principle of identity, a cornerstone of conventional *exo*-logic that cannot be consistently upheld here. One needs to evolve new strains of *endo*-logic, that would explicitly allow for endo-causation.)
- Concerning the *unicity* or otherwise of W, here is how this issue can be handled (lack of identity notwithstanding). W is *unique* if we choose to adopt a "B-intrinsic" definition of

two beings B and B' whereby: $\{B' \neq B \Leftrightarrow D_{B'} \neq D_B\}$. It is not so if we endorse instead a "B-contingent" definition which takes into account the diversity of states arising from S_B as it changes itself, by means of B-deeds (or for that matter, of W-deeds).

REALITY LAID BARE

From the four faces of any possible self-breeder, we proceed to the four basic results hereafter. They are what I call the fundamental "onto-theoretical theorems": *(Here B is the generic being - regardless of whether S_B and D_B are identical with n or not.)*

(R1) If $D_B \neq n$, no transition from B to N is available.

(R2) If $D_B \neq n$, no transition from N to B is available.

(R3) If $D_B = n$, the being B can transform itself into naught.

(R4) If a being can transform itself into naught, the converse holds true.

Remarks

- Forewarning. *All the above transitions, transformations and the like, are implicitly taken to be* spontaneous *and* direct. *They are* spontaneous *inasmuch as they are strictly endo-determined or endo-causal. They are* direct *in that they involve no* needful *intermediary which would be yet another self-breeder, different from those explicitly brought into play. What this really means will I hope become clearer in due course.*

- **(R1)** translates formally as: $\forall B, B = (S_B, D_B), \ \{D_B \neq n \ \Rightarrow \ \neg \exists T_B^N\}$. [The string of symbols " $\neg\exists$" means: "not (there exist)", *i.e.* "there is no". As for T_X^Y for instance, it refers to any *spontaneous* and *direct* transition from a being X to another self-breeder Y. By definition it is to be achieved by X-deeds, that is, by deeds arising from X exclusively: it is wholly *X-deeded.*]

  (R1) underlines that whatever the *stuff* B, it cannot wipe itself out. No matter what, it is firmly "hooked" to existence, in the sense that N is unflinchingly out of its reach. [This plainly results from the stable or unwavering character of the exo-determined or deterministic part, D_B, with respect to B-deeds, or to S_B. It is impossible for B to *endo*-modify D_B into n. One ever has: $B = (S_B, D_B) \neq (n, n) = N$, since $D_B \neq n$.]

- **(R2)** is reciprocal to **(R1)**. It asserts the arrant impossibility, for any *stuff*, to spring forthwith up from naught. The whole enigma of ontogenesis hinges precisely on that! It underscores the shortcoming, or the ontological weakness, of any determinism (interpreted as exo-determinacy); which cannot stand on its own (it is not self-sufficient). This is why an element of nondeterminism - pointing to some deeper endo-determinacy or endo-causation - must be added to reality, if the riddle of existence is ever to be made intelligible.

  As before, **(R2)** can be recast as: $\forall B, B = (S_B, D_B), \ \{D_B \neq n \ \Rightarrow \ \neg\exists T_N^B\}$. [Again, the impossibility bears exclusively on *spontaneous* and *direct* transitions. It does not rule out indirect (*i.e.* combined) transitions between N and any *stuff* B ($D_B \neq n$). These more intricate - *i.e.* not fully *N-deeded*, here - transitions are indeed possible; they happen to involve W. (More about this shortly, with the "two bangs".)]

- **(R3)** reads: $\forall B, B = (S_B, D_B), \ \{D_B = n \ \Rightarrow \ \exists T_B^N\}$. It says that the *ur-stuff*, W, can spontaneously turn itself into naught. [In case B = N, **(R3)** is a truism.] The *ur-stuff*, or ontological *whiff*, can wipe itself out clean. In plain language: it can commit suicide. (This "suicide" is thorough: it leaves no corpse nor anything behind!) We thus awaken to the most welcome idea that the *ur-stuff* can - unlike any *stuff* - touch or graze naught.

Hence its decisive ontological worth: it spans a crucial part of the path which links up aught and naught. [I call such a path - which will soon turn out to be a two-way lane - a ***genesic path***, or ***gen-path*** for short. It is meant to be explanatory and causal.]

- No in-depth treatment of **(R3)** is possible without a formal framework. Even though I don't provide it, I shall hint at one key item, namely: *N is no other than a contingent state of W.* (To wit: $\{N \in W\}$.) It is so because the (local) field of N-potentialities, Π_N, belongs "forthwith" to Π_W (*i.e.* $\{\Pi_N \subset \Pi_W\}$), in the sense that N-deeds are free from any exo-causal yoke. (More generally: $\forall \omega \in W, \Pi_\omega \subset \Pi_W$.) The point is that N is not the barren state of utter deedlessness it's cracked up to be.

(Recall that, on shedding a deed - say, ∂, $\partial \in \delta_B$ - the being B changes its inner or contingent state. *It brings about a new state, call it b', which differs from the state b that B would be in, were it not for ∂. In short:* $\{b \in B, \partial \in \delta_B, b' = \partial(b); b' \in B, b' \neq b\}$. *Moreover:* $\forall b \in B, \{b \in \Pi_b \subset \Pi_B\}$; $\forall b', b' = \partial(b), \{b' \in \Pi_b, b' \in \Pi_{b'} \subset \Pi_B, \Pi_{b'} \neq \Pi_b, \Pi_{b'} \cap \Pi_b \neq \varnothing\}$. *As an aside, that N-deeds do exist at all is no more befuddling that the existence of ω-deeds for any W-state ω. This would ensue naturally from a more thorough study of what a deed stands for. Roughly speaking, the reason is that, as a* creative spurt, *a deed begets something brand new: in a way, it always brings about something for nothing. It does so regardless of or beyond its actual "ontological womb" - be it aught or naught. The only contextual aspect that truly matters is the optional element of exo-causation, which drastically changes the overall picture.)*

- **(R4)** claims that the self-annihilation skill (in the ontological sense of a downright 'snuffing out' of existence) goes with that of self-creation; and vice-versa. This, along with $\{$**(R1)**, **(R2)**, **(R3)**$\}$, shows that we are facing an idiosyncratic attribute of the *ur-stuff:* W is blessed with the reverse and equivalent properties of "W-thanasia" and "W-genesis". This points anew to the paradoxical nature of W (which can freely "decide" to vanish totally, only to bounce back, unscathed, into existence).

All in all, the *whiff* ever fluctuates between aught and naught. Touching naught, as far as it is concerned, does not mean the terminal kiss of death. Its ontological status, to put it mildly, is rather equivocal: it does not "exist fully", on account of its ability to graze naught and stay there, at whim. On this ground, I also dub it a ***be-able*** or ***beable*** (after the late John Bell, who coined the word in another context).

- The transitions of the type T_X^Y brought into play in $\{$**(R1)**... **(R4)**$\}$ are to be defined as:

$$(1) \qquad T_X^Y = \{T^{(X)}{}_x^y \mid (x, y) \in X.Y, y = \partial(x), \partial \in \delta_x \subset \delta_X\}$$

(Actually T_X^Y is a tad subtler. Rather than $\{y = \partial(x), \partial \in \delta_x \subset \delta_X\}$, a more accurate, "multi-deeded", expression will - as shown in (6) - turn out to be: $\{y = \partial_{y\otimes x}(x)\}$. Now, **(R1)** and the like assert that, given $\{B = (S_B, D_B), D_B \neq n, \alpha \in B\}$, any β such that $\{\beta \in B', B' = (S_{B'}, D_{B'}), D_{B'} \neq D_B\}$ is "off-limit" or *unreachable*. This formally reads as:

$$(2) \qquad \forall (\alpha, \beta), \alpha \in (S_B, D_B), \beta \in (S_{B'}, D_{B'}), D_B \neq n, \{D_{B'} \neq D_B \Leftrightarrow \neg\exists T_\alpha^\beta\}.$$

In a nutshell: *a given* stuff *cannot reach any contingent state of another 'fair' being, when their exo-causal parts are dissimilar.* (No *onto-transition* is available between them.) 'Stuff-zapping' or 'stuff-hopping' does not come spontaneously! (This once more boils down to the fact that D_B is by definition *stable* with respect to the B-deeds: no deed shed by B - and really, by S_B - can jostle D_B into becoming a different compound of self-breeding processes.)

- Incidentally, we as a rule have hardly any qualm about "W-thanasia" (*i.e.* the existence of the transition T_W^N), but our intuition is decidedly ruffled when it comes to "W-genesis" (*i.e.* T_N^W). "How - we may wonder - something that *does not even exist yet* can nevertheless bootstrap itself into being? Baloney! This cannot make sense." *(Really, what does not make sense here is, first and foremost, the* time-arrowed *language used, which conveys a misleading and mistaken image.)*

- The above results {**(R1)**, **(R2)**, **(R3)**, **(R4)**} can be summed up in one formula, which is:

(3) $$\forall B, B = (S_B, D_B), \ \{\exists T_N^B \Leftrightarrow \exists T_B^N \Leftrightarrow D_B = n\}.$$

If all this is not totally wide of the mark, then we reaped the worthwhile insight that exo-causation is *the* bulwark against naught. It is the key to securing existence in the strong, unswerving and clear-cut sense of being a world apart from naught.

Beable versus full-blown being: what a striking case of "less is more"! *Less* by way of actuality (*viz.* of actual, steadfast existence); and (boundlessly and fathomlessly) *more* by way of virtuality. (Unlike B if $D_B \neq n$, W fills the whole field of potentialities: it is "reality-wide".)

EXISTENCE COMES WITH TWO BANGS

In light of the foregoing, a pattern readily springs to mind. Granted {**(R1)** to **(R4)**}, it takes little thinking to see that a two-step "rundown" of ontogenesis emerges almost naturally, in which ontogenesis becomes a twofold process, made up of both a *wee bang* and a *big bang*. *(Existence doesn't come with a bang, but with two!)*

A ***wee bang*** is a transition between naught and a beable. (It is no other than T_N^W.) Similarly, a ***big bang*** is any transition T_W^B between W and B, where $D_B \neq n$. (Needless to say, this notion is not to be confused with the cosmological Big Bang - the latter regards *our* peculiar world whereas the former bears on *any* ontologically 'fair' universe.) The ontological *big bang* is the coming into being, straight from W (and *not* from N!), of a concrete universe - name it U - composed of one or an arbitrary number of ontological *stuffs*, noted B_1, B_2, B_3, etc.... I shall jot such a "many-stuff" universe down as: $U = B_1 \oplus B_2 \oplus B_3 \oplus ....$

On the whole, one gets the following:

- *wee bang* : T_N^W [reversible *and* direct *transition between N and W*]
- *big bang* : T_W^B or T_W^U $(U = B_1 \oplus B_2 \oplus B_3 \oplus ...)$ [*transition from* W *to* U; *and in fact: from* W *to* W⊕U, *for* W *is not necessarily wiped out in the process!*]

The resulting two-step evolution reads:

$$\{\text{"B-genesis"} \equiv T_W^B \otimes T_N^W\}.$$ ("⊗" *symbolizes an appropriate composition law.*)

A *wee bang* is a *reversible* swing or fluctuation. It is reversible with respect to both W and N - it is at once *W-reversible* and *N-reversible.* (Actually N-reversibility is but a consequence of W-reversibility since N is just a particular occurrence, or state, of W.) A *big bang*, on the other hand, is *U-irreversible* and *W-reversible.* (The *big bang* transition is *exogenous* to the universe U, *i.e.* to any one of its *stuffs* B_i (i = 1, 2, 3,...); it is *endogenous* to the *ur-stuff.* It falls irreversibly outside the preserve of U, while it can still be undone by a suitable W-deed.)

The *gen-path*, tagged $\Gamma = \Gamma(B)$, of any *stuff* B ($D_B \neq n$) - or more generally, that of any deterministic universe U, $\Gamma' = \Gamma(U)$ - is therefore *indirect* or *combined*, since it always breaks down into two sub-paths, one being the *wee bang* and the other, the *big bang*. This is summarized by the following formula, which is the *(genesic)* ***inclusion formula***:

(4) $$\Gamma(B) = T_W^B \otimes T_N^W \quad (D_B \neq n).$$ *(The notation* $T_W^B \otimes T_N^W$ *is to be read backwards: from the right to the left,*

from N towards B through W.)

Such a $\Gamma(B)$, which links up naught with B and yet shuns any ontologically 'unfair' transition of the form T_N^B, embodies *prima facie* a consistent answer to the problem of ontogenesis. ($\Gamma(B)$ is 'fair', because: *a)* it complies with **(H1)**, which rules out pure chance; *b)* it is not a direct transition, in keeping with **(R2)**.)

The *wee bang* is the missing link of ontogenesis!

However, at least one point must be clarified. I'll do it right away, without working out all the details. It arises from a looming ambiguity, as we now see. On the one hand, it turns out that N is but a contingent state of W ($N \in W$; **(R3)** and **(R4)** bear this out). This beclouds the meaning of T_N^W and blurs the distinction between T_W^B and T_N^B. *(Aren't they one and the same thing?)* On the other hand, T_N^B - *but not* T_W^B - is adamantly ruled out. How to reconcile the two? How can we make sure that T_W^B is not T_N^B by another name? (Likewise, how can we tell, say, T_N^W, T_W^W and T_N^N apart?)

To sort all this out I must give some more details about my approach. To that end, recall that by (1) a transition T_X^Y is really a set of individual *onto-transitions* $\{T_x^y, T_x^y = T^{(X)}{}_x^y\}$ triggered by X-deeds (*viz.* by deeds ∂ stemming from X: $\{\partial \in \partial_x \subset \partial_X\}$; with plainly:

$\delta_X = \{\partial \mid \partial \in \delta_x, x \in X\}$). The upshot is to shift a X-state x ($x \in X$) into a state y.

Moreover:

- All the $T^{(X)}{}_x^y$ considered are *direct*, which reads:

$$(5) \qquad \forall (x, y) \in X.Y, \ \Big\{\exists T^{(X)}{}_x^y \Leftrightarrow \exists \{\partial, \partial', \partial'', ...\partial^{[n]}\} \in \delta_x.\delta_{x'}.\delta_{x''}..... \subset (\delta_X)^{n+1}: \\ x' = \partial(x), x'' = \partial'(x'), x''' = \partial''(x''), ..., y = \partial^{[n]}(x)\Big\}.$$

(This all but forbids any ∂ such that: $\partial \not\subset \delta_X$. Indeed, such an "un-X"-deed ∂ ($\partial \not\subset \delta_X$) could well be part of a particular pathway going from x to y: this wouldn't really matter as long as *there exists a* direct *transition* $T^{(X)}{}_x^y$ *- in the above sense - that binds them.*)

- Given two *stuffs* X and Y, any $T^{(X)}{}_x^y$ is *reversible*, that is to say:

$$\forall (x, y) \in (X, Y), \{\exists T^{(X)}{}_x^y \Leftrightarrow \exists T^{(Y)}{}_y^x\}; \quad \text{since:}$$

- *either* X = Y; in which case, roughly speaking one draws the expression of $\{T^{(Y)}{}_y^x, T^{(Y)}{}_y^x = T^{(X)}{}_y^x\}$ from that of $\{T^{(X)}{}_x^y\}$ by putting something akin to: $\{\partial^{[n]}, ...\partial'', \partial', \partial\}$ instead of: $\{\partial, \partial', \partial'', ...\partial^{[n]}\}$ *(no further details here)*;
- *or* $X \neq Y$; in which case both $T^{(X)}{}_x^y$ and $T^{(Y)}{}_y^x$ simply do not exist.

(It is worth noting that: $\neg\{T^{(X)}{}_x^y = T^{(x)}{}_x^y\}$, *since for the opposite to hold true, the above definition of* $T^{(X)}{}_x^y$ *should read:* $\{\delta_x.\delta_{x'}.\delta_{x''}..... \subset (\delta_x)^{n+1}\}$ *instead of:* $\{\delta_x.\delta_{x'}.\delta_{x''}..... \subset (\delta_X)^{n+1}\}$.)

- Define and label "$\partial_{y\otimes x}$" any combined X-deed such that:

$$\{(x, y) \in X.Y; \{\partial, \partial', \partial'', ...\partial^{[n]}\} \in \delta_x.\delta_{x'}.\delta_{x''}..... \subset (\delta_X)^{n+1}; x' = \partial(x), x'' = \partial'(x') = \partial'(\partial(x)), \\ x''' = \partial''(x'') = \partial''(\partial'(\partial(x))), ..., y = \partial^{[n]}(x) = \partial^{[n]}(\partial^{[n-1]}(\partial^{[n-1]}(...(\partial(x))...))) = \partial_{y\otimes x}(x)\}$$

- With this, one finally arrives at the full definition of T_X^Y:

$$(6) \qquad T_X^Y = \{T^{(X)}{}_x^y \mid (x, y) \in X.Y, X = (S_X, D_X), Y = (S_Y, D_Y), y = \partial_{y\otimes x}(x)\}.$$

As for the transitions T_N^B and T_W^B, where $\{B = (S_B, D_B), D_B \neq n\}$, one can now write:

$$(7) \qquad T_N^B = \{T^{(N)}{}_N^b \mid (N, b) \in (W, B), b = \partial_{b\otimes N}(N)\}.$$

$$(8) \qquad T_W^B = \{T^{(W)}{}_\omega^b \mid (\omega, b) \in (W, B), b = \partial_{b\otimes\omega}(\omega)\}.$$

Is the difference between (7) and (8) a genuine or a spurious one? It is genuine, since formula (7) is not on a par with formula (8). The reason is, N is merely a state - like b but unlike W or B ($N \in W$). So $\partial_{b \otimes N}(N)$ goes with the proviso: $\{\{\partial, \partial', \partial'', ...\partial^{[n]}\} \in \delta_N.\delta_{\omega}.\delta_{\omega'}.... \subset (\delta_N)^{n+1}\}$, instead of: $\{\{\partial, \partial', \partial'', ...\partial^{[n]}\} \in \delta_{\omega}.\delta_{\omega'}.\delta_{\omega''}.... \subset (\delta_W)^{n+1}\}$ as would have been the case were N a being - as W in (8). (The reader will have realized that the gist of the matter is that, in (7), one reads $T^{(N)}{}_N{}^b$ instead of $T^{(W)}{}_N{}^b$.)

Herein lies the root of the discrepancy between $T_N{}^B$ and $T_W{}^B$. The nub is that, generally speaking, for most couples $\{(\omega, \varpi), (\omega, \varpi) \in W^2, \omega \neq \varpi\}$, one has: $\{\Pi_{\omega} \neq \Pi_{\varpi}\}$; or equivalently: $\{\delta_{\omega} \neq \delta_{\varpi}\}$. One accordingly gets, for most W-states ω (for which: $\omega \in W$, $\omega \neq N$): $\{\delta_N \neq \delta_{\omega}\}$ (with: $\delta_N \subset \delta_W$, $\delta_{\omega} \subset \delta_W$). Whether one has: $\{\delta_N.\delta_{\omega}.\delta_{\omega'}.... \subset (\delta_N)^{n+1}\}$, or: $\{\delta_{\omega}.\delta_{\omega'}.\delta_{\omega''}.... \subset (\delta_W)^{n+1}\}$ hence *does* make a difference. *(It goes without saying, by the way, that the* $\{\omega, \omega', \omega'', ...\}$ *appearing on both sides are totally unrelated.)*

$T^{(N)}{}_N{}^b$ is all but impossible, as we know from **(R2)**. (Remember: $\{\forall (B, b), B = (S_B, D_B), D_B \neq n, b \in B, b \notin \Pi_N\}$.) Fortunately, it is not so for $T^{(W)}{}_N{}^b$ - thanks precisely to *wee* and *big bangs*. Now, it takes something very special to elicit a *big bang*. My (educated) guess is that the W-deeds involved are far too sophisticated to belong to δ_N. It ought to take a comparatively high degree of sophistication for W to yield a deterministic universe. (This is what **(R2)** implicitly contends.) A *wee bang* is thus required: to trigger a *big bang* - and therefore spawn a slew of exo-causal beings - W must first 'hoist' itself into a state ω such that:

(9) $\exists \partial, \partial \in \delta_{\omega}$: $\{\exists (B, b), B = (S_B, D_B), D_B \neq n, b \in B, \partial(\omega) = b\}$.

This leads me to introduce two new notions:

- The ***nimble area***, noted $\Pi_{n\text{-}a}$, is the cluster of all the W-states ω from which a *big bang* is possible ($\Pi_{n\text{-}a} \subset \Pi_W$). It is defined as:

(10) $\Pi_{n\text{-}a} = \{\omega \mid \omega \in W, \exists (B, b), B = (S_B, D_B), D_B \neq n, b \in B, \exists T^{(\omega)}{}_{\omega}{}^b\}$

- The ***dumb area***, noted $\Pi_{d\text{-}a}$, is the part of Π_W ($\Pi_{d\text{-}a} \subset \Pi_W$) from which no *big bang* is ever possible:

(11) $\Pi_{d\text{-}a} = \{\omega \mid \omega \in W, \forall (B, b), B = (S_B, D_B), D_B \neq n, b \in B, \neg \exists T^{(\omega)}{}_{\omega}{}^b\}$

(Please take note that we deal here with $T^{(\omega)}{}_{\omega}{}^b$*, and* not *with* $T^{(W)}{}_{\omega}{}^b$*.)*

One gets, using conventional symbols: $\{\Pi_{n\text{-}a} \cap \Pi_{d\text{-}a} = \varnothing\}$ and $\{\Pi_{n\text{-}a} \cup \Pi_{d\text{-}a} \subseteq \Pi_W\}$. *(However, let us not be fooled: these notions and relations are much trickier than it seems. For instance, problems such as Gödelian undecidability are likely to crop up fairly swiftly when dealing with them....)*

A given ω ($\omega \in W$), is likewise said to be *nimble* iff $\omega \in \Pi_{n\text{-}a}$ and *dumb* iff $\omega \in \Pi_{d\text{-}a}$. One similarly defines the (dual) *nimble* and *dumb* entities $\delta_{n\text{-}a}$ and $\delta_{d\text{-}a}$. My hunch is that - unsurprisingly - N is positively *dumb!* That is: $N \in \Pi_{d\text{-}a}$. I even surmise, rather more strongly, that:

(12) $\Pi_N \subset \Pi_{d\text{-}a}$

The reason for (12) would be that around N, W is far more smoke than fire. It would be exceedingly flimsy, or well-nigh insubstantial. It would hardly be more than a ghoulish whiff of an illusion..... Interestingly enough, this cuts the riddle of W-genesis down to size: it shrinks it into a lesser enigma! (This "lesser enigma" is embodied by the transition T_N^{ω}, where: $\{\omega \neq N, \omega \in \Pi_N\}$.)

We now understand why a *wee bang* cannot be a transition of the type $T^{(N)}{}_N{}^{\omega}$. In order for $T^{(N)}{}_N{}^{\omega}$ to be a *wee bang* - that is, to be part of the *gen-path* $\Gamma(B)$, $\Gamma(B) = T_W^B \otimes T_N^W$ or more explicitly, $\Gamma(B) = T_{\omega}^b \otimes T_N^{\omega}$ (with $\omega \in W$, $b \in B$, $D_B \neq n$) - the contingent state ω ought to be *nimble*. By (12) this rules out the case $\omega = N$; and more generally, the case $\omega \in \Pi_N$.

In light of the foregoing and given (7) and (10), one defines a *wee bang* as:

(13) $T_N^W = \{T^{(\bullet)}{}_N{}^{\varpi} \mid (N, \varpi) \in W^2, \varpi \in \Pi_{n\text{-}a}\}$;

where $T^{(\bullet)}{}_N{}^{\varpi}$ is any string of "single-deeded" W-transitions such as $\{T^{(\omega''')}{}_{\omega'''}{}^{\omega''''}\}$, with:

$$\{\omega = N, \omega' = \partial_{\omega}(\omega), \omega'' = \partial_{\omega'}(\omega'), \omega''' = \partial_{\omega''}(\omega''), \ldots,$$
$$\omega'''''''' = \partial_{\omega'''''''}(\omega''''''') = \varpi; \varpi \in \Pi_{n\text{-}a}\}.$$

Granted the duality (mentioned earlier on) between states and deeds, the string $T^{(\bullet)}{}_N{}^{\varpi}$ is equivalent to the string of the above W-deeds $\{\partial_{\omega'''}\}$. I shall henceforth call $T^{(\bullet)}{}_N{}^{\varpi}$ - which is formally an equivalence class of *combined* individual transitions such as $T^{(\omega''')}{}_{\omega'''}{}^{\omega''''}$ - an *ur-transition*. More generally, one defines the ***ur-transition*** $T^{(\bullet)}{}_{\omega}{}^{\varpi}$, from a given W-state ω, as (the equivalence class of) any transition of the type:

(14) $T^{(\bullet)}{}_{\omega}{}^{\varpi} = T^{(\omega'''')}{}_{\omega''''}{}^{\varpi} \otimes T^{(\omega''')}{}_{\omega'''}{}^{\omega''''} \otimes \ldots \otimes T^{(\omega')}{}_{\omega'}{}^{\omega''} \otimes T^{(\omega)}{}_{\omega}{}^{\omega'}$; with $\varpi \in \Pi_{n\text{-}a}$.

One accordingly defines $\partial_{\omega' \otimes \omega}$ by:

$$\forall(\omega, \partial_{\omega}, \partial_{\omega'}) \in W.\delta_{\omega}.\delta_{\omega'}, \quad \omega'' = \partial_{\omega'}(\omega') = \partial_{\omega'}(\partial_{\omega}(\omega)) = \partial_{\omega' \otimes \omega}(\omega).$$

Lastly, $\{\omega'' = \partial_{\omega' \otimes \omega}(\omega)\}$ becomes: $\{\omega'' = \partial_{\bullet}(\omega)\}$ when $\omega'' \in \Pi_{n\text{-}a}$. (Such a $\partial_{\bullet}$ is christened an ***ur-deed***.)

Remarks

- As we know, the notation $\{T^{(\omega')}{}_{\omega'}{}^{\omega''} \otimes T^{(\omega)}{}_{\omega}{}^{\omega'}\}$ runs backwards. It pinpoints any evolution of W which starts off at ω, ends up at ω'' and goes through the halfway W-state ω' (with: $\{(\omega, \omega', \omega'') \in W^3, \omega' = \partial_{\omega}(\omega), \omega'' = \partial_{\omega'}(\omega')\}$). This expression can also be shortened into: $T^{(\omega'' \otimes \omega)}{}_{\omega}{}^{\omega''}$; it further becomes $T^{(\bullet)}{}_{\omega}{}^{\omega''}$ when $\{\omega'' \in \Pi_{n\text{-}a}\}$, and - given (13) and (14) - is a *wee bang* when moreover $\{\omega = N\}$.
- Incidentally, we shall say that the ontological whiff, W, is by definition *skew-reversible* because:

  a) $\forall \omega \in W, \exists \varpi \in W: \neg \{\exists \partial_{\omega}, \partial_{\omega}(\omega) = \varpi\}$ (put otherwise: $\varpi \notin \Pi_{\omega}$; one obviously gets: $\varpi \neq \omega$, and, more significantly: $\{\forall \omega \in W, \Pi_{\omega} \neq \Pi_W\}$);

  b) $\forall(\omega, \varpi) \in W^2, \exists \{\omega, \omega', \omega'', \omega''', \ldots, \varpi\}: \varpi = \partial_{\omega'''''''' \otimes \omega}(\omega)$ (short for: $\varpi = \partial_{\omega'''''''' \otimes \ldots \otimes \omega'' \otimes \omega' \otimes \omega}(\omega)$. (Plain reversibility would require: $\{\forall(\omega, \varpi) \in W^2, \varpi \in \Pi_{\omega}\}$; and, as a consequence: $\{\forall \omega \in W, \Pi_{\omega} = \Pi_W\}$.This is a far more stringent condition!)

I hope that the meaning of T_N^W (as opposed to T_N^N) is by now clear, on the understanding that any transition T_N^{ω} such that $\{\omega \in \Pi_N\}$ - and more generally: $\{\omega \in \Pi_{d\text{-}a}\}$ - falls short of allowing any *big bang* to take place.

But are we any wiser concerning the nature of a *big bang?* Not quite. We just know that in order to make sense, the *big bang* T_W^B, say, ought to be defined as: $T_W^B = \{T^{(\omega)}{}_\omega{}^b \mid \omega \in W, \omega \in \Pi_{n\text{-}a}, B = (S_B, D_B), D_B \neq n, b \in B\}$. (The presence of $T^{(\omega)}{}_\omega{}^b$ implies as we know that: $\{\exists\partial, \partial \in \delta_\omega: b = \partial_\omega(\omega)\}$.) This still gives us no clew as to what the underlying phenomenon consists in (this phenomenon is the onslaught of determinism).

Fortunately, the earlier notion of *sham-determinism* will come to our rescue. Recall that a spell of *sham-determinism* shows up whenever the *ur-stuff* ushers itself into a phase of thorough (albeit reversible) *deedlessness.*

Admittedly, full-blooded exo-determinacy takes a little more than *sham-determinism.* What is lacking, then? My guess is that exo-determinacy (i.e. *true* determinism) is the outcome of a process of *withdrawal* or ***withholding***, whereby W manages - in this former offshoot of itself which is to become a deterministic *stuff* B or a deterministic universe U - to chop off a measure of its primeval *full* endo-determination.

Additional remarks

- This process of *withholding*, carried out 'at will' by W (from any state ω, $\omega \in \Pi_{d\text{-}a}$), can just as easily be undone: on the whole, it is W-reversible. (*Unless*, that is, W drifts - *skew-reversibly* - into a state ω' such that $\omega' \in \Pi_{n\text{-}a}$. *Could this very fact turn sham-determinism into a more genuine streak of exo-determinacy?)* On the other hand it is distinctly B-irreversible, or U-reversible if we consider the broader case of a *big bang* giving rise to a "many-stuff" universe U.
- It is all-important in these matters to distinguish sharply between exo-determinacy as "seen" from W and as "seen" from B or U. As "experienced" by the resulting deterministic being (B or U), the *withholding* manifests itself as an unyielding stiffened process of self-breeding. It goes with a 'hamstrung' (or even 'blotted out', if $S_B = n$) knack of self-determinacy: no amount of B-deed can override D_B ($D_B \neq n$).
- By and large, this is not the whole story. But at least *sham-determinism* gives us an auspicious lead to elaborate on. It shows that a fickle self-breeding process can bring forth a steady one, by dint of sheer deedlessness. This is something which seemed all but impossible at first sight.

I believe that we can push our thinking about the *big bang* much beyond that point, although I shall leave it here at least for now. At this stage let us go back to the overarching question: *Why is there aught rather than aught?* My proposed solution is roughly as follows:

a) The first half of the answer lies in the fact that an ontologically relevant naught is by no means a staunch, stark, barren emptiness. Instead, it is 'something' (so to speak!) subtler and richer. It is a particular inner state of a larger whole - of no less, that is, than the fully endo-causal *whiff* or *beable.* And this *beable*, as long as it wafts or dawdles in the neighbourhood of naught, is so dim and spooky that it is hardly anything at all!

b) The second half of the answer rests on the insight that the *whiff* can eventually "outwit naught", as it were. It does so by turning itself into a more 'heavyweight' *beable*, by bootstrapping itself past a threshold beyond which it can eventually brew and strew an exo-causal world. This world, conditional on something external to it, is thus partly or wholly deterministic. It embodies *the* real leap into full-blown existence - achieved at the hefty price of a drastic loss in the range of available possibilities....

BACK TO OUR WORLD

Ontology is a way of thinking about reality whose primary goal is to understand why reality is there at all, and whose secondary goal is to throw new shafts of light on the inner working of actual things. As such, it should pave the way for a broader and deeper understanding of *our* particular world. So the question is: Have the preceding speculations any practical weight or relevance regarding our physical world?

The answer is positive, at least on three counts. *For one,* it tells us couple of things about what a speck of matter (call it a 'microparticle', or particle) *cannot* be; *for two,* it may help to elucidate what quantum mechanics stands for; *for three,* it is poised to put the space-time arena in its right perspective.

The first aspect is straightforward. It follows from (**A2**) that any actual being (*e.g.* a photon, an electron, etc.) is a self-breeder. As an on-going bootstrapping process, it is thereby *inherently dynamical.* So a particle can be seen as an instantiation of some inward "intrinsic motion" - without which we'd get only naught.

The message, in short, is that no intrinsically static *object*, or *being-thing* - endowed with both *solidity* and a clear-cut "thisness" - can possibly exist. Instead, reality is made of more elusive *being-events.* "Things" are, deep down, streams of events.

This, to begin with, rules out *classical* particles - those teensy-weensy snooker balls, occasionally set in motion but inwardly still, or motionless. Such beings just cannot exist! Being ontologically 'unfair', they are fatally flawed (and as anyone knows, the classical or newtonian view failed to stand the test of time).

Quantum waves on the other hand come much closer to measure up to the ontological yardstick of inner motion: even a standing wave is the outcome of a dynamical process. True to them however, they are no classical, well-behaved waves. By a weird quirk of their own, they qualify as *parallable.* (I portrayed the *parallable* quantum wave at some length in my *ANPA 17* paper.)[25]

Along the same line of thought, one gets a fresh inkling of what the so-called "wave packet" (or wave function) of quantum mechanics is really up to. As ontologically fair, this packet ought to be an on-going self-breeding - or, for that matter, self-reproducing - process. Luckily enough, this is clearly borne out by Richard Feynman's path integrals approach, which describes the wave-packet evolution by drawing on Huygens' principle.[9] [28] (This evolution shuts down during the onslaught of what is known as a 'wavepacket collapse', for reasons soon to be uncovered - a deed is unleashed!)

Going one step further, it requires scant insight to see that the idea of ***psychomatter*** (as sketched out in my former *ANPA* papers)[23] [25] fits the overall scheme B = (S, D). Recall that psychomatter, as I view it, exists under two guises, namely: that of ***matter*** proper, and that of ***paral.***

Matter is the state psychomatter is in, whenever its 'psi' component[23] [24] is latent or, so to speak, 'fallow'. *Paral* is psychomatter whose 'psi' part is active, or 'unfallow'. If one uses q and π as labels for specks (*i.e.* microparticles) of matter and paral respectively, one has:

$$\text{matter: } q = (\psi, \varphi); \qquad \text{paral: } \pi = (\eta, \zeta)$$

As expected, these are self-evident instantiations of B = (S, D), where $\{\psi, \eta\}$ are the respective *endo-causal* (*i.e.* "psi"!) parts whilst $\{\varphi, \zeta\}$ are the apposite *exo-causal* parts. One can then think of *our* universe U_0 as made up of (at least) two different *stuffs*:

(15) $U_0 = [q]\oplus[\pi]$.

(This formula is obviously a special case of $\{U = B_1 \oplus B_2 \oplus B_3 \oplus ...\}$. *It reads* $U_0 = B_1 \oplus B_2$; *where* $B_1 = [q]$ *and* $B_2 = [\pi]$ *denote* matter *and* paral *respectively, whose particles or specks are labelled in turn* q *and* π.*)* One writes accordingly:

$$[q] = ([\psi], [\varphi]) \quad \text{and} \quad [\pi] = ([\eta], [\zeta]).$$

This provides some extra backing to the psychomatter hypothesis (not on concrete, universe-specific grounds; but on an abstract, ontological level). It also suggests that quantum indeterminacy - often taken as sheer, hollow randomness - is really an effect of endo-causation. More to the point, the wave-packet collapse *(which I dubbed a* paralling, *or a* paral phase *since it turns* matter *into* paral) would precisely consist in the release of a deed.

Such deeds - arising from the endo-determined *("psi")* part, ψ, of any speck, q, of matter - would underpin the (bogus) 'particle' side of the wave-particle duality.(25) This well-researched duality has commanded a lot of attention. Yet and unsurprisingly, it is a poorly understood feature of quantum physics.(5) (15) (27) (29)

We finally move on to the third aspect, regarding the space-time arena, branded Σ, of aught. At this juncture a point must be made clear. It bears on the status of space-time; namely: *Space-time is not a primary datum, it never stands on its own.* Rather than being a pre-requisite to aught, it is a shadowy side-effect of it.(6) (30) Take out any self-breeder, and you shall automatically do away with space-time; recast the *exo-causal* content of a given *stuff*, and you shall modify the affixed space-time.

In other words, space-time is "aught-dependent": it is widely acknowledged, thanks to Einstein, that space-time is not a mere passive or inert frame, that would pre-exist and be alien to matter. The truth is quite the opposite: matter and space-time are inextricably bound up.(4)

My approach concurs in that view. The following piece of reasoning (of which I give only the most basic outline) should make it plain. In the world of naught, there is flat nothing - no matter, no energy, no space, no time. In the fickle world of unmitigated endo-determinacy, there is nothing steady, not a single shred of dependable reality. One thus jumps to the conclusion that a space-time structure, hinging heavily as it does on some measure of dependability at the heart of reality, all but cannot exist against such an untoward backdrop.

Weed out of reality any modicum of steadfastness, and any hint of space-time is forthwith nipped in the bud; it is left with no firm anchor on which to unfold. In short: *the* ur-stuff *begets no space-time.* (To be accurate, this is true *unless* one chooses to take into account whatever "sham" space-time goes with *sham determinism* - to wit: with any spell of reversible deedlessness.)

Going back to the general formula $\{B = (S_B, D_B)\}$, when $\{D_B = n\}$ there is absolutely nothing, no stable element, on which to root a lasting information framework or network, that is mandatory to foster a steady space-time structure. Contrariwise, exo-determinacy affords a reliable basis on which specific space-times do arise. As we see, space-time is wholly embedded in exo-determinacy; it is D_B-yielded or D_B-begotten ($D_B \neq n$).

That space-time is no footloose, free-wheeling reality, should be further evidenced by the *(I presume, likely)* fact that, in *our* universe, gravity is an effect of the *parallable* quantum waves which permeate *(and warp the fabric of)* space-time. This conjecture - which I believe deserves to be taken in earnest - has as yet to be investigated.

If true, it would go a long way towards explaining why there seems to be no quantum gravity (in the conventional quantum field-theoretic sense)(4) (11); and towards pinpointing where such oddities as Mach's principle stem from.(25) (26)

In summary: endo-determinacy, by its very nature, is too "shifty" to yield a proper space-time, which thrives on steadiness exclusively. This leaves a lone option, namely: *The space-time arena to which B attaches is D_B-dependent* (where, as usual: $B = (S_B, D_B)$, $D_B \neq n$). Change D_B, and chances are that you will accordingly change the B-related space-time arena, Σ_B. *(This is easy to figure out. Imagine that some sly scientist, by tampering with photons, succeeds one day in modifying the photon-borne speed of light: by meddling with the deterministic features of some quantum particles, he will outright alter our relativistic space-time frame.)*

It entails that if any two self-breeders do not belong to the same space-time, their exo-causal parts, say D_B and $D_{B'}$, will be different. (That is: $\Sigma_{B'} \neq \Sigma_B \Rightarrow D_{B'} \neq D_B$.) The converse generally holds true as well. So, two beings which possess differing deterministic parts will generally evolve in dissimilar arena, *viz.* in dissimilar spatio-temporal frames.

This of course applies to *matter* and *paral.* Recall that $U_0 = [q]\oplus[\pi]$: our world comprises two ontological *stuffs*, namely, *matter*, $\{[q] = ([\psi], [\varphi])\}$, and *paral*, $\{[\pi] = ([\eta], [\zeta])\}$. Their respective properties are unmistakably very different. In other words, one has: $\{[\psi] \neq [\eta]\}$ and $\{[\varphi] \neq [\zeta]\}$.[25] Since $\{[\zeta] \neq [\varphi]\}$, we typically expect to observe: $\{\Sigma_\pi \neq \Sigma_q\}$.

Now, what strictly relates to *matter* ([q]) is relativistic: Σ_q is faithfully depicted by Einstein's theory. But, in light of the foregoing, we see that there is no *a priori* reason why *paral* ([π]) and *paral*-related phenomena should also comply with relativity (both special and general). On this account, Einsteinian relativity is, ironically enough, itself relative. Isn't this about right?

This is cogently underlined by the so-called "E.P.R paradox", whereby under certain circumstances (linked to what I dub *supralness* - a.k.a. 'non-separability' or 'quantum entanglement'), correlations are observed between events separated by *space-like* intervals.

Besides, this is further exemplified by the arguable *instantaneity* of the wave-function collapse. The striking feature here is that all these (*paral*-related) events are "light-cone-blind": they blithely fall foul of relativity.

Now we can begin to address a thorny question linked to the cosmological Big Bang. It reads: "What was there *before* the Big Bang?" In fact the answer is well-known nowadays. It says: "There was no before". Why? Simple enough: there was nothing by then, no matter, no energy; *hence neither space nor time.* It makes little sense to ask questions about a "before" in such an utterly timeless background! Remember, the birth of the universe is also the beginning of space and time.

(To be more specific: any space-time Σ_B, B-induced or B-related as it is, is strictly B-relative. So the above reads: "There was no B-before." Or else: "The former question is B-meaningless." This is a distinctly more accurate, but also more restricted, statement.)

Really, this is not the whole story. According to the current theory, *before* the Big Bang itself something was a-lurking - and this "something" is the quantum vacuum. This is the state-of-the-art view. But let us not be gulled by a name: this vacuum is *not* naught, not even remotely. It is very much aught. It is fraught with physical laws, awash with energy, and seething with activity (in which particle-antiparticle pairs are constantly being created and undergoing mutual annihilation).[2]

Such a vacuum can hardly be taken in earnest: it is not stark emptiness, bereft of anything exo-causal (such as a physical law or a unitary *superforce* of sorts...). It is glaringly structured. In other words: it is brimming with deterministic, or exo-causal, contraints. And therefore, this 'ersatz naught' does not fit the ontological bill!

If we are to believe most quantum cosmologists, "The entire cosmos simply comes out of nowhere, completely in accordance with the laws of physics. (...) The world of physics routinely produces something for nothing. Quantum gravity suggests we might get everything for nothing." (Paul Davies)

Why should naught bother to be "completely in accordance with the laws of physics"? As David Darling put it once, "Either there is nothing to begin with, in which case there is no quantum vacuum, no pre-geometric dust, no time in which anything can happen, no physical laws that can effect a change from nothingness into somethingness; or there is something in which case it needs explaining." *Back to square one!*

What was there, before the Big Bang? At the very least, there was a *big bang*, and also a *wee bang*. Before our universe came into being, there was an ubiquitous *beable* (that keeps bouncing back and forth from naught...). Now, to conclude, here are the main points of this Section:

- the quantum "object", or microparticle, is intrinsically dynamic in nature: it is not a solid entity but is rather an on-going stream of "events";
- in keeping with the overall formula B = (S, D), what we call "matter" is in fact a *bi-dimensional* stuff; it comprises an endo-causal part - whereby it occasionally displays nondeterministic traits - along with an exo-causal, or deterministic, part (I dub it *psycho-matter*);
- much like water that can freeze into ice, *psychomatter* can, given the right conditions, exist as *matter* proper or as *paral*, whose qualities are dissimilar;
- the nonrelativistic character of the wave-packet collapse (or *paralling*) boils down to its being *paral*-related (relativity holds its sway over *matter* - not over *paral*);
- quantum vacuum, unlike what some made it out to be, is at stark variance with naught: it is exo-determined; as such, it is on the same ontological footing as our universe, and needs explaining just as well (it does not lead towards a "rational scientific theory of all existence").

OPEN QUESTIONS

Rather expectedly, the above raises far more questions than it settles. Here is a quick glance at few of the many question-begging issues involved. The first one, already mentioned, has to do with logic. It seems that logic cannot be eschewed; and what therefore strikes me as deeply puzzling is its ontological status.

Wouldn't even the Almighty be tied up and stymied by consistency, with no possibility ever to churn out *actual* square circles? *Why* then should logic as such be so unassailable? It is only a fake result of our flawed and limited human *rational* perspective? Is it on the contrary because logic - or consistency - is somehow over and above actual reality?...

These are far-reaching questions, and by no means easy ones. Furthermore, the acknowledged awkward possibility - in *our* universe - of niceties such as "time travels" and other unruly time loops(6) put the very idea of consistency in jeopardy. It wreaks havoc to ordinary causation, which shows that overall consistency cannot be taken for granted without prior vindication. As we know, "Any attempt to consider anything as simultaneously subject to influences in both temporal directions is liable to collapse in complete incoherence." (M. Dummett, *in* (10), p 153.)

Besides - and to make things worse - recall the *paradoxical* nature of self-reference. It is potentially shattering, since *endo*-causation is very much self-referential. *(However, self-reference is a recipe for doom* provided *it is specifically wed to* exo-causation. *As one can expect, it is arguably less threatening and does not spell much trouble when blended with* endo-causation. *This is all too normal, since endo-causation is its own turf!)*(21)

Can we still trust the world - and heedlessly believe that it is logically sound? Can it be safely claimed, in like manner, that *existence comes at a price, and* that this "price" can be worked out - at least partially - on sheer logical grounds?

Another point worth mentioning bears on the wave-packet collapse. On the one hand, I wrote that 'stuff-zapping' or 'stuff-hopping' is not spontaneously allowed. (Remember: "D_B is by definition *stable* with respect to the B-deeds: no deed shed by B - and really, by S_B - can jostle D_B into becoming a different compound of self-breeding processes.")

On the other hand and as I wrote, too, the wave-packet collapse consists in the release of a deed *and* it happens to turn *matter* into *paral* (this is why I also dub it a *paralling*, or a *paral phase*). This is a blatant case of spontaneous 'stuff-zapping'! How to reconcile the two?

Here is how: $\{[q] = ([\psi], [\varphi])\}$ and $\{[\pi] = ([\eta], [\zeta])\}$ must be thought of as two *local* faces of the same *global stuff* - again, just like water and ice. (As we see, (15) is only an approximation of the truth.) The broader issue raised here is that the concept of *stuff* (along with the attendant ones of exo- and endo-determined parts) is slightly more complicated than has been reckoned hitherto.

For example, here D_B and S_B ought to be broken down into $\{D_q, D_\pi\}$ and $\{S_q, S_\pi\}$, depending on the context. This ties-in with our earlier considerations on *local* versus *global* fields of B-potentialities, whereby: $\{\forall (b, b') \in B^2, \Pi_b \subset \Pi_B,\ \Pi_{b'} \subset \Pi_B, \Pi_b \neq \Pi_{b'}\}$. (Here $\Pi_b \neq \Pi_{b'}$ translates into: $\Pi_q \neq \Pi_\pi$).

On another level, the problem is really that of understanding the detailed nature both of S_B, and of the relationship between S_B and D_B. Once again, it can be tackled only within a formal framework (I therefore leave it out).

Another knotty conundrum is to get to the bottom of the "ontological kernel", *i.e.* of the *whiff* W. Basically, it is something that stirs... nothing. But why, and how? What are the trappings of *beable*-ness? What is the root cause of endo-causation? *(The answer here is in a way straightforward: the root cause is plainly endo-causation* as such, *which bears within itself a fundamental self-sufficiency.)* And lastly, what process lies at the heart of the onset of exo-determination?

Underpinning these questions is the hackneyed and highly controversial notion of free will. To my mind - which admittedly goes against the tide - free will is a legitimate and fitting idea. It is implicit to the notion of endo-causation, which can be seen as arising from some inward 'willed' agency. A deed is a creative spurt, it can roughly be regarded as an 'act of will' since it is a matter of "pulling the strings from within"; and this points pithily to some underlying free will.

Incidentally, I believe that the speculative endeavour outlined in this paper can be pushed somewhat farther afield - and far enough to give us a clew as to how it could become humanly feasible to alter our world in a quite startling way, by letting the freedom genie out of the bottle. Or that is: by no less than changing the exo-determined profile of psychomatter!

This is *in principle* possible, by creating carefully crafted conditions (these conditions, of which I shall say no more here, are typified by the *spontaneous symmetry breaking* situations encountered in quantum field theory).(21) (31) In so doing, we would tinker or meddle with determinism - with sweeping repercussions!

To sum up, beneath any degree of endo-causation there could be a shred of "free will", and ultimately a *degree of awareness* (a.k.a. *d.o.aw.*, or *doaw* for short). The nature and *raison d'être* of such a *doaw* is up for grabs; I have a feeling that it is the utmost riddle that is enshrouded in the broader mystery of ontogenesis. A *doaw* would be of varying intensity - and this would explain why there is a *threshold* which splits, as I suggested beforehand, Π_W into (at least) two sub-areas, $\Pi_{d\text{-}a}$ and $\Pi_{n\text{-}a}$. *(More broadly, it would underpin most of the* topological *features linked to W and B; such as that, already seen, of* skew-reversibility. *Other thresholds can be conceived of, virtually* ad libidum *and* ad infinitum - *since the ur-stuff can stretch itself way beyond whatever height of inner sophistication we may imagine...)*

The ur-stuff, W, is *terra incognita* - borderless, inchoate, an unchartable world unto itself. It is the Fathomless, and can only be dwarfed by our stunted efforts to approach it. *Is it God by another name?* Would God, therefore, be mighty enough to bootstrap himself into existence? This seemingly goes against most religious views however, inasmuch as they allege that God has always existed, and did not come into existence from nothing.[(12) (13) (16) (18) (19) (20)]

Why has God always existed? It is, we are told, because God's nature is to be a permanent and unmoving "prime mover". To be sure, this does not square with the fickle and self-breeding nature of the *ur-stuff* - nor does the widespread idea that "God does not depend for his existence on himself or on anything else." (Swinburne, *in* [(18)], p. 147.)

Many other aspects are on the contrary exciting areas of convergence. But I am not a theologian, and I shall forbear to develop and comment on a comparison between the available images of God and the ontological *ur-stuff* - lest I write too many howlers and inaccuracies.

All in all, I agree with Euan Squires, who wrote: "I personally am willing to take that step: the physical Universe did not create consciousness but, in a way that I do not comprehend, was created by it."[(29)] This view is close to mine, insofar as the *beable* (that - or *who?* - is universe-yielding, too, by means of big bangs!) is a straight matter of untrammelled *doaw. (Other facts point in the same direction, such as: the* anthropic principle, *which lays stress on the erstwhile unforeseen dizzying* unlikelihood *of a life-bearing world; and the* primeval paralling[(20)], whereby in our infant universe - right after the Big Bang - the "psi" content of *psychomatter* was a-kindled on a very wide scale.)

Not all the contents of reality are amenable to rational understanding. This is yet another reason why my approach is at best very tentative and limited. It could nevertheless encapsulate or at least point to a grain of truth. My hope is that this grain, however small and trifling it be, is still worth the candle.

The only certainty is uncertainty. So let us acknowledge that our grasp of reality will ever rest on partial and shaky grounds. For all our efforts and achievements, the human perspective is bound to remain narrow, biased and short-sighted.

But this is what the staid language of reason says, and we rather go for the language of passion. Against tantalizing odds, we keep trying and striving, for the mere sake of it.

Herein lies the matchless, awe-inspiring, mind-broadening beauty of this forlorn endeavour! *Understanding reality* is an abiding human quest, an ennobling one that will stay with us for a long time to come.

REFERENCES AND BIBLIOGRAPHY

(1) Baggott, J. (1993). *The Meaning of Quantum Theory. (reprint)* Oxford: O.U.P.
(2) Carrigan, R. & Trower, W. (eds) (1989). *Particle Physics in the Cosmos.* New York: W. H. Freeman and Co.
(3) Castello, M. & Zartarian, V. (1994). *Nos Pensées Créent le Monde.* Paris: R. Laffont
(4) Cohen-Tannoudji, G. & Spiro, M. (1986). *La Matière-Espace-Temps.* Paris: Fayard.
(5) Cushing, J. (1994). *Quantum Mechanics.* Chicago: The University of Chicago Press.
(6) Earman, J. (1995). *Bangs, Crunches, Whimpers, and Shrieks.* Oxford: O.U.P.
(7) Ferrier, J. (1992). Plaidoyer pour un Univers des Deux Mondes, sans Oublier leur Interface. In

Bulletins de l'Académie du Var, Toulon.
(8) Ferrier, J. (1994). Le Point sur I. In *Bulletins de l'Académie du Var*, Toulon.
(9) Feynman, R. & Hibbs, A. (1965). *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill.
(10) Flood, R. & Lockwood, M. (eds) (1995). *The Nature of Time*. Oxford: Basil Blackwell Ltd. *(reprint)*
(11) Freund, P. (1989). *Introduction to Supersymmetry. (reprint)* Cambridge: C.U.P.
(12) Gale, R. (1991). *On the Nature and Existence of God*. Cambridge: C.U.P.
(13) Harris, J. F. (ed.) (1992). *Logic, God and Metaphysics*. Dordrecht: Kluwer Academic Publisher.
(14) Hottois, G. (1989). *Penser la Logique*. Brussels: De Boeck-Wesmael.
(15) Isham, C. (1995). *Lectures on Quantum Theory*. London: Imperial College Press.
(16) Kenny, A. (1992). *The God of the Philosophers*. Oxford: Clarendon Press *(4th impression)*.
(17) Largeault, J. (1993). *La Logique*. Paris: P.U.F., coll. "Que sais-je ?"
(18) Miller, B. (1992). *From Existence to God*. London and New York: Routledge.
(19) Polkinghorne, J. (1994). *Science and Christian Belief*. U.K.: SPCK.
(20) Rañada, A. (1994). *Los Científicos y Dios*. Oviedo: Ediciones Nobel.
(211) Ransford, E. (1986). Où sont Passé les Enoncés Indécis de la Physique Théorique? In *La Jaune et la Rouge* (N° 419). Paris.
(22) Ransford, E. (1991). Le Big Bang et la Genèse du Monde. In *La Jaune et la Rouge* (N° 462). Paris.
(23) Ransford, E. (1994). From Quantum Physics to the Conscious Brain. In *Entelechies*, Proc. of *ANPA 16*.
(24) Ransford, E. (1995). Peeking at the Conscious Brain: New Clues, New Challenges. In the *ANPA West Journal*, Vol. 5, No 2.
(25) Ransford, E. (1996). Elementary Particle or 'Wavicle'? Seeing through the Quantum Fog. In *Philosophies*, Proceedings of *ANPA 17*.
(26) Roscoe, D. (1996). Gravity out of Inertia. In *Philosophies*, Proceedings of *ANPA 17*.
(27) Schiff, L. (1988). *Quantum Mechanics*. Singapore: McGraw-Hill Int. Co. (3rd Edit.)
(28) Simon, B. (1971). *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms*. Princeton: P.U.P.
(29) Squires, E. (1985). *To Acknowledge the Wonder*. Boston: Adam Hilger .
(30) Stephani, H. (1990). *General Relativity*. Cambridge: C.U.P.
(31) Sterman, G. (1993). *Quantum Field Theory*. Cambridge: C.U.P.
(32) Zaslawsky, D. (1982). *Analyse de l'Etre*. Paris: Les Editions de Minuit.

THE CINEMATIC MODEL of QUANTUM TOUCHING

by

Viv Pope

(This follows from the paper: 'The Gateway to Paradox'.)

Verbatim:

In the ANPA Newsletter, Number 16, Clive Kilmister raises the question of some 'Fundamental Problems in Quantum Physics'. Foremost among these is the very strange way 'photons' and elementary particles behave in modern versions of the Thomas Young type of two-slit experiment.

Now I don't want to be a spoilsport but in all honesty I have to say that I cannot see this problem as being due to anything more than that same sort of over-attachment to institutionalised ideas that I described in my talk here yesterday. Emancipation from this historicism (against the sheer 'poverty' of which Popper has warned us) allows us to cut away what I yesterday called the 'Gordion Nut' and perceive a very simple solution to these problems. This is to dump the unnecessary and nonsensical idea of light-interaction being conveyed by clairvoyant, space-travelling 'photons'. In other words, all we need to do is to accept that 'photons' simply and logically *cannot be*. In the first place, anything travelling at the speed of light, as Einstein's 'photon' is supposed to do, cannot have mass, because if it did, then according to Einstein's equations, that mass at speed c would have to be infinite, which is impossible. In the second place, anything travelling at that speed c cannot register any proper time or proper distance between the emitter and absorber. Logically and literally, therefore, it can have no material existence or duration. So if, in the context of Einstein's theory, his postulate of the 'photon' as something that exists in the nature of things is not a complete and utter contradiction, then I don't know what is!

Cut away that postulate of the 'photon', then (the Gordion Nut)and what have you got? What you have is the plain and simple, sufficient conclusion that every quantum of physical interaction, of light, gravitation or whatever, is intrinsically an irreducible, proper-time-instantaneous quantum jump.

This quantisation of interaction means, logically, that there can be no permanent and continuous 'space', like the traditionally imagined etherial 'soup' with particles floating in it like peas. What we are accustomed to call space (void or vacuum), as I have proposed (e.g., in the Internet Homepage, see paper 1), is a quantum *discretum* consisting of prodigious numbers of transient interconnections between masses which, at any quantum instant, form automatically paired and instantaneously (i.e., reciprocally) balanced angular momentum relations. The distance at any instant between a pair of masses is thus the length r of the moment-arm on the ends of which the two bodies are proper-time-instantly balanced; and space is the statistical, three-dimensionally extended network of such quantum lengths.

And because angular momentum is balanced and conserved, any loss of angular momentum at one place (in quanta $mvr = nh/2\pi$) has to be immediately accompanied by a gain in precisely that amount of angular momentum somewhere else, simply by the law of conservation and without raising any question of how that balancing-influence 'travels'.

This concept of an overall instantaneously extended angular momentum nexus answers to the 'space' of Newtonian physics, except, of course, that its lengths are now discrete and transient, not continuous and permanent in the way Newton envisaged, and its 'inertial motions' are orbital (i.e., non-Euclidean) not rectilinear – which produces the essence of general relativity at a stroke.

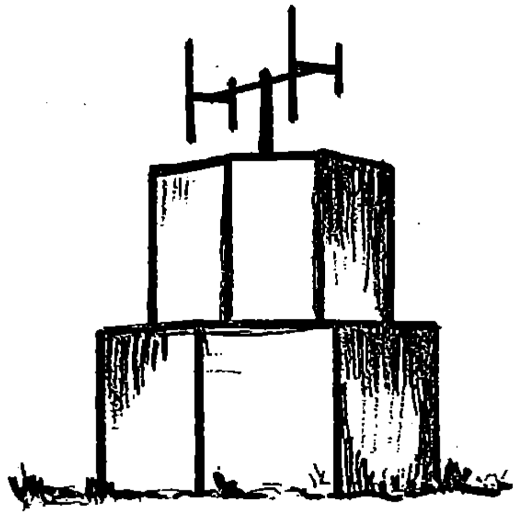
Any *change* in that angular momentum constitutes an *event*, and *time* is our conventional (i.e., statistical) measure of the *sequences* of these events. Each event in the sequence may thus be regarded, by a cinematographic analogy, as a 'quantum still' in which, by definition, nothing *moves* but in which all the various bits are simply *there*, instantly interconnected, or co-occurring (as Manthey calls it), like the objects depicted in a still photograph or frame. All *action*, as in a movie film, is then a function of the temporal *sequencing* of these quantum stills relative to some observational frame.

Like all other actions, light-interaction, at its measured speed *c*, is a cinematic *sequencing* of these proper-time-instantly extended elements, or quantum stills. Those objects whose positions do not change from one still to another simply *endure* in the sequential or 'longitudinal' dimension of the film. Those whose positions do change not only endure but, in doing so, also register distance in the three proper-time-instantaneous dimensions that are 'lateral' to that of the comparatively 'longitudinal' time-dimension. The space-time resultant of these two *inter-dependent* and therefore orthogonal distance-time components answers precisely to relativistic 'time-dilation', as I have demonstrated here, at ANPA, on more than one occasion. Expressed as distances *s* travelled in that relative or dilated time, all motions tend to the finite limit *c* as in the so-called 'Einstein separation' – which is really no separation at all but a sequence of quantum immediacies, or 'touchings'.

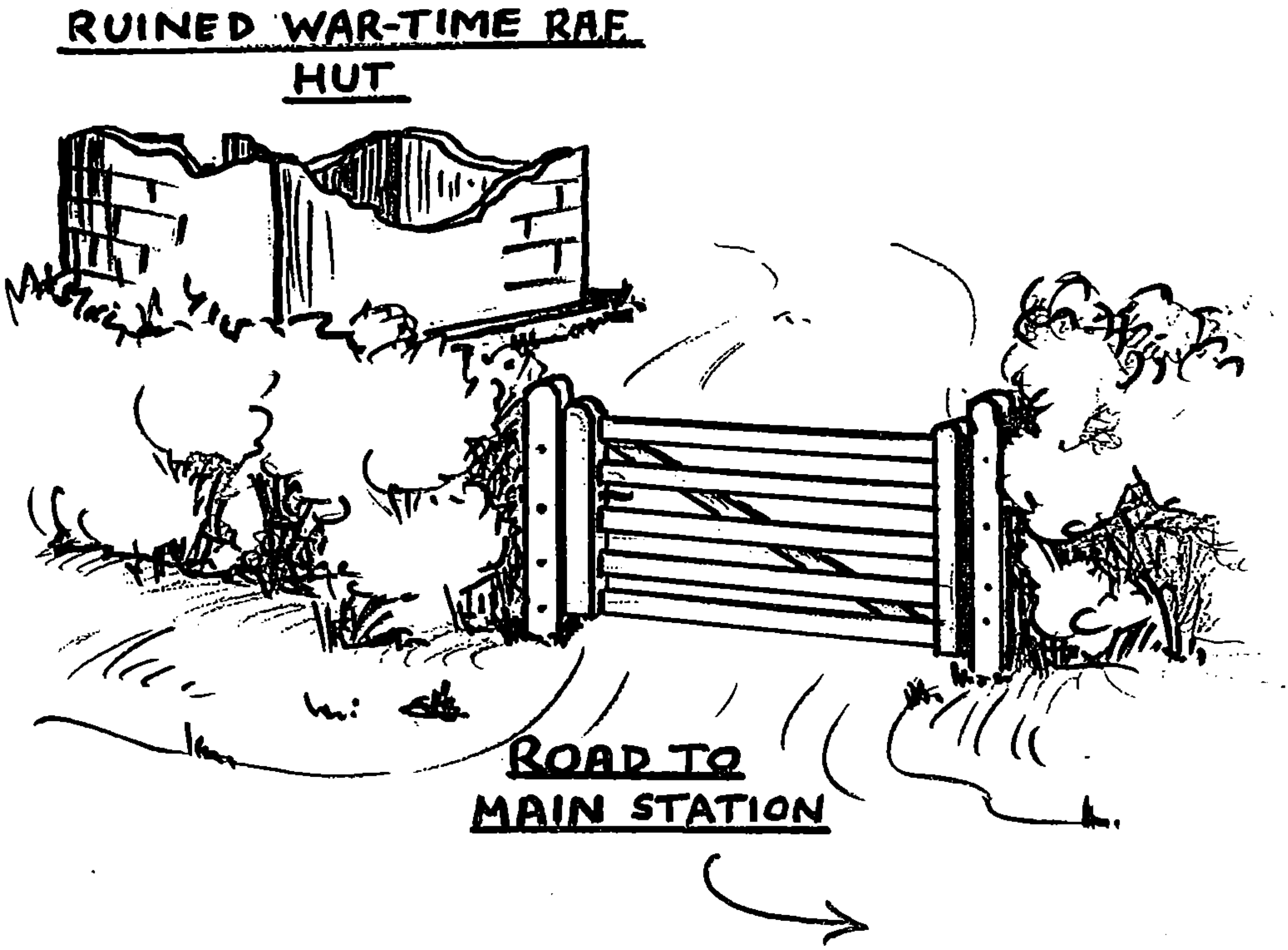
The advantage of this 'Cinematic Model' of light-propagation is that it reveals how we can think of light in a commonsense way, without any contradiction whatsoever, as *both* instantaneously extended (in its angular-momentum 'quantum-still' aspect) *and* time-retarded (in its other, cinematic or sequential aspect). What 'interferes' in the two-slit experiments are therefore not 'photons' and particles, miraculously going through the two separate slits at once, clairvoyantly anticipating detection-traps as so many mystics of science now so often and so jubilantly claim. As may be recalled from the discussion of this subject in Newsletter 16, my suggestion is that the 'interference' is plainly that of *the paths themselves*, as proper-time-instantaneously extended wholes (quantum stills) which, measured through the slits, have lengths that are geometrically congruent or incongruent in terms of the quantum parameters. Thus I recommend that all talk of 'photons', 'virtual particles'and 'spooky superluminal' influences guiding them to their destinations and so on is made redundant and should at the earliest opportunity, be crated and consigned, with concepts like those of æther, caloric and phlogiston, to the backrooms. of science-history.

[This paper was the prelude to a vigorous discussion, all of which (including the first paper and ensuing discussion) were videoed by Jon Blay, of Armadillo Communications. Copies are obtainable from Armadillo. tele. 01714390400].

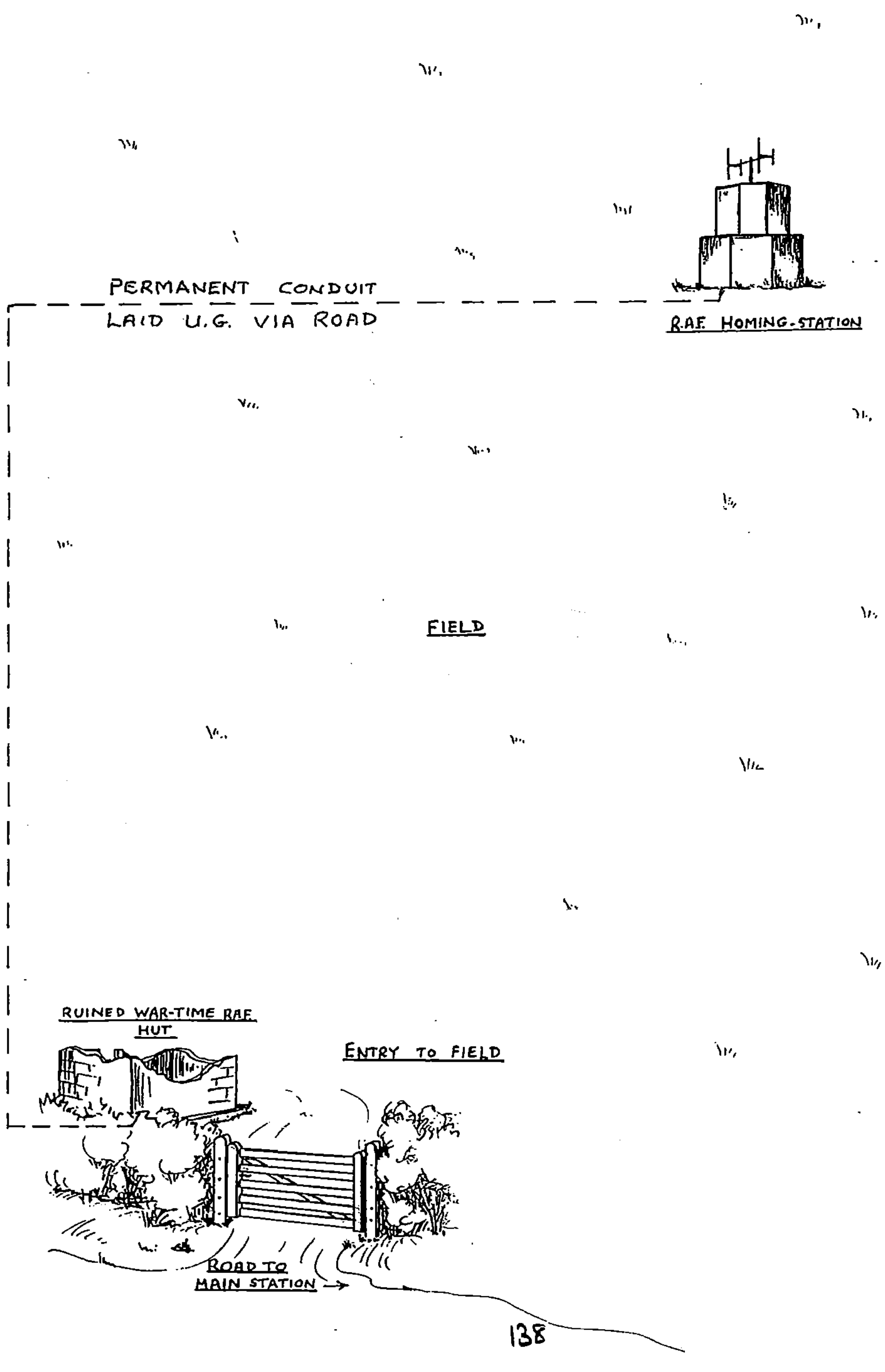
* * *



R.A.F. HOMING-STATION


RUINED WAR-TIME RAF
HUT
ROAD TO
MAIN STATION

O.H.P. (3


PERMANENT CONDUIT
LAID U.G. VIA ROAD
R.A.F. HOMING-STATION
FIELD
RUINED WAR-TIME RAF HUT
ENTRY TO FIELD
ROAD TO MAIN STATION

Distributed Computation, the Twisted Isomorphism, and Auto-Poiesis

Michael Manthey
Computer Science Department
Aalborg University
Fr. Bajersvej 7E
9200 Aalborg, Denmark
manthey@cs.auc.dk

Abstract [1]

This paper presents a synchronization-based, multi-process computational model of anticipatory systems called the Phase Web. It describes a self-organizing paradigm that explicitly recognizes and exploits the existence of a boundary between inside and outside, accepts and exploits intentionality, and uses explicit self-reference to describe eg. auto-poiesis. The model explicitly connects computation to a discrete Clifford algebraic formalization that is in turn extended into homology and co-homology, wherein the recursive nature of objects and boundaries becomes apparent and itself subject to hierarchical recursion. *Topsy*, a computer program embodying the Phase Web, is currently being readied for release.

Keywords. Process, hierarchy, co-exclusion, co-occurrence, synchronization, system, auto-poiesis, conservation, invariant, anticipatory, homology, co-homology, twisted isomorphism, phase web paradigm, Topsy, reductionism, emergence.

Introduction

Anticipatory systems (Rosen, 1985) display a number of properties that, together, differentiate them strongly from other kinds of systems:

- They possess *parts* that interact *locally* to form a coherently behaving *whole*.
- The way in which these parts interact differ widely from system to system in detail, yet wholes with very different parts seem nevertheless to resemble each other *qua* their very wholeness.
- It is impossible to ignore the fact that such systems are *situated* in a surrounding environment. Indeed, their interaction with their environment is so integral to what they are and do makes their very situatedness a defining characteristic.
- A critical behavior shared by these wholes is the ability to *anticipate* changes in their surrounding environment and react in a way that (hopefully) ensures their continuing existence, ie. *auto-poiesis*.

Attempting to get a handle on anticipatory systems *computationally* can mean different things to different people.

[1] Invited paper, CASYS'97 First International Conference on Computing Anticipatory Systems, Liege (Belgium), August 11-15, 1997. ©June 5, 1997.

Suppose, for example, that the mathematical description of the phase web described in §2 were programmed directly, with all the do-loops, data structures, and algorithms this traditionally implies. While the result might be a good *simulation* of an anticipatory a system, I personally would be dissatisfied because I seek a system description which *by the very nature of the computation itself* would produce *actual* behavior. That is, while the output of such a traditional program is all well and good, the detour through an *a priori* mathematical description obscures both the mechanism and the process by which this output is produced.

Another way to say this is that for me, computation is just as fundamental as mathematics, but the two have different strengths. The strength of a computational description is that it must exhibit actual *mechanisms* and the processes engendered thereby. I seek a computational formulation that can be seen to *inevitably* produce systems with the properties listed above, without any external or *a priori* guiding hand, indeed, with no need to appeal to mechanisms beyond what it itself embodies.

This is a tall order! However, I believe I have succeeded to a reasonable extent, not least because the resulting purely *computational* system- descriptive apparatus has (ironically, in view of the preceding comments) a very clean mathematical formulation (presented in §2). Those familiar with the various attempts to describe computation mathematically know that the two are fractious bedmates, so I view this denoument as a sign that there is something very right about it.

In contrast to many, the approach presented here emphasizes *structure* so strongly that the algorithmic component that for most people is the sine qua non of computation is nearly non-existent. This emphasis is ultimately the reason why the approach offered here - called *the phase web paradigm* - differs from all others I am familiar with, and correspondingly, why its mathematics comes out so differently (algebraic topology, namely, rather than logic).

But how can one even *have* computation without an 'algorithm'?! The answer is that the classical concept of an algorithm is a specification of a *process* that is to take place when the algorithm is unrolled into time. The phase web paradigm is however focused entirely on the process aspect, and thereby essentially obviates the need for the *a priori* existence of a defining algorithm. One might compare this to the theory of evolution based on natural selection: this is a process-level theory, for which the existence of some *a priori* algorithm is problematic.

Of course, one still writes programs, but in pure process terms. However, since an anticipatory system in general grows/learns, this programming is ultimately sculptural rather than specificational in character.

The next section introduces the basic computational model, which is described at greater length in [www]. The mathematical translation of this computational model follows, and the paper closes by relating all this back to anticipatory systems and auto-poiesis.

1 The Computational Model

The goal of this section is to sketch the essentials of the phase web's computational model.

The principal problem computer science has faced over the last two decades is the digestion of the phenomenon called "parallelism", and virtually all contemporary research is colored by issues arising from it. This means that we are already in a decidedly *process*-oriented context. The computational concepts I use - synchronization, co-occurrence, exclusion - are well-established and used by researchers in the field. I prefer the term "concurrency" to "parallelism" because the latter is tainted by associations to interleaving the events constituting several parallel processes to achieve a formally sequential process (equivalent to disassembling a living cell to form a long end-to-end chain of molecules, and then not even realizing that it's dead).

I have been particularly concerned with what are called *distributed* systems, that is, systems which - like an ant hill - exhibit globally coherent behavior via solely local decision-making on the part of its constituents. I have been looking for some small set of seed concepts out of which *any* kind of "ant hill" may be built. My goal all along has been to apply the understanding gained from this search to construct an entity that can learn from its experiences and behave in an increasingly sophisticated way on the basis thereof.

As a starting seed, it appears from very general considerations that a necessary condition for the ability to profit from experience is the ability to draw *distinctions*. In a sequential context, this demand is met by the *if-then-else* construction or equivalent. In the concurrent context of the present work, the fundamental distinction I have cooked everything down to is that between *occur together* versus *exclude each other*. That is, can two situations co-occur in experience versus they cannot self-consistently do so. (The following sub-section therefore treats the computational mechanism - synchronization - that addresses such relationships.) The overall approach is to express knowledge of self and surround as patterns of exactly these two *complementary* synchronization forms, and to express behavior via their manipulation.

The second seed concept is that of symmetry, by which I mean several things:

- A general symmetry I like is "outside is as inside", that is, the *boundary* separating what is outside from what is inside an entity can be drawn arbitrarily, at least in principle. In practice this means that the representation of internal relationships should have the same form as the representation of external relationships.

- A specialization of symmetry is the physicists' use of group-theoretical symmetries, which cogently summarize such varied relationships as conservation laws, Lorentz (ie. relativistic) invariance, and particle properties. It has turned out, though after the fact, as it were, that the phase web's group symmetries are very much akin to those of quantum mechanics.

- A third aspect of symmetry is the requirement that the form of a part of a whole is the same as the form of the whole, that is, this is a hierarchical requirement. When combined with the ability to harvest observations (cf. *occur together*), which is a requirement for learning from experience, this symmetry leads to the ability *internally* to explicitly represent internal states and relationships, which in turn supplies the desired self-reflective component.

The third seed concept is that of goal-directed behavior, by which is meant that an entity can explicitly represent to itself the *goal* or intention of its activity. It is hard to see how this can be avoided; the teleological element it introduces is however elastic. Goals can be either introduced from the outside or generated internally.

Besides the above concepts, the phase web paradigm is also the product of a two broad constraints: 'mechanism' and what I call 'bio-engineering plausibility'. By mechanism is meant that an a priori and purely mathematical explanation is eschewed in favor of a process-oriented one: the former have been tried (eg. propositional calculus, Newtonian physics) without particular success. The phase web and Topsy are, in contrast, pure process, and this is what led to the mathematics we present later, and not the other way around.

By bio-engineering plausibility is meant that the mechanism proposed for a computationally-based entity is profitably constrained by requiring that this mechanism can conceivably be embodied in biological systems as well. After all, the best examples we have of anticipatory systems are biological. The information flowing across the boundary from outside the organism to inside should, for example, be concrete, should be 'grounded': molecular polarity, touch, sound waves, retinal pixels, etc. It should perhaps also be noted that although a biological system constantly creates and destroys its constituents, this is not modelled in the computational model for reasons of efficiency (but could otherwise be).

1.1 Synchronization

As late as the 1960's main-frame and mini-computers, and again with personal computers from the early 1980's until recently, one had *one* computer on which ran *one* program. The coordination between this computer *cum* program complex and the outside world (ie. "input/output") was deeply buried in technicalities and generally considered vastly uninteresting. However, when one began, with the advent of timesharing, to harbor *multiple* programs on the same machine, the issue - and profundity - of coordinating the interaction of otherwise independent processes gradually became visible.

With multiple interacting processes, a number of new phenomena (at least to software people) appeared, eg. concurrency, non-determinism, deadlock, communication; and as well, pair of critical new concepts - *sharable resources* and the necessary *mutual exclusion* of processes using same. Issues concerned with process interaction and communication came into the foreground. All of these things appear in the concurrent world, and none of them in the *sequential* world of single non-interacting processes.

In order to deal with these things, it was found necessary to introduce a new primitive operation into computing, that of *synchronization.*[2] Viewing an 'event' as the execution of (say) a single computer instruction, the role of computational synchronization is to allow the programmer to specify before-after relationships between events belonging to otherwise separate processes.

This allows processes that otherwise are unknowing of each other's existence to cooperate. Arbitrarily complex inter-process synchronization relationships can be built up from primitive before-after relationships. Such synchronization is the foundation on which is built all modern software: your personal computer's operating system, local networks, air traffic control, on-line

[2]Not to be confused with the synchronization-via-photon-exchange exercises performed in relativistic analysis, although the two are of course related.

databases, the Internet and WWW, ... everything.

Synchronization possesses a singularly interesting property: it doesn't really compute anything! It has the same relationship to the programs that invoke it as the pieces in a board game have to the game itself. That is, synchronization relationships *obtain* while simultaneously being conceptually invisible to the processes (ie. game actions) that depend on them. Thus, from the point of view of a program, synchronization is not a *value*-returning function at all, even though textually it often looks like one. This may be clarified by the following.

Definition. An *event* is a change of state of a system. A process is a *sequence* of such events.

A sequential process with the states $s_1 \rightarrow s_2 \rightarrow \ldots \rightarrow s_\ell$ is typically modelled by the composition of *functions*: $s_\ell = f_\ell(f_{\ell-1}(\ldots f_2(f_1(s_1))\ldots)$. In a typical computational process, the f_i would be arithmetic operations. While this functional form suffices when there is only one process present (ie. traditional programming), analyzing systems with *multiple* processes encourages the dissolution of this very tight functional binding of states to allow us to see the intermediates states as *pre-condition, event, post-condition.* In this way, the fact that a given pre- or post-condition can be caused in more than one way is more readily visible.

The concept of synchronization then allows us to express a multi-process computation explicitly in terms of 'when' a given pre- or post-condition (ie. state) obtains, namely whether before or after (or concurrent with) some other state. At this point, the functions f_i begin to fade into the background, since only their result is visible to other processes. The phase web paradigm takes this to its logical extreme: its processes contain *no* arithmetic functions at all, but rather *only* sequences of synchronization operations.

The synchronization relationships between processes often possess an invariant, which I have argued elsewhere (Manthey,1992) corresponds to a conservation law. Conservation laws are group symmetries, not functions. This can be seen as the core of the phase web approach, in that the structure, organization, and operation of a system is expressed in terms of such invariants. We return to this several times in the course of this paper.

By virtue of its before-after focus, synchronization also introduces an explicit notion of *time*, which notion is automatically *relative* to events in other processes. It is however important to understand that this 'time' is something much more primitive than that of ordinary usage. [So any decent computational theory of physics must build such things as ordinary time (and space) up from the relationships obtaining between otherwise isolated primitive synchronizations. Conventional theories face their own version of this. I would say that I establish plausibility that this is possible in the phase web.]

Let us now look at the mechanism by which synchronization is achieved.[3] The two operations wait and signal operate on an entity called a 'binary synchronizer' or 'binary semaphore', denoted S. S contains a single bit of local state (denoted s) which can take on two mutually exclusive values, denoted 1 and $\bar{1}$. Define now wait and signal on S as follows:

[3] The story that follows is, at bottom, one of several possible standard computer science stories, colored by the demands of context.

	S.s=1	S.s=$\bar{1}$
wait:	S.s ← $\bar{1}$; return	continue waiting
signal:	return	S.s ← 1; return

The effect of these definitions is to ensure that a given sequential computation (ie. process) will stall (namely when s=$\bar{1}$) until some other computation signals it (which sets s to 1). Furthermore, a successful wait sets s to $\bar{1}$, thus ensuring that no other computation can follow 'on its heels'. Notice that

- no 'value' is returned by either operation. Rather, each computation simply proceeds on its way after executing wait or signal as if nothing had happened;
- no information is exchanged between waiting and signalling computations;
- the effect of the synchronization cannot be 'observed' locally (cf. preceding item) but will be globally visible as a correlation between events in the system as a whole (Manthey,1992);
- the overall effect is to *order* events - namely the respective wait and signal events - belonging to two *different* processes, such that (presuming S.s=$\bar{1}$ initially) the wait in the one process will always be after the signal in the other. No more and no less.

These definitions are depicted in Figure 1, in which S_o (open) corresponds to 1 and S_c (closed) corresponds to $\bar{1}$. The two processes are denoted by the thick and thin lines, and the two stars indicate their starting positions (=states). Following the lines and obeying the rules for wait and signal, it it easily seen that state {a,$\bar{b}$} excludes state {$\bar{a}$,b}. This state-oriented view is the one we take in this paper.

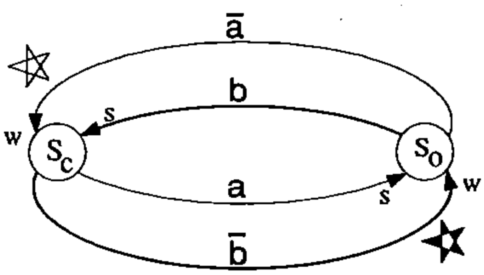


Figure 1: *Synchronization can ensure that certain states (here a,b) exclude other. Note that the synchronization stick, initially 'in' the rightmost synchronizer, is conserved.*

In the figure example, one can conceptualize the alternating mutual exclusion between the two computations in terms of a single 'synchronization token' - I call it a 'stick' - that is passed between them like a hot potato. Such a stick represents the fact that a particular state obtains. At all times there is exactly *one* stick present in Figure 1, either in one of the semaphores or implicitly owned by one of the computations by virtue of the state it is currently in. Such a conserved stick, which necessarily must move on a cyclic (ie. closed) path, reflects the existence of a so-called resource invariant.[4] [Incidentally, the term 'mutual exclusion' is often abbreviated to 'mutex'.]

The preceding discussion has concentrated on the *mutual exclusionary* effects that can be expressed by synchronization. To express the fact that two states can, in contrast, *co-occur*, we need only require that the initial state of the leftmost synchronizer in Figure 1 be open instead of closed. This will allow the co-occurrence of states {a} and {b}, that is, the state {a,b} can now occur. [*Reader exercise*: show that this possibility is unstable or fleeting, and that the system can decay into the earlier mutex form. This instability is the lot of the typical co-occurrence.]

We have thus seen that a synchronizer, which is an archetypic computational synchronization mechanism, can be arrayed to express both of the distinctions we are after - co-occurrence and exclusion. This particular pair of distinctions has the following properties:

- The elements of a co-occurrence are *indistinguishable* in time, in that by definition they occur neither before nor after each other. Thus, within a co-occurrence there is literally no "time" at all: a co-occurrence is a "now".
- Following Leibniz, co-occurring indistinguishables (namely, synchronization sticks) contain the germ of the concept of space. More generally, co-occurrence can be extended to encompass such static 'structural' aspects as form, situation, pattern, and the like.
- Two successive events of a given process by definition exclude each other.[5] Combining this with viewing "time" as a 1-1 mapping of the events constituting a given computation to a local time axis, we see that mutual exclusion contains the germ of sequential time. In general, every process constitutes a local *relative* time frame, which frame obtains meaning only via synchronization - that is, establishing before-after relationships - with other processes' frames.
- Just as co-occurrence contains the germ of the concept of space, exclusion's time-like aspect can be extended to express such dynamic concepts as action, transformation, intention, and the like.
- As a pair, co-occurrence and exclusion over the same states exclude each other, thus conceptually closing on each other and leading one to believe that they form a complete and minimal set of distinctions.

With (Rosen, 1991) in mind, we next investigate a little more closely the relationship between synchronization and Turing's model of computation.

1.2 Escaping from Turing's Box

An implicit claim of the Turing model is that a single sequence of computational events can capture all essential aspects of computation, that is, that *computation consists only of state transformations*. To refute this claim, consider the following gedanken experiment:

[4](Manthey,1992) argues the interpretation of this concept as the computational analog of quantum number conservation laws, and uses it to explain how the EPR 'paradox' is not a paradox at all.

[5]A consequence of the computational assumption of discreteness. Once can rightly say that synchronization is the handmaiden of discreteness.

Co-occurrence

***The coin demonstration** - Act I. A man stands in front of you with both hands behind his back, whilst you have one hand extended in front of you, palm up. You see the man move one hand from behind his back and place a coin on your palm. He then removes the coin with his hand and moves it back behind his back. After a brief pause, he again moves his hand from behind his back, places what appears to be an identical coin in your palm, and removes it again in the same way. He then asks you, "How many coins do I have?".*

It is important at the outset to understand that the coins are *formally* identical: indistinguishable in every respect. If you are not happy with this, replace them with electrons or geometric points. Also, I am not trying nefariously to slide anything past you, dear reader, in my prose formulation. What is at issue is the fact of indistinguishability, and I am simply trying to pose a very simple situation where it is indistinguishability, and nothing else, that is in focus.

The indistinguishability of the coins now agreed, the most inclusive answer to the question is "One or more than one", an answer that exhausts the universe of possibilities given what you have seen, namely *at least* one coin. There being exactly two possibilities, the outcome can be encoded in one bit of information. Put slightly differently, when you learn the answer to the question, you will per force have received one bit of information.

***The coin demonstration** - Act II. The man now extends his hand and you see that there are two coins in it. [The coins are of course identical.]*

You now know that there are two coins, that is, *you have received one bit of information.* We have now arrived at the final act in our little drama.

***The coin demonstration** - Act III. The man now asks, "Where did that bit of information come from??"*

Indeed, where *did* it come from?! Since the coins are indistinguishable, seeing them one at a time will never yield an answer to the question. Rather, *the bit originates in the simultaneous presence of the two coins.* We have called such a confluence a *co-occurrence*, and shown how it is computed in the preceding section. In that a co-occurrence, by demonstration a bona fide computational entity, is 'situational' rather than 'transformational', the assumption that computation is purely transformational is shown to be false.

To very briefly dispose of the most common counter-arguments:
Q: Whatever you do, it can be simulated on a TM.
A: You can't 'simulate' co-occurrence sequentially, cf. the coin demo.
Q: But you can only check for co-occurrence sequentially - there's always a Δt.
A: This is a technological artifact: think instead of constructive/destructive interference - a phase difference between two wave states can be expressed in one bit.
Q: One can simply define a TM that operates on the two states as a whole, so the "problem" disappears.
A: This amounts to an abstraction, which hierarchical shift changes the universe of discourse but doesn't resolve the limitation, since one can ask this new TM to 'see' a co-occurrence at the new level. In general, this type of objection dodges the central issue - what is the *mechanism* by which indistiguishables can be observed.
Q: Co-occurrence is primitive in Petri nets, but these are equivalent to finite state automata.
A: The phase web in effect postulates *growing* Petri nets, both in nodes and connections. All bets are then off.

[At this juncture, I hasten to mention that we are dealing here with *local* simultaneity, so there is no collision with relativity theory. Indeed, Feynman (Feynman,1965 p.63) argues from the basic principle of relativity of motion, and thence Einstein locality, that if *anything* is conserved, it must be conserved *locally*; see also (Phipps),(Pope& Osborne).

I ought also to mention that I am well aware that Penrose (1989) has argued that computational systems, not least parallel ditto, *in principle* cannot model quantum mechanics. However, I believe that his argument, together with most research involving (namely) parallelism in my own discipline, is subtly infected with the sequential mind-set, going back to Turing's analysis, and truly, earlier. An analogy with the difference between Newtonian and 20^{th} century physics is, to my mind, entirely defensible. The coin demonstration is my reply to such arguments, which I do not then expect to hold.

Notice by the way how the matrix-based formulations of QM neatly get around the inherent sequentiality of $y = f(x)$-style (ie. algorithmic) thinking, namely by the literal co-occurrence of values in the vectors' and matrices' very layouts; and thereafter by how these values are composed *simultaneously* (conceptually speaking) by matrix operations. Relating this now back to the phase web paradigm, if we assign an (arbitrary) ordering on sensor names, then co-occurrences become vectors, etc. Instead of the matrix route, I've taken the conceptually compatible one of Clifford algebras, which are much more compact, elegant, and general, cf. (Hestenes).

Returning to our discussion of Turing's model, we see from the coin demonstration that there is information, *computational information*, available in the universe *which **in principle** cannot be obtained sequentially.* Thus we have in the coin demonstration a compelling argument that, at the very least, the Turing model of computation fails to capture all relevant aspects of computation: it is semantically incomplete, and the thing it ultimately lacks is *space-time* - space: co-occurrence, time: mutual exclusion. Synchronization operators represent precisely the way computations can express space-time relationships and give them semantic content.

This can be taken further. Suppose we replace the coins by synchronization sticks, which are surely indistinguishable. We can then say that the information received from observing a co-occurrence is indicative of the fact that two states (represented by their sticks) do not mutually exclude each other.

<u>Co-Exclusion</u>

The block demonstration. *Imagine two 'places', p and q, each of which can contain a single 'block'. Each of the places is equipped with a sensor, s_p respectively s_q, which can indicate the presence or absence of a block.*

The sensors are the *only* source of information about the state of their respective places and are assumed *a priori* to be independent of each other, though they may well be correlated. The two states of a given sensor s are mutually exclusive, so a place is always either 'full', denoted (arbitrarily) by s, or 'empty', denoted by $\tilde{s}$; clearly, $\tilde{\tilde{s}} = s$.

Suppose there is a block on p and none on q. This will allow us to observe the co-occurrence $\{s_p, \tilde{s}_q\}$. From this we learn that having a block on p does not exclude not having a block on q. Suppose at some other instant (either before or after the preceding) we observe the opposite, namely $\{\tilde{s}_p, s_q\}$. We now learn that not having a block on p does not exclude having a block on q. What can we conclude?

First, it is important to realize that although the story is built around the co-occurrences $\{s_p, \tilde{s}_q\}$ and $\{\tilde{s}_p, s_q\}$, everything we say below applies equally to the 'dual' pair of co-occurrences $\{s_p, s_q\}$ and $\{\tilde{s}_p, \tilde{s}_q\}$. After all, the designation of one of a sensor's two values as '$\sim$' is entirely arbitrary. It is also important to realize that the places and blocks are story props: all we really have is two two-valued sensors reflecting otherwise unknown goings on in the surrounding environment. These sensors constitute the *boundary* between an entity and this environment.

Returning to the question posed, we know that s_p excludes $\tilde{s}_p$ and similarly s_q excludes $\tilde{s}_q$. Furthermore, we have observed the co-occurrence of s_p and $\tilde{s}_q$ and vice versa. Since the respective parts of one co-occurrence exclude their counterparts in the other co-occurrence (cf. first sentence), we can conclude that the co-occurrences *as wholes* exclude each other.

Take this now a step further. The transition $s_p \to \tilde{s}_p$ is indicative of some *action* in the environment, as is the reverse, $\tilde{s}_p \to s_p$. The same applies to s_q. Perceive the transitions $s_p \leftrightarrow \tilde{s}_p$ and $s_q \leftrightarrow \tilde{s}_q$ as two sequential computations, each of whose states consists of a single value-alternating bit of information. By the independence of sensors, these two computations are completely independent of each other. At the same time, the logic of the preceding paragraph allows us to infer the existence of a third computation, a *compound* action, with the state transition $\{s_p, \tilde{s}_q\} \leftrightarrow \{\tilde{s}_p, s_q\}$, denoted $s_p\tilde{s}_q$ or equivalently $\tilde{s}_p s_q$. In effect, by combining in this way two single-bit computations to yield one two-bit computation, we have lifted our conception of the actions performable by the environment to a new, higher, level of abstraction. This inference we call *co-exclusion*, and can be applied to co-occurrence pairs of any arity > 1 where at least two corresponding components have changed.[6]

Notice by the way that the same reasoning applies to $\{s_p, s_q\} \leftrightarrow \{\tilde{s}_p, \tilde{s}_q\}$, denoted $s_p s_q$ or $\tilde{s}_p\tilde{s}_q$. The two actions $s_p s_q$ and $\tilde{s}_p s_q$ are, not surprisingly, *dual* to each other, so co-exclusion on two sensors can generate two distinct actions. [As will be seen later, co-excluding the orientations of the duals produces a "complete" simplex at the next level up.] Like co-occurrence, an action defined by co-exclusion also possesses an emergent property, in this case generally comparable to spin $\frac{1}{2}$. This will be made clearer in the mathematical discussion below.

It sometimes troubles people that the elements of the co-occurrence (say) $\{s_p, \tilde{s}_q\}$ don't seem at all indistinguishable - on the contrary, s_p is clearly distinct from $\tilde{s}_q$! The confusion is understandable, and derives from confounding the *value* of a sensor with the synchronization *stick* that represents the fact that the value (= process state) obtains for the moment. The difference is clearer in the implementation, where the sticks for the respective states of the sensor processes s_p and s_q are represented by the tuples [p,1] and [q,$\tilde{1}$], which tuples can be thought of as making precise exactly *which* state's stick is being referred to. The processes accessing such tuples in fact know *a priori* the exact form of the tuple (ie. state) they are interested in, so no information is conveyed by accessing such tuples (which is as it should be, since synchronization must not convey information between processes). Summa summarum, the sensor values are not what are distinguished, but rather the sticks representing the associated sensor-process states, and these sticks are indistinguishable *in time*.

Finally, relative to the co-exclusion inference itself, it provides a very general (and novel [Manthey US]) way for an entity to learn from experience: simply observe co-excluding co-occurrences, since these then will represent an abstraction of experience. Furthermore, this is

[6]Greater arity is one way to exceed the binary limitation of ± 1 to obtain more nuance, though this will not be described further here. Also, the term 'inference' is to be taken in its generic, not its formal logical, sense: co-exclusion is more nearly inductive in its thrust.

also neurologically plausible, in that co-occurring synapse firings combine to exceed the nerve's threshold. The repetition required by neural systems to 'remember' is however short-circuited in Topsy: once is enough.

1.3 How Topsy Works

The trick now is to turn all these observations about co-occurrences and co-exclusion-based actions into something that can run on a computer, ie. Topsy. First, a few general observations:

- Even though I have made much of true concurrency, it is entirely okay to implement Topsy on an ordinary sequential computer, in that one may simply accept a certain Δt slop in co-occurrence detection. This of course means that information deriving from co-occurrences occurring at a granularity less than Δt will not be available - fair's fair.

- It's useful to think of processes as interacting by communicating with each other via some medium. In the case at hand, the medium is the computer's memory, but it could be wires, micro-waves, QM's spooky action-at-a-distance, or whatever. The determining distinction for present purposes is, rather, whether a given communication reaches all ("broadcast") or just a few ("point-to-point") of the other processes. For the phase web paradigm and hence Topsy, it is critical that the propagation regime be *broadcast*, so any process that might be interested in a given synchronization stick, even only potentially, will have access to it.

- A very neat way, due to (Raynal), to capture the distinction between truly distributed system architectures and their imitators is that whereas the imitators implicitly interpret a sent communication as a 'request' for information and a received communication as a 'reply' containing same (which is really the same old sequential $y = f(x)$ paradigm disguised as communication), processes inhabiting a truly distributed system interpret a communication sent as an 'announcement' of local state (ie. a stick), and received communications as other processes' ditto. Each process decides locally if/when/how it will react to the announcements of other processes. The request-reply regime is inherently centralizing, whereas the announce-listen regime is inherently distributive. It is a fact that virtually all contemporary distributed systems are, in this sense, imitators, quite despite appearances.

- I introduce the concept of a *goal* on-the-fly: a goal is an *explicit* expression of a state that the computation in which it occurs desires to reach. Their use in computing goes back to the 1960's in AI (if not earlier), and is also found in eg. the language Prolog. Goals may seem unusual, since they are at best implicit in traditional 'imperative' languages (and also in Prolog), but in fact there is nothing new here. Rather, the important thing to note is that, by being explicit, goals allow a program using them to 'remember' what it is supposed to be doing, and thus to recover from blind alleys. Furthermore, in being explicit, they allow the program to reason about them, and thus eg. reason about and resolve conflicts.

Topsy is formally connected to its environment by binary *sensors* and *effectors*, and these together constitute its *boundary*. Sensors are simple two-state processes, which two states are denoted $\{s, \tilde{s}\}$. Effectors are viewed as things that influence one or more sensors, and are therefore described as $s \rightarrow \tilde{s}$ and vice versa.

Each sensor state is, in the program, converted to a corresponding synchronization token, ie. the state s is converted to the token $(s, +1)$, and $\tilde{s}$ is converted to the token $(s, -1)$. Similarly, if an effector is in a state where it carry out the transformation $s \rightarrow \tilde{s}$, this is converted to the token $(s, +1, -1)$. A goal for this effector would, similarly, be expressed by the token $(!, (s, +1), (s, -1))$. In fact, *all* program states of interest are treated like this. In this way, all relationships between the processes constituting Topsy can be expressed via synchronization relationships alone: there is, as it were, no "data"... just processes announcing and listening for various synchronizational states.

Since an action is defined by co-excluding sensory processes, it expresses both a 'static' sensor-based aspect - deriving from its defining pair of co-occurrences - and an 'active' transformational aspect, deriving from the complementarity of these same co-occurrence pairs.[7] These two aspects suggest how to build up a running action, namely divide the code for an action into a half devoted to each side of the exclusion.

Thus, once the required pair of co-excluding co-occurrences $(s_p, \tilde{s}_q)$ vs. $(\tilde{s}_p, s_q)$ has occurred, a multi-threaded[8] action embodying the two transitions $(s_p, \tilde{s}_q) \rightarrow (\tilde{s}_p, s_q)$ and $(\tilde{s}_p, s_q) \rightarrow (s_p, \tilde{s}_q)$, is instantiated as a new entity; in a running Topsy system, there will be from hundreds to millions of these. One half of an action keys on the co-occurrence $\{s_p, \tilde{s}_q\}$ and the other on $\{\tilde{s}_p, s_q\}$. Since these co-occurrences exclude each other, only one of these halves will be activated at a time. When one of these pre-conditions occurs, and at least one associated goal is present, the action "wakes up". For example, when $\{s_p, \tilde{s}_q\}$ obtains, along with (say) the goal $s_p \rightarrow \tilde{s}_p$, the action fires and issues a goal for $\tilde{s}_q \rightarrow s_q$ as well. Thus a cascade of transformation goals propagates and activates other actions.

Actions carried out at the boundary (effectors) affect the environment, causing the sensors to reflect this new situation. This new situation bubbles up (see below) through the current aggregration of actions, orienting them to the new reality, and old goals are accordingly retracted and new ones issued. The seeming anarchy is controlled by the invisible hand of the dynamically nested synchronization invariants that the actions represent.

1.4 The Cycle Hierarchy

We have now at our disposal co-occurrences, co-exclusion and actions, and goals, and proceed to show how these can be combined recursively to yield a hierarchical structure. The basic claim here is that the ability to express the complexity and nuance of anticipatory behavior is to be found via the growth and interplay of hierarchical relationships. This growth, of course, occurs naturally and automatically via co-exclusion on sensory experiences.

The hierarchy is called the 'cycle hierarchy' because (1) the basic unit of its construction is co-excluding processes - the 'actions' described above - (2) whose internal conservation of synchronization sticks yields a basic cyclic structure (cf. Figure 1), (3) which cyclic structure

[7] It would really be better to call actions 'things', since traditionally a 'thing' is namely characterized by both aspects. One can also toy with the speculation that 'syntax' (ie. form) is based on the static, whereas 'semantics' (ie. function) is based on the active.

[8] A *thread* is CS jargon for a process possessing a relative minimum of own state.

is compounded recursively to yield a hierarchy of cycles of cycles.

The cycle hierarchy reflects a *weakly* reductionistic stance, in that it requires that any higher level phenomenon - which may well be emergent - be grounded in the structure and behavior of lower levels. This is in contrast to the endemic 'subroutine call' or 'function composition' hierarchy most people (especially scientists and engineers) unconsciously invoke in such discussions. This latter hierarchy is *strongly* reductionistic, in that it allows *no* place for phenomena that cannot be modelled by the sequential composition of lower level activities.[9] The basis of the cycle hierarchy in co-occurrences offers an interesting alternative to the reductive question of ultimate constituents, namely that one's hierarchical descent collides with the boundary to the environment. One is thus ultimately referred to "the rest of the universe", a result reminiscent of Leibniz's monadology.

Finally, although the following exegesis of the phase web's hierarchical structure presumes that the hierarchy is well-nested, ie. like one pancake on top of another, this is by no means necessary: co-exclusions can span over sensors from multiple levels (Figure 4a is a little misleading in this respect). Indeed, cycles in the hierarchy itself can be used to express self-propagating internal processes.

This overall sketch of hierarchical properties now behind us, we show how such hierarchies can be constructed in the first place. The basic insight is:

> GIVEN that every action possesses an innate polarity based on the orientation of its transformations, $\{s_p, s_q\} \rightarrow \{\tilde{s}_p, \tilde{s}_q\}$ vs. $\{\tilde{s}_p, \tilde{s}_q\} \rightarrow \{s_p, s_q\}$, which distinction maps to ± 1, co-occurrences of such action polarities can themselves be subjected to the co-exclusion inference, producing a meta-level of description/abstraction.

In other words, any action, whatever its arity, possesses two locally global states, corresponding to the two possible transitions it can accomplish. These two states exclude each other, which in turn means that this property of an action can be reflected in a two-valued sensor, a so-called *meta*-sensor. [A meta-sensor is in other respects just like a primitive sensor.]

Meta-sensors themselves can be co-excluded to produce meta-actions, which in turn - being, again, actions - possess the same polarities. These meta-polarities can again be mapped to a meta-meta-sensor, which can again be co-excluded to produce meta-meta-actions, etc. The result is a cycle hierarchy.

Notice that the two complementary co-occurrences whose co-exclusion defines an action also neatly specify the respective pre- and post-conditions for that action - for example, when the environment is in state $\{s_p, \tilde{s}_q\}$, the action's pre-condition is precisely $\{s_p, \tilde{s}_q\}$ and its post-condition is $\{\tilde{s}_p, s_q\}$; and vice versa.

When an action's pre-condition obtains, and if a goal to invert (at least) one of an action's constituent sensors co-occurs herewith, we say that the action is *relevant*. The action will then fire, ie. volunteer and broadcast goals to invert the actions's remaining constituent sensors, and in so doing attempt to achieve said goal from the micro-perspective of that action.[10]

---

[9] To adopt the third possibility, that of emergent phenomena in no way grounded in lower levels, is of course to abandon any consistent notion of cause and effect and therefore rational thought in general. To those readers who see red when the word 'emergent' is uttered, I note that the concept of emergent phenomena has a counterpart in the global properties found in mathematics, eg. curvature.

[10] That is, a given co-exclusion, say $\{s_p, \tilde{s}_q\} \leftrightarrow \{\tilde{s}_p, s_q\}$, reflects a particularized micro-view of reality that says,

Relevance can be similarly volunteered, on the reasoning "if s_p can be changed to $\tilde{s}_p$ then an action $\tilde{s}_p s_r$ can volunteer that $s_r \rightarrow \tilde{s}_r$ is possible, and therefore is relevant as well. Thus volunteering is a way to achieve the associative behavior characteristic of anticipatory systems.[11]

Volunteered goals will in general cause other relevant actions to fire, until a goal referring to an effector causes that effector to propagate the desired effect across the boundary to the environment on the other side thereof. This will ultimately change some sensor(s), setting off a wave of changes in the associated relevance relations, reflecting the new state of the environment. This interplay between the state of the environment and Topsy's goals occurs continually, with current goals changing dynamically in reaction to the environment's response to the effects of earlier goals.

Besides volunteering, one other implicit and dynamic mechanism is necessary, namely a means for propagating relevance and goals from level to level. This is accomplished by reflecting an action's relevance in an associated meta-sensor, whence the same thing will take place for meta-meta-sensors, etc. We call this process the *bubbling up* of sensory impressions.

Similarly, a goal to invert a meta-sensor will be reflected by the associated meta-effector's issuing goals to the level below. Since a given meta-sensor represents in one bit the state of a co-occurrence, ie. *more than one* sensor, a meta-effector fans goals out, level by level, on their way down toward the primitive effectors at the environmental boundary. We call this process the *trickling down* of goals.

The hierarchy-construction process leads to a number of features and properties deserving mention:

- The meta-sensors and meta-effectors of a given level form the *boundary* between that level and the level below. It follows that the boundary constituted by the primitive sensors and primitive effectors is, conceptually, entirely arbitrary.
- Since the environment is formally unbounded in its complexity, it follows that the hierarchy must be as well. And it *is* formally unbounded, in that if we abandon the pancake restriction, the number of entities that can be co-excluded increases hyper-exponentially: $3, 7, 127, 2^{127} - 1$. This is an instance of *the combinatorial hierarchy* (Bastin and Kilmister), (Parker-Rhodes), (Manthey, 1993).
- Co-exclusion over meta-sensors is inherently introspective and self-reflective, in that meta-sensors themselves explicitly express internal, situated states. The capture of an internal relationship by a co-exclusion elevates what was previously implicit and 'unconscious' to an explicit object.
- We have seen that sensory impressions S bubble up and goals G trickle down. A given *meta*-level $n+1$ is built over $S_n \times S_n$, and serves to further *classify* sensory impressions. When level $n+1$ is based only on level n, we say the hierarchy is a *flat* or *pancake* hierarchy.

"given the goal $s_p \rightarrow \tilde{s}_p$, if s_p is to change to $\tilde{s}_p$, then this means that $\tilde{s}_q$ *must* change to s_q", and so 'volunteer' the goal $\tilde{s}_q \rightarrow s_q$, which goal, like the first, is visible to all other actions.

[11] In a traditional frame-based AI systems, this is called 'spreading activation', but it should be apparent that, although the effect of the two processes is analogous, the mechanisms are quite different.

- One can also consider hierarchies built by co-excluding over $G \times G$ and $S \times G$, ie. 'meta'-sensors sensing goal-co-occurrences, and 'meta'-effectors and 'meta'-actions manipulating goals for $G \times G$; and analogously for $S \times G$.
  - For conciseness, we write (eg.) $G \times G$ for $G \times G \times \ldots \times G$.
  - $G \times G$, captures relationships between goals, and, via hierarchical expansion, can express the structure of arbitrarily complex purposive *activities*. Since $G \times G$ actions are grounded in goals, which themselves are primarily internal, their hierarchy is increasingly less grounded in environmental reality, wherefore I have dubbed such actions *icarian*.
  - $S \times G$, expresses the interplay between up-bubbling sensory impressions and down-trickling intentions. Since $S \times G$ in a sense 'covers' both $S \times S$ and $G \times G$, I consider $S \times G$ to be the most profound, and call such actions *morphic*. Note that morphic actions provide a means for expressing the *self*-generation of goals given sensory situations (read self-choice), and in the other direction, the self-generation of sensory situations given goals (read imagination).

    Thus the basic phase web mechanisms of co-occurrence and co-exclusion, re-applied, can create three distinct *types* of hierarchy. In addition, entities belonging to each of these can themselves be similarly combined ad infinitum. This should provide sufficient expressive power for even the most demanding application.

- Bubbling up in an $S \times S$ hierarchy corresponds roughly to integration ($\int$), whereas trickling down in the corresponding dual goal hierarchy corresponds to differentiation (∂). The $R \times G$ actions connecting them correspond then to the meeting of a goal and a currently obtaining state, leading to 'action'. This is elaborated in the mathematical section

- One can draw an analogy with Huygen's principle as recently elucidated by Jessel (Bowden, in press), which says that any radiating primary source, can, when surrounded by an arbitrary boundary, be simulated by a finite number of appropriately tuned secondary radiators placed on that boundary. Thus hierarchical ascent can be compared to approaching the original primary source. That the cycle hierarchy is at the same time formally unbounded leads to a meta-physically satisfying outcome. below.

- I believe, though without being able to demonstrate it, that moving upward in the morphic hierarchy corresponds to a shift to a more powerful system in the context of Gödel's incompleteness arguments.

Finally, the initial discriminatory basis for the hierarchy construction - the tensions between *excludes* and *co-occur* and *co-exclude* in *time* - seems to blur in their interplay the traditional distinction between epistemology and ontology. This obtains because, while co-occurrences and co-exclusion-based actions together constitute the universe of 'ontological objects', their discovery (ie. epistomology) invokes the very same properties. Only after one has built up considerable structure - corresponding to traditional space-time - would one seem to be able to clearly separate the two.

2 The Mathematical Model

This section presents, very informally, the most important mathematical aspects of the phase web paradigm. In general, the vector orientation of the present approach is unique in computing, which has traditionally been logic-oriented.

The point of departure is to view sensor states as vectors instead of scalars, as is conventionally done.[12] Figure 2a shows a single sensor's states so expressed, and Figure 2b the way two such vectors can indicate a state, eg. of an action.

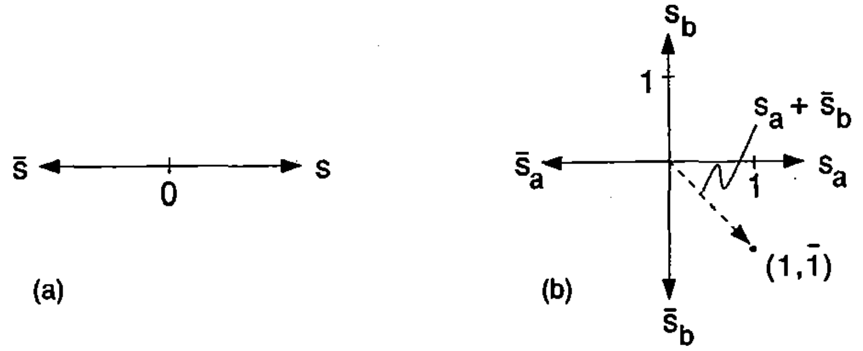


Figure 2: Sensors as vectors.

The sensor state $s = 1$ indicates that sensor s is currently being stimulated, ie. a synchronization stick for that state is present, whereas $s = \tilde{1}$ indicates that s is currently *not* being stimulated, and hence no stick for state s is present. Thus the two states of s are represented by the respective semaphore values introduced in the definition of wait and signal in §1.1.

That the sensors *qua* vectors are orthogonal derives from the fact that, in principle, a given sensor says nothing about the state of any other sensor. A state of a multi-process system such as that depicted in Figure 2b is then naturally expressed as the sum of the individual sensor vectors. For example, the state $(s_a, \tilde{s}_b) = (1, \tilde{1})$ is written as the vector sum $s_a + \tilde{s}_b$, which also introduces the visual convention that a vector component written without a tilde is taken to be bound to the value 1, and vice versa. Since such states represent co-occurrences, it follows that co-occurrences are vector sums. Note how the commutativity of '+' reflects the lack of ordering of the components of a co-occurrence.

The next step is to find a way to represent actions mathematically. (Manthey,1994) presents a detailed analysis of the group properties of both co-occurrences and actions, concluding that the appropriate algebraic formalism is a (discrete) Clifford algebra, and that the state transformation effected by an action is naturally expressed using this algebra's vector product. A prime characteristic of this product is that it is anti-commutative, that is, for $(s_1)^2 = (s_2)^2 = 1$, $s_1 s_2 = -s_2 s_1$.[13] The magnitude of any such product is the area of the parallelogram its two

[12]A *scalar* is simply a magnitude, whereas a *vector* is a magnitude together with a direction (orientation). The operations on vectors ($+, \cdot, \wedge$) ensure that one's intuitive expectations for how things combine are maintained.

[13]The Clifford product ab can be defined as $ab = a \cdot b + a \wedge b$, ie. the sum of the inner ($\cdot$) and outer ($\wedge$) products, where $a \wedge b = -b \wedge a$ is the oriented area spanned by a, b. The vector cross product $a \times b$ familiar to many is a poor man's version of $a \wedge b$ introduced by Gibbs. The basis vectors s_i of a Clifford algebra may have $(s_i)^2 = \pm 1$, and while here we choose $+1$, reasons are appearing for choosing -1. As long as they all have the same square,

components span, and the *orientation* of the product is perpendicular to the plane of the parallelogram and determined by the "right hand rule".

Applying the Clifford product to a state, one finds - using the square-rule and the anti-commutativity of the product given above - that

$$(s_1 + s_2)s_1 s_2 = s_1 s_1 s_2 + s_2 s_1 s_2 = s_2 + \tilde{s}_1 s_2 s_2 = \tilde{s}_1 + s_2 \tag{1}$$

that is, that the result is to rotate the original state by 90^o, for which reason things like $s_1 s_2$ are called *spinors*. Thus *state change* in the phase web is modelled by rotation (and reflection) of the state space, and the effect of an 'entire' action can be expressed by the inner automorphism $s_1 s_2(s_1 + s_2)s_2 s_1 = \tilde{s}_1 + \tilde{s}_2$, which corresponds to a rotation through 180^o.[14] It is interesting to note that $(s_1 s_2)^2 = -1$, that is, the s_1, s_2-plane is the so-called *complex* plane, and thus that $i = \sqrt{} - 1$ is intimately involved.

One of the felicities of Clifford algebras is that one needn't designate one of the axes as 'imaginary' and the other as 'real'. Rather, the i-business is implicit and the algebra's anti-commutative product neatly bookkeeps the desired orthogonality and inversion relationships.

The above spinors are just one example of the vector products available in a Clifford algebra - any product of the basis vectors s_i is well-defined, and just as $s_1 s_2$ defines an area, $s_1 s_2 s_3$ defines a volume, etc. Not least because they are all by nature mutually perpendicular, the terms of a Clifford algebra

$$s_i + s_i s_j + s_i s_j s_k + \ldots + s_i s_j \ldots s_n \tag{2}$$

themselves also define a vector space, which is the space in which we will be working. [The term (eg.) $s_i s_j$ above, for $n = 3$, denotes $s_1 s_2 + s_2 s_3 + s_1 s_3$, that is, all possible non-redundant combinations.]

At this point it is perhaps worth stressing that this vector space is the space of the *distinctions* expressed by sensors, and as such has no direct relationship whatsoever with ordinary 3+1 dimensional space. The latter must - at least in principle - be built up from the primitive distinctions afforded by the sensors at hand. This too is treated as a discrete space, rather than the usual continuous ditto.

A Clifford product like $s_1 s_2$ reflects both the emergent aspect of a phase web action (via its perpendicularity to its components) and its ability to act as a meta-sensor (since its orientation is ± 1).

One might therefore expect that the co-exclusion of two such meta-sensors, say $s_i s_j$ and $s_p s_q$, would be modelled by simply multiplying them, to get the 4-action $s_i s_j s_p s_q$. This turns out however to be inadequate, since although by the same logic the co-exclusion of (say) s_i and $s_i s_j$ in Topsy expresses explicitly a useful relationship (eg. part-whole), the algebra's rules reduce it from $s_i s_i s_j$ to s_j, which is simply redundant.

Instead, we take as a clue the fact that goal-based *change* in Topsy occurs via trickling down through the layers of hierarchy, and draw an analogy with differentiation. In the present decidedly geometric and discrete context, differentiation corresponds to the *boundary operator* ∂. Informally, define $\partial s = 1$ and let

$$\partial(s_1 s_2 \ldots, s_m) = s_2 s_3 \ldots s_m - s_1 s_3 \ldots s_m + s_1 s_2 s_4 \ldots s_m - \ldots (-1)^{m+1} s_1 s_2 \ldots s_{m-1}$$

it doesn't matter for what is said here.

[14]Some readers might recognize this when written in the form $as = s'a$.

that is, drop one component at a time, in order, and alternate the sign. Using the algebra's rules, one can show that

$$\partial(s_1 s_2 \dots s_m) = (s_1 + s_2 + \dots + s_m) s_1 s_2 \dots s_m$$

which is exactly the form of equation (1) for what an action does.

The boundary operator ∂ has a straightforward geometric interpretation. Consider an ordinary triangle ABC specified in terms of its vertices A, B, C, whence its edges are AB, BC, CA. Then

$$\partial(ABC) = BC - AC + AB$$

Since specifying the triangle's edges in terms of its vertices means that edge AC is oriented oppositely to edge CA, we can rewrite the above as $AB + BC + CA$, which is indeed the boundary of the triangle (versus its interior).

To find co-exclusion in this, we exploit the geometrical connection further. Take exppression (2) expressing the vector space of distinctions, segregate terms with the same number of product-components (*arity*), and arrange them as a decreasing series:

$$s_i \xleftarrow{\partial} s_i s_j \xleftarrow{\partial} s_i s_j s_k \xleftarrow{\partial} \dots \xleftarrow{\partial} s_i s_j \dots s_{n-1} \xleftarrow{\partial} s_i s_j \dots s_n \tag{3}$$

Here as before, $s_i s_j$ is to be understood as expressing all the possible 2-ary forms (etc.), and hence the co-occurrence of pieces of similar structure. Each of the individuals is a *simplicial complex*, and the whole mess is called a *chain complex*, expressing a sequence of structures of graded geometrical complexity in which the transition from a higher to a lower grade is defined by ∂. Furthermore, the entities at adjacent levels are related via their group properties - their *homology*, which I here assume is trivial.

Still on the scent of co-exclusion, it turns out that there is a second structure - a *cohomology* - that is isomorphic to ("same *form* as") the homology, but with the difference that arity (complexity) *increases* via the δ (or *co-boundary*) operator,[15] precisely opposite to ∂ (cf. equation (3)):

$$s_i \xrightarrow{\delta} s_i s_j \xrightarrow{\delta} s_i s_j s_k \xrightarrow{\delta} \dots \xrightarrow{\delta} s_i s_j \dots s_{n-1} \xrightarrow{\delta} s_i s_j \dots s_n \tag{4}$$

Building such increasing complexity is exactly what co-exclusion does. [I note that a Clifford algebra satisfies the formal requirements for the existence of the associated homology and cohomology.]

Figure 3, due to (Bowden,1982), illustrates these relationships (eqns. 3,4). I call this a *ladder diagram.*

The left side of the ladder is the homology sequence generated by ∂ over the representation of actions as Clifford products. The downward flow of decomposition of the structure into simpler pieces (ie. the crossing of successive boundaries) corresponds to the trickling down of goals described earlier.

The right side of the ladder is similarly the cohomology sequence generated by δ from sensory impressions. The upward flow of composition of structure to form more complex structure corresponds to the effect of co-exclusion, up through which increasingly complex structure sensory impressions bubble.

---

[15]More precisely, $(\sigma_p, \delta d^{p-1}) = (\sigma_p \partial, d^{p-1})$, where σ_p is a simplicial complex with arity p, and d^p the corresponding co-complex.

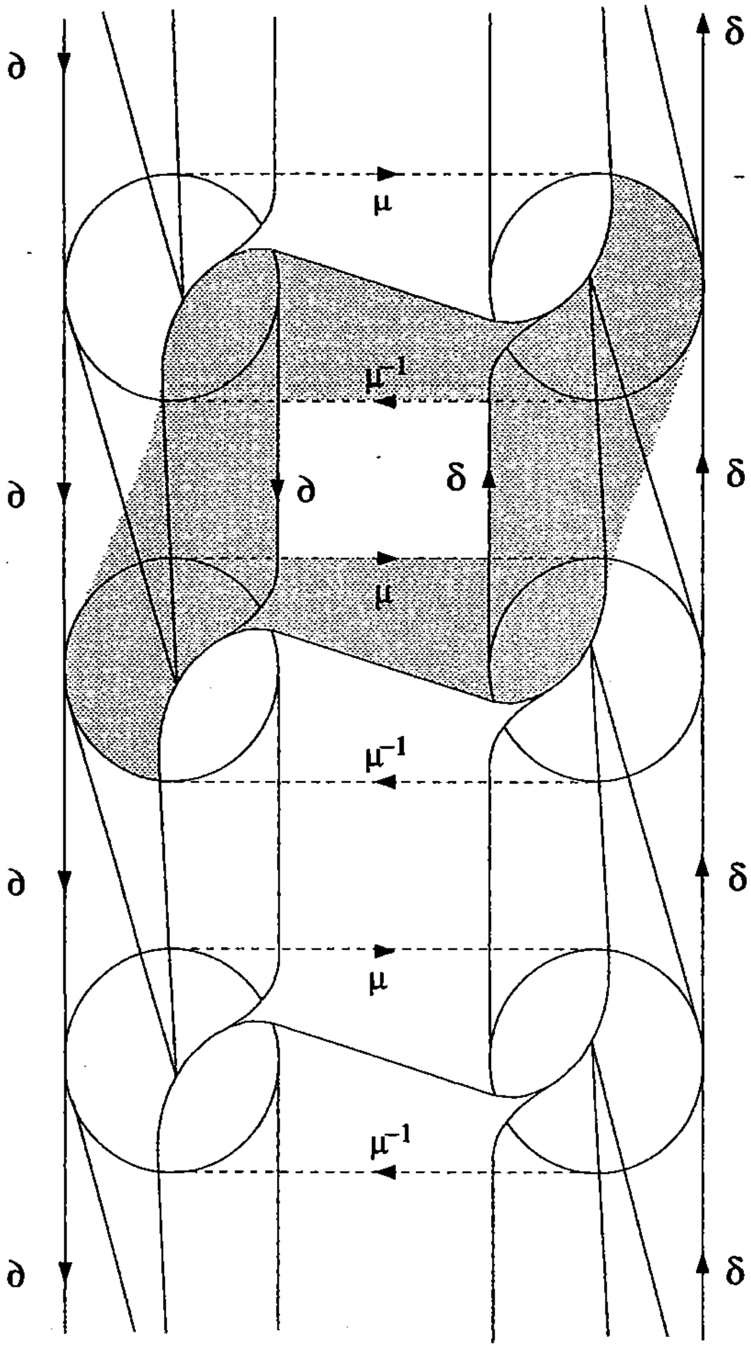


Figure 3: Ladder diagram, illustrating homology-cohomology relationships.

The circles represent all the entities (Clifford algebra terms) at the particular level of complexity. The larger of the two circle-halves holds those entities which will map to zero with the next hierarchical transition (∂ or δ) - called the *kernel* of the group - as indicated by the pointed 'beak'.

The rungs of the ladder, besides denoting the location and content of cycle hierarchy levels, also express the fact that there exist isomorphisms (μ, μ^{-1}) between the structures at either end of a given rung.[16] The shaded portion, which can be seen to repeat in both directions, expresses the so-called commutation relationships that obtain. That is, if one chooses a particular group element and follows the transforming arrows around the interior box, one not only arrives back where one began, but also back at the exact same *element* one began with! One says that the isomorphisms *commute*, and one may also take longer paths, though always obeying the box-arrows (otherwise the commutation relation generally won't hold).

The shaded shape points out a unique property of the homology-cohomology ladder, one that even most topologists seem unaware of, namely that the isomorphisms μ, μ^{-1} are *twisted*, that is, the kernel of the group at one end of a rung is mapped by μ (respectively, μ^{-1}) into the non-kernel elements of the group at the other end. This property was discovered by (Roth) in his proof of the correctness of Gabriel Kron's (then controversial) methods for analyzing electrical circuits (Bowden), and turns out to have profound implications. For example, the entirety of Maxwell's equations and their interrelationships can be expressed by a ladder with two rungs plus four terminating end-nodes (Bowden), and (Tonti) has - independently - shown similar relationships for electromagnetism and relativistic gravitational theory. Roth's twisted isomorphism (his term) thus reveals the deep structure of the concept of boundary, and shows that the complete story requires both homology and cohomology.

I interpret the twisted isomorphism to be expressing a deep complementarity between the concepts of action and state, between exclusion and co-occurrence. In the running Topsy program, μ, μ^{-1} connect goals' trickling down to sensory states' bubbling up. Think now about this: suppose you follow only the goal/homology side down. As boundary after boundary is crossed, all that happens is that a larger goal is split into successively narrower subsidiary goals. Imagine that you are following a ladder structure that describes the entire Universe. When and where does the actual *change* occur?!

The answer is that it never does, as long as you stick to the homology side of the ladder. Similarly, if you stick to the cohomology side, states never turn into goals: there is eternal stasis. It is the mappings μ, μ^{-1} that allow dynamic and change, converting something that doesn't exist, even conceptually, on the one side to something that constitutes the conceptual universe of the other. But this is what morphic actions do explicitly, so μ, μ^{-1} correspond to actions over $S \times G$.[17]

Summarizing, we have seen how morphic actions correspond to μ, μ^{-1}, and it should be clear that the conversion of meta-sensors to meta-actions[18] via co-exclusion ($S_n \times S_n$) corresponds to δ; similarly, icarian actions ($G_n \times G_n$) correspond to ∂; non-pancake hierarchies are, of course,

[16]Strictly speaking, μ/μ^{-1} should be indexed by level, μ_m/μ_m^{-1}.

[17]Given that I associate wave properties with the concept of co-occurrence and particle properties with exclusion and action, this means that μ, μ^{-1} express wave-particle duality, but in a hierarchical structure that itself expresses no difference between the microscopic and the macroscopic.

[18]Linguistic and conceptual purity would demand, since we're on the δ/state side of the ladder, that I write 'meta-object' or 'meta-state' instead of 'meta-action', but this distinction is blurred outside of a mathematical context, so I don't.

more complex mathematically. Figure 4 illustrates two possibilities, both non-pancake.

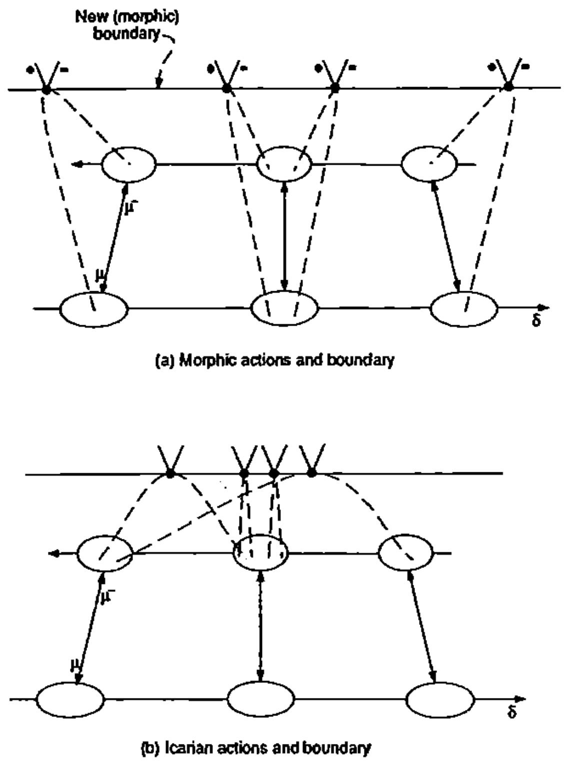


Figure 4: Morphic vs. icarian hierarchies.

In conclusion, the attempt to put the phase web and Topsy on a firmer mathematical footing turns out to lead to the same mathematical structures as underpin both contemporary physical and (I understand) bio-physical theory. Things could hardly have turned out better for a novel computational approach to expressing information, learning, behavior, and in general, the structures and mechanisms pertinent to anticipatory systems.

3 Modelling Auto-Poiesis

Our goal now is to describe auto-poiesis (Maturana and Varela, 1987) using the apparatus of the phase web. We have two ways to express the latter at hand, a computational way and mathematical way. The former provides many important details that the latter lacks, but since this lack is an advantage in the present context, we will couch the description with reference to the mathematical version.

Auto-poesis invokes three interdependent activities: (1) sensing the environment and categorizing the information so gained; (2) acting on this information in a manner consistent with the structure of said environment; and finally, (3) acting in such a way as to be able to maintain, and even improve, the ability to carry out these three activities.

It should be stressed that the unpredictability of the environment effectively precludes a completely pre-programmed solution, if for no other reason than that the combinatorial complexity of the impulses coming from the environment together with the desired response, if explicitly listed, would simply take up too much 'room'; that is, such a solution is, from a bio-engineering point of view, implausible. This conclusion obtains even when, as is the case with (say) insects, the top-level behavior is apparently very rigid, since even walking about or flying cannot be similarly rigid. Accordingly, we assume that the solution must be non-deterministic and adaptive in character.

We now take up each of the above three activities.

3.1 Sensing and Categorizing

The key problem here is the sheer combinatorics of the sensory input: n sensors taken $1, 2, 3 \ldots n$ at a time yields 2^n possible combinations. Since even a simple cell can be construed to interact with its environment over most of its surface, clearly n is very large. Moreover, this analysis does not consider that it is not least the *order* in which the various sensory combinations occur that is important, in which case things grow factorially.

The concept of *hierarchy* is therefore a critical bio-engineering *tool*. This is so because, as already noted by (Simon, 1967), a hierarchical organization reduces complexity logarithmically. However, even Simon's analysis implicitly invokes a subroutine-call/functional hierarchy, and overlooks the fact that the multiplicity of simultaneous impulses from the environment requires multiple, likely overlapping, hierarchies, one for each context so invoked.

This is where the bubbling-up of sensory impressions characteristic of the cycle hierarchy (ie. the δ-side of the ladder) comes to the rescue: the δ structure contains implicitly all *possible* hierarchies[19], and the bubbling-up process orients the structure to *which* hierarchies are relevant in the given context. The particular orientation thus achieved is at the same time *ipso facto* a *categorization* of the the sensory stimuli in that context.

Therefore, we model the categorization of an organism's sensings as a simple meta-hierarchy, where each level builds essentially directly on the level below. How many such levels there might be for this is presumably dependent on the organism's complexity. The top level of such a hierarchy will contain the highest-level categorizations, and presumably most contexts will "light up" several, though distinct, top-level category nodes.

3.2 Environmentally Appropriate Behavior

Once these categorizations of the environment are available, we can address the issue of choosing an appropriate response. A given response can in general be expected to span several modalities (ie. distinct "effectors").

Even though the response is not required to be especially prescient (this is the responsibility of (3) above), it is nevertheless a fact that even the most trivial 'intelligent action' requires a fair amount of organization and coordination to achieve. For example, bio-engineering economy requires that the same components be used to respond to similar situations, and ensuring that this functional overlap does not get in the way of the correct response in the *particular* situation leads to the need for the aforementioned coordination. The example in (Manthey,1996) illustrates this clearly.

[19]Modulo, of course, the distinctions the organism is in fact capable of.

Another way to put this is that, given that the categorization process selects one or several categories out of many, the generation of appropriate behavior can be characterized as the ability to control the transition from the current category-set to an intended new ditto. This in turn requires a structure that spans over the existing categorization structure. We write 'span over' (when 'span' alone would have been sufficient) to emphasize that we are not speaking here of simply adding levels on top of the categorization meta- hierarchy, but rather a *morphic* hierarchy (cf. eg. Figure 4a) that literally spans *over* the entire categorization structure, more or less from top to bottom.

This superstructure, besides exhibiting clearly the greater flexibility and generality of the ladder hierarchy compared to functional ditto, is genuinely self-reflective. This obtains, as should be clear from the figure, because it is the mechanism and dynamics of the categorization process *itself* that are being abstracted over.

We choose a morphic hierarchy primarily because a meta-hierarchy (which might otherwise be considered as a candidate here) cannot itself generate its own goals (being constructed over $S \times S$), whereas this is a principle characteristic of a morphic hierararchy (which is constructed over $S \times G$). The morphic superstructure, consisting (say) of two-four levels, has the responsibility for issuing goals to the underlying meta-hierarchy. It is these goals that control the transition from the current situation to a desired new one. [In this connection, it is perhaps appropriate to emphasize that the structures we are describing are entirely comfortable with unexpected reactions to their effector-born manipulations from the environment: the goal-driven regime ensures that the organism will adaptively pursue the achievement of its intents in the face of non-deterministic outcomes.]

3.3 Auto-Poietic Behavior

The goal of auto-poietic behavior is to ensure the continued existence of the organism. To this end, it seems obvious to simply iterate the logic of the preceding construction. That is, just as we above used a spanning hierarchy to self-reflect the categorization process in order to control the transitions between categorizations, we now introduce a second spanning hierarchy to self-reflect over the transition-control process.

This second spanning hierarchy must, like the preceding, be able to generate goals, although in this case the goals are meant to control the long-term behavior of the organism. That is, it is entirely conceivable that in all but the most primitive organisms, there are a number of more or less mutually exclusive reactions that could be generated in a given situation. The preceding morphic hierarchy cannot 'intelligently' choose among these possibilities because their mutually exclusive properties are not explicitly visible to it(self). Thus, another way to describe the utility of the superstructures we are building here is that they make explicit various relationships that are entirely implicit in the structure over which they brood.

Whether this second self-reflective level should be morphic or icarian is at this point a matter of speculation. Icarian structures, because they are built solely over $G \times G$, that is the co-occurrence or mutual exclusion of goals themselves, are naturally suited to controlling longer sequences of actions, ie. actions that must be carried out in a *particular* order in order for the whole sequence to be successful. To accomplish this, icarian actions interact with their root boundary by retracting and (perhaps later, re-)issuing goals 'belonging' to the underlying hierarchy, here the morphic one just described. In that goals are purely internal to the organism, the danger exists that a deep icarian hierarchy can become so involved in pursuing its own goals that it loses sight of the actual environmental situation and feedback. In contrast, a morphic hierarchy avoids this - it's by definition (more) directly connected to the environment via its S-component - at the expense of tending toward hard-wired reaction and a weaker ability to manage the interplay between goals.

In either case, one can argue that this auto-poietic superstructure should be rooted not only in the underlying morphic hierarchy but also in the original 'primitive' level. This will help the overall structure to maintain a closer connection to environmental reality. One can also consider literally feeding-back the results of this stage into the lower level(s), again in the errand of greater cohesion of the whole.

Of course, there is still no guarantee that this second self-reflective level will succeed in the long term, and the same applies no matter how many such levels exist, so at some point, one must appeal to natural selection and its accompanying phenomena for the final auto-poietic judgement.

Howsoever, we claim that for an organism to display auto-poietic behavior, it must possess, as a minimum, the above-described three level-structures - one to categorize the current state of the environment vis a vis the organism, one to connect these categorizations to immediate reactions, "tactics", and one to oversee its long-term behavior, "strategy".

4 Summary and Conclusion

We have presented a situated computational model - the *phase web* - that we believe capable of describing the true structure and behavior of anticipatory systems. It is a *pure process* model whose fundamental departure from traditional algorithmic thinking allows it to meet Rosen's (1991) criticism of the latter, and which can express emergent phenomena. Being a *computational* model, it is able to propose explicit mechanisms and processes for acquiring and using, adaptively, information from its environment. The key insight is to express the desired activities in terms of *patterns of synchronization* among the events constituting an organism. These patterns fall into two distinct categories - co-occurrence and mutual-exclusion - that together are capable of expressing the concepts of event (synchronization itself), process (via exclusion), information (cf. the coin demonstration), space (via co-occurrence), time (via exclusion), action (via co-exclusion), structure (via hierarchy), self-reflection (via co-occurrence and exclusion over internal events), intent (via goals), etc. As a result, the model need not appeal to mechanisms outside of itself, that is, the modelling tools themselves exhibit logical closure, and are, apparently, complete.

After a brief description of how a program embodying these concepts (ie. Topsy) actually works, we showed how this same model can also be described in terms of algebraic topology. The key identifications making this possible are: (1) a binary sensor can be viewed as a vector, and the set of sensors connecting an system to its environment as an orthonormal basis; (2) the sum of sensor vectors captures the concept of co-occurrence and their (Clifford) product the concepts of exclusion and action; (3) the co-exclusionary property of complementary co-occurrences allows the composition of sensory-object abstractions from environmental stimulation and corresponds to the co-boundary operator δ; (4) this induces a hierarchy of co-boundaries and co-chain complexes, which in turn, via Roth's twisted isomorphism, (5) induces a corresponding and *isomorphic* homology - ie. a hierarchy of boundaries and chain complexes - that expresses the *de*composition (via the boundary operator ∂) of goals on meta-sensors into sub-goals/sub-actions, leading ultimately to externally-directed effects on the environment; and finally, (6) this *ladder hierarchy* yields three distinct types - meta, morphic, and icarian - corresponding (very roughly) to classification, situated reaction, and goal-interaction.

Finally, the preceding section showed how to apply this model to the description of auto-poietic behavior. We concluded that an auto-poietic system must as a minimum contain three distinct hierarchies, one to classify sensory input in a (presumed) non-deterministic regime; one to self-reflectively react to the current, now classified, situation in an appropriate and controlled manner; and one to self-reflectively choose among the possible, now identified, reactions, and in so doing pursue the overall goal of achieving and maintaining auto-poiesis.

We claim as well that this cannot be achieved within reasonable bio-engineering constraints without invoking the ladder-hierarchical structures we have described. We claim further that these hierarchical structures must be interconnected essentially as described in order to obtain the self-reflectivity without which very little of this behavior can be achieved.

Finally, we claim that these are not burdensome constraints at all, but rather just those that are needed, both in terms of modelling apparatus and technique, to describe that which is so very, very special about anticipatory systems.

Acknowledgements

Thanks to IEEE for permission to reprint the coin and block demonstrations. And special thanks to Peter Marcer, Arne Skou, and my students.

References

Bastin, T. and Kilmister, C.W. *Combinatorial Physics*. World Scientific, 1995. Isbn 981-02-2212-2.

Bastin, T. and Kilmister, C.W. "The Combinatorial Hierarchy and Discrete Physics". Int'l J. of General Systems, special issue on physical theories from information (in press).

Bowden, K. "Physical Computation and Parallelism (Constructive Post-Modern Physics)". Int'l J. of General Systems, special issue on physical theories from information (in press).

Bowden, K. *Homological Structure of Optimal Systems*. PhD Thesis, Department of Control Engineering, Sheffield University UK. 1982.

Feynman, R. *The Character of Physical Law*. British Broadcasting Corp. 1965.

Gelernter, D. "Generative Communication in Linda". ACM Transactions on Programming Languages, 1985.

Hestenes, D. and Sobczyk, G. *From Clifford Algebra to Geometric Calculus*. Reidel, 1989.

Hestenes, D. *New Foundations for Classical Mechanics.* Reidel, 1986. The first 40 pages contain a very nice, historical introduction to the vector concept and Clifford algebras.

Manthey, M. "Synchronization: The Mechanism of Conservation Laws". Physics Essays (5)2, 1992.

Manthey, M. "The Combinatorial Hierarchy Recapitulated". Proc. ANPA 15 (Cambridge, England), 1993.

Manthey, M. "Toward an Information Mechanics". Proceedings of the 3rd IEEE Workshop on Physics and Computation; D. Matzke, Ed. Dallas, November 1994. Isbn 0-8186-6715X.

Manthey, M. et al. Topsy - A New Kind of Planner. Working paper, 1996. See [www].

Manthey, M. US Patents 4,829,450, 5,367,449, and others pending. My intent is to license freely (on request) to individuals and research or educational institutions for non-profit use. See [www] for licensing information.

Maturana, H.B. and Varela, F.J. *The Tree of Knowledge - the biological roots of human understanding*. New Science Library; Shambhala, 1987. Isbn 0-87773-403-8.

Phipps, T. "Proper Time Synchronization". Foundations of Physics, vol 21(9). September, 1991.

Penrose, R. *The Emperor's New Mind.* Oxford University Press, 1989. Isbn 0-19-851973-7.

Pope, N.V. and Osborne, A.D. "Instantaneous Relativistic Action-at-a-Distance". Physics Essays, 5(3) 1992, pp. 409-420.

Parker-Rhodes, F. *The Theory of Indistinguishables - A Search for Explanatory Principles Below the Level of Physics*. Reidel. 1981.

Raynal, M. *Algorithms for Mutual Exclusion.* MIT Press, 1986. Isbn 0-262-18119-3.

Robert Rosen, *Anticipatory Systems*. Pergamon Press, 1985.

Rosen, R. *Life Itself - A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life.* Columbia University Press, 1991. ISBN 0-231-07564-2.

Roth, J.P. "An Application of Algebraic Topology to Numerical Analysis: On the Existence of a Solution to the Network Problem". Proceedings of the US National Academy of Science, v.45, 1955.

Simon, H. The Sciences of the Artificial. MIT Press 1982 (reprinted from ca. 1967).

Tonti, E. "On the formal structure of the relativistic gravitational theory". Accademia Nazionale Dei Lincei, Rendiconti della classe di Scienze fisiche, matematiche e naturali. Serie VIII, vol. LVI, fasc. 2 - Feb. 1974. (In english.)

www. Various phase web and Topsy publications, including (soon) code distribution, are available via *www.cs.auc.dk/topsy.*

A Pure Process View of Anticipatory Systems

Michael Manthey
Aalborg University
Aalborg, Denmark

ABSTRACT [1]

This paper presents a *process-oriented* computational view of the concept of an anticipatory system. That is, the focus is not on algorithms that produce "results", but rather on how to exploit the fundamental properties of processes to express the structure and behavior of anticipatory systems. Four ways such systems can anticipate the future are via conservation laws, goals, hierarchy, and self-reflection. This paper explains, descriptively rather than rigorously, how this is done in a pure process regime (elsewhere referred to as the phase web paradigm).

KEYWORDS

Process, hierarchy, co-exclusion, co-occurrence, synchronization, system, auto-poiesis, conservation, invariant, anticipatory, homology, twisted isomorphism, phase web paradigm, Topsy, reductionism, emergence.

Introduction

This paper presents a *process-oriented* computational view of the concept of an anticipatory system. That is, the focus is not on algorithms that produce "results", but rather on how to exploit the fundamental properties of processes to understand the structure and behavior of anticipatory systems.

To *anticipate* is to implicitly introduce the time-oriented concept of the future, and brings with it the complementary concept of the past. This we all know, and often is where we stop. But there is a third time-oriented concept that is correspondingly neglected, namely that of the present. Of course, the concept of the present as the 'current state' is well known, but 'the present' takes on a new role and significance in a pure process regime.

The anticipatory ability of systems which so fascinates us can be captured conceptually in various ways. The four that I will discuss here are

- ***Conservation laws.*** Given that some conservation law applies, we can characterize a future state of the system, whatever it may turn out to be, as a state satisfying the requirements of such a law. Of course, more than one such law can apply simultaneously.
- ***Goals.*** The concept of a goal implicitly points to the future. When a system explicitly includes a denotation of some future state as a part of its current state, and uses it to guide its current activity, we call such a desired, future state a *goal*. Of course, more than one goal may be present, and these are not required to be either mutually exclusive or

[1]Keynote address at InterSymp '97, 9th International Conference on Systems Research, Informatics, and Cybernetics, special Focus Symposium "Modeling Anticipative Systems". Baden-Baden, Germany, August 1997.

consistent. Goal-oriented computation is not new, nor should it cause the teleologically sensitive reader to over-react.

- ***Hierarchy.*** One can view a system at various levels of abstraction. Choosing a node at any level above the lowest, a change in the state of this node implicitly makes a statement about changes at the level(s) below it, which changes precede the change in the original higher-level node. Thus hierarchical abstraction implicitly includes the notion of the future.

- ***Self-reflection.*** If a dynamic entity can 'look at itself', this implies that it is able to represent states other than it's current one, and thus it hereby can implicitly represent its future.

Since *we* can understand how a system can express, and thereby anticipate, its future by all of these means, it only seems reasonable to ask if it is possible for a computational representation to do the same. The remainder of this paper explains, descriptively rather than rigorously, how this can be. It is worth mentioning that to my knowledge, no other contemporary descriptive framework achieves this goal.

The Pure Process World

It should perhaps be noted at the outset that the systems we discuss here are massively parallel, though asynchronous (thus differentiating them from cellular automata), and inherently non-deterministic. In that they are 'symbolic' in character, they are also different from neural nets. They are also inherently capable of learning (cf. co-exclusion, described later) from experience, and it appears that this learning is, via its symbolic and hierarchical character, speed-wise logarithmic compared to that of neural nets.

Definition: A *process* is a sequence of events, wherein each successive event establishes the pre-condition for the event following it. An *event* is the establishment of a new state, ie. an event's occurrence is marked by the coming into existence of a corresponding new (though not necessarily novel) state.

The usual view, even among computer scientists, is that computing is the study of algorithms - you put some data into a program, and the algorithm embodied therein produces some new data, its output. Since this can be represented as $y = f(x)$, this is a *functional* view of computation, and is captured theoretically by Turing machines. Such a program is usually assumed to be a *sequential* computation, but can also consist of several sub-algorithms which run cooperatively and in parallel. Even so, the overall situation is still $y = f(x)$: the only thing that has changed is that the computation takes less time according to the clock on the wall.

As demonstrated for example by (Rosen, 1991) the functional view is entirely inadequate for describing anticipatory systems. The pure process view is different and deeper: in the last analysis, there are simply and *only* processes, which processes consist of events that either *co-occur* or that *exclude* each other (= *mutual exclusion* = *mutex*), and 'data structures' are replaced by these inter-process relationships. Thus the pure process view states that:

- processes are all there is;

- what the processes are "actually" doing (ie. their algorithmic aspect) is uninteresting);
- what *is* interesting is which events *can* co-occur and which *must* exclude each other.

In the extreme, there are only two operations a process can execute in this analysis, the *synchronization primitives* $wait(e)$ and $signal(e)$, where e is an event. An example of a simple process might thus be $wait(e_1)$; $signal(e_2)$, where e_2 indicates that this process has observed e_1. Moreover, out of these two primitives we can build both co-occurrence and mutual exclusion.

Since a process can be viewed as the unrolling of an algorithm into time, and what the process is actually doing is uninteresting (ie. just waiting and signalling), we have thus distanced ourselves quite considerably from the algorithmic $y = f(x)$ view. Moreover, the essential property of a *sequence* of events is that, for the given process, they occur one after the other, that is, the occurrence of one event in the sequence conceptually excludes the others. In other words, the process concept itself can be reduced to mutual exclusion relationships between events. Complementarily, the concept of many processes going on concurrently can be reduced to those events that can co-occur. Thus the conceptual 'pure process' universe consists solely of events e, and the operations on them, $wait(e)$ and $signal(e)$. Aggregating these yields the additional concepts of co-occurrence, mutex, and process.

We will imagine a system to be connected to its surrounding environment via a set of primitive sensors $S = \{s_i\}$. The set S constitutes thus the *boundary* between our system and its surround. Furthermore, each s_i takes on either the value 1 - meaning that the sensor is currently being stimulated - or the value $\tilde{1} = -1$, meaning that the sensor is not currently being stimulated.

We will further view each $s_i \in S$ as a unit vector orthogonal to every other s_j, $i \neq j$, and hence S forms an initial basis for a (discrete) vector space that will be the foundation for our description of a system. For a given sensor s, the state exclusion property (ie. that a process can only be in one state at a time) requires that $s + \tilde{s} = 0$. This in turn leads to the interpretation that '0' means 'cannot occur'. No sensor ever has the value 0. A sensor is the simplest possible process: two states $\{s_i, \tilde{s}_i\}$.

Mapping successive process steps 1-1 to the integers, we see how even at this very primitive stage the exclusion concept is connected to time. The 'present' of such a two-state process is its current state, and the other state is both its past and its future. This 'time' is however extremely primitive, and much structure must be built up before one can speak of 'ordinary' time. Nevertheless, it is clear that mutual exclusion, by distinguishing one event time-wise from another, contains the germ of the concept of time.

Oppositely, the concept of co-occurrence of two events implies that they are indistinguishable in time, in that each is neither before nor after the other. Thus a co-occurrence directly captures the concept of 'the present' as a dynamic and fleeting entity, a real and pregnant *now*. Furthermore, as Leibniz pointed out, indistinguishability contains the germ of the concept of space (and again, much structure must be built up before we can speak of 'ordinary' space).

With the mutual exclusion and co-occurrence of events in hand, we can now bootstrap ourselves up another conceptual level, as follows. The co-occurrence $s_1 + \tilde{s}_2$ excludes the co-occurrence $\tilde{s}_1 + s_2$, since s_1 excludes $\tilde{s}_1$ and $\tilde{s}_2$ excludes s_2. Equivalently, we can write $(s_1 + \tilde{s}_2) + (\tilde{s}_1 + s_2) = 0$ and draw the same conclusion. It turns out that this observation allows us to infer the (Clifford-algebraic) vector-product relationship $s_1 s_2 = -s_2 s_1$. We call this inference *co-exclusion.* [2]

[2]Rearranging the parentheses yields $(s_1 + s_2) + (\tilde{s}_1 + \tilde{s}_2) = 0$, which is also a valid co-exclusion. The two

The co-exclusion inference has several pregnant qualities, the most important of which is that the product $s_i s_j$ is *oriented*, that is (like the more familiar *cross*-product $s_i \times s_j$), can be said to 'point' either *up* or *down*, that is, can have the values ± 1. Which is to say that the two orientations of $s_i s_j$ can function as a sensor, what we call a *meta-sensor*. Moreover, we can now imagine co-exclusions of meta-sensors to yield meta-meta-sensors, etc. That is, the co-exclusion inference leads us to a 'pure process' *hierarchy* concept, and at that, one that can grow both dynamically and unboundedly.

Before pursuing this hierarchical aspect, it is important to note that the co-exclusion inference also yields a concept of *action* (Manthey, 1994), in that the co-occurrence (say) $s_1 + s_2$ can represent the current state of a process, and $\tilde{s}_1 + \tilde{s}_2$ the (mutually excluding) next state of this process via the inner automorphism $s_1 s_2 (s_1 + s_2) s_2 s_1 = \tilde{s}_1 + \tilde{s}_2$. So we herewith increase our conceptual pure-process vocabulary to include that of action (in this case, actions at the same level).

Objects and Hierarchy

The pure-process hierarchy concept we just uncovered is nevertheless not functional, in the sense that lower level entities are not just simple subroutines of higher-level entities. Rather, the relationship is topological. In fact, $(s_1 + s_2) s_1 s_2 = \tilde{s}_1 + s_2 = \partial(s_1 s_2)$, where ∂ is the so-called *boundary operator* of homology theory, that part of topology that treats the relationship between parts and wholes, and is (I understand) used in theoretical biology to analyze similarity and emergence of form. ∂ captures actions that *cross* level boundaries, that is, boundaries that are simultaneously both conceptual part/whole boundaries *and* function/sub-function boundaries. Thus one view of the limitations of the sequential/function-composition paradigm is that it confounds these two forms of action, and in the final analysis restricts itself to a single level, which is what leads to its reductionistic properties.

Homology theory has a dual, called co-homology, whose *co-boundary* operator δ takes (so to speak) $\tilde{s}_1 + s_2$ to $s_1 s_2$, that is, a kind of inverse to the boundary operator ∂. In fact, δ turns out to capture the co-exclusion inference! Furthermore, J.P Roth (see (Manthey, 1997)) showed that there exists what he called a *twisted isomorphism* between an homology and its corresponding co-homology, wherein the kernel of ∂ maps to the non-kernel of δ and vice versa. The final picture is a hierarchy of chain complexes (the homology) coupled to its dual co-chain hierarchy (the co-homology) via the twisted isomorphism: a self-mirroring hierarchy, where structure (captured in the form of co-excluding co-occurrences) is mirrored as action (captured in the form of changes to the current co-occurrences) and vice versa. See Figure 1. Since the transition between levels is mathematically marked by rotations of the vectors (sensors, co-occurrences, and actions), we call the overall approach the *phase web* paradigm (Manthey, 1994), (Manthey, 1997).

One way to see just how different this pure-process hierarchy is from conventional functional hierarchy is to examine mutual exclusion more closely. Two processes that exclude each other have the following form:

co-exclusions are *duals* of each other. We note that $s_i s_j = s_i \cdot s_j + s_i \wedge s_j$, ie. the sum of the inner and outer products. This product is defined for any arity; arity 2 (shown) gives the complex plane, in that $(s_1 s_2)^2 = -1$, taking $s_1^2 = s_2^2 = 1$.

P1: repeat wait(e1); x1; signal(e2) forever
P2: repeat wait(e2); x2; signal(e1) forever

where initially (say) *e1* obtains and *e2* doesn't (and only *P1* and *P2* signal them). Examination of the interaction between the two processes reveals that *x1* and *x2* will *never* co-occur. Imagining *P1* and *P2* as two runners in a relay race, they continually exchange a single 'baton' between them, and the initial condition says that *P1* begins with the baton. It follows that the number of batons (technically known as synchronization tokens) is conserved. This conservation idea (technically known as a resource invariant) generalizes to all such mutex situations, and is computing's version of a *conservation law*.

Thus our action concept implicitly includes the invariance through time of the structure it represents, which we interpret as the foundation of the very concept of an *object*: something that persists in time. So the δ-constructed co-homology hierarchy is a hierarchy of object inclusions - much more subtle than a simple tree - up through which the environment's sensory impressions bubble. Where exactly they bubble depends on the experienced sensory co-exclusive relationships the particular hierarchy represents.

Inversely, *goals* for actions trickle down through the co-existing, (twistedly) isomorphic homology hierarchy. These goals can be to change the state of the (co-ex defined) objects, either purely internally, or via the exertion of influence on the environment via *effectors* on the outermost boundary.

Thus the overall dynamic picture is one of sensory impressions bubbling up, causing multi-faceted classification, while simultaneously goals are trickling down and fanning out. Moreover the software implementation of this (which, incidentally, motivated the mathematical picture, and not the other way around) compactly, efficiently, and automatically adjusts (and presumably, learns) in accordance with this full-blown, breadth-first, massive concurrency. This efficiency obtains because the underlying structure is built out of mutex relationships, so this usual bugaboo is the engine, rather than the bane, of existence.

Three further properties of this *ladder* hierarchy deserve mention. The first is that co-occurrences of object orientations (ie. states), which objects may reside anywhere in the ladder, may themselves be co-excluded. The result is the base of a new ladder that is *orthogonal* to its underlying ladder. This new ladder can then grow 'upwards' just as the underlying one did (and perhaps continues to do). More to the point, the new ladder constitutes an explicit self-reflection of the behavior of the underlying ladder, expressing (in general) how the parts of the underlying ladder play together.

More precisely, the ladder hierarchy possesses two fundamental types, sensors S and goals G. The levels in the ladder pictured in Figure 1 were constructed as the cartesian product of the sensors belonging to the level immediately below, ie. $L = S_{i+1} = S_i \times S_i$. S_{i+1} is the *meta*-level of level S_i. But we can generalize the product, eg. $S_{i+1} = S_i \times S_{i-1} \times S_{i-3}$, yielding (the base of) a meta-hierarchy over a meta-hierarchy.

The second property of the ladder hierarchy is a further generalization of this idea, in that we can represent the presence or absence of a particular goal as a sensor, and construct two new *kinds* of hierarchy, namely $G \times G$ and $S \times G$. We call the former *icarian*, and the latter *morphic*, hierarchies. Morphic hierarchies express the ability to translate between states and goals, that is, a morphic action can generate a goal given a state, or generate a state (perhaps wholly internal) given a goal. An icarian hierarchy, in contrast, expresses the interplay among goals, and thus can be very useful in the carrying out of complex plans with many, potentially conflicting, goals. Finally, the products above may be m-ary; and new hierarchy types can be produced by combining these three, ad infinitum.

An example application of these ideas is a description of auto-poiesis, where meta-hierarchy L_x represents an organism's physiological processing, including acting on the environment, morphic hierarchy L_b the coordination of same, and icarian or morphic hierarchy L_z the attending to the survival of the organism. This is of course quite oversimplified, but nevertheless nicely illustrates the representational power of ladder hierarchies.

The third property of ladder hierarchies is that they can be circular, in the sense that L_b can be built over L_a, L_c can be built over L_b, and L_a can include elements of L_c (as they appear). Of course, L_a must contain other initial elements (eg. physical-boundary sensors) to get the process started. A possible example is speech understanding, where (roughly) L_a is based on primitive phonemes and performs syntax analysis, L_b stands for the semantic analysis, and L_c the situational analysis, which feeds back to L_a's syntactic analysis. The final meaning of an utterance is here imagined to be drawn from the upper levels of L_c. The continuing story of speech understanding has namely indicated that this type of circular analysis is indeed necessary.

Finally, it seems natural to conclude that, because (eg.) $s_1 s_2$ is orthogonal to both s_1 and s_2, and furthermore induces a boundary-crossing, ladder hierarchies express emergent phenomena and properties of the systems they model.

Summary and Conclusions

The pure process view replaces algorithms - and their concomitants, data structures and 'function' - with synchronization relationships among the events constituting its processes. Its primitive concepts are events and the co-occurrence/mutual exclusion of same, and (aside from the ability to *name* things) no more is needed. With these, we have seen how time (via mutex), state (via co-occurrence), conservation laws (in the form of synchronization invariants), goals (simple self-reflection of internal states), the future (via goals), and structure and hierarchy (via ∂ and δ) arise naturally, integrally, and via experience with the environment. Finally, the latter structure and hierarchy can itself be subjected to the same treatment (ad infinitum), and we speculate that two such iterations are the minimum necessary to capture the auto-poietic phenomena characteristic of the simplest anticipatory systems.

References

Manthey, M. Toward an Information Mechanics. Proc. of the 3rd IEEE Workshop on Physics and Computation. Dallas TX, November 1994, pp. 95-109. Isbn 0-8186-6715-X.

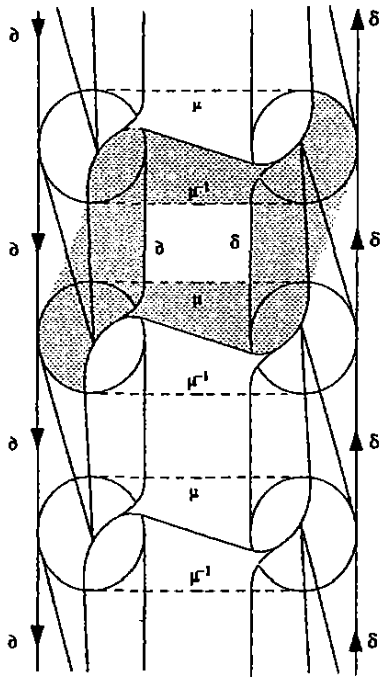
Manthey, M. The Phase Web Paradigm. Int'l Journal of General Systems, special issue on physical theories from information, K. Bowden, Ed. (in press)

Rosen, R. Life Itself - A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life. Columbia University Press, New York, 1991. Isbn 0-231-07564-2.

Further information and software: *www.cs.auc.dk/topsy* and *manthey@cs.auc.dk*

Figure 1. The homology/cohomology hierarchy, wherein each level is a chain/ co-chain.

Sensory impressions bubble up the right (co-homology) side, whereas goals trickle down the left (homology) side. Each circle represents the (meta-)entities defining a given level. μ is the twisted isomorphism connecting action (left) to structure (right).



FROM THE SELF-CREATION TO THE CREATION OF THE SELF

Peter J Marcer
53 Old Vicarage Green
Keynsham
Bristol BS18 2DH

Abstract

Quantum holography provides a mathematical specification of an evolutionary quantum cosmology from the Self-Creation to the creation of the self in excellent accord with existing standard scientific models and known scientific fact.

> Rig Veda x. 129 .... There was not then what is nor what is not. There was no sky, and no heaven beyond the sky. What power was there? Where? Who was that power? Was there an abyss of fathomless waters? There was neither death nor immortality then. No signs were there of night or day. The ONE was breathing by its own power in deep peace. Only the ONE was: there was nothing beyond. Darkness was hidden in darkness. The all was fluid and formless. Therein, in the void, by the fire of fervour arose the ONE. And in the ONE arose love, Love the first seed of the soul. The truth of this the sages found in their hearts: seeking in their hearts with wisdom, the sages found that bond of union between being and non-being.

INTRODUCTION

The description of the physics of quantum holography [Schempp, 1986][Schempp, 1992] of which for example functional magnetic resonance tomography is already an important realisation, utilizes the usual quantum mechanical formalism based on the transactional interpretation of quantum mechanics [Cramer, 1986]. It is therefore the obvious way to make

3 dimensional space (and time) subject to quantum principles. Something so far which cosmologists have been unable to achieve.

Quantum holography (see Appendix) incorporates signal theory into quantum physics by means of the spatiotemporal quantization procedure compatible with the transactional interpretation. It is therefore ideally suited to the 'informatic/organic' self organization of the cosmos to be described.

THE DEFINITION OF THE SELF CREATION AND THE CONSEQUENT COSMOLOGICAL EVOLUTION

The weight of current scientific evidence today is that the origin of the Universe can be traced back to a single event usually referred to as the Big Bang. That is, to a unique point, time's beginning. Under such circumstances therefore in relation to quantum mechanics there would be

a. temporal coherence which is the requirement that wavelets generated from a motherwavelet or source, all travel in the same time or at the same frequency and

b. spatial coherence, the requirement that all wavelets move together in phase as if they started at a single point in space.

These are the two types of coherence necessary for holography [Schempp, 1992], both of which are required at least to some extent for obtaining a stationary interference pattern or hologram. It would ensure, in the strictly bounded system the scientific evidence indicates the Universe to have been, that spontaneous phase conjugate adaptive resonance [Noboli, 1985, 1987] appropriate to quantum holography would occur from this initial isotropic hologram state.

That is, there would be a resonance or self-creation of energy as a prelude to an on-going process of adaption where this adaption is phase conjugate adaptive resonance defined in terms of the emitter/absorber model of quantum holography. The basis of the Self-Creation and the consequent adaptive Quantum Holographic Cosmological Evolution is therefore defined.

The Self-Creation or initial spontaneous phase conjugate resonance is predicted by the quantum holographic model to consist of a three dimensional quantum mechanical coherent whole or a single quantum which is then subject to a unique incremental evolution over time as a result of a sequence of further adaptive resonances defined in terms of the creation/annihilation operators of the emitter/absorber model of quantum holography. The evolution may be envisaged as a sequence of frames of a three dimensional movie where adaption occurs from frame to frame. And because in holography, phase is the quantity of physical significance, and in quantum holography this phase is the Berry or geometric phase [Berry, 1988][Anandan, 1992], known to maintain a complete historical record of i) how long the quantum system has been away from its original quantum state, ii) where it has been in space and iii) of what other quantum states it has visited, then such a record constitutes the essential basis which makes the postulated evolution possible.

THE PROPERTIES OF THIS EVOLUTIONARY COSMOLOGY

This specification of quantum holography requires

a) a unique, self-organised evolution without end, which consists of a self-guided or optimally controlled evolution of the uncertain quantum cosmos postulated,

b) with all the basic features of reality as we perceive it, namely

i. a fixed unalterable past;
ii. an arrow of time as a consequence of the evolution;
iii. where the current increment of the evolution is the present advancing in the direction of that arrow;
iv. where there are three dimensional objects located relative to one another in three dimensions of space - for these are the features of any holographic image; and
v. where the special relativistic invariance of the model [Schempp, 1992][Cramer, 1986] says that space-time is a universal property of the relative motion of those objects, and

c) a self-created cosmos which

i. consists of many worlds/galaxies simultaneously visible to one another across space-time; and which

ii. has the potentiality to evolve sentient, conscious, DNA specified life forms [Marcer and Schempp, 1996] like ourselves capable through quantum holography of understanding this evolution;

iii. how it came to be;

iv. how it brought the DNA lifeforms into existence [Marcer and Schempp 1996], and

v. how such understanding is possible via a brain/mind working by quantum holography, a model [Marcer and Schempp, in press][Marcer and Schempp, in preparation] of the brain/mind which provides a mathematical specification of consciousness as brain/mind interaction, the mind and of the self.

QUANTUM REALITY IS REALITY AS WE PERCEIVE IT

It is a model of evolution and of the brain/mind where the problem of the relation of the ideal and the real, i.e. the world in the head to the world outside the head which is the distinctive feature of modern philosophy requiring solution in cognitive science, is solved by the physical condition of phase conjugation [Marcer, 1995]. This is the condition in holography whereby the object wave is returned in real or virtual form along its path so that the object image and the original object coincide and where in holography one records not the optically formed image of an object as in photography but the object image bearing wave itself. The Self-Creation is defined therefore in terms of the quantum self interference taking place, specified by means of the holographic trace transform, i.e. in terms of its own self image.

This evolution with the format of a 3 dimensional movie ensures as stated above that quantum reality in keeping with the reality we perceive has all the features b) i-v.

This is a universe therefore by b)i where time travel is not possible and the logically inconsistent phenomena associated with time travel cannot therefore occur, but where instantaneous travel in space may be a potentiality. It is a universe where the reality we all commonly perceive and quantum reality coincide, so that quantum mechanics does not, as is often stated to be the case, put reality on the rocks.

Moreover since this Universe at its conception is a single quantum or quantum mechanical whole then everything in the subsequent cosmological evolution that follows the Self-Creation will be in synchrony with it. This explains the nature of quantum non-locality, which is essentially the only mystery of the theory and how it differs from classical physics. It also explains why non-locality and non-local quantum interference which are the basis for the non-classical mechanism of quantum holography are of such importance. It says that the Self-Creation with its associated stationary holographic interference pattern which is updated with every consequent adaptive absorption and emission, essentially constitutes the Universe's self reference frame. That is, this reference frame is the frame against which ultimately everything else must be measured, but which cannot be measured itself because there is nothing to measure it against. This reference frame or dynamic pattern of energy is what is often referred to as the dynamic quantum vacuum. Its direct unobservability is therefore in no way mysterious, and the current holographic pattern relative to this reference frame constitutes the afterglow or historical record of the Self-Creation.

THE UNIQUENESS OF THE COSMOLOGICAL EVOLUTION POSTULATED

Mathematically the adaptive time asymmetric evolution arising from the emitter/absorber model of quantum holography can be shown to be a UNIQUE birthordering [Marcer, 1989][Marcer, 1989a][Marcer, 1990] of the field automorphism constituted from all the time symmetric mappings appropriate to the model which correspond up to a unitary isomorphism to the infinite-dimensional irreducible unitary linear representations of the Schrodinger type U_v and $\overline{U}_v$ of G the Heisenberg nilpotent Lie group, the symmetries of which specify the quantum coherence and the phase conjugate adaptive resonance taking place. Thus the evolution is without end and so there is no Big Crunch. Furthermore as a consequence of the non-unimodular affine Lie group G_+ $t \rightarrow at + b$ of transformations which concern U and $\overline{U}$, [Schempp, 1992] i.e. U_v and $\overline{U}_v$ summed over the frequency v, it is seen that in relation to U (Cosmos) and $\overline{U}$ (Cosmos), i.e. those of the quantum Universe that this Universe will be in

176

continuous dilation or expansion. That is this Universe corresponds to the physical parameters of one that will expand forever.

More remarkable still is that in accordance with the conditions of quantum coherence appertaining at the Self-Creation leading to the self-organised cosmological evolution, this incremental evolution corresponds to an optimal control of the Schrodinger equation [Dahleh, Pierce and Rabitz, 1990] governing it at every step. That is, it is a self-guided evolution of the uncertain quantum cosmos [Rice, 1992], so that in relation to the appearance of order, this order will emerge in an optimal fashion! Thus in a truly logical sense, the Universe in which we live is not only this model cosmology says unique and without end but consists of the best of all worlds possible at any point of adaption or time. That is, the evolution may be said to be correcting itself by a learning process in order to approach perfection - by an optimal process that continues indefinitely, i.e. it is asymptotic. Such optimal control must admit an Hamiltonian defining its energy which is unmeasureable for reasons already stated, is time reversal asymmetric, and concerns an infinite in the limit number of emissions and absorptions n of the quantum holographic model. This is in excellent accord with Berry's Hypothesis [Berry, 1986] in relation to the Berry or geometric phase that it concerns the Hamiltonian of an unknown dynamical system without time reversal symmetry of which the phase space trajectories are chaotic and the eigenvalues of which are the imaginary parts E_m , m = 1,2 ..... of the zeros of the Riemann Zeta function

$$\zeta(z) = \sum_n 1/n^z = \prod_p (1 - p^{-z})^{-1}$$

where z = x + iy is a complex number and p is prime

Taking n to be the number of emissions and absorptions of the proposed cosmological evolution this unknown dynamical system would be that of the Universe itself. In this context the well-known Riemann Hypothesis that the zeros of the Zeta function all lie on the line x = 1/2 so that $\zeta(1/2 + iE_m) = 0$ for all m, must be interpreted as saying that each stage of the evolution is a unique fermionic state and so by the Pauli exclusion principle the quantum holographic information processing taking place will indeed have the canonical labelling it needs to constitute computation as it would in the case where phase conjugate adaptive resonance takes place. That is, the zeros of the Zeta function specify the actual acts of creation or emergence of the objects of the evolution. And thus since the quantum holographic model determines an embedding of the complex plane C: z = x + iy in G the Heisenberg group [Schempp, 1992], the E_m specify the phase at which each emergence occurs. The fact that the Riemann Zeta function specifies the distribution of prime numbers, shows the great significance of the role that the primes, numbers and counting play in the descriptions of this Universe, as indeed they do in our own. That is, the two are one and the same. Similarly the fact that the primes are linearly independent, i.e. form a statistically independent ensemble and that each represents a unique property of number, confirms the quantum holographic encoding and decoding taking place during the evolution is optimally efficient so that in the Shannon sense the mutual information code coefficients between each creative act is zero. This is as it must be for every truly creative act in a process of emergence, where each such act must lead to some unique property without precedent. It is consistent with the thesis herein that each quantum coherent whole is more than the sum of its parts because it concerns such an emergent property.

Consider the fact known from renormalisation group theory that any material phase transition is only predictable quantum mechanically. Take for example the simple phase transition point between water and steam at which on all scales from the molecular level upward there is steam containing droplets of water and water containing bubbles of steam. That is, it is steam/water hologram. It may therefore be inferred on the basis of renormalisation group theory, that the mapping of the holographic Universe onto itself as a process of phase conjugate adaptive resonance, is a process of emergence taking place through a sequence of material phase transitions where each emergent property is a consequence of such a transition. Matter in the quantum holographic Universe can be expected to take new properties that never before existed at various stages of evolution.

The Self-Creation as a process of phase conjugate adaptive resonance as defined above, therefore constitutes a quantum computer constructor universal attractor capable of perfectly simulating any physical process or algorithm [Deutsch, 1985] and of constructing replicas of itself on the appropriate geometric scales.

THE BASIC CONFIGURATION OF THE QUANTUM HOLOGRAPHIC UNIVERSE

At the basic level of the Cosmos as a coherent whole, the quantum holographic computation or evolution is a perfect simulation which concerns the generation of wavelets from the mother wavelet or Huygens' principle of secondary sources. That is, as axiomized by Jessel [Jessel, 1954], one may replace or perfectly simulate the source of a field of waves necessary for the quantum holography taking place, by a corresponding set of secondary sources. And thus on the basis of the emitter/absorber model, the Quantum Holographic Cosmos must be envisaged as a continual exchange between such a source and a suitable set of secondary sources. That is, it may be postulated for an observer such as ourselves constituting a secondary source many times removed, as being between a Source or great attractor and a set of secondary sources, i.e. the galaxies, which like the Universe itself, their proto-holographic object image, may be observed to be in expansion. Thus in the quantum holographic interpretation of quantum mechanics, the many worlds hypothesis is literally true and the many worlds are the galaxies so that the many worlds are connected and visible to one another.

CHARACTERISTIC BEHAVIOURS

The behaviour of the evolution appropriate to the emitter/absorber model of quantum holography [Schempp, 1992][Schempp, 1993] concerns topological computation described by means of shift-register action on the nilmanifold of the Heisenberg nilpotent Lie group G, the three dimensional representation of which is

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \quad \text{written as } (x,y,z) \text{ for convenience}$$

and where x,y,z can be interpreted as having their usual meaning as spatial dimensions/co-ordinates, and where each emission/absorption, the adaption taking place concerns an increment of time dt.

That is characteristic behaviour is specified in terms of Lie transformations so that spatio-temporal Lie group invariances determine size constancy, shape constancy etc. and Lie group prolongations generate the differential invariances characterising higher form structure [Hoffman, 1970]. Similarly the spatio-temporal Lie invariant pathcurves will specify inertial motions. These concern the action of the Lie derivative L_X

$$L_X = X_1\partial/dx + X_2\partial/dy + X_3\partial/dz + X_4\partial/dt$$

where X is the vectorfield $X = (X_1(x,y,z,t), X_2(x,y,z,t) X_3(x,y,z,t), X_4(x,y,z,t))$

and may be found by solving the Pfaffian system

$$dx / X_1 = dy / X_2 = dz / X_3 = dt / X_4$$

They are the straight lines in space-time, i.e. the geodesics which arise as a consequence of general relativity.

Thus the inertial behaviour or motion of objects follows that predicted by general relativity, and as the Universe is a holographic quantum coherent whole, then Mach's postulate that the mass of any object is a consequence of the all other masses follows as a consequence of quantum non-locality in relation to the gravitational field. Gravitational mass is therefore equal to inertial mass as defined by the spatiotemporal Lie invariant pathcurves above. In quantum holography this equality is however an equivalence relation saying that the two fields namely the unified field and the gravitational field have in relation to the condition of phase conjugation, the same properties, i.e. that the gravitational object image of an object is its inertial mass. Mach's equivalence principle therefore follows.

Thus the spatio temporal and other Lie invariances define the objects of the Universe and their properties.

SUMMARY OF CONCLUSIONS OF RESEARCH INTO QUANTUM HOLOGRAPHY SO FAR

Extends but does not overturn validated scientific understanding.

Provides a mathematical model of the Self-Creation and its source.

Provides a mathematical model of the unique quantum cosmological evolution, where this is open, known as the Einstein-deSitter; and almost certainly asymptotic, i.e flat (in fact a dually flat Riemannian manifold).

Such a universe is compatible with Mach's principle, General Relativity and the Kaluza-Klein model, where holography implies both that the universe is isotropic, but has sources or 'seeds' for the galaxies which are secondary sources of the virtual source which is that of the Universe or Self-Creation.

In this cosmos, the many worlds are real, as is the Feynman sum of histories of objects. It is therefore appropriate to think of the virtual parts of complex quantities as physical real, and not as simply representations in the mind of the observer as would be the interpretation in classical physics (or in the Copenhagen interpretation).

Quantum reality is reality as we perceive it.

The Berry phase of the quantum universe is the Jungian universal unconscious, just as the Berry phase of the individual brain/mind is its personal unconscious: the complete record of the quantum coherent system history.

Explains how the three dimensional morphology and dynamics of DNA encodes the three dimensional morphology and dynamics of organisms, i.e. it says that DNA is the quantum chemical medium for quantum holographic chemical encoding of life.

Explains the workings of the simplest living cells the prokaryotes.

Explains the workings of neurons.

Explains how all the senses may be transduced.

Explains locomotion; how birds and insects fly, how they navigate, etc.

Explains the working of the brain as a conscious system.

Explains the mind, consciousness and the self.

Defines death and health validating alternative medicine as an extension of current medical practice.

Provides a new model of the heart; and the related cranial-sacral fluid system.

Explains the nature of the immune system.

Explains that human beings are instantiations or creations of the self as specified at their conception and why each such self is unique.

Explains Gaia in terms of quantum ecology or the biosphere as a quantum coherent whole.

It constitutes an entirely new paradigm of information processing which is quantum constructor computer universal, and where such constructor computers are fully distributed, synchronously partitioned massively parallel processors employing quantum parallelism, geometric encoding and decoding rather than logic/binary encoding and decoding, and phase gates rather than amplitude or bit gates. That is, the principles concern whether signals in the form of physical waves are in phase or out of phase so as to either reinforce or attentuate each other, and form physical wave interference patterns of energy or holograms, not representations of bits.

Explains why and how language and logic are enhancements to the primary capabilities of three dimensional perception and cognition, and why number, the primes and counting play such a fundamental role in relation to the unique quantum cosmological evolution.

This new paradigm of quantum holography already applies to functional magnetic resonance tomography, which has provided the means by which Schempp has validated the theory of quantum holography defined in terms of the Heisenberg nilpotent Lie group. This theory defines the well-understood processes of classical or optical holography in terms of the formalism of quantum mechanics so that it applies to any kind of wave. And since holography is carried out by means of interference with a non-object bearing reference beam, which may be part of the original beam providing the object illumination, quantum non-local interference/entanglement is fundamental to quantum holography and in agreement with the transactional interpretation of quantum mechanics.

Other existing applications are to synthetic aperture radars, etc.

APPENDIX

Quantum Holography

Quantum holography [Schempp, 1992] describes the processes of holography by means of the standard quantum mechanical formalism in terms of the mathematics of the Heisenberg nilpotent Lie group of which the three dimensional representation G is

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{written for convenience as } (x,y,z)$$

where x, y and z can be given their usual geometrical meaning. The standard quantum mechanical commutators of the symmetries of G are the structure relations of g, its Lie algebra as was known to Herman Weyl in 1928. The classic uncertainty principle takes the form of the Robertson relation

$$\Delta U_v(P) \, . \, \Delta U_v(Q) > \tfrac{1}{2} |U_v(Z)| \quad (v \in R, \, v \neq 0)$$

where U_v is up to a unitary isomorphism a unique infinite-dimensional irreducible unitary linear representation of the Schrodinger type of G in the standard Hilbert space $\mathcal{H} = L^2(R)$, Δ is the stand root mean square deviation, v is the frequency, and {P,Q,Z} is the canonical basis of g given by the matrices

$$P := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad Q := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and } Z := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so that the Heisenberg commutation relations read as follows,

$$[P,Q] = PQ - QP = Z \, ; \, [P,Z] = 0 \, ; \, [Q,Z] = 0$$

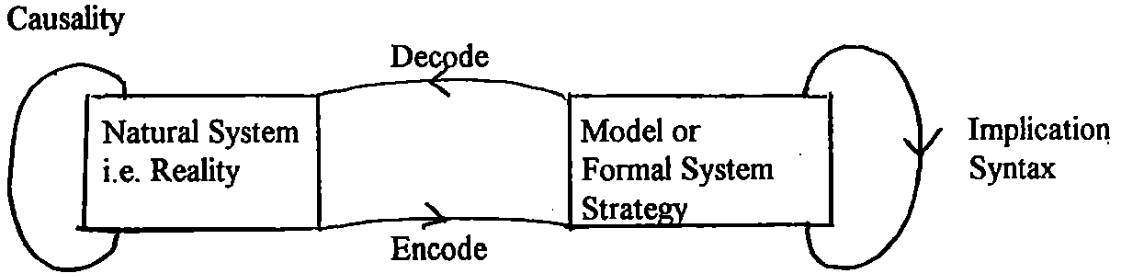
In quantum holography, extensively validated by Schempp in respect of the operation of functional magnetic resonance tomography [Schempp, 1994], the symmetries of G define the quantum non-local coherence/entanglement and phase conjugation necessary for the holography. These descriptions of the processes of the holography in principle apply to any kind of wave [Marcer, 1996], and are in excellent accord with the transactional interpretation as explained by Cramer [1986].

The transactional interpretation is explicitly non-local and thereby consistent with recent tests of Bell's inequality, yet is relativistically invariant and fully causal. Experiments conducted at DRA Malvern by Rarity [1994] show that quantum non-locality/entanglement is demonstrable over distances of kilometers, and thus confirm the potential applicability on such scales of quantum holography.

As an information processing paradigm [Marcer, 1996] it concerns the geometric encoding and decoding of holographic interference patterns or holograms as distinct from the logic binary encoding and decoding of binary/digital information. It thus defines a paradigm where information as experiential knowledge is processed, i.e. concerns semantics as well as syntax!

Quantum holography as a general strategy for modelling or semantics

Figure I below sets out the relationships concerning the modelling of natural systems.



In the above diagram

a) The decoding/encoding processes of quantum holography are geometric, where

b) the mapping onto reality or means by which the natural system is perfectly simulated, is phase conjugation [Pepper David, 1985]. That is, the object image bearing illumination in real or virtual form is returned such that the object image and object coincide. Noting that in holography one records via holograms, not the object image as in photography but the object image bearing wave itself. This is done by means of interference with a non-object bearing beam, which may be part of the original beam providing the object illumination. Quantum non-local interference/entanglement is therefore fundamental to quantum holography.

c) The symmetries of the Heisenberg Lie group which specify the nature of the quantum non-local coherence/interference and phase conjugation of the holography taking place [Schempp, 1992], therefore constitute the formal system strategy describing the modelling which takes place analogically. That is, quantum holography constitutes a means by which reality, i.e. real physical processes may perfectly simulate reality in accordance with Deutsch's Church-Turing Principle [Deutsch, 1985] governing universal quantum computation. Quantum holography employing geometric encoding and decoding must therefore be quantum computer constructor universal, i.e. able to perfectly simulate any algorithm, physical process, morphology or dynamics. Any form of causality can, therefore in principle, be perfectly simulated. This includes Turing computation which Deutsch [1985] proves quantum computers may simulate; indeed this is what 'digital hardware' does. And since no two objects may occupy the same position in space time the quantum holography taking place (by phase conjugation) is canonically labelled, and so constitutes computation in the formal sense. It is topological computation [Schempp 1993], the formal system strategy says, described by means of shift register action on the nilmanifold of the Heisenberg Lie group so that such action is defined in terms of Lie transformations.

It further follows i) because the natural Lie diffeomorphism is exponential, that quantum holography is able to deal with exponential complexity or even exponential towers of complexity, and ii) that the renormalisation problem of quantum physics does not arise since for any Lie group an element X^{-1} maps onto X indicating a topological transformation in a geometric system and not an infinity.

d) In a quantum holographic transducer performing the above modelling analogically, acoustic illumination may be converted to say electro-magnetic so that the two processes are equivalent and the coding in Figure 1 commutes as it must for the diagram to be valid. That is, the electro-magnetic output will have the same properties as the acoustic input, exactly mimicking it in the domain of transduction. Thus the equivalence classes constitute the nature of the syntax/implication in the diagram.

e) In the case where quantum holography utilising phase conjugation applies, Figure 1 shows that the formalism of quantum mechanics need not, as is widely held to be the case, lead to models of the physical world at variance with the world as perceived. In such circumstances quantum holography takes place by phase conjugate adaptive resonance, so that computation proceeds analogically by selection and adaptation/learning. The computers are therefore fully distributed synchronously partitioned massively parallel processors, without the need of co-ordinating signals between geometrically separate actively computing machine subsystems, and which affect changes of functionality by instantaneously shutting down such activity and replacing it by that in another set of subsystems again without the use of signals needed in a classical machine from a central control or switching centre.

Moreover, in the domain of phase conjugation, superresolution [Schempp, 1992][Leith and A. Cunha, 1989] where noise serves to sharpen images, is possible and hence there can be 'stochastic resonance' as is observed in actual neural systems [Moss and X Pei, 1995]. It might be objected that actual neural systems do not employ phase conjugation. However two simple perceptual experiments serve to show that they do.

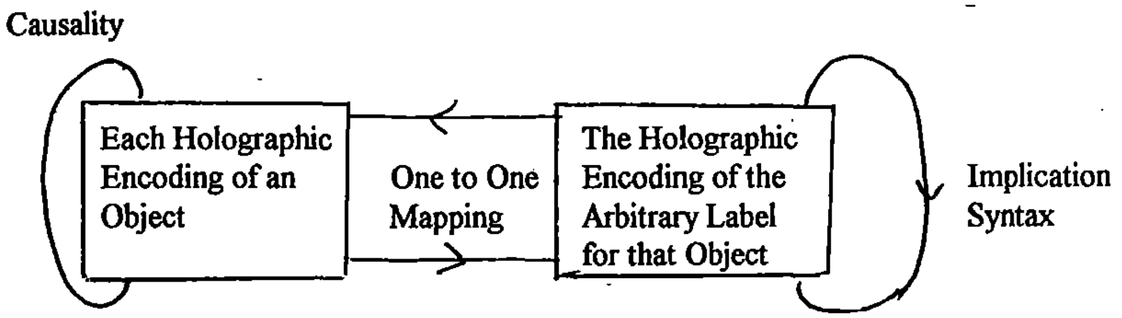
(i) Snap one's fingers and ask where the brain creates the acoustic object image of the snap. Listen. It is outside the head coincident with the snap, as a phase conjugate image must be, now

(ii) reach out for a nearby object. Verify that not only do the visual and tactile object images the brain creates coincide, but this coincidence is also with the object itself, i.e it is phase conjugate with the object in every geometric particular with reference to the scale of resolution of which the senses are capable. There is therefore good reason to believe that the brain and its sensory apparatus employ both phase conjugation, and geometric encoding and decoding, i.e. quantum holography with respect to perception and cognition.

Phase conjugation must be of vital importance to the processes of image formation of the brain or imagemaker since survival is completely dependent on locating objects where they actually are, whether it be by vision, sound or touch etc., that is of sensing the world the way it actually is, and also for communication between imagemakers.

Quantum holography as a general strategy for language/grammar

Figure II below sets out the relationships concerning the explicit arbitrary labelling of objects/processes in quantum holography as it applies to the natural systems in Figure I when phase conjugation takes place. Grammar is defined as the essential logical rules governing the naming or labelling procedure for sets or here equivalence classes.



In the above diagram

a) There is one to one mapping so that each holographic encoding of an object or icon maps to and from, the holographic encoding of the unique arbitrary label for that object or icon, so making the implicit iconic phase conjugate labelling of Figure I explicit. Thus a shared/common arbitrary labelling of objects by means of their phase conjugate object images provides a basis for a natural language among those who share such a labelling. In the above diagram therefore the one to one mapping satisfies Wittgenstein's principle [1975] that for each fact, there can be only one proposition that answers to it, and that the sense of the proposition call only be expressed by repeating it.

A postulated instantiation of a morphology and dynamics that the diagram expresses is therefore (in quantum holography) the two hemispheres of the human brain joined by the corpus callosum [Marcer and W Schempp, in press]. That is the right hemisphere containing the holographic encodings of the real world (or more precisely the receptacle for those) is the artistic brain and the left, containing the arbitrary labels (or receptacle for these) for the real world objects will be the logical brain since an essential aspect of such arbitrary labellings will include number and sets, and their logical relationship or mappings one to the other, i.e. the implication/syntax of the diagram. In this natural language situation, the phase conjugation of quantum holography will quite naturally guarantee that actual causality holds since in these circumstances the objects of the holography are perfectly simulated. This is imperative to survival.

REFERENCES

ANANDAN J. (1992), "The geometric phase" Nature, 360, 26, November, pp 307-313.

BERRY M. "Quantum chaos and statistical nuclear physics" Springer Lecture Notes in Physics, No. 263, Eds. T.H. Seligman and H. Nishioka, Springer, Berlin, pp. 1-17, 1986: "Riemann's Zeta Function: A model for quantum chaos?"

BERRY M. (1988), "The geometric phase" Scientific American, December pp 26-32.

CRAMER J.G. (1986), "The transactional interpretation of quantum mechanics" Review of Modern Physics, 58, 3 pp 647-687, July.

DAHLEH M., PIERCE A.P., and RABITZ H. (1990) "Optimal Control of Uncertain Quantum Systems" Physical Review A, 42, 3, pp. 1065-1079, August.

DEUTSCH D. (1985), "Quantum Theory, the Church-Turing Principle, and the universal quantum computer" Proc. Roy. Soc. Lond. A400, pp. 97-117.

HOFFMAN W.C. (1970), "Higher Visual Perception as Prologation of the Basic Lie Transformation Group" Mathematical Biosciences, 6, 437-471.

JESSEL M. (1954), "Une formulation analytique du principe de Huygens" Comptes Rendus, 239, 8th December, pp. 1599-1600.

LEITH E. AND CUNHA A. (1989), "Holographic methods for imaging through inhomogeneity" Optical Engineering, 28, 6, pp 574-579.

MARCER P.J. (1989) "ANPA has much to celebrate" Proceedings of ANPA 11 University of Cambridge, September, 14-17, pp. 44-45.

MARCER P.J. (1989a) "A model for an Evolutionary Epistemology based on the Axioms for Universally Extending Continua" Proc. ANPA 11 as above.

MARCER P.J. (1990) "A Grand Unification for Wavefield Physics" Proc. ANPA 12, University of Cambridge, September, 13-16, pp. 136-163.

MARCER P.J. (1995) "How quantum mechanics maps onto reality" ANPA 17 Proceedings in press.

MARCER P.J. (1996), "A novel paradigm of information processing for the new millenium" Proceedings symposium E-C-H-O, 21-23 August, editors Ehresmann A.C., Farre G.L., and Vanbremeersch J.P. pp.127-132.

MARCER P.J. and SCHEMPP W. (1996) "A mathematically specified template for DNA and the genetic code in terms of the physically realisable processes of quantum holography" Proceedings of the symposium on Living Computers, University of Greenwich, March 9th 1996, (in press) editors P.J. Marcer and A. Fedorec.

MARCER P.J. and SCHEMPP W. (in press) "The Brain as a Conscious System", International Journal of General Systems.

MARCER P.J. and SCHEMPP W., (in preparation, already presented at two symposia) "A model of the neuron working by quantum holography".

MOSS F. and PEI X. (1995), "Neurons in parallel" Nature, 376, pp. 211-212, 20th July.

NOBOLI R., (1985), "Schrodinger Wave Holography in Brain cortex" Phys. Rev. A.32, 6, pp 3618-3626.

NOBOLI R. (1987), "Ionic Waves in Animal Tissues" Phys.Rev.A, 35,4, pp 1901-1922.

PEPPER DAVID M. (1985), "Non-linear optical phase conjugation" Laser Handbook 3, North Holland, Amsterdam.

RARITY J. (1994), "Dreams of a quiet light" Physics World, June, pp. 46-50.

RICE S.A. (1992) "New Ideas for Guiding the Evolution of a Quantum System" Science, 258, 16th October, pp. 412-413.

SCHEMPP W. (1986), "Harmonic Analysis on the Heisenberg group with applications in signal theory" Pitman Research Notes in Mathematics series 147, Longman Scientific and Technical London.

SCHEMPP W. (1992), "Quantum Holography and Neurocomputer Architectures", Journal of Mathematical Imaging and Vision, 2, pp 279-326.

SCHEMPP W. (1993), "Bohr's Indeterminacy Principle in Quantum Holography, Self-adaptive Neural Network Architectures, Cortical Self-organisation, Molecular Computers, Magnetic Resonance Imaging and Solitonic Nanotechnology" Nanobiology, 2, 109-164.

SCHEMPP W. (1994), "Phase Coherent Wavelets, Fourier Transform Magnetic Resonance Imaging and temporally encoded synchronized neural networks" Pitman Research Notes in Mathematics Series 147, Longman Scientific and Technical, London.

WITTGENSTEIN L., Philosophical Remarks, Oxford.

THE HIERARCHY AND PHYSICAL SPACE

Ted Bastin,
The Mill House, Pen Bryn
Sarnau, Cardiganshire A44 6QL

PROLOGUE

Here are a few statements which seem incontrovertible, if crude, but whose consequences for any representation of *process* are often overlooked.

1 The combinatorial hierarchy is well-known to be comprehensible if and only if one thinks of it as a *process* where we operate at some point and not at just any point. If we want to survey the whole thing as we do in conventional physical theory then we have to install apparatus to make sense of that.
2 This difference from conventional physical theory appears again because at each step it has to be decided whether we are taking a set of elements as individuals or as a single entity.
3 CWK removes this level-dependence but only by giving the whole construction a process character which is built in.
4 Computer models in general, and therefore those for use in physics, automatically have the process character.
5 Inevitably in using such models we are saying that the activity of the computer represents an essential aspect of the world directly.
6 The use of computers to make calculations using the equations of, say, quantum mechanics are not models in the above sense at all. You may say that the equations are part of a theory, but in that case they are part of the classical sort of theory and not of the process sort. Computers used that way merely speed up the calculations and do not represent the world directly.
7 If there are other computer models for physics which are of the process sort it is not much use talking about them in the abstract: they should be trotted out and examined for physical seriousness.
8 All the arguments in the combinatorial hierarchy can be reduced to computer operations -obviously. (Notice the important sense in which this is not true of, say, classical mechanics.) In fact that reduction is a useful criterion for making sure that one has not slipped in conventional physical concepts. The converse is not true. It is easy to make up computer models which have no physical significance.
9 Two computer models have been extensively discussed. One is Program Universe which is a simplified form of the combinatorial hierarchy, and the other is Pope's cinematograph theory. (Presumably one can regard a cinematograph as a sort of computer). Both these models represent time as a series of instantaneous stills or CLICKS. Perhaps it would be better not to use the word time, but to say that the state of the world advances in a sequence of jumps, and that if there is any idea of extension being imagined then each jump affects everything everwhere. I do not list Manthey's work here because in my opinion it is not a model but is a different way of looking at the hierarchy theory.

10 Now to the besetting problem of combinatorial physics. The conventional physical concepts like charge and length and and spin and momentum and velocity and mass and all the rest are defined in classical mechanics and little work has been done to redefine them in a truly independent manner. In order to get a way of talking about the process world we have to redefine all these in a way compatible with the process method. This redefinition can never be a matter of once-off identification: each concept has to be constructed on its merits.

I. Process Theories

The combinatorial hierarchy is well-known to be comprehensible if and only if one thinks of it as a constructive (or perhaps a deconstructive) *process*. The recursive relation of levels has to be taken as a prescription of this process character. The simplest, though perhaps the crudest, way to think is to invent a computer program which performs the construction.Whatever way we we think, we are clearly claiming that the construction, or the activity of the computer, represents something that is actually going on. The universe is supposed to act in essentials in the same way as the mathematics or as the computer. Classical physical theories are quite unlike this and in these remarks I want to try to get a clearer vision of this tremendous difference.

Early on in such an endeavour we find we want to talk about time. I suppose that an inevitable first step is to point out that if things happen after one another - sequentially- then some time-like idea has already been imported. I have to break in at this point to prevent a misunderstanding that always causes quite a lot of friction. I use 'sequential' to include any sort of bit of program including concurrent ones. The computer people, I think, contrast 'concurrent' with 'sequential' -presumably because they never meet anything which is not sequential in my sense of the word. However current physical theory is not sequential and I need a general word for that which is.

Etter attacks the problem: presumably he is sympathetic to computing models, but I'm not sure. He says 'process' has a general meaning meaning of 'something that proceeds or moves along', but he says he will use another connotation where 'a process of a certain kind refers to that which is allowed to happen, or which can happen, under certain specified conditions'. It follows from what I have been saying that the hierarchy mathematics entails process in Etter's general sense. If this is disputed it were best to have the dispute out in the open now. I think that for some of us who have been connected with PU, the going-on-an-on-ness of the computer model is what they want to represent about the world.

Etter's central theme is the disparity between composition and extension. (or system and process). He compares these again (respectively) to collections of autonomous parts and to map-making. I want to get him back into his general use of 'process' by speaking not of map-making but of map-reading. You sit in the car and guide the driver by constant reference to the map.

Etter defines a LINK which leads from one moment to the succeeding moment, each represented by 'variables'. Each variable has some sort of background provided by a probability distribution. When the variables are linked, two things happen. (1) a joint probability distribution is formed, and (2) one of the variables which we may term the 'first' disappears and gives place to the other. That is then available to form a new link with something else. So what trace in the world is left of the transaction? Answer: that which had been the joint probability distribution.

Now the immediate target for LINKS is Etter's ostensible aim of rebasing quantum theory. The probability distributions are in the first place those which appear in that theory. However I think his intention is wider than that. In any case I shall suppose that whereas quantum probability is strictly the map-making kind of process, the introduction of the LINK turns it all into the map-reading kind of process.

II. Physical Space

Alright then; let us suppose we can escape from limitations of the quantum theory, with its god-like detachment from the probability distributions, and become part of the process, by using the link concept. We then have to readdress the question 'what are these probability distributions now, in this wider sphere?' Well; I want to associate them with the probability distributions of unknown entities which CWK and I postulate to start off the hierarchy algebra. (See Combinatorial Physics). We maintain that the simplest way to proceed in the construction is to suppose that prior probabilities of encountering these entities are equal, and that this 'principle of indifference' is needed to get us our magic numbers. However the principle is one which the real world will always be liable to transcend, because there is no way to get empirical knowledge except through deviations from this exact equality; even though the deviations may be small enough not to perturb the calculations too much. The deviations are the source of empirical knowledge about the world.

Let us ask what physics requires if we are to understand the *continuity* of things. This question has to be answered equally whether we think of a mathematical continuum or of a rational manifold, so I am not discussing continuity as opposed to discreteness. I am bound to make an incursion into the ideas of topology, and I observe that topology presupposes an insideness /outsideness relationship as a meaningful idea which can be taken over without discussion. We can't do that in combinatorial physics. There is actually no sense to attach to the idea of *insideness*, so we could not take over the topological language without scrutiny even if we wanted to. The intuitive sense of insideness or mapping *onto* appears on the path towards establishing limit points, and this argument all has to do with establishing the real numbers. My idea is to put the horse before the cart. I argue that we MUST provide the limit point or fixed point in the constructive process in the of physics, and only after that can we proceed to things like insideness.

III. The Fixed Point

We tend to think that physical operations can be represented by rational numbers on a scale of measurement. It seems to me that this assumption requires a sort of steering by the human mind. Look at it this way. If we suppose that successive measurements or observations can be represented by rationals then how do we know how to do the narrowing down which keeps the operations on course to reach a result of successively better approximation? We know mathematically of course, but what is the physical concommitant? If we go back to the mathematical limiting process then we are arguing in a circle. The problem which we encounter when we try to present our experience without the appeal to intuitive space and its insideness has ties with what confronted Brouwer in his proof of the fixed point theorem.

Kilmister has helped me take the fixed point theorem back to its basics, and I use his treatment.

Theorem: EVERY MAP OF A CLOSED N-DIMENSIONAL CELL INTO ITSELF HAS A FIXED POINT.

Note 'closed' and 'into'. Take the case of two dimensions (n=2). Take any room (the cell) Make a plan of it -not necessarily preserving scale or straight lines. Just any sketch. 'Into' means that the scale is less than 1 ; 1. So take this map and throw it onto the floor of the room. That constitutes a map of the floor into itself. Next consider the the piece of floor covered by the plan; this piece of floor is shown by some part of the plan, so this piece of the floor is also mapped into itself by the process. Repeat this argument; at each stage the part of the floor you are talking about gets smaller. So this sequence of nesting regions on the floor are shrinking up to the fixed point. All the complicated arguments are just making this intuitive argument respectable. This way of doing it shows what the theorem is really about -the completeness of the continuum. If one did this in a plane with rational coordinates, the nests need not have a common point -the old root 2 trouble.

Topology seems to have deliberately thrown away the ladder it climbed up on so that, for example, the fixed point theorem no longer seems to depend on the repetition of an onto-mapping. They would like to cover up the appeal to an intuition of *insideness* with technical language. However the technical language must presuppose it, and that presupposition is very important for us. I should stress at this point that I am not trying to reintroduce the continuum: I am only stressing that something has been forgotten in any application of mathematics to the world whether we take nesting processes to a hypothetical infinite limit or not. Moreover, 'nesting' is only one special case of 'relatedness' which we need in describing a measurement. The progressive steps have to be related, and giving them particular mathematical form like increasing smallness presupposes that done. It does not do it.

I conjecture that the situation to which I have drawn attention is what provoked Brouwer to renounce topology and develop time dependent mathematics. The fixed point theorem loses pride of place to Brouwer's Spread Theorem.

There is a different argument which bears on this. What happens to the observer. Could one start from different points in the hierarchy and think of each starting point as the construction made by a different observer? Relativity seems to indicate the need for something of that kind. And in that case would one have many hierarchies with the obligation to reconcile them, or just one with different paths through it? I was never quite sure what answer was given in Program Universe to these questions about the observer. One has a strong feeling that one has slipped into a naive and anthropomorphic way of thinking in entertaining these questions, and yet one does not quite see how to avoid them. I hope that this observer talk is just an alternative, and perhaps not very satisfactory, way of referring to the fact that in our limiting processes there must be a way of identifying a number of different sequences as referring to the same thing (which we can speak of as the limit point). We reconcile our feeling that we, as observers, take part in the process by speaking as though there can be an indefinite number of participants able to *identify* a given process.

So the picture which emerges is that measurement imposes an order of some sort, and our historical preconceptions make us see the quantities in terms of a unit of measurement. This unit will be seen a nesting limit of the sort discussed in the fixed point theorem. Hence we are changing the logical order and seeing the matter of identification of a sequence of measurements as the fixed point given our intellectual requirements for a physical space or we might say 'continuum'.

IV. Units

Attention now switches to the unit. We have no identification for it at this stage, since what we know about the scale-constants and coupling constants is all dimensionless. Of course we need them there because they give us the dimensional constants we have to work with (apart, of course, from giving the numerical ratios of them). At this stage these dimensional quantities are *conventional* in the sense that we can take the properties of each to conform to its conventional place in physical theory. What's in a name? In particular one of them has to be the carrier of the property of having a limiting value. Having fixed that one, the others are free to have discrete values. Guess what! the particular one that has this property is velocity, and its limit is the velocity of light.

This must look like verbal sleight-of-hand: we seem to have conjured something physical and important out of mere words. I actually think not, and I have faith in my argument. Of course it shows the velocity of light in to be a very different kind of thing from what the conventional language suggests, and we have no alternative and appropriate language as yet. It isn't just a scaling constant in Pope's sense, since that way of speaking suggests that it has nothing to do with velocity, but I agree with him that it is in a different logical category from the velocity of sound, and that is left quite unclear in the current language.

Our identification of c leaves the fine-structure constant still only partially tied down. We need to do something about mass. Noyes thinks we should use the pion cross-section. The bounds which show up in measurement apply to dynamical concepts in general. In particular they apply to mass. Noyes some time ago spotted an interpretation of the fine-structure constant seeing the 137 as the limit to the number of electron/positron pairs that can stably form a unit. To go further on this road Noyes sees the pion as the bridge concept. Just as he got a lot of mileage out of saying that there must be some number of electrons which could be contained without the combination becoming explosive, and saying that that number must be 137, so he extends the angels-and-pin argument to electron/positron pairs to get the enclosure which first brings in mass to give the pion. Noticeably the mass, $274.5m_e$, of the pion, is close to 2x137, which fits in nicely. We emphasize again that at this stage it is a matter of convention whether we load our vision of the physical counterpart or resolution of the combinatorial fixedness onto mass or onto velocity or momentum. Our argument is in a way only repeating a very elementary principle of physics which concerns units of measurement. To have a quantity which is measureable we need to define a unit. This is normally done from experimental practice: now we are drawing attention to the underlying logic.

I should wish to say that Noyes regards the pion as the critical construction which links combinatorial quantities with extended quantities (though that is probably to put my words into his mouth). He thinks of a critical length which is associated with a critical mass defined by counting and being the largest number, 137, of positron/electron pairs that can be stable. Thus for Noyes, the meeting place of space-time and the energy-momentum concept is the finite mass of the pion (from the E,p side) and the Compton wave-length of the pion (from the t,q) side. (See his Helsinki paper, P.2). We quote for reference the Weinberg ratios given in the Helsinki paper:-
Bohr radius = $(h^2/2\pi)/m_ec^2 = 5.10^{-9}$cm;
Electron Compton wave-length = $(h/2\pi)/m_ec = 3.9\text{x}10^{-11}$cm.
Nuclear scales = Classical electron radius = $e^2/m_ec^2 = 2.8\text{x}10^{-13}$

Then $e^2/m_ec^2 = \alpha.(h/2\pi)/m_ec = \alpha^2(h/2\pi)^2/m_ec^2$.

What I have been saying adds a bit to what I said two years ago about entelechies. There I suggested that we needed a concept more general than physical time where different orderings could be potentially present -a bit like the Feynman diagrams with the background of 'real' time thrown away; or rather left to be constructed as seems best in whatever circumstances. Naturally if this kind of freedom is ascribed to physical time it will spread to space. Fortunately it is no longer the case that any departure from a euclidean topology with bending and compression allowed (as in general relativity) is greeted with total incredulity. I read that fractal theories are being seriously discussed for use in the wilder reaches of cosmology. In these theories different appearances of identifiable patterns on different scales have to be treated on the same footing, and in consequence unique scaling goes out of the window. You cannot say what scale you are operating on. This change is not quite so extreme as the one I am suggesting because the patterns do presuppose some spatial order. However the Rubicon is crossed.

The projected work of building dynamics remains to be done except for the vital starting point, and that starting point must guide every further step. In his physical prophecies, Etter (p.61) takes Stein's random walk method for understanding the Lorentz form given the finite velocity of light, and demythologizes it to a statistical theorem -though still given the finite velocity of light. He finishes up with (pp' - qq')/ pp + qq'). Here c is 1, and p and q are probabilities of opposite outcomes such as moving one way or the other or getting heads and tails. The Ur-velocity is then the probability of heads minus the probability of tails, and Etter's contention is that the linking process gives something we can interpret as the relativistic velocity addition law.

Back to our general dynamical problem of defining the classical dynamical concepts. Identification of the limit we call the velocity of light, even when given some statistical trappings like what Stein and Etter suggest, does not give us carte-blanche to jump straight into the continuous ideas. What I hope is to be able to use the suggestions implicit in the limit processes as I have presented them to get a picture of ranges of physical variables to fill in -in between the scale constants.

A Very Brief Note on Link Theory

By Thomas L. Etter

Link theory is

Essentially equivalent to:

SQL
The theory of constraint satisfactions.
Tensor network theory (Penrose, Kauffman)
Older Etter theory

Easily translatable into:

Relational algebra (Pierce, Schroeder, Russell-Whitehead)
The theory of random variables

Has as special cases:

The general theory of input-output systems, including
- Algorithms and computers
- Neural nets
- Cellular automata
- Markov chains and diffusion processes

Quantum mechanics

What link states accomplish:

Classify systems so as to give clear formal meaning to causality and interference.
Distinguish outer (process) structure from inner (system) structure

What link theory isn't:

An alternative empirical theory.

What link theory is:

An alternative method of analysis. The concept of link state enables us to make a number of distinctions that can't be made using the standard vocabulary of the equivalent theories. This greatly clarifies the statement of several chronic problems in the philosophy of science, and raises a host of new questions that were previously unimaginable.

AN INVESTIGATION, ARISING FROM THE THEORY OF QUANTISED VARIABLES, INTO INTER-PARTICLE FORCES

Geoffrey Constable

ABSTRACT

The theory of quantised variables and Schrödinger's equation are used to develop the Quantised Structure Model for particles of half-integral spin. The development of this model relies upon the hypothesis that, within the structure of such a particle, speed is quantised.

Predictions concerning long-range forces between particles are made using the principle of indistinguishability. The existence of the electrostatic force and the value of the fine structure constant are two such predictions.

A further prediction concerns the existence of a weak inverse-square law force, analogous to the force of gravity. It is shown that the scale of this force varies inversely with particle mass, the force between two electrons being more than expected from the force of gravity and the force between two protons being less. The force between two particles of identical charge turns out to be repulsive while that between two particles of opposite charge is attractive. Such a force between two atoms is of the same sense as gravity and may be of the same scale provided that matter is a compound of electrons and pions.

1) INTRODUCTION

The theory of quantised variables has been developed to its present level through a series of papers presented to ANPA. The object of this paper is to explore the possibility that particle structure is quantised and the consequences, with particular reference to inter-particle forces, that would flow from such quantisation.

In the first paper submitted to ANPA it was argued that all variables are quantised in some manner, and that fundamental variables (such as mass, length, time) possess both maximum and minimum values and thus vary incrementally rather than continuously. Such increments may be small - well below the level at which they might be observed using conventional measuring techniques - but their existence can be inferred from the quantised behaviour of other variables.

If one fundamental variable is divided by another, the result is a 'derived variable' that also possesses quantisation, such quantisation being of a 'reciprocal' nature. Such quantisation was examined in the second paper submitted to ANPA. Electrical resistance, for example, is a derived variable (being the quotient of voltage and current). Given special circumstances, low temperatures and high sample purity in particular, quantised levels of electrical resistance are observed of the form R/2, R/3, R/4.... ., indicating that what is being observed is one quantum of voltage divided by two or more quanta of current.

If experimental conditions are refined still further (for example, if the purity of the materials used is increased) other levels of resistance are observed of the form 2R/3, 3R/4 and other simple fractions, indicating that 'n' quanta of voltage have been divided by 'm' quanta of current (n and m being integers).

Gravitational potential is another derived variable, being the quotient of two fundamental variables - mass and distance. As was shown in the third ANPA paper, there is much experimental evidence that the red-shifts of distant galaxies are quantised and that such quantisation accords with differences of gravitational potential of the form $c^2/2$, $c^2/3$, $c^2/4$.... . This leads to the proposal that such shifts are not solely attributable to Doppler Shift but may arise from the quantisation of gravitational potential and the consequent frequency shifts caused by General Theory of Relativity.

Speed is yet another derived variable (length divided by time). All experimental evidence thus far indicates that speed varies continuously in a non-quantised manner from zero up to its observed maximum, the speed of light. The theory of quantised variables suggests, however, that circumstances should exist in which speed displays reciprocal quantisation. As the maximum value of distance is cT_o and the maximum value of time is T_o (T_o being the age of the universe - see paper 1), circumstances should exist in which speed displays reciprocal quantisation of the form c/2, c/3, c/4..... .

(Note: as shown in paper 1, the quotient of the maximum values of two variables is identical to the quotient of the minimum values of the same two variables).

It has been determined by experiment that sub-atomic particles possess many unusual properties that differ from those observed in the 'macro' world - including quantised angular momentum, quantised electric charge and the ability to behave as both particles and waves. Is it possible that the structure of such a particle might provide the special circumstances that are needed for the reciprocal quantisation of speed to occur? This paper considers the likelihood of such a suggestion and its consequences in relation to the quantisation of particle structure. Phenomena that could be accounted for by such quantisation - including inter-particle electrostatic and gravitational forces - are examined.

2) PARTICLE STRUCTURE

The nature of sub-atomic particle structure remains uncertain. The 'standard theory' of mesons and baryons is based on the concept of quarks and gluons - a concept that is widely accepted and is supported to some degree by experiment, including the observation (during scattering experiments) of point-like concentrations of charge within the proton of magnitude e/3 and 2e/3 (e being the charge of the electron). On the other hand, quark theory has weaknesses that have yet to be resolved. For example, it has proved impossible to determine precise values for quark masses. Furthermore, the magnetic moments of some known particles as predicted by this theory differ significantly from values determined by experiment.

Turning to the lepton, scattering experiments indicate that such a particle is a 'point-like' mass that is diffused into some form of spinning 'charge cloud'. However, there are some features of leptons that are not explained by such a concept.

For example, one important feature of the lepton is that it possesses a half-integral spin or angular momentum. If such angular momentum is to be attributed to the particle spinning about its own axis, the minimum radius that can be possessed by a spherical and homogeneous particle is its Compton Radius R, which is equal to $\hbar/mc$ (m being the mass of the particle).

This observation can be explained as follows. The moment of inertia of a relativistically spinning sphere is $1/2 \times mr^2$, r being the radius of the particle.

If $r = \hbar/mc$, and the angular momentum is $\hbar/2$, it is simple to show that the peripheral velocity of the sphere at its equator will be 'c', the speed of light. If the particle radius were less than its Compton Radius, this speed would have to exceed c, an impossibility.

On the other hand, evidence from scattering experiments indicates firmly that the electron (for example) is point-like and has a radius less than 10^{-17} cm, some six orders of magnitude smaller than its Compton Radius. Plainly, current understanding of particle structure, and of lepton structure in particular, is less than complete.

3) THE LOCATION OF A POINT-LIKE PARTICLE

The anomaly referred to above requires us to accept that a point-like particle (eg an electron) cannot achieve its angular momentum by spinning about its own axis and can only do so by orbiting at tangential speed 'v' at a radius 'r' about a mean location. Given that such a particle has semi-integral spin,

$$\hbar/2 = mr^2 x\, v/r = mrv \qquad (1)$$

If the tangential velocity is not to exceed 'c', the minimum radius of such an orbit must exceed $\hbar/2mc$, or half the Compton Radius.

Both the Uncertainty Principle and Schrödinger's Equation express (in slightly different ways) the concept that the location of a particle is not precise but is extended and probabilistic. Irrespective of location, however, we are certain that angular momentum will be conserved. Consequently, (see equation 1) v is inversely proportional to r.

Thus by (1), the kinetic energy 'E' of the particle is given by

$$E = 1/2\, mv^2 = \hbar^2/8mr^2 \qquad (2)$$

As is well known, the uni-dimensional location (x) of a particle is described by the time-independent version of Schrödinger's Equation:

$$-\frac{\hbar^2}{2m}\left(\frac{d^2\psi}{dx^2}\right) + V(x) = E(\psi) \qquad (3)$$

where m is the mass of the particle, V(x) is the potential energy of the particle at a point 'x', E is the 'momentum-based' energy of the particle, and y is the Schrödinger 'wave function'.

Should the potential function possess spherical symmetry, it can be shown that Schrödinger's Equation can be written in a polar form:

$$-\frac{\hbar^2}{2m}\left(\frac{d^2u}{dr^2}\right)+V(r)=E(u) \qquad (4)$$

where $\psi = u/r$.

Schrödinger's Equation is commonly used to determine the 'eigen states' of particle energy that arise should the potential V adopt various forms and boundaries. However, we now use this equation to describe the location of a particle in free space. No constraints (potential barriers and otherwise) are imposed except those that might arise from the particle's own presence and properties.

Although the precise location of a point-like particle may be probabilistic, we can postulate that such a particle will occupy only one location at a time. In other words, when a particle is in one location, it is not in another, and vice versa.

Thus the potential energy of a single point-like particle in free space is at all times zero - an arrangement that possesses spherical symmetry. The particle cannot have gravitational or electrical potential energy due to interaction with 'ghost' particles left in previous locations.

On the other hand, the presence of angular momentum that is both conserved and quantised causes the kinetic energy E of the particle (which becomes the total 'momentum-based' energy of the particle) to vary inversely as the square of r, as described by equation (2).

Using equations (2) and (4), and putting V(r) equal to zero,

$$-\frac{\hbar^2}{2m}\left(\frac{d^2u}{dr^2}\right)=\frac{\hbar^2 u}{8mr^2}$$

$$\text{or} \quad -\left(\frac{d^2u}{dr^2}\right)=\frac{u}{4r^2} \qquad (5)$$

This equation is remarkable in that it is independent of 'm', the mass of the particle. Thus, this equation is valid irrespective of whatever changes to the rest mass of the particle might take place over the full range of 'r'. This finding is important because at small values of 'r' the speed 'v' of the particle is not small in comparison with the speed of light. Any relativistic changes of mass that arise in consequence, and their effects upon momentum and energy, can be ignored.

Furthermore, if the angular momentum of the particle is to be conserved, the kinetic energy of the particle will have to vary - as pointed out above. We have postulated that the potential energy is zero for all values of 'r' so, if the energy of the complete system is to be conserved, the mass of the particle will have to vary appropriately with changing values of 'r'. No matter: such variation can also be ignored.

(5) is a special form of the equation sometimes referred to as the Cauchy or Euler Equation:

$$x^2 y'' + axy' + by = 0$$

Ignoring, for the moment, factors associated with boundary conditions and the avoidance of infinities, a solution to this equation can be obtained by making a substitution of the form

$y = x^m$, (or $u = r^m$),

which yields from (5) the auxiliary equation:

$$-m(m-1) = \tfrac{1}{4}$$

or $m = \tfrac{1}{2}$.

Thus the solution to (5) is u = (C1 + C2) √r, C1 and C2 being two constants.

or $u = \sqrt{Ar}$ u , where A is a third constant. (6)

This fact is used as a 'hint' when boundaries and infinities are considered later.

4 THE QUANTISATION OF SPEED WITHIN THE STRUCTURE OF A POINT-LIKE PARTICLE

It was argued in (3) above that, in order to conserve angular momentum, the tangential speed of an orbiting point-like particle varies according to the reciprocal of the orbit radius.

Let us investigate the possibility that this tangential speed is quantised reciprocally, as outlined in (1), according to the values c/2, c/3, c/4 ..... . Such a possibility has to be presented at this moment as a hypothesis, although there are findings given later in this paper that support its likelihood.

As angular momentum is quantised and conserved, by (1) and our hypothesis

$\hbar/2 = mrv = mrc/n$, where n is an integer eg 2, 3, 4..... . (7)

Or $r = n\hbar/2mc = nR/2$ where R is the Compton radius $\hbar/mc$ (8)

Thus r is no longer continuous but is confined to values that define probability points and are an integral number of half Compton Radii. Under these circumstances, equation (5) - a continuous differential equation - may be unsuitable for our purposes. We convert it into a difference equation that is suitable for dealing with discontinuous variables, such as 'r' has become in the light of our hypothesis.

Fortunately, the need to undertake computer simulations of Schrödinger's equation has produced the type of difference equation that is needed. In such simulations it is common to use the approximate relationship

$$\frac{d^2u}{dr^2} = \frac{u_{(n-1)} - 2u_{(n)} + u_{(n+1)}}{\Delta r^2}$$

$$= -\frac{u_n}{4r^2} \quad \text{by equation (5)} \qquad (9)$$

We use the hint given earlier and try the solution $u = \sqrt{Ar} = \sqrt{\frac{AnR}{2}}$ for equation (9).

Putting $\Delta r = R/2$

$$\sqrt{\frac{R(n-1)}{2}} - 2\sqrt{\frac{R(n)}{2}} + \sqrt{\frac{R(n+1)}{2}} = -\sqrt{\frac{R(n)}{2}} \div 4n^2$$

This can be written as

$$\sqrt{n\left(1-\frac{1}{n}\right)} - 2\sqrt{n} + \sqrt{n\left(1+\frac{1}{n}\right)} = -\frac{\sqrt{n}}{4n^2}, \text{ or}$$

$$\sqrt{1-\frac{1}{n}} - 2 + \sqrt{1+\frac{1}{n}} = -\frac{1}{4n^2} \qquad (10)$$

The function $\sqrt{1 \pm \frac{1}{n}}$ can be expressed as

$$1 \pm \frac{1}{2n} - \frac{1}{8n^2} \pm \frac{1}{16n^3} - \frac{5}{64n^4}\ldots\ldots$$

Substituting in equation (10), and ignoring terms above $1/n^3$, yields

$$-\frac{2}{8n^2} = -\frac{1}{4n^2}$$

Thus $u = A\sqrt{r}$ is a solution to difference equation (9) for all large values of n. [It is a close approximation to being a solution even where n is small - minor discrepancies in this instance being, perhaps, attributable to (9) being an approximate form of (5)].

We now have no need to consider boundary conditions attributable to the discontinuous nature of 'r', since such as apply are described within the difference equation.

It is well-known that the angular momentum of a fermion is either 'up' or 'down'. We choose to select axes so that the motion of the point-like particle is confined to the XY plane, the angular momentum that arises being aligned along the OZ axis.

In these circumstances, the probability distribution of such a particle in the XY plane will be a set of discrete concentric circles, each of radius nR/2.

As $\psi = u/r$, see equation (4) $\qquad \psi = \sqrt{A/r}$, and $\psi^2 = A/r$

A must have the dimensions of length, so we choose to write it as $R/2B$, where B is a dimensionless constant to be determined later when, in a normalising process, the sum of the probabilities associated with all possible particle locations ($\Sigma\psi^2$) is equated with unity.

$$\text{Thus } \psi^2 = \frac{R}{2rB} \qquad (11)$$

To summarise, the analysis given above indicates that the probability distribution of the location of a point-like particle is discontinuous and is proportional to the reciprocal of the distance of the particle from its mean position. Furthermore, if predictions of speeds exceeding that of light are to be avoided, a point-like particle has to be located at a distance from its mean position that is greater than half its Compton radius.

5 FURTHER IMPLICATIONS OF THE QUANTISATION OF ANGULAR MOMENTUM.

Consider a point-like particle with angular momentum $\hbar/2$ that is located (momentarily) on the Y axis (x = z= 0). Its angular momentum is given by

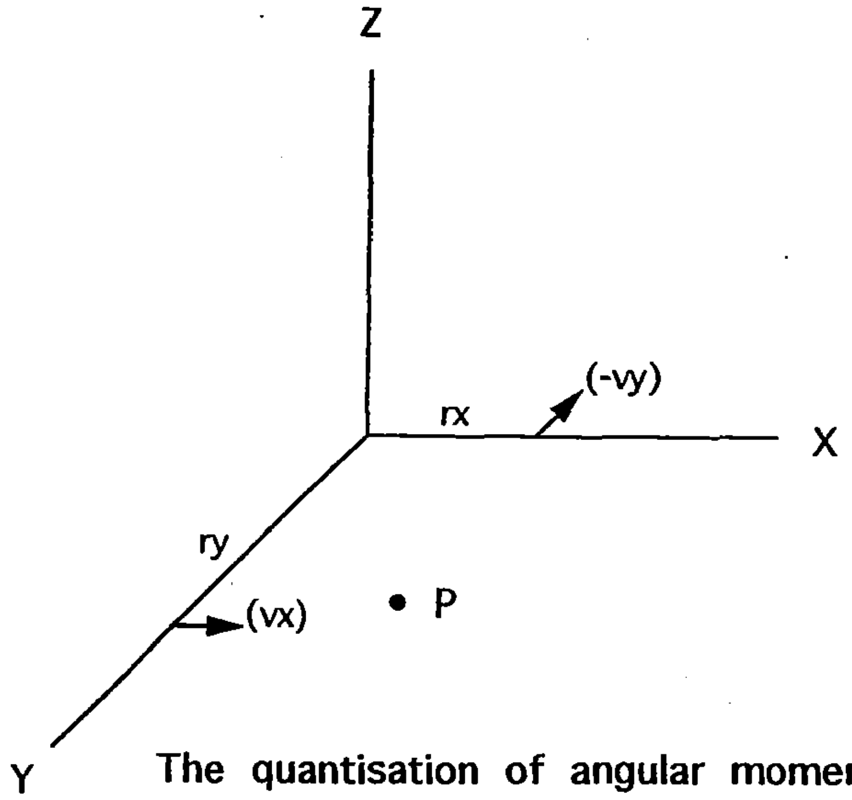
$\mathbf{L}_z = m\mathbf{r}_y\mathbf{v}_x = \hbar/2$ see Fig 1

A similar particle located on the x axis has angular momentum given by:

$$\mathbf{L}_z = m\mathbf{r}_x(-\mathbf{v}_y) = \hbar/2,$$

the minus sign arising from the fact that $\mathbf{v}_y$ is now aligned along the - Y axis.

Fig 1



The quantisation of angular momentum

At an intermediate point P the angular momentum is given by the addition of two vectors:

$$\mathbf{L}_z = m\mathbf{r}_y\mathbf{v}_x + m\mathbf{r}_x(-\mathbf{v}_y)$$

This expression presents us with a problem. It possesses two terms which, when summed, are equal to a single quantum of angular momentum. As neither can have a value of less than, $\hbar/2$ we are forced to the conclusion that either one or the other must be zero.

It is difficult to imagine how **v** can be zero at the intermediate positions referred to above. After all, one quantum of angular momentum is present at all times. Hence either $\mathbf{r}_y$ or $\mathbf{r}_x$ must be zero. In other words, the particle can be observed only when it is located on the X or Y axis, the orientation of such such axes being determined by the perspective of the (quantum) observer.

Such a conclusion places additional restraints upon the model derived in (4) above. The set of concentric circles located in the XY plane now reduces to sets of equidistant probability points, one point of each set being located on one of the four axes in the XY plane.

6) THE QUANTISED STRUCTURE MODEL

These considerations lead to the conclusion that a particle (lepton) in general possesses a structure that consists of 'elements' at 'probability points' that are located along quantised x or y coordinates, the quantum of length being half the Compton Radius. Each element is at its probability point only momentarily but, in accordance with our hypothesis, possesses a quantised transverse speed that is inversely proportional to the distance between that point and the mean position of the particle. As defined by equation (11), the probability that an element is located at its point is again inversely proportional to the distance between that point and the mean position of the particle. This model is illustrated by Fig 2.

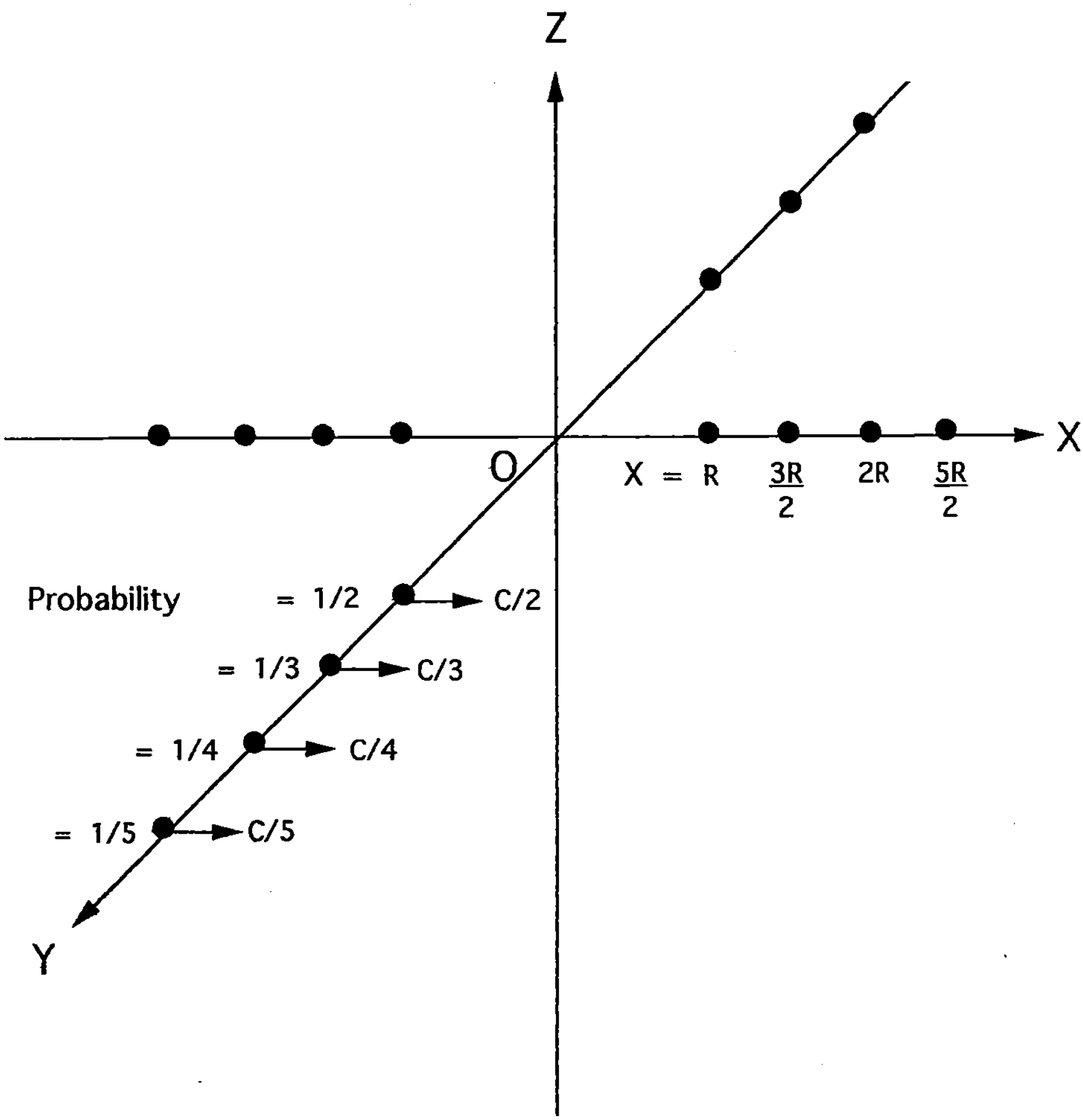


Fig 2 The quantised structure model

Such a proposal is radical. It is necessary, therefore, that its implications be tested to see whether they lead to predictions that are consistent with observation.
Many such tests can be proposed. In this paper consideration is given to the following:

a) can the probabilities defined by the quantised structure model described above be normalised in a satisfactory manner;

b) can the wave/particle duality of some particles be explained by the quantised structure model; and

c) does the model shed any light on long-range forces between particles?

Such tests are considered in turn.

7) THE NORMALISATION OF PROBABILITIES

According to the quantised structure model, the probability of a point-like particle being located at a distance 'r' from its mean position (the origin) is given by:

$\Psi^2 = R/2rB$ (see equation 11):

Hence, the sum of probabilities for any one axis is:

$$R/2B\left(2/2R + 2/3R + 2/4R \cdots\cdots 2/nR\right)$$

$$= 1/B\left(1/2 + 1/3 + 1/4 \cdots\cdots 1/n\right) \qquad (12)$$

If n is infinitely large, the sum of this series is also infinite. Fortunately, as indicated by the theory of quantised variables, all primary variables have maximum values. The maximum value for distance is cT_0 (T_0 being the age of the universe - see paper 1). Since we are dealing here with a radius, not a diameter, it is reasonable to guess that, in this context, the maximum value of the distance $nR/2$ is $cT_0/2$, whence

$$n = cT_0/R$$

We denote this number as N.

All factors required to calculate 'N' are known precisely with the exception of T_0, which is believed to have a value between 10 and 20 billion years. A value of 15 x 10^9 years (or 4.73 x 10^{17}s), which is considered to be the most probable value by many cosmologists, is used for the following calculation. On the basis that we are calculating 'N' for an electron and that the value used for R is that of the Compton radius of the electron,

$N = 3.672 \times 10^{38}$

In order to calculate the normalisation constant B, we make use of the formula that defines Euler's Constant, 0.5772157:

$$\text{Lim (n tends to infinity) } 1 + \tfrac{1}{2} + \tfrac{1}{3} + \ldots \tfrac{1}{n} - \ln(n) = 0.5772157$$

$$\text{or } \left(\tfrac{1}{2} + \tfrac{1}{3} + \ldots\ldots \tfrac{1}{n}\right) = \ln.N - 0.4228 = 88.3765$$

Note: this series has to start with 1/2, not 1 since, as explained earlier, an element of an electron cannot be located at one half Compton Radius from the mean position of the electron unless the tangential velocity is 'c' - an impossibility.

Thus, from equation (12)

B for one axis = 88.3765 (13)

Thus, the theory survives the first test as set out in (5), and normalisation can be undertaken. Furthermore, none of the terms expressed in equation 12 are infinite. We are not, therefore, troubled by infinities (as would be the case should r = 0 be a possibility).

8) THE WAVE/PARTICLE DUALITY

It is well-known that particles can display wave-like characteristics. Beams of electrons are diffracted by substances possessing crystal lattices to provide images that display concentric peaks and troughs of intensity. Other particles, eg protons, neutrons and alpha particles, have been shown to exhibit similar properties.

Other wave-like properties of particles include the production of diffraction patterns from twin or multiple slits, and even from straight-edges. In short, every wave-like property exhibited by light seems to be reproducible by particles. The conclusion that particles behave both as particles and as waves seems to be incontestable.

There are, however, some other possibilities. If, for example, a particle has a structure based on an array of points, diffraction and interference phenomena as referred to above will also occur. Such a particle, however, could not reasonably be described as wave-like. Arguments such as this demonstrate that it is sufficient for a particle structure to possess some measure of periodicity if wave-like properties are to be exhibited. The quantised structure model proposed in this paper (fig 2) possesses such periodicity and can, therefore, be expected to lead to wave-like characteristics.

The relevance of this model to the wave-like performance of particles is reviewed as follows.

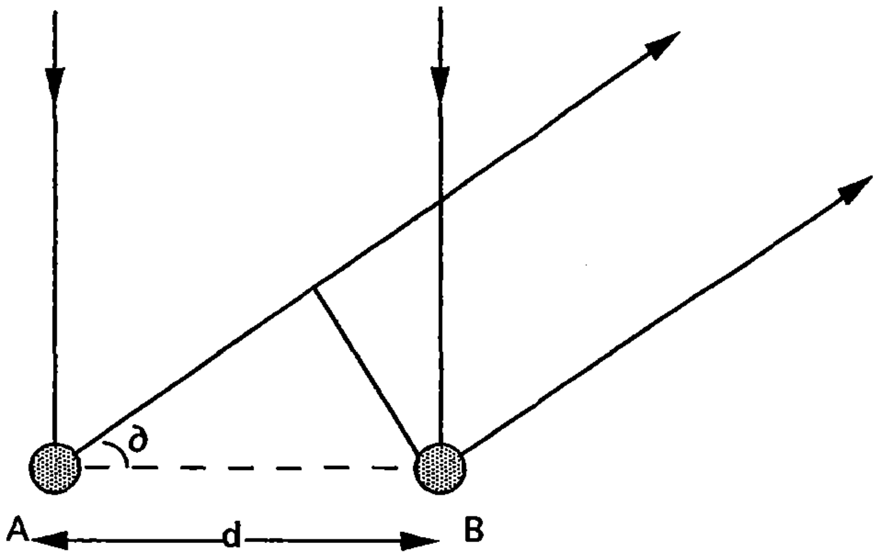


Fig 3 Conventional Explanation of Electron Diffraction - $d \cos \partial = nL$

a) The conventional explanation for electron diffraction relies on Fig 3. For reinforcement to take place, it is argued that an incoming electron has to be reflected from a crystal lattice such that $d \cos \delta = n\lambda$, the pitch of the lattice 'd' being considerably larger than the 'wavelength' λ of the electron. Unfortunately, no convincing explanation is given to the student on how one electron can be reflected simultaneously by two atoms (the Compton radius of the electron being considerably smaller than the radius of any atom) and then, subsequently, on how the two emerging but widely separated rays are able to interfere with each other. The student is left to imagine some mechanism for 'reinforcement at a distance'.

The quantised structure model is helpful in this regard. The spread of the probability points of both the incident and reflected 'rays' is sufficiently wide to bridge the lattice pitch 'd'. Thus, a further explanation of 'reinforcement at a distance' is not needed.

b) It is a common feature of electron diffraction photographs that the interference bands are in the form of sharp and narrow peaks - see fig (4). The conventional explanation is that electrons penetrate through several layers of the crystal lattice, thereby reducing the spread of angle over which reinforcing interference can take place and narrowing the width of the resulting fringe. Interestingly, electrons of low energy - ie those with insufficient energy to penetrate beyond the first atomic layer of a crystal - also produce photographs with thin sharp reinforcement interference fringes, a situation that is described as 'rather surprising' in standard texts.

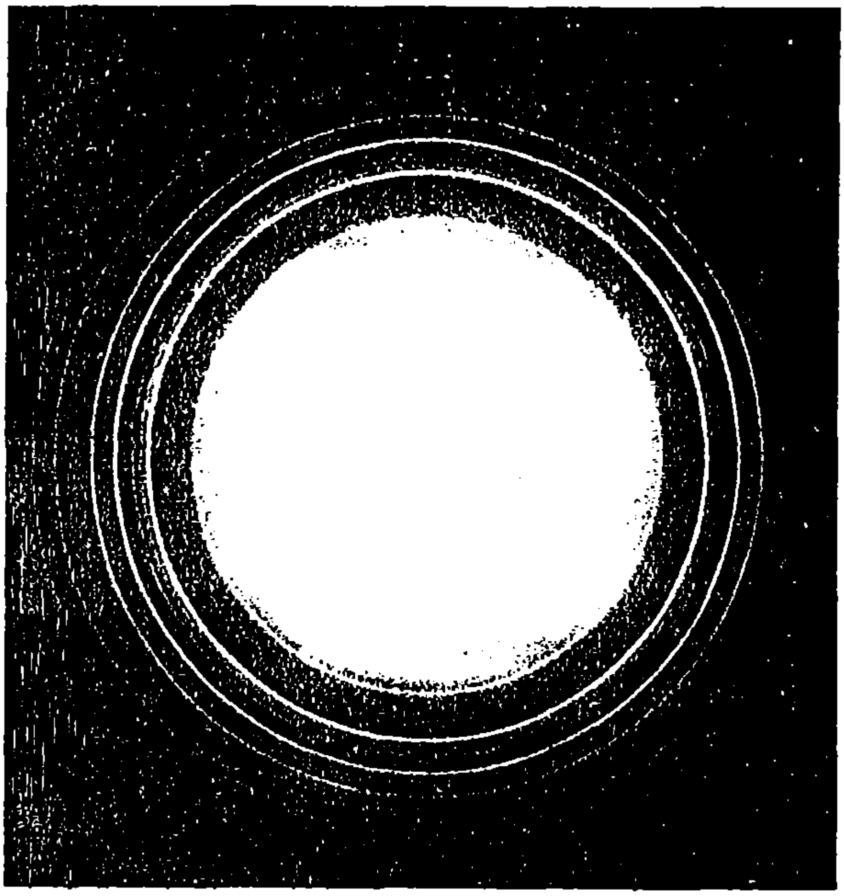


Fig 4 Typical Electron Diffraction Pattern

The quantised structure model makes such surprise unnecessary. As the location of the electron is spread over an array of sharply defined points, the resulting interference peaks will be sharp and narrow, however few layers of atoms are penetrated.

9 INTERACTION AT A DISTANCE BETWEEN TWO IDENTICAL ELECTRONS

One test of the proposed quantised structure model is that the properties of this model should be consistent with known inter-particle forces. The possibility that one particle can act remotely upon another is examined as follows.

Imagine two point-like particles (eg two electrons) A and B of identical charge separated by a distance 'd', where $d = pR/2$ (p being a large integer and R being the Compton Radius of the electron). Such an arrangement is shown in Fig (5).

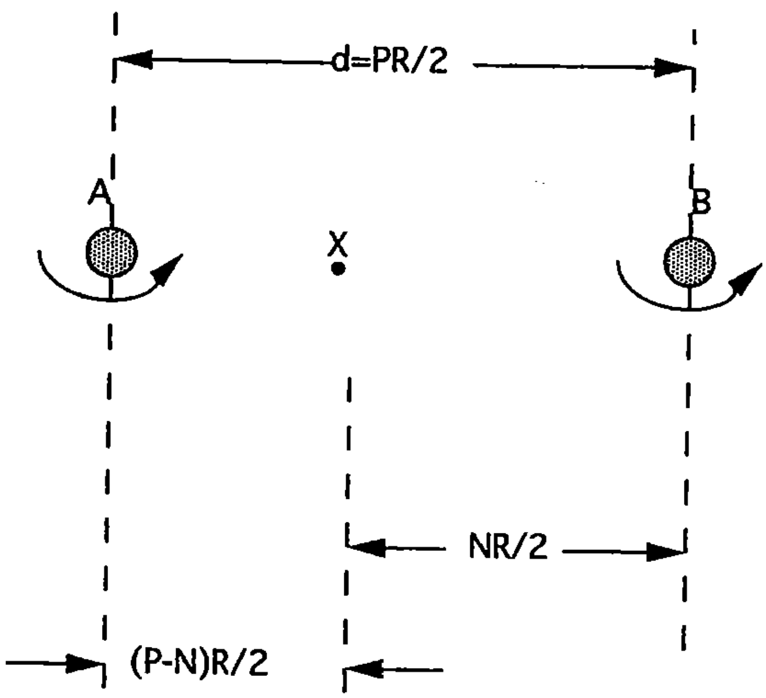


Fig 5 Interaction between two electrons
Region between particles A and B

An intermediate point 'X' will occasionally be visited by elements of electron B. Due to the principle of indistinguishability, the transient presence of an element of electron B at X will act on electron A as though it were a part of its own structure.

According to the quantised structure model, the transverse velocity of the element of electron B when at X is c/n - the distance from X to B being nR/2. As the two electrons possess angular momenta of identical sense, this velocity will be in the reverse direction to that of an element of electron A when it is also at X. If electron A were in isolation, the centrifugal forces of its various elements would on average be in balance, and there would be no net force on the particle. The presence of elements of electron B in the region between the two particles will reduce the net centrifugal force on A. The outcome is that there will be a net repulsive force on electron A, ie a force that drives electron A away from electron B.

The acceleration of the element of B towards A is given by v^2/r, or

$$\left(c/n\right)^2 x \left(2/(p-n)R\right)$$

The probability of an element of electron B being at location X is 1/nB, B being the normalisation constant as defined by equation (13). Note: we can assume that the 'X' axes of the two electrons are co-incident. As explained earlier, one electron being the quantum observer will 'perceive' the other electron as being so aligned, and vice versa.

The mass of the electron is M_e and, having been factored by the appropriate probability 1/nB, the resulting force $\Delta F1$ on A is given by:

$$\Delta F1 = \frac{2M_e c^2}{RBn^3(p-n)} \qquad (14)$$

The positive sign indicates that this force is repulsive in nature.

To calculate the cumulative force that arises from this effect in the region A to B, this expression has to be summed from n = 2 to n = P. Fortunately, expression (14) can be turned into a set of partial fractions.

$$F1 = \sum_{n=2}^{n=P} \frac{2M_e c^2}{BR}\left(1/p^3 n + 1/p^2 n^2 + 1/pn^3 + 1/p^3(p-n)\right) \qquad (15)$$

Should an element of electron B be located to the right of the mean location of electron B (see fig 6), further forces will be exerted on electron A. In this region elements of both electrons have velocities that are aligned in the same direction and, therefore, add. The centrifugal force on A is increased and thus A experiences a force towards B. As this force is attractive in nature, it is marked by a minus sign.

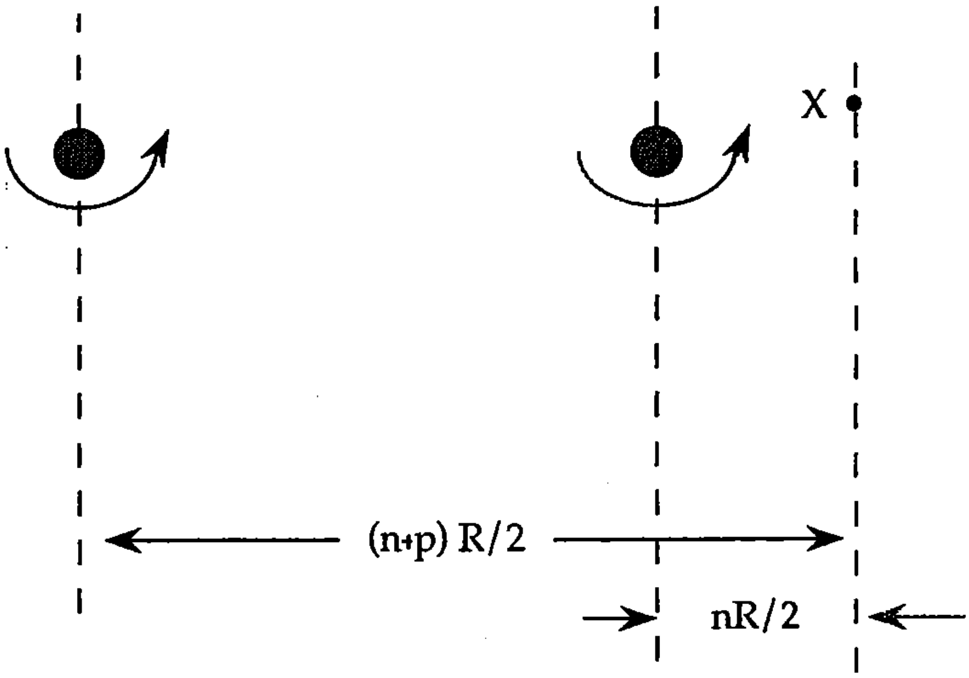


Fig 6 Interaction between two electrons
Region to the right of particle B

The acceleration of the element of B as 'perceived' by electron A is

$$-\frac{c^2}{n^2}x\frac{2}{(n+p)R}$$

The resulting force Δ F2 , using similar arguments, is given by the expression

$$\Delta F2 = -\frac{2M_e c^2}{BRn^3(n+p)}$$

$$F2 = -\sum_{n=2}^{n=N} 2M_e c^2\left(\frac{1}{n^3 p} - \frac{1}{n^2 p^2} + \frac{1}{np^3} - \frac{1}{p^3(p+n)}\right) \qquad (16)$$

In order to take account of the complete force generated in this region (to the right of B) we must sum for all values of n from n = 2 to n = N.

We must now calculate F3, the force that arises from elements of electron B being located in the region to the left of electron A. Once more, the motion of an element of B adds to the motion of an element of A. However, the additional force drives A away from B and is, therefore, repulsive. We regard it as positive.

Using arguments similar to those given above,

$$F3 = \sum_{n=p}^{n=N} 2M_e c^2 \left(1/_{n^3 p} + 1/_{n^2 p^2}\, 1/_{np^3} + 1/_{p^3 (n-p)} \right) \qquad (17)$$

to be summed from n = p to n = N.

The total force F upon A is the sum F1 + F2 + F3.

$$\text{Thus } F = \frac{2M_e c^2}{BR} \left\{ \sum_{n=2}^{n=p} \left(1/_{p^3 n} + 1/_{p^2 n^2} + 1/_{pn^3} + 1/_{p^3 (p-n)} \right) \right.$$

$$- \sum_{n=2}^{n=N} \left(1/_{p^3 n} - 1/_{p^2 n^2} + 1/_{pn^3} - 1/_{p^3 (p+n)} \right)$$

$$\left. + \sum_{n=p}^{n=N} \left(1/_{p^3 n} + 1/_{p^2 n^2} + 1/_{pn^3} - 1/_{p^3 (n-p)} \right) \right\} \qquad (18)$$

Terms of the form $1/_{pn^3}$ cancel to zero as, approximately, do terms of the form $1/_{p^3 n}$ and $1/_{p^3 (p+n)}$.

We are left with terms of the form $1/_{p^2 n^2}$ which can be expressed as

$$\frac{2M_e c^2}{RB} \sum_{n=2}^{n=N} 2/_{p^2 n^2}$$

But substituting $$p = \frac{2d}{R},$$

$$F = \frac{M_e c^2 R}{Bd^2} \sum_{n=2}^{n=N} 1/n^2$$

With the further substitution $R = \frac{\hbar}{mc}$

$$F = \frac{\hbar c}{Bd^2} \sum_{n=2}^{n=N} 1/n^2$$

It is well-known that the sum of the infinite series $1 + 1/2^2 + 1/3^2 ..... = \pi^2/6$

Thus $1/n^2$ (summed from n = 2 to n = N, N being large) $= \pi^2/6 - 1 = 0.644931$

As B = 88.3765, the net repulsive force can be calculated:

$$F = \frac{\hbar c}{d^2} \text{ x } 0.644931/88.3765$$

$$= \frac{\hbar c}{d^2} \text{ x } 1/137.0324 \qquad (19)$$

The (repulsive) electrostatic force F between two identical electrons is, by definition,

$F = e^2/d^2$ (e being the electronic charge).

As $e^2 = \hbar c/137.0360$ (The fine structure constant α being 1/137.0360)

$$F = \frac{\hbar c}{137.0360 d^2},$$

in astonishingly good agreement with (19) above.

If electrons A and B are of opposing spin, similar arguments can be used to show that force F still obtains, but is attractive rather than repulsive.

The number given in equation 19 depends upon the logarithm of an estimate of the age of the universe. We would not, therefore, expect this number to be identical to the actual fine-structure-constant. Due to this logarithmic relationship, comparatively large variations in T_0 causes only small changes to the predicted value of the fine-structure-constant. In fact, multiplying the value of T_0 by a factor of two would change the calculated value of the fine structure constant by less than 1%.

Nevertheless, this prediction of the value of the fine-structure constant and, hence, of the charge of the electron suggests that the original hypothesis of this paper may have substance. That the existence of electrostatic force between electrons seems to arise from considerations that are entirely non-electrical in nature is surprising, and may pave the way for further study.

In particular, the work of ANPA in establishing links between the properties of numbers and those of nature may receive some measure of support from these findings. The fine structure constant, as shown above, is given by the expression

$$1/\alpha = 137.03 = \left(\frac{\sum 1/n}{\sum 1/n^2}\right)_{n=2}^{n=N} \quad \text{where } N = \frac{cT_0}{R} = 3.672 \times 10^{38}$$

R being the Compton radius of the electron and T_0 (the age of the universe) being estimated at 15 billion years.

It is immediately apparent that the value of N is similar to that of the final number in the combinatorial hierarchy developed by ANPA, ie $2^{127} = 1.7x10^{38}$. Why this should be so has yet to be explained. There may, however, be a clue in that the expression defining N can be rewritten as

$$N = \frac{cT_o}{R} = cT_o x \frac{mc}{\hbar} = \frac{mc^2}{\hbar / T_o} = m \Big/ \frac{\hbar}{c^2 T_o}$$

In other words, N is related to the ratio between the mass of the electron and the minimum mass that can exist - see paper 1. The thought that the mass of the electron might be defined in terms of the combinatorial hierarchy and the minimum mass is intriguing, and may merit further consideration.

10 LONG-RANGE WEAK FORCES BETWEEN ELECTRONS

A detailed examination of equation (18) reveals that the terms that include $1/p^3$ do not, as assumed in the previous section, sum precisely to zero. The small residual force that results is referred to as Fg so, from (18),

$$F_g = \frac{2M_ec^2}{BR}\left\{\sum_{n=2}^{n=p}\left(\frac{1}{p^3 n} + \frac{1}{p^3(p-n)}\right)\right.$$

$$+\sum_{n=2}^{n=N}\left(-\frac{1}{p^3 n} + \frac{1}{p^3(p+n)}\right)$$

$$\left.+\sum_{n=p}^{n=N}\left(\frac{1}{p^3 n} - \frac{1}{p^3(n-p)}\right)\right\}$$

but, $\sum_{n=2}^{n=p}\frac{1}{p^3 n} + \sum_{n=p}^{n=N}\frac{1}{p^3 n} = \sum_{n=2}^{n=N}\frac{1}{p^3 n}$

and $\sum_{n=2}^{n=p}\frac{1}{p^3(p-n)} + \sum_{n=2}^{n=N}\frac{1}{p^3(n+p)} = \sum_{n=2}^{n=N+p}\frac{1}{p^3 n}$

Thus,

$$F_g = \frac{2M_ec^2}{BR}\left(\sum_{n=2}^{n=N}\frac{1}{p^3 n} - \sum_{n=2}^{n=N}\frac{1}{p^3 n} + \sum_{n=2}^{n=N+p}\frac{1}{p^3 n} - \sum_{n=2}^{n=N-p}\frac{1}{p^3 n}\right)$$

$$= \frac{2M_e c^2}{BR} \sum_{n=N-p}^{n=N+p} \frac{1}{p^3 n}$$

If N is much larger than p, a condition that is likely to be satisfied for most laboratory experiments, but not for all calculations in cosmology,

$$\sum_{n=N-p}^{n=N+p} \frac{1}{p^3 n} = 2p \frac{1}{p^3 n} = \frac{2}{p^2 n}$$

Which gives $$F_g = \frac{4M_e c^2}{BRp^2 N}$$

As $$N = \frac{cT_o}{R}$$

$$F_g = \frac{4M_e c}{BT_o p^2}$$

As $$p = \frac{2d}{R}$$

$$F_g = \frac{M_e cR^2}{BT_o d^2}$$

But $$R = \hbar / mc$$

Hence, $F_g = \dfrac{\hbar^2}{M_e cT_o Bd^2}$, which can be written as

$$F_g = \frac{\hbar^2}{M_e^3 cT_o B} \frac{M_e^2}{d^2} \qquad (20)$$

Several observations are needed at this point.

a) This force resembles the electrostatic force between two similarly charged particles described in (9) above in that it varies inversely as the distance 'd' between the two particles.

b) Due to the positive sign of Fg, this force is repulsive in nature.

c) Due to the inclusion of the term R/cT_0, this force is weaker than the electrostatic force by a factor of some 10^{38}.

d) The force (between two identical particles) varies inversely with particle mass.

e) Using arguments similar to those shown above, it can be shown that the force between two identical but oppositely charged electrons (ie an electron and a positron) is equal in scale to that described by equation (20), but is attractive in sense.

f) Inserting appropriate values into equation (20) shows that the value of the force constant for electrons, denoted as G_e, is given by

$$G_e = -\frac{\hbar^2}{M_e^3 c T_0 B} = -1.176 \times 10^{-3} \qquad (21)$$

some five orders greater than the gravitational constant ($G_n = 6.673 \times 10^{-8}$).

Thus, this weak force between two electrons is not what one might expect from the law of gravity. The force constant is too large and the sense is reversed.

11 WEAK LONG-RANGE FORCES BETWEEN PROTONS

The proton resembles the electron in that it, too, has half-integral spin. It should, therefore, comply with the arguments already stated, provided that appropriate constants are employed.

Using the analysis given above, the weak force between two protons (Fp) is given by the expression

$$F_p = \frac{\hbar^2}{M_p^3 c T_0 B} \frac{M_p^2}{d^2} = G_p \frac{M_p^2}{d^2} \quad \text{(where Mp = proton mass)}$$

The gravitational force constant G_p for a pair of protons is less than that for a pair of electrons by a factor of 6.189×10^9 - and consequently is less than the gravitational constant G_n.

The gravitational constant is measured by conducting experiments that employ masses consisting of atoms - not collections of protons, or electrons, or any other single sub-atomic particle. We consider, therefore, how equation (20) might be modified to take account of this fact.

Let us imagine that an experiment is carried out to measure the gravitational constant using two hydrogen atoms.

As explained above, the repulsive force arises almost solely from repulsion between the two electrons. The hydrogen atom masses are each $M_e + M_p$ + binding energy which, in total, can be approximated to M_p. We need to calculate the gravitational force constant G_h between two such masses that experience a gravitational force equivalent to that between two electrons

Equation (20) can be rewritten as

$$F_e = \frac{\hbar^2}{M_e M_p^2 c T_o B} \frac{M_p^2}{d^2} = G_h \frac{M_p^2}{d^2} \qquad (22)$$

In this case the force constant is 3.48×10^{-10}, less than the gravitational constant (6.673×10^{-8}) by a factor of some 200 and of the wrong sense.

However, we have ignored the possibility that an attractive force arises from interaction between the electron in one atom and the proton in the other. This is an area in which assumptions have to be made if we are to progress.

First, it is not clear whether elements of an electron disposed around a proton in a manner as described in section (9) would be seen by the proton as 'indistinguishable' elements of common matter, or whether such elements would be rejected by the proton as belonging to another and distinguishable particle. We assume for the moment that the former is the case, such an assumption being supported by the existence of electrostatic attraction between a proton and an electron.

Referring to figs (5) and (6) we propose that an electron is located at point B and a proton at point A. All distances are measured in terms of the electron Compton radius, with the result that the analysis proceeds as before. The fact that the proton is a relatively massive particle is unimportant with respect to this model and calculation; we are concerned solely with the centrifugal force generated by elements of the electron as, through the principle of indistinguishability, they appear to orbit around the proton.

The reverse situation, in which elements of a proton exert force on an electron, is complex and difficult to analyse, and further assumptions would be needed. We rely, instead, upon Newton's third law: namely that action and reaction are equal and opposite. Thus the force exerted by the electron on the proton must equal the force exerted reciprocally by the proton on the electron.

Due to the particles in question possessing opposing charges, the weak gravitational force that results from this effect will be attractive. As such attraction comes from the electron of the first atom acting on the proton of the second and from the proton of the first atom acting on the electron of the second, this attraction will be twice as large in scale as the repulsive force described in equation (20).

It follows that, with this model, the overall force from this source will be attractive in nature and will equal in scale that described in equation (20). As the masses concerned are those of hydrogen atoms, not electrons, the force constant is as given in (21), ie too small by a factor 200 or so, but of the correct sense.

However, it may be incorrect to assume (in the context with which we are concerned) that the proton behaves as a single and homogeneous particle. It is widely assumed, for example, that the proton has some form of structure that involves subsidiary parts. As atoms, in general, are electrically neutral, we explore the scale of mass of the positively charged sub-atomic particle that would be needed (in conjunction with electrons) for the construction of nucleons and atoms, if a gravitational constant of the value that is observed by experiment is to be produced by interaction between atoms. The mass of such a particle is denoted by M_X.

As in equation (22), the force constant G_X of the gravitational-type force created by interaction between atoms that consist of equal numbers of electrons and M_X particles is given by:

$$G_x = \frac{\hbar^2}{M_e M_x^2 c T_o B}, \text{ which we put equal to } G_n \text{ or } 6.673 \times 10^{-8}$$

But by (21) $\qquad G_e = 1.176 \times 10^{-3}$

The mass of the electron, M_e, is 0.511 MeV/c^2, which enables us to calculate M_X by using the relationship

$$M_x = M_e\sqrt{G_x/G_n}$$

from which we calculate that M_x has the value $\frac{135.66}{2} M_eV/c^2$.

This value can be compared immediately with the masses of the uncharged and charged pion, which are 134.96 and 139.6 MeV/c^2 respectively.

Such agreement may be simple coincidence. Alternatively, it may be an indicator that, at least as far as gravitational effects are concerned, matter should be viewed as a compound of electrons and pions. (The missing factor of two might be accounted for by error in one of the assumptions given in this paper, or by the fact that pions posses integral rather than half-integral spin).

The existence of the pion was predicted some 50 years ago by Yukawa, who calculated the particle mass that would be needed if an exchange of particles were to account for inter-nucleon forces. That pions have a fundamental role with regard to the structure of matter is, therefore, a concept that has been accepted for half a century. That gravitational attraction may result from such a role could turn out to be a mere extension of Yukawa's predictions.

An Analysis of 900 Rotation Curves of Southern Sky Spiral Galaxies: Are the Dynamics Constrained to Discrete States?

D. F. Roscoe
School of Mathematics,
Sheffield University, Sheffield, S3 7RH, UK.
Email: D.Roscoe@ac.shef.uk
Tel: 0114-2223791, Fax: 0114-2824292

June 12, 1997

Abstract

One of the largest rotation curve data bases of spiral galaxies currently available is that provided by Persic & Salucci (1995; hereafter, PS) which has been derived by them from unreduced rotation curve data of 965 southern sky spirals obtained by Mathewson, Ford & Buchhorn (1992; hereafter, MFB). Of the original sample of 965 galaxies, the observations on 900 were considered by PS to be good enough for rotation curve studies, and the present analysis concerns itself with these 900 rotation curves.

The analysis is performed within the context of the hypothesis that velocity fields within spiral discs can be described by generalized power-laws. Rotation curve data is found to impose an extremely strong and detailed correlation between the free parameters of the power-law model, and this correlation accounts for virtually all the variation in the pivotal diagram. In the process, the analysis reveals completely unexpected structure which indicates that galactic dynamics are constrained to discrete states.
keywords: spiral galaxies, rotation curves

1 Introduction

The following analysis is performed within the context of a prediction arising from a theory of weak-field slow-motion gravitation in material distributions that motions in spiral discs conform to the power-law structure

$$V_{rot} = AR^{\alpha}, \quad V_{rad} = BR^{\alpha}, \quad \alpha \geq -1, \tag{1}$$

where V_{rot} and V_{rad} are the rotational velocity and radial velocity respectively, and for constants A and B; since one of these can be absorbed into the scaling of the problem, it can be assumed that there are only two free parameters, (A, α) say. A crucial result, from the point of view of reconciling the with the observations, is the constraint $\alpha \geq -1$, a result which immediately removes any mystery associated with the existence of 'flat' rotation curves.

The foregoing solution was derived purely from an analysis of the dynamics, with mass-conservation being ignored. However, the additional constraint of mass-conservation can do no more than impose an additional constraint on the space of solutions (1). This amounts to a correlation being imposed on the free-parameters, (A, α), of the model, and it can be shown that the existence of a perfect correlation would imply the model is exact for the physics. However, rotation-curve data is extremely noisy, and so we cannot expect perfect correlations; it follows that, since perfect correlations cannot be expected, the whole argument revolves around the *quality* of any correlations uncovered.

From the point of view of the second part of the following analysis (§6), it is important to know that it was stimulated as a consequence of a trial investigation using a very small independent sample provided in by Rubin, Ford & Thonnard (1980, RFT hereafter). This trial was sufficient to give a quantitative element to the hypothesis concerning the (A, α) correlation and, additionally, gives rise to an hypothesis concerning the distribution of $\ln(A)$. Referring to these hypotheses as H_1 and H_2 respectively, they can be stated as

- H_1 : $\ln(A)$ and α are linearly correlated in a negative sense;
- H_2 : When linear scales are assigned on the basis of the assumption that $H = 50/km/sec/Mpc$ then $\ln(A)$ is constrained to take values which lie within ± 0.15 of an integer or half-integer; this is equivalent to the hypothesis that the allowed values of $\ln(A)$ have a periodic structure, with period approximately 0.5.

2 The Data

The data given by PS is obtained from the raw $H\alpha$ data of MFB by deprojection, folding and cosmological redshift correction. For any given galaxy, the data is presented in the form of estimated rotational velocities plotted against angular displacement from the galaxy's centre; estimated linear scales are not given and no data-smoothing is performed.

The analysis proposed here requires the linear scales of the galaxies in the sample to be defined which, in turn, requires distance estimates of the sample galaxies from our own locality. This information is given in the original MFB paper in the form of a Tulley-Fisher (TF hereafter) distance estimate given in km/sec, and assumes $H = 85km/sec/Mpc$ for the conversion. We have assumed:

- that the MFB method of presenting TF distances in km/sec, including their use of $H = 85km/sec/Mpc$, gives an accurate estimate to the cosmological component of the redshift in the sample galaxies. This assumption is actually central to MFB's analysis since this analysis was primarily designed to give accurate determinations of peculiar velocities in the sample;
- that the criteria by which RFT selected, observed and processed their very much smaller sample ensured relatively accurate determinations of the corresponding cosmological redshifts.

Given these assumptions, then nominal agreement between the RFT and MFB linear scales can be obtained by converting the MFB distances, as quoted in km/sec, to a linear scale using the RFT value of $H = 50km/sec/Mpc$.

An analysis of the distribution of morphological types in the PS data base shows that the great majority of the selected galaxies are of types 3,4,5 and 6, with only two examples of types 0,1,2 and a tail of 31 examples of types 7,8,9. To maximise the homogeneity of the analysed data, the distribution tails - consisting of the morphological types 0,1,2,7,8 & 9 - were omitted, and the remaining 867 galaxies partitioned into the classes $\{3\}$, $\{4,5\}$ and $\{6\}$. These contained, respectively, 306, 177 and 384 galaxies. A separate analysis

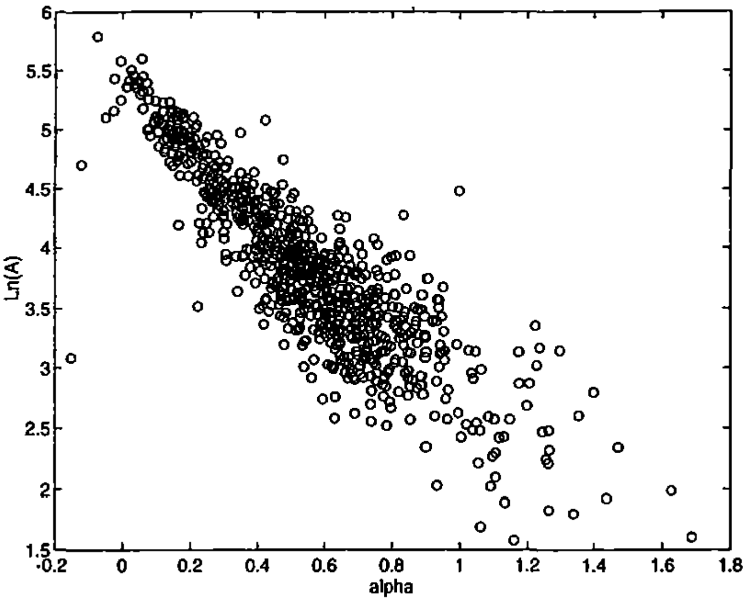


Figure 1: Plot of $\ln(A)$ against α for whole sample

was performed on each of these three partitions.

3 Is There Any Correlation Between A And α?

The basic assumption is simply that rotation velocities behave as $V_{rot} = AR^{\alpha}$ and the discussion of §1 concluded there should be a correlation between A and α. Since the regression constants arising from a linear regression of $\ln(V_{rot})$ on $\ln(R)$ give estimates of $\ln(A)$ and α, our basic analysis performs a linear regression on each of the 867 rotation curves, and records the pair $(\alpha, \ln(A))$ for each galaxy. Fig. 1 gives the scatter plot of $(\alpha, \ln(A))$ for the full sample and shows that there exists an extremely strong negative $(\alpha, \ln(A))$ correlation. The corresponding figures for the individual galaxy type-classes (not shown) are similar in all respects, each occupying similar areas in their respective $(\alpha, \ln(A))$ planes and each displaying the same fan-like structure going from a broad spread of points at the bottom right-hand of the figure to a narrow neck at the top left-hand of the figure. For the remainder of this paper, discussion will be restricted to the type-class {6} (that is, late-types) galaxies, since the conclusions arising here are broadly repeated in the two remaining classes.

4 A Consequence of Linear Correlation Between α and $\ln(A)$

A tentative interpretation of Figure 1 is that

$$\ln(A) = a_0 + b_0\alpha, \tag{2}$$

where a_0 and b_0 are constants which might differ between galaxy type-classes. It is easily shown how this implies that all the rotation curves in any given type-class,

$$\ln(V_{rot}) = \ln(A) + \alpha\ln(R),$$

intersect at the fixed point $(-b_0, a_0)$ in the $(\ln(R), \ln(V))$ plane. If this point is denoted as $(\ln(R_0), \ln(V_0))$, then (2) is more tranparently written as

$$\ln(A) = \ln(V_0) - \alpha\ln(R_0) \tag{3}$$

Whilst the idea of a single intersection point for all rotation curves in the class seems rather extreme, it is unambiguously deduced from the most obvious interpretation of Figure 1. The most direct test of the statement is simply to plot the corresponding rotation curves in the $(\ln(R), \ln(V))$ plane, and to observe their actual behaviour. This is done in figure 2 for the type-class 6 - the plots for the type-classes 3 and 4,5 are similar in all details. The figure gives a very clear sense of convergence in a relatively small region of the $(\ln(R), \ln(V))$ plane. As a secondary means of illustrating this apparent convergence, we have calculated the coordinates of intersection between all possible pairs of rotation curves in the plane and have plotted the

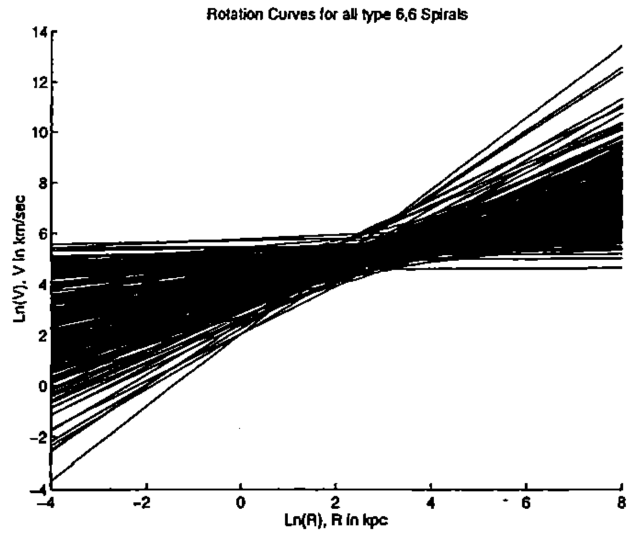


Figure 2: All Type 6 Rotation Curves in $(\ln(R), \ln(V))$ Plane

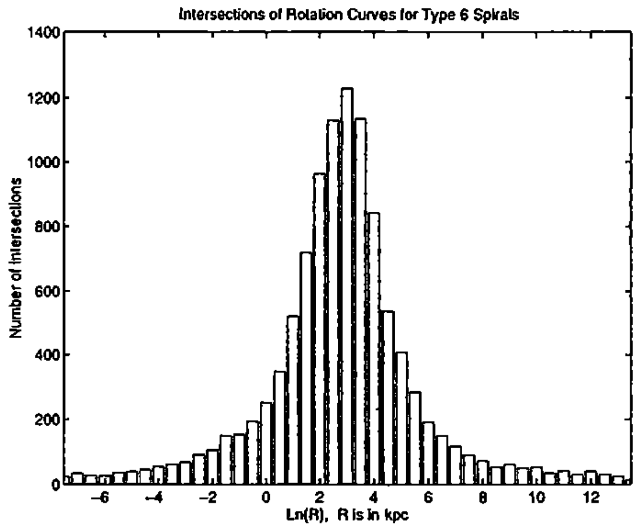


Figure 3: Frequency of Intersections on $\ln(R)$ axis

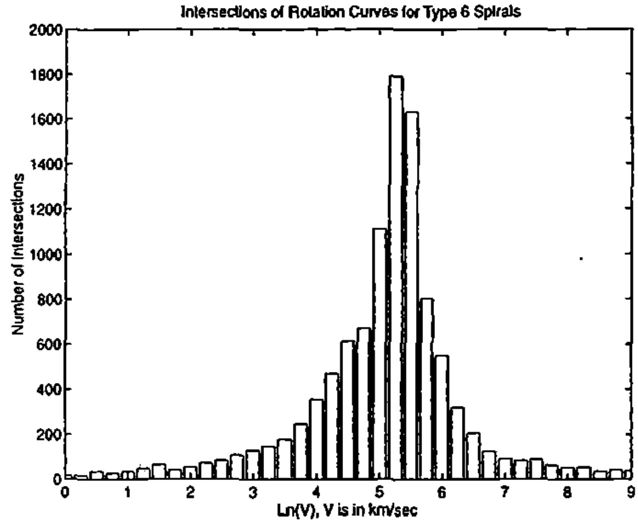


Figure 4: Frequency of Intersections on $\ln(V)$ axis

frequencies of intersection for type-class 6 along the $\ln(R)$ and $\ln(V)$ axes respectively in Figures 3 & 4. It is clear from these figures that there is a very sharp peak of intersection points at $(\ln(R), \ln(V)) \approx (3, 5.2)$. The results for the type-classes 3 and 4,5 are similar.

A powerful geometric test of the convergence statement can be formed from the realization that, if the rotation curves really do converge on a single point in the $(\ln(R), \ln(V))$ plane, then all of the rotation curves will *transform into each other* under rotations about the convergence point, $(\ln(R_0), \ln(V_0))$, in this plane; that is, the individual rotation curves in the set of all rotation curves associated with a given $(\ln(R_0), \ln(V_0))$ are equivalent to within a rotation about this point in the $(\ln(R), \ln(V))$ plane.

5 Testing H_1: Is the Linear Assumption Reasonable?

Suppose that, as Figure 1 suggests, all the rotation curves in a given type-class pass through $(\ln(R_0), \ln(V_0))$ in the $(\ln(R), \ln(V))$ plane. Then an arbitrarily chosen straight line passing through this point can be

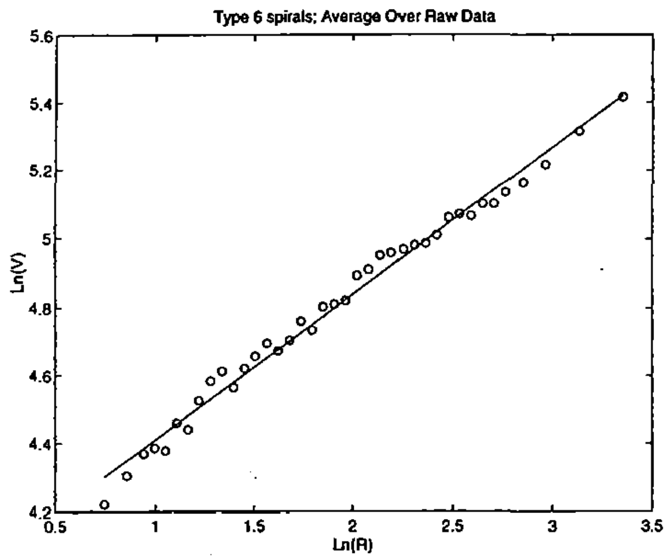


Figure 5: Type 6 spirals; Average over raw data

defined as a standard 'reference line' into which every rotation curve in the class can be transformed by a simple bulk-rotation about $(\ln(R_0), \ln(V_0))$. Since, according to this idea, the rotation curves are reduced to equivalence by the rotation, then the process of forming an 'average rotation curve' from the set of rotated such curves should greatly reduce the internal noise associated with the individual rotation curves, and we would expect the resulting average curve to be a very close fit to the standard reference line, referred to above.

For the analysis, the galaxies are grouped into type-classes $\{6\}$, $\{4, 5\}$ and $\{3\}$ and, for brevity, we only give the results for type-class $\{6\}$ here - the results for the other type-classes are similar. The fixed point in the $(\ln(R), \ln(V))$ plane is estimated (using a minimisation procedure not described here) as (3.615,5.416). Figure 5 shows the plot which arises from averaging all rotation curves in each of the type-classes *directly* in their raw state (that is, without rotating about the fixed point $(\ln(R_0), \ln(V_0))$). Figures 6 shows the plot which arises when, for type-class $\{6\}$, the rotation curves are rotated about the fixed point so that

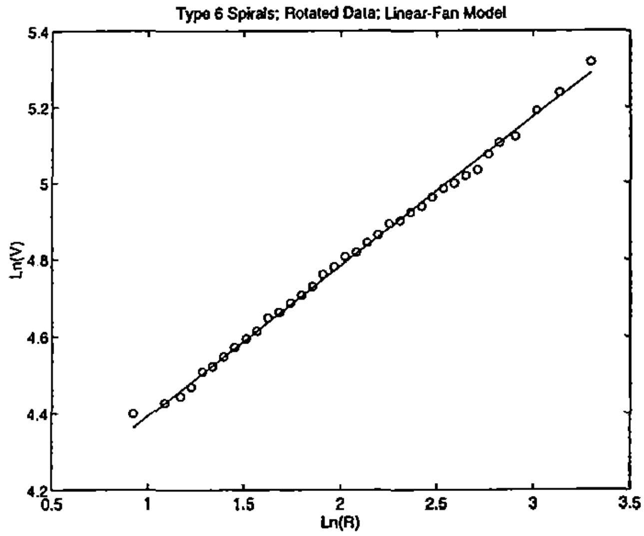


Figure 6: Type 6 spirals; Averaged rotated data testing H_1

they coincide *in a least-square sense* with the reference line defined to pass through $(\ln(R_0), \ln(V_0))$, and fixed (arbitrarily) so that $\ln(A) = 4$, and the results averaged over all the rotation curves in the class. The scatter present in Figure 5 is virtually eliminated, providing the strongest possible evidence for the idea of the equivalence of rotation curves with respect to rotations about particular fixed points in the $(\ln(R), \ln(V))$ plane.

6 Testing H2: An Analysis of the Ln(A) Distribution

A preliminary analysis of the RFT data, mentioned in §1, gave rise to the hypothesis H_2 that the $\ln(A)$ parameter was constrained to lie within ±0.15 of an integer or half-integer. On the original RFT data, this was observed to occur at a rate which had a by-chance probability of approximately 0.002 - which is sufficiently small to make it natural to pose the question 'is this an artifact of the RFT data, or does it reflect some underlying constraint imposed on the physics of rotation curves?' We have analysed the distribution

Table 1: Distance Scale = TF3			
Galaxy Types	Sample Size	Number Of Hits	Probability of Chance Event
0..9	900	586	0.909×10^{-3}
0..5	485	294	0.410
6..9	415	292	0.731×10^{-5}

of the $\ln(A)$ values obtained from the PS data base, and present the results in the following.

The specific question posed was *how many of the* $\ln(A)$ *values lie within* ± 0.15 *of either an integer or half-integer value?* Because these ranges cover 60% of the real line, they are referred to as 60% bins in the following. The range of ± 0.15 was chosen simply because this is the question raised by the RFT data. The details of the data-reduction used in this analysis are not given here but, briefly, they amount to a *prior-decided* means of minimizing the effect of the nuclear bulge on the rotation-curve calculations. The results of the analysis are condensed into Table 1 and, in this table, a 'Hit' is defined to be when a particular $\ln(A)$ lies within one of the 60% bins. From the table, we see that when the analysis is performed over the whole sample, then the number of hits is 586 from a possible 900 trials; the odds of this being a chance event are about 1:1100. As with the RFT data, this is an interestingly small probability, but is not sufficiently conclusive. However, when the data is partitioned into two classes consisting of the type-classes $\{0,1,2,3,4,5\}$ and $\{6,7,8,9\}$ a totally different picture emerges: specifically, on the class $\{0,1,2,3,4,5\}$ there are 294 hits out of 485 trials. The odds of this occurring by chance are about 4:10; this is exactly what is to be expected if there is no effect. By contrast, on the class $\{6,7,8,9\}$, there are 292 hits out of 415 trials. The odds of this occurring by chance are about 1:137,000 which is *extremely* small. There remains the possibility that there is a significant effect on the $\{0,1,2,3,4,5\}$ data which is masked either by being phase-shifted with respect to the effect on the $\{6,7,8,9\}$ data, or by having a different period.

This possibility was tested by running both sets of data through a power spectrum analysis program which searches automatically for periodicities at arbitrary phases. This secondary analysis confirmed the positive results on the late-type spiral data, and found no evidence of any effect of the early-type spiral data.

There are two broad questions which immediately arise:

- Firstly, the analysis concerns the distribution of the $\ln(A)$ values. However, A arises originally in the power law $V_{rot} = AR^{\alpha}$ from which it is clear that its value for any given rotation curve depends on the definition of the linear scale - which, here, has been defined on the basis of the assumption that $H = 50km/sec/Mpc$. This fact makes it highly implausible that $\ln(A)$ is constrained to take on exactly integer or half-integer values.
- Secondly, if the effect on the $\{6,7,8,9\}$ class is real, why does it appear to be absent from the $\{0,1,2,3,4,5\}$ class?

In answer to the first question, it is sufficient to note that a 'Hit' is defined to occur when a given $\ln(A)$ value is within the 60% bin. The 60% bin, by definition, occupies 60% of the real line so that, although the two parts of the bin are centred on integer and half-integer values respectively, there is plenty of room for $\ln(A)$ to have a preference for discrete values without these necessarily being exactly integer or half-integer values. In answer to the second question, it is sufficient to hypothesise that the effect does exist for early-type spirals, but with a different period. The effect was spotted in the late-type spirals purely because of a fortuitous choice of the Hubble coefficient in the original analysis of Rubin et al.

If $\ln(A)$ is constrained to occupy discrete values, as the odds against a chance result of 1:137,000 appear to indicate, then, since $\ln(A)$ and α are known to be extremely strongly correlated, α must also be constrained to occupy discrete values. However, since the range of α is much less than that of $\ln(A)$, it is to be expected that any α-periodicity will be much less than the 0.5 of $\ln(A)$; this means that the 60%-binning method used to test the explicit hypothesis H_2 is not appropriate for testing α. Correspondingly, we used the method employed for the secondary analysis of H_2, the power-spectrum analysis program, as the primary means of searching for α-periodicities. This indicated good evidence for a period in α of 0.245, with a phase-shift of about 0.14 from the origin; noting that the predicted periodicity is about one-half of the $\ln(A)$ period, an alternative test became possible: specifically, if the α-values were multiplied by two (giving a period of 0.49), and shifted by 0.28 (double phase), then a significant effect should be detectable with the 60%-binning analysis. This analysis then recorded 286 hits from 415 trials, which has a by-chance probability of $0.104 \times 10^{-3} \equiv 1:9600$. However, it is to be noted that the $\ln(A)$ result was a prior quantitative prediction made on RFT data, and is therefore to be given a much greater weight. The latter results on the α data can only be considered as confirmatory, and used as a basis of formulating a specific hypothesis to be tested against independent data.

The wider implications of the present positive result (1:137,000 for a chance happening) obtained for the type-class $\{6, 7, 8, 9\}$ are that galactic dynamics are constrained to occupy discrete states.

7 Conclusions

The most interesting part of the forgoing analysis from the point of view of the ANPA programme is that of §6, which appears to strongly suggest that the dynamics of large scale structures exhibit discrete structure. The purpose of the earlier parts of the analysis is simply to emphasize that the power-law model reveals substantial other structure in the data. One can use the metaphor of spectacles: The wrong spectacles for the short sighted man add nothing to what he sees; but the right spectacles reveal a whole new world. The power-law model has revealed a considerable structure in rotation curve data and has led to the further hypothesis that - maybe - regular discreteness (quantization in some form) exists on very large scales.

References

[1] Mathewson, D.S., Ford, V.L., Buchhorn, M. 1992 A Southern Sky Survey of the Peculiar Velocities of 1355 Spiral Galaxies *Astrophys J. Supp.* **81** 413-659.

[2] Persic, M., Salucci, P., 1995 Rotation Curves of 967 Spiral Galaxies *Astrophys. J. Supp.* **99** 501-541.

[3] Rubin, V.C., Ford, W.K., Thonnard, N. 1980 Rotational Properties of 21 Sc Galaxies With A Large Range Of Luminosities And Radii From NGC 4605 (R=4kpc) To UGC 2885 (R=122kpc) *Astrophys. J.* **238** 471-487.

A REPRESENTATIONALIST THEORY OF CONSCIOUSNESS

Adam Parker-Rhodes, 5 Muswell Hill Rd., London N10 3JB.
gingco@dial.pipex.com

ABSTRACT.

Modern ideas about the nature of consciousness centre on the idea of representation. This paper outlines a hypothetical account of its evolution. Von Neumann correctly predicted on logical grounds that living things as self-reproducing systems contain structures that are coded representations of the mechanisms whose function is to read and replicate those coded representations. The DNA-protein system has this form. This implies a necessary link between structural representation and the evolution of self-organising systems. During evolution, the living things that survive and reproduce best are those that have most efficient means of self-maintenance. The resulting meta-stable systems can be described as having a distributed or 'functional' self-representation containing information about the form of their own stable state. The emergence of motile organisms leads to the evolution of neural, sensory-motor feed-back loops. Animals that posses these reflexes are described here as **pre-sentient**. Simple neural connections evolve into specialised central nervous systems with extensive neural-nets. Such organisms can learn adaptively and seek information pre-emptively. The sensory-motor information they gather is still in the form of functional rather than structural or depictive representation. They show 'curiosity' and it is argued that they should be thought of **sentient**. At a later stage some animals, notably mammals, are able to reduce the risks of trial-and-error learning by testing out actions using internal sensory-motor simulation. It is suggested that dreaming is linked to this mode of cognitive function. Such simulation does amount to a form of depictive representation and creatures using it are said to have **percipient** consciousness. Humans, as a result of their highly cross-linked cerebral cortices, acquire the ability to imitate and so cognitively simulate each other's sensory-motor activity. This capacity for shared simulation is called **pre-sapient** and makes possible the cultural development of languages. A being with language has two autonomous, but interactive and mutually regulating systems of representation. This leads to a major advance in cognitive capacity and what we experience as **sapient** consciousness. Postmodernist writers are suggesting that the unlimited multiplication of systems of representation is leading us into a new relation to culture and language. Perhaps this should be called **post-sapience**. To understand this, and other problems such as those of quantum mechanics, we need to develop a general logical theory of representational systems.

1. INTRODUCTION

Theories of consciousness have recently become a rather popular genre. Most accounts have tended to be unsatisfactory because they tend to treat consciousness as something unitary, requiring a single cause to explain it. If this is so, it suggests that consciousness is a property of all matter that cannot be analysed further. If consciousness is an unanalysable phenomena then, like the more substantial idea, 'Mind', it seems to be too mysterious to be grasped. If the problem is treated in this way then MacGinn[1] is right in saying human consciousness can never understand itself. Another, even more serious difficulty with recent accounts is that they have made no serious attempt to define what is meant by 'consciousness'. The tacit assumption here is that we are all already familiar with it by direct personal acquaintance. The result in practice, has been that each explanation seems to be talking about something different. Some accounts of consciousness are so far removed from naturalistic contexts that they even leave open the question of what *use* consciousness is to the organisms that posses it.

I shall treat the problem in an evolutionary fashion, showing how human consciousness could have emerged from inanimate matter by a series of distinct steps, each proceeding from the one before. This approach has a number of advantages. If it works, it provides an account that is both materialist, and so respectably scientific, and a plausible *story* of consciousness. More importantly, by showing it emerging in a series of steps it distinguishes different biological and functional levels of consciousness. This provides a series of distinct concepts for talking about different kinds of consciousness, helping to overcome the problem of definitions mentioned above.

2. CONSCIOUSNESS CONCEIVED AS REPRESENTATION

An explanation of consciousness must describe our habitual use of First Person experience and patterns of reflection in objective Third Person terms. It must do so in a way that makes sense of our existing preconceptions about ourselves. Whatever kind of 'scientific truth' it achieves, it will be of no real use to us if it does not tally with our common introspective experience and the insights of existing 'folk psychology', which are largely Cartesian. Its terms are the dualistic ones of Mind and Matter, Subject and Object, Appearance and Reality, fact and value, etc.; terms that are all derived from the fundamental dualism implicit in the concept of representation. Mind, the representing, and matter the represented cannot, according to this conception exist in the same logical space. In conventional representationalist terms, knowledge is a representation of the world; perception provides pictures of 'external reality'; memory is an inner store of the representations of past events; communication is the passing of coded representations from one mind to another. One is right to be suspicious of these simplistic notions, and a deeper intuition tells us that they are all in some sense false. They do none-the-less serve us well enough as a natural attitude and a normal way of talking. Representationalism has, after all been inseparable from the project of Western science, and any schema that is as powerful as that must express some fairly profound kind of truth. The story of consciousness, as the Modern Mind understands it, has to be the story of the evolution of representation.

The essence of Western Representationalism, a position that Wittgenstein later rejected, is well illustrated by a few lines from the*Tractatus*:

4. A thought is a proposition with a sense.

4.0. A proposition is a picture of reality. A proposition is a model of reality as we imagine it.

4.014. A gramophone record, the musical idea, the written notes and the sound waves all stand to one another in the same internal relation of depicting that holds between language and the world. They are all constructed according to a common logical pattern.

In this paper I shall try to give a scientific account of consciousness, though finally, I do not believe that reductive and naturalistic explanations of this kind will ever be able to provide more than temporary answers. All systematically self-reflexive thought leads to paradox and the Representationalist theory of representation is finally just as self-contradictory as its intellectual complement, Hermeneutic Relativism. Elsewhere I shall deal with the deeper matter of how this account, like the *Tractatus*, is finally self-deconstructing. We cannot today describe the workings of our minds except in terms and concepts derived from the idea of representation. There are those, postmodernists, pragmatists, relativists etc., who seek to reject this image of the 'Mind as the Mirror of Nature' , but the difficulties of presenting any other model seem overwhelming. The need to use the terms 'Image', 'model' at once shows this. How can there be a philosophy that does not aim to describe how things are?

To be of use, an evolutionary theory of consciousness has to be a history of the evolution of representation. It is argued by some (*e.g.* Rorty[2]) that representation is not a 'natural kind' - *i.e.* that it does not exist except as a concept derived from human practices: pictures, writing etc. Others hold that all representation is inseparable from 'intentionality' (in the philosophical sense) and so from consciousness itself. If either of these were true, any attempt to derive consciousness from natural representation would be hopeless. I shall show that this is not so.

3. EVOLUTION

It is often suggested that evolution is an entirely blind and random process without direction or progress. This is a misleading attitude and one that makes it particularly hard to understand how consciousness could have emerged from inanimate matter. One way in which evolution advances is by the organised aggregation of elements to produce an organism of a higher order of complexity. To deny that this is 'progress' is like saying that bricks, houses and cities are all objects of the same type. The way in which evolution is directional also shows itself is cases of convergent evolution. The emergence of consciousness was doubtless random and accidental in its details, but its final appearance was as no less pre-determined than that of legs, wings and eyes, and other functional structures that evolution has had time to create in may different ways.

Each stage of evolution changes the world and leads, as it were dialectically, to its logical successor. This is in a sense pre-ordained, but it is not by a simple mechanical determinism, because it could have happened in any number of different ways or not at all. Nor need we assume that it is a goal-directed process based on some priore image or conception. The interaction of evolving organisms produced an ever more complex and varied environment in which survival depends on a corresponding increase in the internal complexity of organisms. There is an inner necessity to the process that leads from the structural representation intrinsic to all self-reproductive systems, to the doubly self-reflexive cognitive systems operating in brains like ours. If we had not become conscious, then given the chance, some other species would have.

4. INSENTIENT REPRESENTATION

In the early '40s von Neumann[3], investigating the idea of self-replication, concluded that if any system is to be properly reproductive, *i.e.* not simply a process of auto-catalysis but something genuinely able to build copies of itself from simple components, then it has to contain a representation of itself. The system must have two distinct parts: a general constructor able, in principle, to build many different kinds of things, and a representation of that general constructor; a blue-print or plan that the constructive apparatus could follow in order to make copies of itself. It is logically necessary that the plan, an inner representation of the system, must be read in two ways: firstly by translation into the form of the constructor it represents, and secondly by direct replication to provide a fresh copy of the plan to be used by the new general constructor. Any organisation that does not have this general form, leads to logical paradoxes of infinite regress in which representations contain representations of themselves. A few years after von Neumann's death, the discovery of the mechanism of the DNA code and cellular apparatus for constructing proteins, proved that his deduction was entirely correct.

We can sum up von Neumann's deduction by the relation:

Reproduction $\Rightarrow$ Replication + Representation.

Though the implications of this for other systems of representation are not obvious, it does allow us to say with certainty that if one accepts the ordinary account of evolution, then despite Rorty *et al.*, representation *is* a logically necessary part of Nature prior to all human conscious and culture. There is in fact a close structural analogy between the DNA or RNA-protein code and the codes which translate (or *de-scribe*) phonetic script into spoken words. It is surely remarkably that the former should be the key to biological life and the latter the paradigmatic example of representation in the WesternRepresentationalist tradition.

In what follows I shall take these two cases as my first models for and paradigm cases of, representation in general. It is easy to see a general similarity between them and such easily understood examples like pictures, maps and various representational systems in science, as well natural examples such as the language of bees. In all these cases there are two (or more) structurally different and spatially distinct media each embodying its own kind of structural elements, and a systematic code whereby elements in one medium map to, and correspond to, those in the other.

This paradigm case gives an informal idea of what it is for something to be a representation; necessary conditions but not sufficient. Not every case of structural parallelism amounts to real representation, and considerable care must be taken over this. Representation always occurs as a functional relation within some more inclusive system: a cell, a hive, a culture or a consciousness, and given this last there is always the problem of whether the representation is intrinsic and 'real' or whether it is only created in its relation to an observing mind. As we have seen, one cannot sustain the argument that *all* representation links back to that in human minds or culture, but there are many examples where this does happen. Each case must be analysed in its own context. (If we had agreed definitions or a general theory of representation then this paper could have been much shorter.)

DNA only represents protein by virtue of the roles within a cell. It is a representation because of its functional relation to replication and reproduction in a larger system that is logically and ontologically prior to it. Reproduction or self-replication is the defining attribute of life. It is the primary and more concrete process in which representation is an internal and secondary echo. Wherever we look, being a representation is not a property of anything *in itself.* It is something that arises, as Wittgenstein said, from being in an *internal* relation of representing. (If things are 'internally' related then each is what it is only by virtue of the fact that they do stand in that relation.) To illustrate this; there may have existed in the time before true life started, a chemical environment in which randomly structured molecules of RNA (or perhaps clay crystals) occasionally catalysed the production of strings of amino acids with corresponding structures. If this was all that happened we would have to regard the latter as no more than mere *imprints* of the former. As we shall see several times, imprints and true representations are different kinds of things. For the conscious hunter, a foot-print can be said to 'represent' the passing of prey. In itself it is just a random mark.

In the absence of an accepted formal theory of representation we will have to proceed in this kind of *ad hoc* way. I shall attempt to give a coherent but informal account of the different kinds of representation that arise, by comparing and contrasting each case with earlier paradigm cases (*e.g.* DNA-protein or phonetic script).

The evolution of life begins with the random reactions of organic molecules in a far-from-equilibrium chemical environment. The frequency of each 'species' depends only on chemical reactions proceeding in a simple linear way, until the appearance of molecules, or small sets of them, with a capacity for catalysing their own creation. This is the first step to self-replication and full reproduction in von Neumann's sense. A purely RNA system probably evolved into one in which RNA mediated in DNA-protein systems of ever increasing complexity. The frequency of different forms (*e.g.* molecular associations or organisms) in any dynamic population depends on their relative rates of creation and destruction. In simple systems the rate of creation will depend only on the basic chemical and physical properties of the forms, but with the first occurrence of self-reproduction this changes radically. Frequencies of different forms now depend on the nature of the self-representations that the forms contain. Their rates of creation depend on their effectiveness as replicators. The rates of their destruction is inversely related to their over-all stability, first chemical and physical stability and later the meta-stability of homeostatic systems.

The survival of different life-forms depends then, on two principle factors: replication and self-maintenance. These two aspects of self-organisation exhibit a kind of deep symmetry. Multiplication and survival are both kinds of reproduction, one a reproduction in space, the other a reproduction in time. Spatial reproduction creates multiple copies of a structure; temporal reproduction repeatedly recreates certain configurations of the same structure by detecting changes and returning the altered structure to its ideal form. Self-replication depends structural representation and self-regulation on what I will call 'functional representation'. Reproduction in space is associated with the unlimited expansion of a positive feed-back process, reproduction in time with the continual self-limitation of processes that are controlled by negative feed-back . It is as though the process of life was expressing a balance between the Yin and Yang of self-regulation and self-replication - each with its own principle of representation.

However, despite this nice conceptual symmetry it is actually far from clear just what kind of thing this so-called 'functional representation' is. The idea of error-correction and so of a return to a 'true' state, strongly suggests the presence of representation, and so does the behaviour of systems that act as though they were directed to some predetermined or remembered goal. Yet cybernetic systems have nothing in them that is at all like the coding of structures that we see with DNA.

Ordinary self-regulating systems, are *not* 'depictive' representations, and, as we shall see, brains do *not* contain the kinds of depictive representations that we think of as occurring in human consciousness. We form ideas and images in our minds that seem to us to be as clearly representational as real maps and pictures. Verbal statements are representations of what we think and can be consistently interpreted to yield other even more explicit kinds of representation. Our reflective consciousness and self-consciousness seem inseparable from structural or depictive representation, yet none of this is in any way like the obscure and distributed functional 'representations' whose role is inferred from the behaviour of cybernetic machines or the physiological systems of our brains and bodies. Thus at this stage in the story of consciousness the question is that of understanding how mechanisms in which representation is present only as an invisible and implicit function can give rise to beings that have our kind of explicit and communicable representations of cognitive activity. The question is not whether the distributed self-representation of physiology counts as representation at all, but how its evolution could give rise to the real structural representation that our undeniably Cartesian experience demands. How could simple, mechanical, self-organising systems have evolved into complex organisms that have full representation in the depictive sense, and give rise to what we experience as consciousness?

5. PRE-SENTIENCE

Plants, fungi and singled organisms are truly insentient and have to be regarded a little more than complex chemical systems; organised molecular factories dedicated to producing more of the same. The first step on the evolutionary path from these, towards consciousness, occurs with the appearance of what I shall call *pre-sentience.* This emerges with the development of organisms with sensory-motor systems,*i.e.* animals.

The simplest motile organisms, often with flagella, move randomly. When variables vital for their survival, such as temperature or the level of oxygen, depart too far from the ideal, they start to move and carry on doing so until conditions improve. Even far more advanced animals like pupa in the Mexican jumping bean 'seek' the shade by the same kind of random trial-and-error (or better, hunt-and-stick) movement. Such organisms need no sense organs or nervous system. The internal temperature, nutrient levels, etc. are a direct measure of external levels and corrective 'action' is turned on and off directly by the state of these vital variables. The organism's metabolism itself provides the only 'sensors' it has. Pre-sentience occurs when this kind of function gives way to true sensory-motor systems. With these, movement in controlled directly by input from specialised senses and indirectly by the 'setting' placed on the senses by the state of vital variables.

The properly pre-sentient organism has at least some simple senses, light or vibration receptors for example, and neural links between these and motor cells of some kind. These constitute simple 'hard-wired' reflexes that lead to, for example, movement towards or away from light, etc. A single reflex arc of two or three neurones is a pure physical expression of the functional principle implicit in any stabilising feed-back loop. The ordinary household thermostat is just such as system. In pre-sentient organisms links between sensory in-put and motor out-put is more or less direct. There maybe some simple comparisons between in-puts, and some internal feed-back along the way, but there is nothing that amounts to a dedicated and flexible information processing system. Everything is hard-wired and ready to run automatically, subject only to in-put and some modulation from the state of the vital variables (as for example when one basic drive takes priority over another.) Motor out-put and sensory in-put form a single system; feed-forward and feed-back are integrated into behaviour that is adaptive but automatic.

This is not to say though that pre-sentient behaviour may not be complex. Starting from the simplest of self-protective and appetitive reflexes, evolution can go on to produce very subtle behaviour from quite simple neural mechanisms. Some ants for example, set off foraging, and having found some food, immediately return home across the sand by the shortest route, as though they had an inner map. This kind of behaviour can however be mimicked by rather simple servo-mechanisms. The Sung Chinese invented the South Pointing Carriage: a cart with an indicator on top that continues to point south irrespective of the direction the cart travels or changes course. Its 'sense of direction' comes only from a system of gears not a cognitive map. Ants manage to navigate using a nervous system with only a few hundred neurones in all. Wasps and bees too, seem to have maps, memories and the ability to calculate, but is not hard to see how such 'skills' can in principle be put together from reflexes, feed-back and the setting of a few neural 'registers'. There is a kind of toy car, controlled by a pair of photo-cells that, like a moth, will follow a moving light. Its behaviour looks eerily intentional. Pre-sentient organisms can display goal directed behaviour of a high order but we do not need to think of it as involving intention or representation. Simple mechanical devices can do the same thing without that.

So long as organisms operate with only a few hard-wired reflexes that link in-put more-or-less directly to out-put, then, unless perhaps one is a Panpsychist, there is no place for consciousness to occur. The first intimations consciousness emerge at a higher and less strictly determinate level of functional organisation than the simply mechanical. To have genuine cognition and the beginnings of what we can call sentience, an animal must not just respond to sensory in-put but actually go out and gather information for its own sake.

6. SENTIENCE

The evolution of pre-sentient organisms with their basic sensory-motor systems lead to a more complex and varied world of predators and prey, an environment that demands a greater ability to predict and control external events. It selects for increasingly complex neural networks and centralised nervous systems. Brains still process information and organise behaviour according to the cybernetic principles of self-regulation, but they do so at a higher, less determinate level than that of simple reflex and feed-back loop. The insentient organism's trial-and-error goal directed behaviour is indeterminate in a limited external sense. A complex nervous system, however has many different internal ways of achieving the same external result and learns (or selects and in a loose sense 'evolves') better ways of doing things. Brains are innately structured to develop and maintain the organisation of adaptive sensory-motor *information* systems. Where previously there were simple sensory stimuli now there are net-works that analysis in-put and discriminate between complex forms and patterns, *i.e.* perceptual systems. Simple reflex responses to in-put give way to the construction of hierarchically organised behaviour with external and internal feed-back on several levels. With the evolutionary emergence of brains comes a new level of neural organisation, a kind of cognitive function: *sentience*. I shall argue that what is described here as sentience can be reasonably thought of as the first level of true consciousness.

When animals develop complex and centralised nervous systems they acquire a completely new order and level of stability and homeostasis. Pre-sentient organisms use their sensory-motor systems to control changes to their internal variables, but this process depended only on specialised, hard-wired structures serving pre-existent needs. They control their biological state but their ability to predict and control their own sensory in-put is still limited and secondary. This is not so with a true central nervous system. To illustrate this, think of a scene, familiar from wild-life films, in which a bird such as an egret sits on a moving branch. The wind moves the branch but the bird's body and neck exactly compensates for this and it keeps its head and gaze absolutely steady and fixed on its prey. The animal's sensory-motor system is working to stabilise the input of information to the brain itself. To take another example, think of the intense orientation reaction of an animal that is attending, with all senses attuned and ready, to something startling, interesting or unfamiliar. Here is an animal seeking to control and augment its own perceptual in-put. These are the characteristics of sentience, and it arrives with a shift from stabilising the body's physical and chemical functions directly to that of stabilising them by means of stabilising the *internal* activity of the nervous system. The sensory-motor information that its cognitive activity processes is not simply information about the environment, nor is it simply the exercise of a skilled habit or behaviour. In real cognitive processing the action and perception are inseparable, a complex of knowing how to perceive and the re-cognition of contexts for action.

In pre-sentient organisms, sensory in-put controls motor out-put in simple servo-mechanisms that stabilise the body's internal organic states. There comes a stage however where organisms work to stabilise their sensory in-put and the state of their information processing systems. This is no longer a case of reflex responses to sensation but of organising sensory-motor information; the active processing of intentional perception. There is of course much more to this 'sentience' than a steady gaze and an inquisitive orientation reaction, but before going on to this we should look further at the idea that perception is a form of stability in a homeostatic information system.

The brain evolves as a highly dynamic and flexible protective system between the organism's basic physical and chemical functions and an otherwise unpredictable environment. Having a central nervous system establishes an autonomous buffer zone of information processing between inner physical variations and outer ones. Where before there was a direct and determinate causal continuity between inner and outer, the body and the environment, there is now a clear functional separation. The brain stabilises bodily functions by way of stabilising its own nervous activity. The sentient animal is one that controls its senses rather than being controlled by them.

In our own experience this shows itself in the normal stability of our perceptual world. We move our heads and the room stays still, obviously, but in order for us to experience this the brain must do an enormous amount of information processing or cognitive work. From the cerebral cortex down to the deeper and older parts of the brain, eyes, muscles, balance organs and proprioception are linked in an automatic control system that exactly compensates for body movements to maintain produce stable perceptual states. Every act of attention, of perceptual orientation, involves the complex activity of a hierarchy of feed-back systems. The development of a nervous system leads to the emergence of a new kind of *informational* autonomy with respect to the environment. The pre-sentient organism is like the still centre of a changing world, the sentient animal is the dynamic centre of a world held still.

Why this should be an advantage is obvious enough. If an animal is to survive in a world of predators and prey, its activity must be flexible enough to deal not only with the predictable changes of the inanimate world but also with the way it is itself responded to by other animals. (Note again the *dialectical* nature of evolution.) It must be able to distinguish between those changes in its sensory in-put that are due to external events and those which are due to its own actions and changes in position; between real happenings and their fixed background. This complex task is handled by nervous system with a hierarchy of re-entrant outputs and the feed-back control of in-puts. The animal is born with, or soon grows, the basic neural architecture of such a system, but it still has to learn about the world. It still has to stock its brain the sensory-motor information needed to survive. It must learn *how* to use its innate perceptual capacities. The pre-sentient organism only learns to respond to stimuli within a fixed array of reflexes and registers, but the sentient animal has neural media in which it can build new adaptive sensory-motor systems. This brings us to what is virtually the defining property of sentience: curiosity.

7. CURIOSITY

In pre-sentient organisms the highest level of interaction with the environment stops with finding the conditions it needs to grow and reproduce. The sentient animal goes further; it stocks its body with nutrients and its brain with information, and it does it do this just in passing, in the process of acquiring conditioned reflexes. Sentient animals seek information for its own sake, collecting it pre-emptively and in advance of an immediate need. They exhibits curiosity; exploring, perceiving, and in their own way, seeking to understand their environment. Pre-sentient organisms have precisely pre-wired nervous systems perfectly adapted for a particular way of life. More advanced organism have such functions too, but they also have, to an ever increasing degree, the ability to function in a flexible and open-ended way, as general purpose information processors. They possess complex neural net-works that cannot perform their proper purpose until they have been extensively trained and allowed to function and inter-act with one another in real life situations. Learning about the environment and learning how to use the brain are two aspects of a single task; that of establishing an effective working relation with the world. In more advanced organisms, we see this in a particularly clear form as play. When a kitten (which is actually more than merely sentient) plays with a cotton-reel or an injured mouse it is learning to anticipate the behaviour of things, learning what to expect and how not to be surprised, learning to anticipate sensory in-put and organising the nervous system so that it can later respond to situations after a minimal amount of neural information processing or cognitive work. Doing this kind of home-work prepares it for the real world and sets up a cognitive system with a maximum of information processing capacity free to deal with whatever cannot be predicted in advance.

The static, sentient world that stands still as we move around is the product of immensely complex cybernetic work by the nervous system. As the phenomenologists have described, when we look around that dynamically constituted world, it is marked out not only by the perceptual acts that shift our attention around in it, but also and equally validly, experienced in terms of the potential acts we could perform in it. The given 'fixed' perceptual world is just the static surface of the dynamic life-world in which we have to survive; a world in which distances are measured out in strides, leaps and stretches; the world in which we know precisely the reach to what lies at hand; in which we can run for cover or leap to safety without putting a foot wrong - a world measured out in the actions of the body and tacitly permeated throughout with the pattern of potential acts. Our real sentience of space is not that of the detached gaze - but of the muscular and sensory motor feelings of what can and cant be done.

The organism inter-acts with its world, gathering information and experience pre-emptively; finding its way around, discovering what is where and how things behave. Through interaction it learns to integrate its instincts and its perceptual and performatory capacities into the skills required by different situations. As it does this it build up the sensory-motor information structures which are its world

8. THE REMEMBERED PRESENT

As an animal learns its way around its environment so it gains more and more perceptual skill and knowledge. It learns when and where to look, listen and sniff out the cues and clues it needs to get things done. Every activity involves perceptual activity, and in a given environment, the animal's home range, the same perceptual acts will be used repeatedly in doing different things. They form a stable system of invariants which an efficient nervous system will naturally store economically, as what we may think of as the animals 'perceptual world', a neurally embodied system of perceptual habits of attention and focusing that forms a stable background to routine habits and novel actions alike. If anything in the animals territory changes it will be carefully explored, internalised and made familiar. By a general 'purposeless' or playful interaction with the new object or phenomenon the animal fits it and its potential behaviour into the structure of a continually updated and renewed perceptual world. Curiosity turns interesting novelties into ignorable background. When a crucial emergency arises it is not overwhelmed by disorganised sensory in-put because, having done it perceptual preparation in advance, it already knows how to get at the sensory information in needs to put together effective actions.

We all know what it is like to be aware of a familiar back-ground. Edelman's[4] phrase 'the remembered present' sums it up well. We take in a familiar setting at a glance, knowing just where and what everything is. Its not that we see everything in detail, only that we know exactly *how* to see it if we should need to. Glancing around my study is like scanning the index of a memory system; not a memory of facts but of the opening moves of familiar perceptual routines. In the normal process of practical thought (*i.e.* thought that is not theoretical, contemplative etc.) there is no distinction between the 'imagined' and the directly perceived. The peripheral awareness or the remembered present, is of things that are at hand and ready to be perceived precisely. It does not stop sharply with edges of the field of vision but extends out to what I could see if I moved my head. Nor need we stop there. I am also 'peripherally' aware of the rest of the house and its surroundings. If I already know what I am looking for in the world around me than there is no sharp distinction between what I focus on by moving only my eyes and what involves moving my head or my whole body. It is almost as though I could see, albeit vaguely, into the books on the shelves, and out through the walls. We tend to draw a sharp distinction between what we can 'perceive' directly before us and what we 'imagine' we could see it we got up and went through the door. We will see later there are excellent reasons for *us* to draw that distinction, but they do not apply to the sentient animal which lacks our reflective abilities and the neural or cognitive functions on which they depend.

9. SENTIENCE AS FUNCTIONAL REPRESENTATION

The pre-sentient organisms does not employ representation in the structural or depictive sense, only the distributed or functional 'representation' that is implicit in any self-organising system that 'remembers' the stable states to which it consistently returns. The sentient animal, however has a complex, stabilised, sensory-motor information system that I have described as embodying its perceptual world. The question now is whether *this* is representation in the depictive sense? Is not true perception necessarily representational? Despite what is assumed by many of those working in the field of Cognitive Science, I believe that it is not. Sentience, the first level of consciousness (or is it pre-consciousness?) does not involvedepictive representation.

When neurophysiologists explore the functions of the cerebral cortex they find only a huge variety of specialised perceptual analysers, *e.g.* for line, edge and colour. They never find a place where these features are reintegrated into a whole. There are no pictures in the brain. There are parallel neural mappings from one part of the nervous to another system, retinal maps and body-maps on the cortex for example, but these are not information, not messages, only a structural aspect of the material medium.

It is easy for us, as analytically minded outsiders to separate the animal from its environment and explain its performatory knowledge of its territory in terms of some inner map, but there is no need for the organism itself to function in that way. Perception is still an integral part of action, and the perceptual world is only a framework of traces left by whatever was invariant in the contexts of past actions. The animal's environment and its way of life are not functionally distinct. If one wishes to conceive of a correlation between the real world and the perceptual world internal to the animal, then a more appropriate metaphor would be some dynamic version of the precise relation of fitting that exists between a key and a lock or the imprinting of a cast to a mould. If the process of perception is an integral part of an animal's capacity for inter-action with its environment, then it relates to it is one of functional fittingness not depictive representation. For the animal, its perceptual and performatory habits on one hand, and the trails and scent marks, refuges and food sources of its territory on the other, fit together to form a single stable system. We tend to see things in terms of maps and images but we should avoid attributing this kind of structural representation to animals that manage perfectly well without it

10. SENTIENT CONSCIOUSNESS

As we have seen the sentient ground of consciousness is unreflective and non-dualistic. But if this is so then in what sense is it sentient or conscious at all? If sentience is defined in such a way that the sentient animal has no distinct inner world of imagination or thought, if its knowledge is all performatory, knowing how to perceive, but without any knowledge that it does so, then what is the nature of its supposed awareness? How are we to define what is going on here? The question is critical because with sentience, the story of consciousness has got beyond its Preface and Introduction and reached Chapter One. Everything that has been said so far seems to suggest that this so-called sentience, the pre-depictive and pre-predicative level of cognition, is at best only a kind of pre-consciousness. Not full conscious in itself but only a level of nervous function that evolutionary more advanced levels of cognitive functioning may (or may not) become conscious of. These doubts are justified. We know that much of our perception is subliminal and unconscious in the sense that it never reaches our kind of reflective consciousness. If fish, amphibians and reptiles are merely sentient then why should not all their experience be on this unconscious level? We are dealing here with difficult matters of definition, a kind of casuistry, but we surely do need to attribute some kind of genuine awareness to animals that do not have our kind of reflective self-consciousness. An organism with properly integrated action and perception is naturally assumed to have some kind of sentient awareness. It may show little of our kind of consciousness and yet it has a well developed central nervous system and we are surely right to attribute to it feelings and some primitive sense of self. How is this justified?

There is no more reason to attribute consciousness to pre-sentient organisms like insects than to any simple machines. Pigeons for example were once thought to be showing a high level of intelligence because they are very good at drawing fine distinctions between things that they saw only in a varied collection of photographs. It now seem that this and much of the apparently intelligent behaviour of birds and reptiles could be mimicked by connectionist, neural-net computers. Could not everything we have described in terms of the sentient animal's curiosity and its perceptual life-world be attributed to an arrangement of entirely mechanical neural net-works?[5] Perhaps, but even if this is so, I believe two crucial distinctions, constitutive of sentience would still remain.

The first originates in the role of the brain as a mediator between the environment and the physical state of the organism's vital variables. The purpose of the central nervous system is to maintain the stability of the organism by separating and co-ordinating the relation between its internal physical environment and the outer one. Now this could not be done unless there was a clear distinction between information reaching it from the two realms. Signals from the outside world must be functionally distinct from those which tell the brain about the organism's own inner state. The sentient animal has a brain precisely in order to mediate and regulate the relation between the sources of perceptual environmental in-put and signals reaching it from the body. The functional distinction between inner and outer are a vital part of its *raison d'être.* The animal necessarily treats inner sensations and outer perceptions as functionally distinct kinds of data because that is the whole point of having a brain. Thus we are right to think of a sentient animal as having 'feelings'. Signals of pain, hunger, cold etc. come from inside the organism and must be dealt with by altering its relation to the external world of perceptual spaces and objects. We do not need to suppose that it *recognises* that inner sensations are essentially distinct from one or another kind of external perceptions, but we must accept that the functioning of its nervous system requires that such a distinction exists. Though we as sapient observers can insist that all sentient beings have feelings, the animal's pleasures and pains are part of its life, properties of the things in a world it need not know it has. It has no absolute sense of inner and outer, self and world are not ontologically distinct but only relative positions on a continuum. As we will see the absolute sense of subjective and objective arises only with cognitive functions that employ genuine depictive representation.

In conclusion, there is no need to assume that sentient animals make even a functional distinction between imagination and perception or thought and reality. All its 'thinking' is concrete and out in the open world of real actions. It may however still have a certain sense of self. Its perceptual world contains a variety of perceptual locations and spaces for action and those which are centred on its own body will necessary be a distinct class among these, the perceptual sub-space that moves when it explores that of the fixed environment, the dynamic centre of its stable world and the sensory reference point by which it judges the value of everything else. There is no need for it have any kind of dualistic Cartesian subjectivity in order for it to have what Merleau-Ponty[6] calls the 'body-subject' as the permanently proximal pole of its world axis.

Whether we wish to call this consciousness or pre-consciousness is not very important. What is of significance is that in one way or another it is something we share with all those animals that have a proper central nervous system. When we are conscious, at the higher percipience and sapient levels, what we are conscious of, or by means of, is the primary sentient level of awareness.

11. EVOLUTION OF SIMULATION

The sentient organism has a good practical memory for its territory and can become very skilled at integrating and applying its instincts, but it has no capacity for detached cognition. It has no thought separate from action. Advanced sentient animals have an extensive performatory or sensory-motor knowledge of their world, the practical knowledge of where things are and what they do, but apart from their learned and innate skills they are entirely reliant on simple trial-and-error. It is from this level of functioning that evolution goes on to produce a radically new level of cognition. Instead of carrying out potentially dangerous trial-and-error searches to satisfy its various needs, why not first use its accumulated perceptual and performatory knowledge as the basis for purely *simulated* trial-and-error? An animal which has an extensive functional, sensory-motor representation of its world has, if the neural mechanisms were there, the potential to carry out actions internally and imagine the results to which they lead. Using past experience to test acts out before performing them overtly has an obvious survival value. A creature with a new and general means of avoiding dangers can lead a more enterprising life and create new environmental niches for itself, (and make the world more complicated for everybody else). Simulative experiments internalise some of the trial-and-error stages of learning and so avoid the inevitable risks that go with real errors. I call creatures that use this kind of cognitive simulation *percipient.*

The sentient animal, we may suppose, surveys its situation and simply 'relaxes' into the course of action that conditioning by past experience presents as most attractive. The percipient creature however looks at the various courses open to it and, inhibiting itself from immediate action, tries out the different routes internally before settling on that which seems to lead to the best outcome. Such cognitive simulation is the first stage of a kind of internal or 'off-line'[7] cognitive activity capable of many degrees of sophistication. Where sentient animals: amphibians, reptiles and probably most birds, use only instinct and the results of trial-and-error learning, mammals can use increasingly sophisticated cognitive strategies that involve the tacit simulation of actions and perceptions.

12. SIMULATION AND PERCIPIENCE

Percipient creatures are animals that can think, *i.e.* can carry on internal cognitive activity that simulates the overt acts and perceptions of their practical life and experience, but which is detached from it. Their brains can function in an 'off-line' mode, that is to say working in a way similar to that associated with real action but somehow disconnected from the normal motor out-put and perceptual in-puts. The simulation of sensory-motor performance involve running through an action in imagination. In ourselves, and doubtless in animals too, this is not done in real time, but efficiently in a rapid and abbreviated way. In an emergency one needs to discover (or recover) the outcome of a performance as quickly as possible. The process is quite literally one of thinking and seeing ahead to perceptual situations that are still only possibilities, literally percipience. The sentient animal looks before it leaps; the percipient one simulates the leap itself in order to imagine its outcome.

To imagine an action is to perform it in a minimal, inhibited way that stops short of overt expression, an inward shadow of the action itself. We do not know exactly how actions, and I include here acts of perceptual orientation, are 'stored' within the net-works of the brain, how the 'engrams' are related to the pattern of strengths and activities of the neural synapses that are their basic medium. We do know, though that our ability to use the neural net-works involved in action is not inseparably connected to actual performance using particular muscles systems. When someone re-learns to write or paint in new ways after the loss of a limb, the same individual signatures and kinetic styles emerge. The imaginary rehearsal of instrumental or dance pieces lead to genuine improvements in real performance. Simulative or 'mental' practice can even improve such basic skills as playing darts, and we can note how this case involves an imagined perceptual image of the dart board. Imagined acts are performed, if not in real-time, then at least sequentially. Forming an imagined visuo-spatial perceptual orientation or frame involves the activity of sustaining an internal stable state.

There is good evidence that real or external action and perception and their corresponding imagined forms, involve, as one would expect, many of the same neural systems. Functional systems that have evolved for perception and trained to distinguish particular things externally can be reactivated from within as images in cognitive simulation. A percipient creature that can summon up perceptual images in a way that is relatively independent of their original sensory-motor context can use them as virtual goals. It does not just survey a scene and think through the consequences of different ways forward (which is little better than mere sentience), it imagines what it wants and then hunts for simulation of the actions that would deliver it. This amounts to a level of homeostatic control as far beyond that of simply sentient animals as they are beyond the simple automatism of pre-sentient organisms. Before we look at the logic of the representation involved in this we should first consider a little more of the mechanism by which the remarkable phenomenon of percipient cognition is actually achieved on a physiological level.

13. DREAMING

Simulative cognition allows the experimental manipulation of internal models of the real life-world, and gives percipient creatures a remarkable a new level of autonomy with respect to the environment. To draw a mechanical analogy, this new order of self-organisation, is not less like a gradual movement to higher cognitive gear than like the introduction of a clutch that makes possible a whole series of gear shifts. A brain that can be disengaged from the immediate demands of its surroundings is a major advance over one entirely ruled by the external world and the internal drives of the vital variables. What justification is there for assuming that only mammals are properly percipient, and if it is such an advantage why did it take so long to evolve?

It seems likely that this clutch mechanism that disengages the brain from the immediate business of living is closely connected with the phenomenon of dreaming. All mammals exhibit the REM sleep which accompanies it in us. While it is going almost the whole brain is in a state of high activity. The sensory, motor and association areas of the cortex, as well as the sub-cortical nuclei that serve these and the deeper centres that mediate emotions and the activation of memory all exhibit the kind of electrical activity associated, in the waking state, with a high level of excitement. Levels of circulating hormones secreted by the brain are raised and sexual arousal is common. While this is going on the skeletal muscles of the body are in a state of almost complete paralysis. Sleep talking and walking never occur during REMEMBER sleep. During dreams the brain is like an engine racing on the spot completely switched off from the body. Mammalian REM sleep is an exaggerated and non-functional exercise in simulative cognitive functioning. In dreams we do what a merely sentient animal can never do, disengage our brains from the real world and engage in pure simulation and fantasy.

Dreaming clearly serves as a vivid model for some aspects of what I am calling percipient thought, but we are still left with question why it should be so necessary. Why should such an odd activity still be needed after a hundred million years of mammalian evolution? To be percipient is to have a brain that can work in two different ways. It can work in the normal sentient way, or it can work out of real-time and without external behaviour, activating its sensory systems with self-generated stimuli. To have a brain with two distinct modes of operation is like having two brains in one.

Using the same hard-ware for two quite different kinds of functions is not an easy trick. To learn and adapt to an unpredictable world requires an enormously flexible central nervous system and much of its function is not fixed to particular cortical locations. As Edelmann[8] has described, the same point on the association cortex can vary in functional commitment from week to week or even day to day. Blind people can come to use areas of 'visual' cortex for auditory processing. The point of this is that in a system of such plasticity a function that is not regularly used in a particular way is liable to be lost. It is easy in everyday tasks , even for us, to revert to simple minded trial-and-error when it would be more efficient to stop and think things through. We are often not as percipient as we should be. In simple creatures such cognitive inertia (or mental laziness) could put the whole system of percipient cognition at risk. A 'percipient' cognitive system that could not keep a very clear and rigidly maintained distinction between the real and the simulated, between thought and reality would be worse than useless. The regular, vigorous and obligatory exercise of off-line functioning in REM sleep ensures that this distinction between the two modes does not break down. Its not what we dream that its important but that we dream at all. It maybe that in some mammals, as sometimes in us, dreaming generates some useful cognitive novelty by the random combination or association of ideas, but it is surely its physiological function not its content that is of primary importance.

14. REPRESENTATION

The cognitive processes of sentient animals do not produce positive depictions of the world but only the negative imprint of a way of life and the pattern of activity implicit in certain habits of attention. The outsider (a cognitive scientist for example) might think that they could deduce the presence of internal representations from the animals behaviour - but they are really no more (or less) justified in this than someone who says that the growth-rings of a cut tree-trunk 'represent' the passage of years. For us they do, but from the point of view of the tree as an organism they do not. If we wish to use ideas about the evolution of representation to understanding our experience as conscious beings, it is only the inside, organismic view that is going to help us.

It might seem that with the evolution of percipient simulation we now reach the stage at which true, structural representation has definitely arrived. Our experience of imagery is certainly a major reason why we find Cartesian dualism so plausible. However we must remember, as we saw earlier, that though the brain contains many different kinds of perceptual analysers that collectively enable us to distinguish between different kinds of things in the world, there is no evidence that these components are ever re-synthesised into a structural or depictive representation in the brain. As we saw when discussing sentience and the idea of the 'remembered present', the wholeness of things is not in our heads but in the world itself and our performatory relation to it. We do not need to synthesis what is already fully integrated. Whatever dislocated, distributed and non-representational pattern of cortical activity is associated with some real experience of the world, that same sensory-motor pattern, or some appropriate part of it, can be re-activated in the process of percipient simulation.

We can say, with considerable justification, that imaginary simulations represent the direct sensory-motor experiences that corresponds to them. This is certainly where we get the idea of thought and perception as pictures in the head. However, both simulation and perception involve the *same* cortical areas and forms of neural activity, illuminated, as it were, in one case from within and in the other from outside. An overt action or perception and its simulation cannot occur at the same time, something more characteristic of functional representation rather than the simultaneity of depiction. The simulative thought is like a faint inner copy of the real performance. But calling the imagined gesture or object a 'representation' of one overtly performed or seen, is like rather like saying that a dark-room lit by faint red light is a 'representation' of the same room in full day-light. Similarly we could pose the question of whether inner speech is a representation of speaking aloud. Isn't this like saying that a whisper 'represents' a shout? As with the case of the tree growth-rings, it depends on your point of view. My whispered rehearsal of speech may serve me as a representation (or simulation) that helps me to reflect on how my words will sound to an audience, but to someone over-hearing me, they are less a representation than a draft of the speech itself. Where sentience seems like representation only from an outsider's view-point, percipience is apparently only a representation when seen from within. It would do no harm to this account to leave it at that, and to allow a full two-sided depictive representation to emerge as a dialectical synthesis at the next stage, sapience. However, I do not think this really does justice to our understanding of ourselves or fellow percipient creatures, the apes especially.

Despite the fact percipient simulation is different in important ways from the simple paradigm cases of DNA and phonetic script, I believe that, even from an internalist perspective, one can argue convincingly that percipient thought is a form of genuine structural representation. There seem to be two approaches here. I will call them the arguments from meta-action and from goal directed action.

15. SIMULATION AS A DOMAIN OF META-ACTION

To summarise the points above, simulated actions are not structurally independent of real actions and so cannot be said to be true representations of them. There is a useful similarity between them because one kind of act is an inhibited or abbreviated form of the other. Simulated actions and perceptions are not representations of real perceptions but the same perceptions activated in a different way. If this is all there was to cognitive simulation, as may well be the case in the simpler mammal, then it is quite right to say that percipience has no more claim to involve full structural representation than sentience does. However, if virtual perception or imagery was *only* of that kind and only an inner echo of the outer function, then it would be of relatively little value to the creature possessing it. To be used fully the content of cognitive simulations have to be the object of a higher level (and quite different kind) of action.

We conceive the basic process percipient cognition as one in which a creature rehearses the courses of action open to it in simulation, envisages their different outcomes and performs the one with the best results . In doing this the creature is not just inwardly repeating outward actions and perceptions but actively manipulating them in relation to one another. It can regularly perform an inner version of the outer trial-and-error activity that in sentient animals is employed sparingly during the exploratory and learning phases of their life and experience. This meta-action is *not* a simulation but the real work itself. The cognitive or mental acts involved in percipient thought are thus on a higher level and of a different kind to the acts performed by sentient animals. They are acts *on* acts; acts which take other sensory-motor operations as their objects. (They can even occur in so-called 'lucid' dreams.) Looked at in this way we can say that simulated acts and perceptions that are compared and contrasted in thought *are,* in a real sense, representations of overt acts and perceptions. Because the acts are now part of system in which they exist on a different level of organisation, the simulated or abbreviated cognitive elements are mere tokens or symbols of those that are used outwardly, and as such they are functioning as genuine structural representations. The telling analogies here are not those given above (*e.g.* with the relation between to two illuminations of the same scene) but with using the relations between operations and objects on one medium to stand for relations between those in another. This representation is 'depictive', but not in the sense that a map depicts a landscape, but as one make practical use of a plan. This relation of representation is not between simple structures *per se*, but is like that which exists between, for example, experimental moves tried out on a map or model and actual manoeuvres carried out on the ground.

16. THE SIMULATION OF OBJECTIVE GOALS

Self-maintaining and self-organising systems are continually moving towards some predetermined state: states of greater stability or safety, and the sequence of states and stages in the cycle of growth and reproduction. This goal directedness is a necessary aspect of the homeostasis that is intrinsic to all living things. Each stage in the evolution of functional organisation has at its highest level of goal directedness, and in percipient creatures this must involve their ability to simulate.

The basic skill in percipient cognition is one of choosing between different potential actions and their outcomes, against the background of the real perceptual world. It is clearly a small step from this to the stage of simulating of perceptual goals. Instead of merely choosing among the courses of action offered by a given situation the creature imagines some ideal situation. These *desires* will doubtless come from the combined effect of past experience and instinctive drives, typically the simulated image or thought of the presence of food, shelter or a mate. The creature's course of action is then directed, under sensory guidance, to successively reducing the difference between its present situation and the imagined one, a process of error correction that continues until its goal is realised. Such intentional action starts with an internal goal and proceeds to brings the objective or outward state of affairs into conformity with it. The evolution of simulation gives the percipient creature the ability to bring about real states of the world that are, as it were, depictive representations of inner, simulated states. The process is like that of person working from a plan or an artist drawing from imagination or memory. It is by a process in which direct or overt perception plays only a mediating role.

We normally think of representations in the mind as proceeding in the other direction, from the outside in. Things in the world cause us to have representations in the mind. When we draw or map from memory, or act according to a specific plan, the raw material for the plan came from the external world, from sentient experience. But as we have seen, sentience alone is not yet depictive representation; there are no representations in the brain as such. During the evolution of consciousness perception only becomes representation through the part it plays in cognitive systems with simulation. And that happens because the world is made to represent simulations, not the other way round. This has nothing to do with the Subjective Idealist idea that the external world is a projection of the Mind. It is rather, that when we explain things in evolutionary terms we have to accept that evolution is not concerned with the 'Truth' but only with the pragmatic matter of survival. It we come to have true representations then this is something that must be seen as proceeding from our needs not those of the world. Philosophers are liable to go on and ask whether our inner simulations really (*really* really) correspond the word, but it is hard, in this context, to attach much sense to such questions.

Once we have a cognitive system in which one component functions as a depictive representation, *i.e.* the mapping of internal imagery onto external reality, then there is no longer any need to withhold this category from the other components of the same system. The mapping and hence the representation proceeds in both directions. So long as we are dealing with percipient organisms, we may now agree with those empiricist philosophers and cognitive scientists who hold that perception is a process of depictive representation. However since the representation is clearly a two way process between organism and environment, their emphasis on theories in which perceptions are *caused* by the effects of external things on the organism seems rather unfair on the contribution made by millions of years of evolution and the activity of an organism's millions of neurones. If, in an interaction, one side does all the work, it is an odd (though understandable) agenda that wishes to attribute the causal role to the other side.

We could not say that the sentient animal's perception was a structural representation of the world, but we can say it of percipient creatures. To a reductionist this will seem odd because, considered purely in physiological terms the basic perceptual mechanisms are essentially the same in the two cases. In discussing percipience, we have taken it for granted that perception *per se*, existed prior to the evolution of the higher level of cognition. Why then, is one a case of structural representation and the other not? The puzzle disappears when one remembers that nothing can ever be a representation in itself. As we saw earlier DNA is a representation only because of its role in a larger system; it is a logical requirement of full self-replication. So with our current examples, percipient perception is representational because of the role it plays in a system that allows a creature to make external and objective representations of inner goal states. It does not make sense to describe sentient perception in this way because it is still only an imprint of sensory-motor experience, without the role in simulative, goal directed action or the meta-level activity that would make it more than that.

17. PERCIPIENT CONSCIOUSNESS

As we saw, the brains of sentient animals perform cognitive functions that involve a necessary distinction between inner sensation and outer perception, but we cannot suppose that sentience itself involves any explicit awareness of this. Its 'remembered present' allows no clear separation between perception and imagination. Any distinction between Self and Other must be relative; between proximal and distal spaces or kinds of experience. Percipient creatures have a higher order of self-awareness. Their brains don't just have two kinds of in-put, but two distinct modes of cognitive functioning. They have to be able to distinguish between the 'in here' of simulation that belongs to the self and an 'out there' direct perceive that does not.

With percipience we are approaching something like our kind of consciousness. Percipient creatures can stand back from the world, not just perceiving but conjuring with imagined possibilities. Their necessary distinction between thought and reality gives them the ground on which to reflect the primary sentient distinction between inner feelings and outer perceptions. They must surely have an explicit sense of the difference between Self and Other. A creature that is only percipient cannot say to itself 'I think therefore I am' - but its form of cognition has the underlying dualistic structure that allows us to do that. Without language to represent that structure and so cannot go on and postulate a representing mind distinct from a represented physical world. The percipient creature has to be, in action, tacitly aware of the distinction between its inner and outer worlds, but there is nowhere for it to make that awareness itself the object of a separate representation. The percipient creature can stop and think, using simulated action and perception, but there is no further medium in which it can reflect on the adequacy of these simulations. It has no higher level of representation and performance in which one might say that it *decided* to stop and think. For that we must go on to the evolution of culture, language and sapient consciousness.

18. SAPIENCE

Human beings are not just cleverer than any other animals, their cognitive power is of an altogether different order . Even allowing for some anthropocentrism, our ability to solve novel and abstract problems, our foresight, self-control and so on, are all vastly superior to the most intelligent apes. If one believes, as many cognitive scientists do, that all higher cognitive processing, in humans and apes alike takes place using innate syntactic codes; what they call the 'language of thought', then the existence of a whole different order of intelligence is difficult to explain. We know that our brains are much larger that those of the other apes, but that is clearly not the whole story - small children can speak well using a brain that is smaller than that of a large ape. Structure is clearly also important, but what is it that makes our thought so incomparably more complex than theirs?

The answer I believe, lies in the fact that only human beings have *two* systems of cognitive representation. The ability of paired systems of representations each able to reflect on and monitor the other results a new order of cognition and consciousness as far above percipience as that is above simple sentience. Biological evolution gave percipient creatures a brain with two ways of functioning, the on-line and the off-line. This dual system provides them with a system of depictive representation and with it a whole new order of cognitive organisation and control. In human beings a relatively small change in neurophysiology has lead to the ability to initiate and develop language and culture, and so to a dual system of representation. The result is a wholly new kind of self-reflective consciousness that I shall call sapience.

19. THE EVOLUTION OF PRE-SAPIENCE

Cognitive science, following Chomsky's lead, tends to ascribe speech to an innate language instinct. It *could* be that this and each of the other major areas of human intelligence, is due to a complex, separately evolved and specific, neural 'module', but this does not seem to be a very useful or persuasive interpretation. Given how closely we are related to the apes, it is more reasonable to suppose that our superior abilities are based on some relatively simple change in brain structure such as might follow from a few minor mutations altering the relative duration of certain stages in its embryological development. An extra stage of cell division that doubles the population of certain neurones, for example, and a change in strength of some nerve growth factor that increases the size and number of certain nerve tracts. The point of this account is to tell a plausible and intelligible story about the evolution of language and consciousness and provide at least some degree of logical and biological detail. The arbitrary introduction of theoretical entities like the 'language acquisition device' does not help here.

The human cerebral cortex is not only relatively larger than that of all other mammals, it is also more complex. It has a larger number of architechtonically distinct areas, and it has more and larger trans-cortical tracts of nerve fibres that connect different areas of cortex together. (The second of these characteristics may be the cause of the first.) The trans-cortical fibres tracts are associated with a special facility in cross-modal pattern matching. Thus when a child is presented with various shapes visually it will be able to match them accurately to similar shapes that are only explored by touch. Chimps take hundreds of trials to master this kind of task[9].

Such trans-cortical connections constitute neurophysiological mappings between different sensory and motor areas, and allow distinctions made in one modality to be made automatically available to another. They are, in effect, the hard-wired basis for of internal depictive representations that 'translate' one sense into another. These are, rather indirectly, the primary neurological basis of our higher representational abilities. Sentient and most percipient animals get by quite well with relatively few inter-sensory tracts. The environment stimulates their senses in coherent synchrony and they have no difficulty leaning the necessary associations between them without the need for high-level, internal short-cuts. Human culture however, is crucially dependant on such connections, and the most important of them is certainly the visual-somatosensory link that mediates the imitation of one person by another.

Human babies are born with an instinctive interest in human faces and an innate ability to start imitating facial expressions and simple body movements[10]. We tend to take imitation for granted, but looked at rightly, it is something of a miracle that an infant with almost no experience of the world can watch someone and translate the visual image of the other's movements into a corresponding sensory-motor performance of their own. Only trans-cortical links can account for this; simple theories of learning by conditioned association certainly cannot. Apes have a certain limited ability to imitate overt behaviour ('aping' or imitation *per* *se*) but human beings are adept at an insightful copying that involves following perceptions as well as action I shall call it 'mimesis' (to use the grander Greek word.) Within a few months of birth the infant's ability to imitate progresses from simple actions to copying perceptual orientations as well. It can follow another person's gaze and acts of pointing, in order to see what they see. They also a develop a very great sensitivity in recognising and picking up other peoples' emotional reactions. They gauge novel situations by instinctively looking to see how other people are responding. Like any sentient animal a child naturally explores its body-centred, sensory-motor world, but unlike other species, it also comes to participate in a shared sentience; the world of a particular way of life and the social practices and habits of attention on which they depend. As a member of a wholly cultural species their most important survival task is that of coming to understand other people, their interests and their world. The innate cognitive basis for this is tacit or internal mimesis, what one might call, *simulative anthroperception*.

Simple imitation requires only the visual-somatosensory link. (Some autistics and Tourette's sufferers do it compulsively.) Full mimesis involves an understanding that depends on using a percipient or simulative identification with the other person as well. The difference is clearly seen in the learning of skills. Young chimps take years to master the method that their parents use to crack nuts with stones. A child at the same level of dexterity can learn, or be taught, to reproduce the skill in hours or even minutes. The key lies in recognising what perceptual discriminations are essential to the task, in knowing what to look at and why.

I call the kind of cognition that depends on full mimesis *pre-sapience.* It depends on the use of one's percipience to tacitly simulate or rehearse other peoples' actions as well as one's own. It is simple cognitive economy to represent other people's intentional acts and perceptions in the same way that we represent our own. Because we experience and rehearse our own acts and understand other peoples' in the same neural medium, each can draw on and potentiate the other, a process that rapidly boot-straps the child's mind into the perceptual, communicational and cognitive practices of its community.

Mimesis, anthroperceptive simulation or pre-sapience (three aspects of the same thing), provide the route, for the child and the species, to move from percipience to the acquisition of language. Infants instinctively imitate other peoples acts of looking and their general orientation reactions. From this first sharing of ways of perceiving they are soon introduced to pointing and other gestures that direct attention. Pack hunting animals use pointing in a limited way and vervet monkeys employ a variety of alarm calls matched to different dangers. Chimps use dozens of different vocal signals, but it takes brains with a strongly cross-linked and differentiated cortex to combine these into a complex social semiotic[11]. Only human beings with their innate understanding of one another's intentions are able to invent easily learned and manipulated systems of auditory communication.

Apart from the followed gaze (aided by the distinct whites and iris), our most basic and instinctive method of sharing perceptions is by pointing things out, and the roots of language are best understood by treating it as a greatly elaborated system of conventional verbal *gestures*[12]. With it, we can indicate not only what is immediately present but also the possible contents of any sentient spaces[13] that are open to percipient imagination. Language allows us to share simulated perceptions that would otherwise remain entirely private, but it is important to realise that from its origins, language itself is primarily a system of indication within shared horizons, and only secondarily a form of representation. Its primary work is not to present depictions but draw the sequences of distinctions that separate what is meant from what is not. Because it starts with the imitation of acts and perceptual orientations, one person copying another gaze etc., it is an *analogue* action system before it is a system of depictive representation. The representation of. the world by language grows out of each person's identification with and mimetic representation of, their fellows. Even the impersonal truths of science depend on the innate, pre-established harmony between one human being and another. As I shall discuss elsewhere, it is a modern tragedy that we have come to identify the structure of the mind with our depictive media rather than with one another.

Structural representation is intrinsically dualistic, but acts of indication or perceptual orientation include in themselves the object or state of affairs that they pick out. If the same objective movement is used to point to two different things then it is a different act of signification in each case. Primary linguistic indication unites, in a single action, the verbal gesture and what it is about. Human mimesis however, easily separates these, reverting to the attitude of simple sentience and letting what is pointed to appear as an autonomous object independent of the human act that picked it out. When gesture gives way to speech, this becomes even more marked and the acoustic image of language is easily taken as quite separate from the motor acts that enunciate it. The simulated images of words and those of situations and things, involve quite different areas of the cortex and can be activated by different cognitive operations. This clear structural and functional separation within the brain is the foundation for the higher level sapient cognition. Despite its origins in our intuitive understanding of one another, we have no difficulty in establishing language as an internal cognitive system that serves as a true structural representation of both real states of affairs in the world and our simulative representations of them. What starts as the innate representation of one person's actions by the percipient simulation of another gives rise within the brain, to a second system of representation of the world in general. The result is a dual cognitive system with an enormously enhanced potential for self-organisation and creativity.

20. LANGUAGE AS REPRESENTATION

It may seem unnecessary to establish that language really does constitute a true system of representation in both communication and internal cognitive activity. We do after all take this for granted in everyday life, but is this necessarily so? It is possible philosophers (*e.g.* the later Wittgenstein) are able to throw serious doubt on the idea that language represents at all. In the present evolutionary account language has been regarded as primarily a system of indication in which the words of an utterance and the things they pick out are 'internally related' in the sense of being inseparable aspects of a single social act. If language is entirely a matter of social practice then words and things are not logically separate but mutually defining.

The point here however is not to solve the (insoluble) problem of the ultimate nature of natural language, but only to establish whether or not it is possible for it to function as a genuine system of representation. The answer is clearly, that it can. Language has a great many different social functions, but one of them is certainly that of providing unambiguous descriptions of objects and situations. Words can be systematically related to things. Language can be used with complete consistency to translate into and out of other representational media, especially in scientific technology. It is a simple matter of fact that given the right conditions and context language can satisfy any reasonable definition of representation that one likes to give. If the DNA-protein system, or that of a phonetic script, are true representation then rules can be drawn up that will allow at least a sub-system of language use to do the same thing. Sets of statements can be made to correspond to sets of objective states of affairs in a consistent way.

If language can work as a system of representation between people then it can also do so within them - *e.g.* as a part of an internal dialogue. Given that we have already established how percipient imagery can be a form of representation within individuals than it is clear that human cognition really does, (as we knew all along,) involve two different and independent forms of structural representation. A language that was only a system of inter-personal indication could greatly enhance social communication and control, but it would not increase individual autonomy, nor would it support the growth of bodies knowledge, like science, that exist independently of individuals.

21. THE DUAL SYSTEM OF REPRESENTATION

The single percipient system, though in principle able to call up and manipulate large areas of the creature's sentient life, is still essentially a closed system in the sense that its performatory simulations and what they simulate are of the same basic type. They are at best analogues of one another. New perceptual information erases old, and simulation will always tend to be of current information. Stray images may occur and recur especially in dreams, but without an independent system of representation they cannot be arranged and controlled except by experience itself. There can be no 'propositional' knowledge and no memory that can be systematically interrogated. The percipient creature's knowledge is organised only by its experience of the world, never by the self-reflective process of a dialectical interaction between two different media. To have one medium of representation is to be tied to the level of practical knowledge, to knowing *how* to do things. Having two media, allows one to have truly representational knowledge, knowledge *that* things are so. Dual media gives an immense increase in cognitive capacity and power. Like a pair of mirrors they open up the prospect of unlimited internal reflection.

22. SAPIENT THOUGHT

Two different and independent systems of representation are each able to reflect on the content and adequacy of the other. Verbal statements can be translated into percipient imagery and tested to see if they are consistent with the accumulated simulations of past experience. Imagined or remembered cognitive simulations can be described in words and tested according to the criteria of sense and logical coherence that govern linguistic discourse. Imagination can test the relation between language and experience; language can pin down and depict the differences between the imagined and the directly perceived. Direct practical experience can decide between the rival claims of verbal and simulative representations. . . . and so on round the semiotic triangle in a virtuous circle of self-regulation. The inter-action of two systems make it possible to think about the difference between facts and beliefs, thoughts and things, appearance and reality.

The double representational system is not just a Mirror of Nature accurately portraying facts about the world. Being sapient gives a unique capacity to stand back and reflect on ourselves and our experience, but in evolutionary terms this is a secondary consequence of a practical, biological autonomy with respect to the environment. Our dual cognitive system lets us enact the role of mental subjects confronting a world of physical objects, but the representational systems that allow this must also be seen as domains of action or performatory media.

Percipience allows the formulation and testing actions. It makes maps because maps are a safe medium in which to plan and try out actions. Sapience improves on this. It becomes possible to perform goal directed actions that are based on the best available picture of how things are and is planned in a dual medium that allows choosing goals, analysing tasks and putting together a consistent description of what is to be done. At our most rational we doubtless do all that, but the role of social language in behaviour starts at a much more basic level. Children with mental disabilities can greatly increase their effectiveness in many tasks merely by learning or being taught, simple methods of verbal self-command. Experimental versions of the same phenomena can even be shown on apes. Self control is greatly aided by the internalised habit of giving orders. In adults these forms of language and attitudes of language use operate on a deep, less conscious level, but they are still profoundly involved in the way we control our actions. Voluntary action is behaviour that we can *tell* ourselves to do[14]. Conscious human action is different from animal behaviour in that it can be controlled, explained and understood by the actor's use of language, the social semiotic. Sentient animals have an innate biological ability to look before they leap. Percipient ones can run tests as well, but even apes have relatively little capacity to restrain their spontaneous impulses. Only sapience and the extra level of control provided by language allow one to really deliberate before acting. Sapience leads to a certain freedom from the immediate demands the body and the states of its vital variables. As sapient beings we choose our goals and values in a context of social discourse[15]. Sapience is only possible because of our innate understanding and attention to the thoughts and feelings of others that makes discourse possible The relation between sapient consciousness and conscience is not merely contingent. All human life has an intrinsic ethical dimension.

Percipient creatures explore and play in their outward behaviour, and doubtless do so in internal simulation too. What they cannot do is explore the representation of the world as such, and think theoretically or creatively. Sapient beings have a systematic ways of generating new ideas, and can change not just their way of acting but also the way of seeing and representing that lie behind these. The sapient being not only has a collection of percipient pictures of the world vastly enlarged by its access to the experience of others, it also has the ability to think creatively by playing with different ways of representing things. To have two system of representation that coded exactly the same information may be a perfect instrument of communication, and in technical situations we can create representations of whatever degree of precision we need, but for internal cognition a 'perfect' dual representation would be perfectly redundant. The vital creative and dialectical power of sapient thought arises precisely because we have a language in which things can be represented in any number of different ways. The dual cognitive system gets its power from the vital mismatch and free play between its two halves. Communication between people can reduce the differences in their information about the world by the method of question and answer. E.g. 'Which is *the way to The Broadway*?' - '*Straight on and left at the top.*' The two expressions describe the *same thing* in different ways that have quite different practical implications. When we think, explicit self-questioning is optional, but its functional equivalent, the interchange between alternative representations is essential. Percipient thought experiments with ways of acting, sapient thought does it with ways of perceiving and representing as well.

23. CULTURE AND THOUGHT

Animals using percipient cognition alone may have cognitive functions analogous to what we can think of as 'memory', 'reasoning', 'imagination' etc., but it is misleading to equate these with what we do. Sapient self-control and the dialectical processes of verbal self-interrogation put our cognitive acts on a completely different level. We can have many different percipient registers and kinds of imagery and imagination, and these can, through their association with language be used to form many kinds of mental models and inner symbolism. The innate capacity for simulation that we see in dreams, and the language we learn from others can be combined into functional cognitive systems in innumerable ways. The human brain is immensely complex, flexible and dynamic, and individual human brains differ in detailed structure to a degree that would be utterly bizarre if it applied to the parts of the body we can see. Each of us constructs and uses our own methods for dealing cognitively with the tasks of our everyday lives and work. Large areas of cortex are dedicated to different cognitive functions in different people (and differently in the same person at different times). There is doubtless a great deal of natural convergence in the ways we perform common and basic tasks, but in the detail and the free play of thought each of us thinks in their own way. This means just what it says and does not imply that we cannot really understand one another. On contrary, this variation only happens because we have the profound and innate mutual understanding needed to develop a complex language and cultures. In this context it is what you do that is important not how you have taught your brain to do it. There is a sense in which any two people in conversation draw on what they have in common to build a new sub-culture, but in the end we share in the common culture of humanity and it is that which defines the nature of our actions.

24. OBJECTIVITY AND SUBJECTIVITY

A sapient beings can be explicitly aware of (or can represent to itself) the fact that its thoughts are of a quite different nature to the things it thinks about. It can, as it were, use its two representational media to triangulate outwards on 'objective' reality that is independent of thought, and inwards on a 'subjective' realm.

Objectivity is in one sense the hypothetical domain that allows our two systems of representation to agree. Regarded as a social activity it finds its highest expression in scientific technology and theoretical practice, and in systems of representation that go beyond the limitations of personal cognition and natural language as a method of indication. In pre-sapient times before the arrival of language our species and its predecessors were already replicating skills and tools. After it appeared there came first tallies, pictures, and maps, then numbers and writing and on to all the modern tables, codes, symbolisms and systems of measurement. With these came meta-representation, the explicit description and logical definition of symbolic systems and ways of representing - and para-representation, the establishment of long chains and loops of consistently linked representational media (especially measurement systems) that bring all kinds of natural phenomena into the orbit of accurately replicatable operations and observations. Such media are is not only a representational but also performatory media, domains of cognitive and practical action. Every map provides a context for action and a surface on which to plan.

Once instituted as a true system of structural representation human language and objective scientific technology can be seen as taking on, in an almost literal sense, a life of their own. The human capacity for mimesis, or indeed for simple imitation, allows the unlimited *replication* of verbal utterances and other representations. Given this, statements whose *representations* are of value, *e.g.* traditional wisdom, scientific facts or whatever, compete with one another to undergo *reproduction* in the medium of human society. Thus the logical relation Von Neumann worked out in the context of any evolving entities (*Reproduction = replication + representation*) re-emerges here in the domain of language and thoughts.

The cognitive dualism leads not only to new levels of self-control and rational action, it also gives a new kind of self-monitoring and internal self-reflection, the triangulation on 'subjectivity'. Because we have two representational media we can reflect on emotions in tranquillity and become aware of feelings and sensations as entities. We can attend to what is subjective when we have words to pin it down and sentences to represent its place in experience. Without language the 'subjective' is always the context, never the content of thought. For pre-sapient creatures an imagined meal may be realised through action, but the food and the satisfaction it gives cannot be thought of separately. The wound *is* the pain, the beloved *is* the joy. Only through discourse can we think about feelings in their own right.

The language of emotions and subjective feelings are of great importance to self-control and social judgement. Sapient consciousness is founded on anthroperceptive mimesis and it is by participating in the acts and habits of attention and indication a child comes to understand itself through understanding others. The different uses of the words 'me' and 'you' are far harder to grasp than 'dog' and 'car'. Subjectivity is a reflective awareness of self that develops from the outside in. To take a simple example, words for colours, that distinguish things on the basis of hue, can be used to reflect inwards and focus, not on things but on our *experience* of them; redness and blueness as such, taken as the names of subjective qualities. In such a way words can be used to explore the basic properties of our perceptual systems. Philosophers often talk of these 'sense-data' or 'raw feels' as though they were in some way fundamental, but in cultural terms and in the history of sapient consciousness they are very late and sophisticated objects of awareness. As recently as a hundred years ago, when Montessori started teaching, few children mastered the colour words before the ages of eight or nine. There is good evidence that human beings posses an unconscious magnetic sense. If this is so, then culture and education may one day find the words and ways of bringing it into the shared horizon of sapient consciousness. It wont be easy; how do you point to a sense of direction?[16]

This process of internal triangulation on our own internal cognitive processes is necessarily limited in scope, and the interpretation of the introspective data we do have is difficult. When we ask ourselves questions like, 'Do I think in language?', 'What imagery do I use?', we cannot expect detailed or reliable answers. We certainly do use language and sensory-motor imagery internally (in classical antiquity the latter was taught explicitly in the *Ars Memoria*[17]), but no cognitive act can be both a thought and its object simultaneously. We are denied a reliable answer to many introspective questions by an intrinsic uncertainly principle: any process as labile as that of thought is necessarily altered by the act of reflection on it. If self-knowledge and control increase it is through practices and models derived from collective discourse.

25. SAPIENT CONSCIOUSNESS

Supposing that my account is correct, and the kind of simulative sensory-motor system with a social semiotic that I describe, really does lead to the behaviour we associate with objectivity, subjectivity, voluntary action etc., is this really enough to account for the phenomenon of consciousness as we experience it? Doubts must remain about whether we have really an explained consciousness.

Many people who have thought deeply about the question of consciousness have decided that it is in principle inexplicable, or that it is a necessary property of all neural nets and organisms, or even of all matter. An explanation is no use if it does not clarify thinking and gives a sense of understanding, so in this sense every consciousness requires its own explanation. There are still some deeply mysterious things here, which I will deal with in another paper and only briefly indicate here.
Before that though there are some other causes of puzzlement that should be dealt with first.

These come from our traditional Cartesian habit of tacitly assuming an idea of consciousness and mind on the model of a *single* system of representation. If mind represents the world, then by the logic of this concept the represented and representing must be ontologically distinct. (Maps and landscapes, like symbols are their referents, are both real - but in quite different ways.) If the mind and the material world occupy different kinds of spaces, or if as Descartes believed, mind has no spatial extension at all, then there is no way they can ever come into contact with each other, and we are trapped in the fly-bottle of sceptical solipsism. From this intensely egocentric or First Person perspective it can seem as though there is no real difference between the question of consciousness and certain other questions that are actually quite distinct. *E.g.* the ultimate ontological question: 'Why does the world exist?', 'Why is there anything at all?' These question really are unanswerable, because they could only be answered from as Cartesian position outside the world - *i.e.* The 'View From Nowhere' that Nagel attributes to Natural Science[18]. From the other side of the dualism comes the ultimate existential question; 'Why do I exist, here and now?' , 'Why am I me, rather than someone else or nothing?'. There may be a causal account of where I come from, but *a priore*, the chances are still infinity-to-one against it occurring. These Cartesian conundrums are easily confused with the problem of consciousness but they are really separate questions.

There are certainly many more problems and paradoxes involved in thinking about the nature of sapient consciousness. Indeed the whole stance from which this paper has been written, is as I suggested at the beginning, open to some profound questioning. This is not the place to go into this in detail, but there is just space to indicate something of why this should be so.

Sapient cognition and consciousness depend on two independent systems of representation, so that when we push philosophical or theoretical reflection to its limit we are necessarily confronted by certain fundamental dilemmas. One form of this is the simple question, Which is to be taken as primary, sensory-motor simulation or language? Which has priority, the biological system of representation based on the simulation of goals or the social system based on mimetic simulation of actions? From a detached systems theoretic stance we know that each is important as the other because sapient thought and consciousness emerges from the inter-action between them, but as sapient thinkers; embodied beings using this kind of cognitive system, such detachment is impossible. When we reflect deeply on any theoretical topic, coherent thought demands that we take one side or the other.

In philosophy and metaphysics this is expressed in two deeply antithetical traditions each with its own approaches to ideas like Knowledge, Meaning and Truth. Since the Renaissance, the Western thought has been mainly dominated by just one of these two poles or attitudes, and it has given us the dualistic, analytical, rational empiricism that has accompanied the growth of scientific technology. The other pole, based on an attitude that is just as natural and equally innate in its foundations, has been consistently repressed or distorted by the dominant tradition. It shows itself in various forms in such philosophers as Spinoza, the German Counter-Enlightenment, Hegel and the Romantics, Nietzsche - and in this century Heidegger, the later Wittgenstein. Today it is expressed inpostmodern and deconstructionist writings.

I hope it will be clear that we are *not* talking about the such dualistic pairs as Rationalism and Empiricism, Idealism or Materialism, Subjectivism and Objectivism etc. These are disputes internal to the dominant tradition. What we are dealing with here is not the Representationalist, First Person / Third Person dualism that has run through this paper, but the dialectical relation between First Person+Third Person & Second Person+Fourth Person, to alter Buber a little, the *I-It* and the *We -Thou.*
This distinction is of course, not new, Vico clearly recognised it in the 18th century, and it has been well understood in the non-dominant tradition and by philosophical hermeneutics, from Schleiermacher to Gadamer, Merleau-Ponty and Ricoeur.

The simplest way to explore the phenomenolgy of the two poles in one's own sapience is to reflect on the attitudes of mind associated with questioning the speaker's and listener's relation to something said. I cannot do more than touch on these very briefly here, but at the risk being of misleading or incomprehensible I will list a series of antitheses that indicate something of the difference in spirit of the two positions and nature of their underlying structure.

As speaker's with something to say, we confront a real or imagined object or state of affairs and attempt to represent it in words. This is a task that only makes sense if the world is something determinate and already given independently of anything we should say or think about it.
As listener's we are always already in dialogue with others. Aspects of a shared world are disclosed through indicative acts and utterances whose sense and relevance depend on the context and social activity of which they are a part.

As Speakers we see ourselves as goal directed agents in a world of subjective and objective things; actors perceived in perceptual world.
As Listeners we enact ourselves as dynamic points of view in a social, historical and natural process; active perceivers in a world of inter-action.

For the Speaker action follows from one perception and is directed to achieving another.
For the Listener perception is the result of action and for the continuity of an activity.

The Speaker's position tends ultimately to solipsism, the Listener's to negipsism, a belief in the illusory nature of the self.
Pure Speakerism leads to a blindly reductive scientism that is profoundly destructive of human values. Pure Listenerism leads to ineffective relativism and inarticulate mysticism.

For the Speaker sapience is based on, and almost lost from sight in, percipience.
For the Listener sapience hardly emerges from a sentient pre-sapience.
In practice full human sapience involves a creative dialectic between these two poles - except when it gets lost in theoretical reflection. Then we are liable to be confused by contradictions or stuck in the partial perspective of one pole or the other.

26. AFTER-WORD: POST-SAPIENCE

We live in an era in which the number of ways and means of representation, replication and simulation is increasing at an explosive rate. Postmodernism is in part an expression of a new awareness of the relation between consciousness and cultural representations. At present this seems almost entirely negative, a 'deconstructive' attack on traditional representationalism, coupled with an understandable refusal to make positive assertions of any kind. (Griffin's[19] so-called 'Constructive Postmodernism' fails to grasp the new relation to language that is the main point.) Language took us from pre-sapience to sapience, and the more apocalyptic and milleniarist postmodern writers suggest that our multi-cultural, mult-media world is reaching a new critical point in the history of consciousness, perhaps even to the edge of some form of post-sapience. There are alarming aspects to this, especially when the seem to imply the rejection of the positive advances of the Enlightenment and a descent intoirrationalism

If we are to make sense of this situation and to do something positive about it we need to find the ways of talking, and ideally a formal system or general logical theory, of representational systems. As outlined in this paper, the evolution of consciousness from first life to the impending emergence of AI, has gone through a whole series of levels and modes of representation: structural and maybe functional, depictive and analogue, performatory and representational media, inter-reflective, meta- & para-representation etc. A proper theory of these is essential.

To link this to ANPA's real interests, it has long been clear to many people that there is a close connection between the problems and paradoxes of consciousness and those of quantum mechanics. I have no doubt that the latter are a direct result, not of contradictions and deficiencies in our understanding of 'Reality' itself, but in our understanding of our own cognitive relations to our experimental and mathematical methods and to the semiotics and symbolisms that make up the practice of quantum mechanics. Vague representationalist talk of 'observation' and the measurement problem understood in the context of objectivity and subjectivity will not longer do. As a first step we need to start making a proper recognition of the epistemological gulf that separates the use of depictive and analogue mathematical expressions.

---

[1]McGinn, C. *The Problem of Consciousness,* Blackwell, '91.

[2] Rorty, R. *Philosophy and the Mirror of Nature,* Princetown U.P., '80.

[3] Von Neumann, J. *Theory of Self-Reproducing Automata.* Illinois U.P. '66.

[4] Edelman, G. M. *The Remembered Present.* Basic Books, '89.

[5] Aleksander, I. *Impossible Minds,* Lecture, Imperial College, London, Dec. '96.

[6] Merleau-Ponty, M. *The Phenomenology of Perception.* Routledge, '62.

[7] Stich,S.& Nichols,S. 'Folk Psychology:..' In Eds. Davis,M.&Stone,T. *Folk Psychology.* Blackwell, '95.

[8] Edelman, G. M. *Neural Darwinism,* Oxford U.P.'89.

[9] Davenport, R.K. 'Cross-Modal Perception In Apes...' *Annals of the New York Academy of Sciences* 280, '76.

[10] Meltzoff, A.N. 'Towards a Developmental Cognitive Sceince...' *Annals of the New York Academy of Sciences* 608, '90.

[11] Halliday, M.A.K. *Language as Social Semiotic.* Arnold, '78.

[12] Sperber, D. & Wilson, D. *Relevance.* Blackwell, '86.

[13] Rosenfield, I. *The Strange, Familiar and Forgotten.* Picador, '92.

[14] Searle, J.R. *Intentionality.* Cambridge U.P. '83.

[15] Taylor, C. *The Ethics of Authenticity.* Harvard U.P. '92.

[16] Abdulla, F. Personal communication. '96.

[17] Yates, F.A. *The Art of Memory.* Penguin '66.

[18] Nagel, T. *The View from Nowhere.* Oxford U.P. '86.

[19] Griffin, D.R. *Founders of Constructive Postmodern Philosophy.* State University of New York. '93.

Some Issues on the Construction of Order in Self-Organizing Systems

Juan M. Alvarez de Lorenzana
Institute of International Relations
University of British Columbia
B.C., Canada.
email: alvarez@unixg.ubc.ca

ABSTRACT: Starting from the general concept of "open system" a description of self-organization (SO) is attempted. Then, using Rosen's modeling scheme, a corresponding description of formal representation for SO is established. A review of analytic and synthetic approaches follows in order to determine their adequacy. The construction of order is considered as pivotal in any formal representation of SO and that element seems to be missing from standard procedures. That sets the stage for using the Combinatorial Hierarchy.
KEY WORDS: self-organization (SO), self-organizing systems (SOS), environment (E), self-construction of order, self-ordering systems (**sos**), multiscaling, combinatorial hierarchy.

1. Introduction

As noted by many people (e.g., Salthe, 1993), self-organizantion is increasingly becoming the bedrock of the so-called *complex systems* which, themselves, range the whole spectrum of reality: from the physical universe to human beings and, beyond that, to constructs of human beings such as technological systems. Evolving systems, are the quintessential expression of self-organization (from here abreviated as SO). But, why SO? Why is it so pervasive? In other words, what is it that SO responds to? What is it that SO solves in each system where it is present? These questions are in need of examination. We tend to acknowledge the existence of SO and, from that acknowledgment, go straightway to the business of describing it as best we can, usually leaving aside the task of indentifying, fully and deeply, what the problem dealt with by SO is. I shall attempt to deal principally with the question: why SO? The outcome will be used to determine what is needed to implement formal representations of SO.

2. Open systems

Let us undertake a first approximation to answer "Why SO?". *SO is a specific characteristic of open systems.* "Open systems" are systems that are open (in space, time and matter-energy) to the environment. In order to have and maintain the condition of "systemness" when being open to the environment, an *inner source* of relational ability has to be at play: we call it *self-organization* (**S O**). Therefore, if we are to have "open systems" (and we indeed do), it is because of systemic SO. From the assumption that SO is a key characteristic of open systems, we can infer that, only systems within which SO takes place can exist in unbounded, finite, environments. If proven, and considering how pervasive open systems are (technological systems are included, considering culture, in this case, as part of the social system's environment), this assertion could account for the universality of SO.

When speaking about "open systems" we should understand: (1) systems that exist (i.e., they are viable) in an environment that is bigger than the system (unbounded, yet finite); (2) the environment is invariant with respect to the system's initial conditions (see III below); (3) the environment is the source for whatever matter-energy the system is in need of.

3. The construction of order as modeling the environment

We do know a little bit about open systems: they exist within environments subject to conditions 1 to 3. We have also stated that open systems are characterized by SO. In other words, we assume, as a working proposition, that SO is a sufficient sign that open systems occur. With that understanding and if the proposition is correct, we should seek the overlap between "the conditions" and whatever SO is, in order to develop a definition of systems based on such an overlap; we name those systems *self-organizing* *systems* (**SOS**). With that established, we should determine the proper context from where to draw a formal representational counterpart to SOS. Such a formal counterpart will be named *self-ordering systems* (**sos**).

Let us run again through the conditions 1 to 3 so we can give a description of a SOS. We change the order in the sequence for convenience.

Condition 2: "the environment is invariant with respect to the system's initial conditions". This means that the parameters of the environment that are fixed with respect to the system's dynamics play the role of initial conditions for that system; that is, the definitional characterization of the system is established *from the environment* and by fixing certain of its parameters (see Alvarez de Lorenzana, 1993). We can say that the relation between *environment* (E) and system (SOS) is that of interacting under a set of constraints (the set of "fixed" initial conditions). Those environmental constraints are *definitionally embedded* into SOS (in the same way that axioms are embedded into theorems). Those constraints are *necessary conditions* for SOS to exist within E. They are necessary because without them there could be no system defined; they are not sufficient because any systemic realization would need, in addition to the initial conditions, systemic laws such as a process dynamics (the equivalent of an operation or inference rule in the case of axiomatics).

Condition 3: "the environment is the source for whatever matter-energy the system is in need of." This condition deals with *boundaries* (initial as well as any other within the system's spatio-temporal relational web). This condition originates and regulates the dynamical unfolding of the system through E. It plays an important role in estimating the *potential* for systemic expansion. The conjunction of both conditions, 2 and 3, establishes the potential of a particular system on a particular environment to exist and span. Beyond that there is still the element of novelty coming from outer-E conditions, that is, phenomena not encompassed by or within, E.

Condition 1: "systems that *exist* [...] in an environment that is bigger than the system..." We cannot entertain the possibility of SO within an environment that is smaller than the system itself (it is like thinking in physics about structure within a black hole). Given, therefore, an environment E that is bigger than the system, its existence is equated to the system's viability. The viability of a system within a given E ("given" in the sense of conditions 2 and 3), can be equated with the ability to systemically aprehend its field of interactions based on its initial conditions. The field of interactions for the system is none other than its E. Systemic aprehension of E is the implementation of an internal (i.e., systemic) representation of E. An internal representation of E by SOS is the construction of a model of E by SOS.

In order for a model to be "sufficient" it has to represent E with respect to each and every definitional constraint of SOS, maintaining invariant what could be termed its definitional closure. The other element of condition 1 is the fact that "E is (much) bigger than SOS." So, it is not just a matter for the model to be relevant and accurate with respect to the definitional import of the system. That has to be attained under the added restriction of having to compress E within the system's modeling resources (e.g. amount of memory) without losing the previous attributes (relevance and accuracy). How can E be compressed into SOS when the environment is much larger than the system (E » SOS)? Before any attempt is made to answer this question we should take stock of Ross Ashby's advice about modeling with systems, the components of which totally interact with each other: "... when a designer attempts to design [...] a system in which all the parts interact fully, complexity at the outputs can often be ignored: it is complexity at the inputs of the system that is to be feared" (Ashby, 1972: 90).

From the other side of the issue there are the critics' views saying that only oversimplified models can be implemented because any approximation to the intricate nature of the relationships of any "real" system, instantly brings about a combinatorial explosion (the one thing that Ashby, and any model maker, abhorrs). If we try to put together the severe warnings about input complexity and the task of any SOS which is to model E » SOS, we are in the presence of a considerable challenge. And yet, the challenge is met and solved each time a SOS comes into being. If the problem is input complexity, the solution has to be along the lines of maximum reduction of input complexity. That is, the very foundation of the modeling process of E by SOS, under the condition of E » SOS, is *reduction of input complexity*.

What, then, determines input complexity? At the risk of suprising some people, the answer is: input complexity is uniquely determined by the initial number of **distinct** system's components and not at all by the richness of the information to be embbeded by E into SOS. Again, the information to be exchange between E and SOS, has no role in the crucial aspect (Ashby again) of input complexity. Reduction of input complexity due to over simplification of E by SOS is not the answer. I said that the "embedded" information, which I have named "structural information" (Alvarez de Lorenzana and Ward, 1985), has no bearing on input complexity. Structural information I equated with semantics, as opposed to "signal information" which I identified with syntax[1].

Now we deal with the issue of how to reduce input complexity in SOS. Let's follow Ashby's advice: "...how much will the informational quantities be increased if the system is changed from one having no interaction between its parts to one having full interaction between them? (Ashby, 1972: 83). His answer: "...when full interaction is occurring, when the system is a 'whole' in the fullest sense, no mere doubling or trebling of the resources, or even a multiplying by a millionfold, is likely to be of any use" (Ashby, 1972: 84). More on this in (Alvarez de Lorenzana, 1989; 1991). Ashby's "full interaction" or "the system is a 'whole' in the fullest sense" is the mathematical equivalent of defining a linear or total ordering over a set of elements. To have such an ordering relation is equivalent to saying that any element (of the given set) is biunivocally related to (is distinct from) any other. A total ordering over a collection of components (Ashby's "parts") provides for a maximum in input complexity. This is the worst case scenario according to Ashby.

The opposite to a total ordering is no initial ordering relations among system components (Ashby's "no interaction between its parts"). Such an unordered collection can only be mathematically represented as a singleton. SOS should maintain a cap on the combinatorial explosion while, at the same time, allowing system's components to interact among themselves and with the environment. Also, it should be clear that, given the fact that E » SOS, the combinatorial explosion will always take place, the choice being only between sooner or later. The only way to minimize the *pace* toward combinatorial explosion is to make the interactions *scale bounded* and to start from the minimal possible scale. Once the system is ordered in relation to the smallest relevant scale, a new, also scale-bounded, relational space initiates its construction with, again, as short a range as possible. The process of relational scale construction continues, until the system reaches a last relational scale that encompasses all possibly feasible interactions between SOS and E, *for the given system's potential.* In this manner, the difficult task of modeling E with ostensibly limited resources by SOS, is achieved. The set of scales of interactions of increasing range, establishes a hierarchy of system dynamics, from the smallest to largest scale. Through the process of constructing a hierarchy of scales of system dynamics, an ordering among systemic components is implemented.

So, the material instantiation of such systemic construction of order is the self-organization of the system itself. *The system constructs its own order in the process of organizing itself.* Moreover, it is through this process of order construction among system components that the system increases its ability to model the environment, which in turn increases its viability.

The modeling scheme for evolutionary systems proposed years ago has some unsettling formal implications: we start with an *undifferentiated* system (Alvarez de Lorenzana & Ward, 1987), that is, a system made out of a collection of indistinguishable (non-distinct) components. No matter how many initially given components there are, from the point of view of ordering relations the collection is a singleton. Mathematically this amounts to the unorthodox situation whereby a finite collection of components, with cardinality Card = |M|, M » 1, has ordinality Ord = 1. So we have a finite collection with **cardinality different from ordinality** which is something not permitted in set theory.

4. Synthetic and Analytic approaches to modeling

Let us break-off from our train of thought for a moment in order to bring in informed views on modeling. I am thinking of Robert Rosen who has an extensive work on the subject (Rosen, 1991, 1993, 1995). In his seminal book of 1991, Rosen divides the area of modeling in two main approaches: synthetic and analytical.

(A) Synthetic models: The very essence of this modeling scheme is the underlying assumption that a system, X, can be comprehensively conveyed by the union of disjoint (non-overlapping) subsystems or fractions of X. Moreover, those "fractions" have to be context-independent so that a particular definition is valid in any environment (Rosen, 1995: 33). The mathematical representation of such a system takes the algebraic form of direct sums (in set theory, union of disjoint sets). This is a clear reductionist view. By accepting fractionation it is implied that none of the connectivity among systemic components has any relevance for the system as a whole.

(B) Analytic models: This approach starts with the whole system as a given and then proceeds to tear it to pieces. The system is defined as a cartesian product or direct product. Any piecewise solution will be implemented by means of quotient sets. Because in this case what is given is the totality of the system, we cannot assume disjointness among subsystems and, any severing of the whole in terms of quotient sets, will not permit us to recover the system in full (at least not necessarily). According to Rosen, the inability to recover all the systemic information from the totality of the pieces into which the system has been cut, is a manifestation of the irreducibility of semantics into syntax.

The big dilemma in the synthetic-analytic dichotomy is the following: (1) If we start with a disjoint union of subsystems ("fractions"), we can construct the system. But, by assuming fractionality, we dismiss connectivity between subsystems as irrelevant. This amouts to saying that whatever systemic information there is, it is all inside the subsystems (it is from that inner information within the "fractions" alone, that a system can be built). (2) If we start with the system as a whole, it cannot be partitioned (analyzed) without losing essential information (the connectivity among parts); so much so that once severed into pieces of analysis, the system as such is lost. Rosen's views on fractionality are that only simple systems, i.e., machines, can be fractionated and built from pieces. In this particular class of systems, syntax and semantics coincide. When that is not the case, we are in the presence of complex systems. In Rosen's own words: "The identity of these two quite different ways of talking about 'states' of X is a direct consequence of supposing that 'analysis' into fractions, and 'synthesis' from these same fractions, are inverse operations" (Rosen, 1995: 35).

5. A third option for modeling

The basic consideration I want to make about what was said in the last section is that the conjunction of the two opposite options given by Rosen does not cover all the possibilities for modeling; in fact, it lives out a major class of systems, precisely the ones that go under the rubric of SOS. Let me use an analogy. Imagine for a moment the following scenario: We have a two-dimensional spatial map of a particular geographic zone in which the flora of that region is plotted. We are faced with the task of determining the kind of ecological dynamic that is present between the different species of flora. We have three different map-options that we give to three different observers (Ob): A, B, C. ObA gets a map as a whole picture. ObB gets a map cut into irregular small pieces as is done with puzzles. ObC gets a map that can only be looked at by means of a *zooming device*; such a device imposes the constraint of an inverse relation between its resolving power and the range of view, that is, if we want to look at the map with outmost detail (a point or pixel), then, the area that we can cover is the point itself and, possibly, a few other points within a short-range radius. If we want to "see" a broader area we will have to do so by diminishing, in a comparable amount, the degree of resolving power, i.e., we would have to give up detail in order to attain a broader view. Now we ask the first two observers, ObA and ObB, to causally explain the relations among the different species of flora present in the map. In this scenario:

(1) ObA will have to do an analytic modeling; ObB, a synthetic modeling.

(2) Although ObB has no whole picture of the map, for him, each and every piece (element) is *distinct*, that is, is distinguishable from and among all others. The fact that, for ObB, the elementary pieces are distinct is a remarkable property and it should be looked at in some detail.

Let us assume that the number of pieces is *n*. If we were to have *n* non-distinct, pieces, we could only define **n** one-dimensional relational spaces (i.e., spaces whose only element is related to itself). If, on the other hand, we are considering (as Rosen seems to do) *n* distinct pieces, we would have defined **one** *n*-dimensional relational space. The dimensionality of this latter relational space turns out to be the same as the dimensionality of the analytic option, the reason being that if we can distinguish any element from any other it is because those elements are embedded within a relational space of the same dimensionality as the number of elements, that is, *n*. So, as long as the observers are working with *n* distinct elements, whether analytically or synthetically, both are *de facto* in an *n*-dimensional relational space. Both, analytical and synthetical approaches, although very different in some respects, are implemented or take place, dimensionally, in one and the same relational space: at least up to the total number of "pieces" where, after all, the connectivities among systemic components are defined.

(3) Dimensionality is of the utmost importance because it tells us that what is giving ground to the condition of "distinctiveness" among initial building blocks (i.e., "pieces" or "fractions") of the puzzle is the fact that the representation takes place within one macro scale only (that is, a scale that encompassed all the possible connectivities among "pieces" or "fractions"). This is equal to assuming that the same ordering relations are present throughout, that is, in both local and global phenomena from the outset (it amounts to assuming that every fraction has relational access to every other "fraction").

(4) ObA has access to the map in a way that ObB does not. ObA shares ObB's total ordering (which comes from the distinctiveness of the pieces) and also carries an

ordering that determines the relative position (coordinate) of all the pieces within the map as a whole. This latter ordering is conveyed through the cartesian product and it amounts to making it very easy to figure out what the relation is between any point and the rest of the picture. ObB on the other hand has to make the puzzle (i.e., construct an ordering) before reaching that "easiness".

(5) ObA has no ability in establishing local and global relationships within the picture because, as with ObB, ObA's map is defined on one macro scale: the scale given by the initial distinctiveness (total ordering) among pieces which, in the case of ObA, are labeled as coordinates. From this example and back to Rosen's options we can conclude the following: Although the two options for modeling (analytic and synthetic) are very different in the way they label the definitional set representing natural systems, and although the analytic one does carry some extra ordering referring to the system as a whole, both methods share the inherent limitation of not having a *zooming device* (such as with ObC) that would allow them to locate specific classes of system states (i.e., classes of coordinates for ObA or classes of pieces for ObB, in Rosen's examples) to specific spatio-temporal systemic levels of a developmental dynamics. Without this "device" in the representational language, we are forced to renounce modeling any natural system carrying the ability of constructing its own systemic order (going from local to global connectedness). The "device" is an *encoding device* that would allow us to transfer a set (or hierarchy) of scale-bound processes taking place within the natural system (SOS) into a set of levels in the formal system (**sos**); without such an ability to transfer information between natural and formal systems, there is no modeling relation possible (that is, the modeling diagram would not commute).

To summarize: the monoscaled (one and only macro scale) nature of both analytic and synthetic modeling, overshadows their differences and reduces their applicability to systems where self-construction of order cannot be formally implemented in a genuine way. We need to formally represent inner systemic scale dependency in order to model systems which are open to the environment and that come into being as a restriction of that environment, and become viable because of their ongoing modeling process of that environment. Another way to state the critical question that we are trying to grapple with is the following: does a process of (systemic) self-construction of order necessarily carry information about the genesis of the process itself? If so, should its formal counterpart carry it too and how? Rosen's two options for modeling pre-empt the previous question in the sense that both assume that all the information is present from the outset. That being the case, nothing is generated and we only need to pick up the information that is already there. And the reason why we only need to pick up what is already there is because we are looking from the outside. Rosen's analytic and synthetic perspective is at all times what, in terms of Otto Rössler (Rössler, 1987; see also Salthe, 1993: 166-173), would be called *exomodeling*: modeling from the point of view of an external observer. That is why, when Rosen talks about the synthetic building blocks (see synthetic model above) he defines them as a direct sum of (disjoint) pieces: he is in front of a *pile of pieces* of a puzzle.

Perhaps the most essential point of this discussion is to clarify why it is that Rosen interprets direct sums as a family of disjoint *but distinct* elements. The answer is again: because he is in the position of an external observer looking from a mono (macro) scale. From the SOS perspective, the only feasible position for an observer, has to be an internal view, that is, the observer is inside the system. What SOS does is -Rössler again- *endomodeling*, that is, modeling from inside the system. By defining the system as a restriction of the environment (condition 2, section 2 above) we are doing two things: (1) we fix the environment in terms of a set of necessary and sufficient initial conditions (equivalent to axiomatic closure in mathematics), that allow the system to be viable in principle; (2) we define the system internally at its maximum degree of potentiality, that is, at its point of indistinguishability: in space, in time and in class membership of components (see Alvarez de Lorenzana 1991; 1993: 305-308). This second condition translates, mathematically, into saying that the direct sum of system components is a singleton, i.e., the components form only one equivalence class. In other words, we don't have a pile of "distinct pieces" of the puzzle but a *bag of indistinguishable components*. By "bag" we mean that we know or can measure the numerosity of the thing, that is, its cardinality. But in terms of ordinality, we can only know or measure what is distinct in that collection and, in an equivalence class, the only distinct thing is the class itself and nothing else. The term "bag" is used in computer science (Scheurer, 1994) or, equivalently the term multiset (Blizard, 1989).

6. The construction of order in self-organization

The previous survey on SO and SOS, should provide a general view of basic requirements for their formal representation. The key element to point out is the role played by the systemic construction of ordering relations: it is the very core of SO. In other words, *self-organization is for natural systems what self-construction of order is for formal systems*. SO is a realization (in the mathematical sense, see Rosen, 1991) of self-construction of order. If we agree on that, then knowledge about the way in which such a construction takes place is essential. The deeper and sharper our formal representation is, the better chance we have in developing a good understanding of SOS and the better corresponding ability to make models for simulation and experiment.

It is my assessment so far that a sound and clear formal scheme for systemic multiscaled phenomena is lacking. This is so whether in statistical or analytical mechanics. Even non-linear approaches seem to be in difficulty when trying to represent causality within systems spanned over three or more scales of interaction. If that is the case, then, none of those options would give us the necessary degree of conceptual sharpness and overall control required for multiscaled SO. Only in the use of algebraic approaches have I found a power of description capable of matching the intricacy and subtlety of the phenomena at play. Let's be clear and say it again, I am referring to SO of complex systems, that is, systems that carry and develop SO and therefore implement self-construction of order to the fullest, in other words, make use of the full range of the hierarchy of scales (that is what the condition E » SOS means). It is precisely the multiscaling nature of systemic phenomena that makes in my view, those formalisms either exceedingly difficult or artificially cumbersome or ad hoc, or not sufficiently specific. There is a serious problem with any realm of phenomena that takes place across several scales. Physics, the "hard" and mature natural science par excellence, is a good example of such a problem; the many formalisms attempting unification of the four basic forces is a sample of the unfinished struggle in the formalization of multiscale phenomena. Non-linear dynamics and chaos work well within the micro-macro paradigm. This two level framework does not allow careful formalization of more than two levels. Remember that scales of interaction exist in SOS as a way to curtail the ever-present risk of combinatorial explosion (carrying, if not avoided, the breakdown of the system or, at best, partial collapse into previous levels). At the same time the viability of the system is based on good modeling of the environment (which means the intrinsic necessity of increasing the scale of system interaction across the environment).

The systems' existence lies in between two opposing needs: the need to expand and the need to control. If it does not expand to favour control, it becomes vulnerable because the control is exerted over too small an area. If it tones down control, favoring expansion, the system also becomes vulnerable because it loosens its grip over the covered environment. The only way to negotiate between those two opposing needs while attaining SO is by means of scaling, not any arbitrary set of scales, but the set that both minimizes interaction growth and maximizes control. There is a very critical balance that has to be maintained and optimized. It is my view, that the algebraic road is the one to chose. It is also my belief that once such a formalism is well established it could make incursions into analytical and statistical mechanics if needed.

7. Implicit assumptions on order in present mathematics

Before ending the description of this viewpoint, we should be aware of the fact that algebra is also marred with implicit assumptions that are questionable. The advantage with algebra is that it has ways to circumvent or overcome, to a considerable degree, its shortcomings. Category theory is an important, relatively new, tool (see Golblatt, 1979; Gago, 1986) that allows to formalize construction of order. The questionable assumptions that I just mentioned are present across the whole realm of mathematics. In particular, I am referring to taking as a given the existence of natural numbers. Kronecker's well publicized phrase declaring that only the whole numbers came from God, all else is the work of Man, expresses the rational grounds for such a claim. According to mathematicias, N is the most elementary entity to be encountered and therefore has to be assumed: "We cannot expect that the natural number sequence can be defined in terms of anything essentially more primitive than itself..." (Stoll, 1979: 57). Such an statement is shocking coming from the SO side of things, where order of any kind is not trivial and therefore has to be constructed. We have seen how difficult the construction of order is in the world of the real where the threat of a combinatorial explosion never fades away. From our vantage point, to assume N as a minimal proposition is not only not necessary but very damaging for any formal representation of self-organization.

References.

Alvarez de Lorenzana, J.M. (1989)."Las Relaciones Internacionales en el contexto de la historia. Un caso de sistemas evolutivos jerárquicos." Ph.D. Dissertation. Madrid: Universidad Complutense de Madrid.

____________________ (1991). "Informal Comments on Indistinguishability, the Combinatorial Hierarchy and Evolutionary Systems Framework." In F. Young (ed.), Proc. 7th Ann. Mtg. ANPA West, Stanford, Palo Alto, CA.

____________________ (1993)."The constructive universe and the evolutionary systems framework." In S.N. Salthe, *Development and Evolution. Complexity and Change in Biology*. Cambridge, Mass.: The MIT Press.

Alvarez de Lorenzana, J.M. and L.W. Ward (1985). "Semantic and Syntactic Information." Proc. 29th Ann. Mtg. *Society for General Systems Research*. Louisville, Kentucky; pp. 78-86.

______________________ (1987). "On Evolutionary Systems." *Behavioral Science*, vol. 32, pp. 19-33.

Ashby, W. Ross (1972). "Systems and Their Informational Measures." In Klir, George (ed.). *Trends in General Systems Theory*. New York: John Wiley & Sons.

Blizard, Wayne D. (1989). "Multiset Theory." *Notre Dame Journal of Formal Logic*, vol. 30, pp. 36-66.

Casti, John (1989). "Newton, Aristotle, and the Modeling of Living Systems." In Casti, John and Anders Karlqvist (eds.). *Newton to Aristotle. Toward a Theory of Models for Living Systems*. Boston: Birkhäuser, pp. 47-89.

Gago, Felipe (1986). *Topos Theory, Why?*. Santiago de Compostela: Universidad, Servicio de Publicaciones Universidad de Santiago de Compostela.

Goldblatt, Rob (1979). *Topoi, the categorical analysis of logic*. New York: North-Holland Pub. Co.

Rosen, Robert (1991). *Life Itself. A Comprehensive Inquiry into the Nature, Origing, and Fabrication of Life*. New York: Columbia University Press.

__________ (1993). "On Models and Modeling." *Applied Mathematics and Computation*. vol. 56, pp. 359-372.

__________ (1995). "The Mind-Brain Problem and The Physics of Reductionism." *Communication and Cognition-Artificial Intellegence*, vol. 12, pp. 29-43.

Rössler, Otto E. (1987). "Endophysics." In Casti, John and Anders Karlqvist eds. *Real Brains. Artificial Minds*. New York: North-Holland.

Salthe, Stanley N. (1993). *Development and Evolution. Complexity and Change in Biology*. Cambridge, Mass.: The MIT Press.

Scheurer, Thierry (1994). *Foundations of Computing. System Development with Set Theory and Logic*. Reading, Mass.: Addinson-Wesley.

Stoll, Robert R. (1979). *Set Theory and Logic*. New York: Dover Pub. Inc. (republication from W.H. Freeman and Co., 1969).

[1] There is, of course, an important relationship between semantics and syntax (as portrayed in that paper and to which I still subcribe). Syntax is a carrier of semantics in a process of increased refinement of discrimination over an initial, undifferentiated, semantically closed system. A good analogy is an axiomatic system in mathematics, where, starting with axiomatic closure, theorems are constructed. Clearly, the theorems of such a system are syntactical constructs. They also convey a certain mathematical "meaning" that escapes syntax but which is specific to that particular system of axioms (e.g., euclidean triangles have a different "meaning" than hyperbolic triangles). The "meaning" of the theorems comes from the specific nature of the set of axioms, but the consistency of the theorems has nothing to do with that particular set of axioms (save axiomatic closure) and everything to do with the rules of inference (syntactic rules), the stuff of which theorems are constructs. So, the "meaning" of the axiomatic structure does not have any inmediate impact on the theorems. The axiomatic structure could be relevant or irrelevant (e.g., in relation to possible usefulness of the theorems), but that aspect of the axiomatic structure does not have an effect on the theorems. This analogy (for a more in-depth look see Rosen, 1991) should clarify any doubts about the necessarily separate realms on which both semantics and syntax operate. Conversely, it also shows the impossibility to reduce semantics to syntax alone because the roots of semantics lie, ultimately, outside the system in question (are metasystemic).