SLAC – PUB – 4008
June 1986
(T/E)

ON THE CONSTRUCTION OF RELATIVISTIC QUANTUM THEORY: A Progress Report*

H. PIERRE NOYES

Stanford Linear Accelerator Center
Stanford University, Stanford, California, 94305

ABSTRACT

We construct the particulate states of quantum physics using a recursive computer program that incorporates non-determinism by means of locally arbitrary choices. Quantum numbers and coupling constants arise from the construction via the *unique 4-level* combinatorial hierarchy. The construction defines indivisible quantum events with the requisite supraluminal correlations, yet does not allow supraluminal communication. Measurement criteria incorporate $c, \hbar$ and m_p or (*not* " and") G, connected to laboratory events via finite particle number scattering theory and the counter paradigm. The resulting theory is discrete throughout, contains no infinities, and, as far as we have developed it, is in agreement with quantum mechanical and cosmological fact.

Submitted to *Proceedings of the 7th Annual International Meeting*
ALTERNATIVE NATURAL PHILOSOPHY ASSOCIATION
King's College, Cambridge, England, 28 June - 1 July, 1985

* Work supported by the Department of Energy, contract DE–AC03–76SF00515.

1. INTRODUCTION

Although a successful challenge to the experimental predictions of quantum mechanics has yet to be mounted, and subtle features such as the supraluminal correlations without supraluminal signaling implied by Aspect's[1] and other EPR-Bohm type experiments have been demonstrated, for some physicists a conceptual unease continues to persist. I present here my current attempt to meet this problem.

What framework do I accept for physics? I believe that most practicing physicists would agree with me that physics is an empirical science. Physics in historical practice has rested on quantitative measurement, or at least on "operational procedures" which lead to "replicable" results. Here I will insist on the stricter standard that the results of experiment be reduced to "counting", taking due account of the (again specified) expected range of uncertainty. I cannot accept the concept "infinite" (or "infinitesimal") as valid in physics; for us finite beings (physicists or no) this great renunciation is (in my opinion) the first step toward acquiring knowledge. Quantum mechanics brought this issue to the fore for physicists; it had been raised for chemists by Dalton and Prout long before, and for philosophers and mathematicians by Democritus and Zeno in antiquity.

For me, any formulation of quantum mechanics as we know it must contain the idea that quantum events are unique and indivisible. In contrast, classical physics is *scale invariant*; any arbitrarily chosen standards of mass, length and time- or any three experimentally independent combinations of those unit standards will suffice. There is in it no place for unique events; the "microscopic" laws are "time reversal invariant". Events (which in classical physics are always in principle decomposable into "microscopic" substructure) acquire what uniqueness they possess due to the imposition of boundary conditions by the analysis of the physicist, or their embedding in a large-number "statistical" background, or their relation to a "cosmological time", or ... Modern physics removes the scale invariance of classical physics by recognizing a limiting velocity, a quantized unit of action (and angular momentum!), and quantized masses. Experimentally the only stable (lifetime $> 10^{33}$ years) "elementary" mass values are those of the proton m_p and the electron m_e in the ratio $\frac{m_p}{m_e} = 1836.1515 \pm 0.0005$, again stable within the stated standard deviation. (According to current standard cosmologies the stable nuclei that could have at least as long "lifetimes" (eg. He^4) are only about 10^{16} years old, or less; whether protons have an "age" is less clear.)

In my approach, I adopt from classical physics the dynamical definition of mass ratios from Newton's third law as articulated by Mach, but generalised to recognise the limiting velocity by using mass invariance $(E^2-(\underset{\sim}{p}\cdot\underset{\sim}{p})c^2 = m^2c^4)$ and 3-momentum conservation $(\Sigma \underset{\sim}{p}_{initial} = \Sigma \underset{\sim}{p}_{final})$. Experimental contradiction of this assumption would be of immediate practical interest for those interested in the exploration of the solar system and beyond! The so far undefined energy (E) in this equation is a global quantity. As Wick saw in the late 30's[2] the easiest way to make a compelling argument for Yukawa's finite range meson theory of nuclear forces[3] is to insist on (relativistic) 3-momentum conservation, but allow energy fluctuations consistent with Heisenberg's energy-time uncertainty principle and Einstein's mass-energy equivalence. This is also a basic principle underlying Heisenberg's and Chew's S-Matrix theory.

There is already a well known conceptual puzzle at this stage in our discussion. If we fasten on macroscopic (gravitational) rather than microscopic (particulate) phenomena as basic, the fundamental mass unit we would choose would be the "Planck mass" $M_{Planck} = \sqrt{\hbar c/G} \simeq \sqrt{1.7\times 10^{38}}m_p$ rather than the proton mass. Contemporary physics meets this problem by using the "equivalence principle". The postulated equivalence of microscopic ("inertial" or 3-momentum conserving) mass ratios and macroscopic ("gravitational") mass ratios allows gravitation, and (perhaps) all other "interactions" along with it, to be "geometricized". But this need not be the only route to "supergravity", or whatever catch phrase becomes current when this paper appears. In my opinion, one of the strengths of the approach to physics developed here is that our theory can have only one type of mass, and that the first approximations to both M_{Planck}/m_p and m_p/m_e are calculated.

Once one accepts quantum events as basic, and the limiting velocity as well established experimentally, the "supraluminal correlations" [1] predicted - but for some people not explained - by quantum mechanics also cry out for conceptual clarity and a deeper insight. I do not believe that this can be achieved by first developing the full technical apparatus of quantum mechanics and then presenting these startling results as a deductive consequence. Etter, McGoveran, Manthey,

Gefwert and I[4] believe that the issue *can* indeed be clarified by invoking a minimal set of postulates that do not depend on the idea of space-time, let alone quantum mechanics. What follows in this paper is consistent with that point of view, and with earlier papers.

Another aspect of contemporary physics that I would like to see emerge at an early stage is that our "universe" start out as simply as possible and evolve by a finite number of steps through recognisable stages into the complex situation that we encounter as we now explore it. When I started on this research I was at least open to the possibility that the universe we are exploring is "indefinitely extensible" in both "space" and "time". That we find great simplicity as we retrodict the past on the cosmological scale, could (as Bastin has often emphasised) simply be a consequence of impoverished data, – i.e. of the successive disappearance of relevant observable points of reference as we extend our horisons. I do not think this problem arises in acute form while retrodicting the last 15 billion years. I have been greatly impressed during the course of my own professional career by the convergence of seemingly disparate and very detailed measurements to a reasonably consistent "time scale" of that length. The past was *different* from the present in the probable range and type of configurations that occurred, but there is no indication as yet that the elementary *possibilities* were significantly different (except, possibly, during the *very* early stages). In the current paper, the very early stages of the evolution are *simpler* and not just different. For those who are more comfortable with a universe that has no beginning and no end John Amson provides a nested hierarchy which can be explored (past ↔ future; small ↔ large) so far as information is available (Appendix VI by John Amson entitled *"Bi-Orobourous" - a Recursive Hierarchy Construction*). So far as I can see, the consequences when this point of view is articulated in the current practice of physics are likely to be practically indistinguishable from those of the approach developed here.

To the best of my knowledge we can retrodict the universe backward in time for only about fifteen billion years. There is an "event horizon" and a preferred coordinate system defined by the radiation that broke away from the matter when the cosmic fireball became optically thin; within the event horizon there are particles whose baryon and lepton number add up to approximately the square of $2^{127} + 136$. Our construction yields all of these observed features as stable consequences of the construction independent of the details. That the theory developed here has a starting point and achieves evolving complexity possessing dynamic and heritable stability in the presence of a "random" background of quantum events, is for me a satisfactory result. The "universe" we construct in this paper will go on increasing in complexity in the future, and hence contains indefinitely extensible possibilities. This theory has a fixed past, an event horizon, and yet and indefinitely extendable (though uncertain) future. It may be that I have found what I was looking for, but I can assure the reader that the steps along the way were taken for more immediate reasons, - so far as I am aware.

In this current attempt to meet these basic requirements when reconstructing quantum theory I have made use of many ideas and techniques conceived and developed by other people[5,6]. In the series of papers on "Concept of Order"[7,8] Bastin and Kilmister presented reasons why distinct "events" should relate to a basic algebraic structure connected to "3+1 space". By 1966 this research, to which Amson, Bastin, Kilmister, Parker-Rhodes and Pask had all contributed, had led[9] to the closed 4-level *combinatorial hierarchy* with the cumulative cardinals $3, 10, 137, 2^{127} + 136 \simeq 1.7 \times 10^{38}$.

The work on the combinatorial hierarchy did not face directly the statistical aspect of quantum mechanics, which I have already indicated I see as fundamental. I therefore start my technical discussion in Chapter 2 by calling on more recent work by Manthey and McGoveran to spell out what current computer practice means by "non-determinism" and "arbitrary choice". As the names of Amson and Pask will indicate, the earlier work had also made use of concepts used in computer science, but before the nondeterminism born of asynchronous communication over a shared memory had come to the fore. I turned that way because Gefwert demanded that anything that laid claim the description "constructive physics" had to be computable. Fortunately the expert I first turned to was Manthey; the result was PROGRAM UNIVERSE.

Our use of a computer simulation to model the theory is sometimes misunderstood. I do not think of the universe as a "big computer in the sky". What the coding does for us is to keep us honest; if we can show that the program is indeed computable, then we have protected ourselves from making all sorts of logical errors. A computer simulation is a specific type of "model". If it succeeds, all that we can say is that, within current experimental error, we have succeeded

in isolating those aspects of experience which act in a manner isomorphic to the action of our model. When the program fails, then we will have isolated a situation from which we might learn some new physics, or possibly something that goes beyond physics. There can well be things in heaven and earth that are not contained in this philosophy. I trust it is clear that I am not a reductionist or a mechanist. Materialism is a separate issue, which will not be discussed here.

PROGRAM UNIVERSE, a peculiarly simple algorithm, *automatically* develops some representation of the hierarchy, necessarily specifies unique, correlated, global events and provides address ensembles for these events labeled by the fixed elements (eventually connected to quantum numbers, masses and coupling constants) provided by the combinatorial hierarchy. The technical details are given in Chapter 2 where we provide a specific construction of the four level hierarchy and the address ensembles; Mike Manthey's coding for this construction is given as Appendix IV.

In order to meet our objective of constructing a quantum mechanical physical theory, we must somehow relate the structure we now have in hand to measurements of mass, length and time in the ordinary sense. We do this in Chapter 3 by noting that quantum events "fire counters" and allow velocities, momenta and energy to be measured by well known techniques to an accuracy only limited by available budget (or space and time available to conduct meaningful experiments). As in Heisenberg's and Chew's S-Matrix philosophy, momentum measurements, and the momentum space formulation of quantum mechanics are a strategically useful place to connect theory to experiment. We make this more than usually explicit by starting from a "counter paradigm" which relates the bit string universe to laboratory measurement. We find that the relativistic version of the "wave-particle dualism" emerges with little effort. We also discover that some fundamental cosmological observations find a ready explanation within this simple framework, independent of the details by means of which it is articulated.

The next step, spelled out in Chapter 4, is to construct from the strings and events provided by PROGRAM UNIVERSE a relativistic quantum scattering theory which, via the counter paradigm, conserves quantum numbers and 3-momenta in a manner consistent with laboratory experience. The basic idea in this scattering theory is to use Faddeev-Yakubovsky equations for the dynamics rather than a Hamiltonian, or Lagrangian or analytic S-Matrix formulation.

Probably the most significant step taken since ANPA 7 is the derivation of the "propagator" for the scattering theory directly from the bit string universe via the counter paradigm. We give a more detailed explanation here than was possible in our conference report[10] .

In Chapter 5 we return to the four level hierarchy labels and connect them to the conserved quantum numbers in the standard model for quarks and leptons. Briefly, level 1 describes the simplest neutrinos, level 2 describes electrons, positrons and the associated electromagnetic quanta, while level 3 gives us two flavors of quarks and the associated gluons related to each other in a color octet. This pattern will repeat until the possibilities close off at the Planck mass, but the coupling between successive generations will be weak because of the combinatorial explosion in possibilities. Thus we can anticipate that the Kobiyashi-Maskawa mixing angles will indeed be small. The count of quantum numbers is correct, and the quantitative or qualitative results so far achieved produce no glaring contradictions.

For completeness we repeat in Chapter 6 the Parker-Rhodes calculation of m_p/m_e as it looks from the present context. The question of whether the result will be stable when we compute the correction to the fine structure constant and the "recoil corrections" is still unanswered.

We will discuss in the concluding section what it might mean if this qualitative-quantitative success persists up to a point where a definite quantitative conflict with experiment counter-indicates the acceptance of the theory.

2. GENERATING AND DISCRIMINATING BASIS STATES; EVENTS

The first problem we must face in constructing our theory is where the "random" aspect of quantum mechanics enters. As one will see in Appendix I, Kilmister allows his generation and discrimination operations to be interleaved in an order which is not explicitly specified. The route we follow below, which depends on the explicit use of a "pseudo-random number generator" in the computer program, is in my opinion a specific articulation of Kilmister's more general discussion. In Appendix II Bastin discusses, among other things, the idea of "inexact matching" which he and Kilmister have explored; this might also end up in something that could be shown to be equivalent to my approach, but has as yet not been articulated far enough to settle the issue.

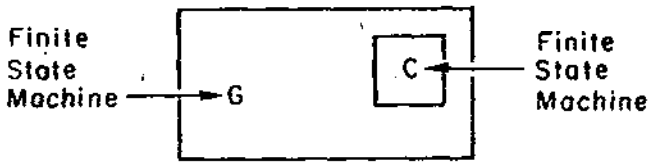
The method of actually writing down a computer program forced Manthey and me to tackle the "randomness" issue head on. For Manthey, the non-determinism born of asynchronous communication over a shared memory – one basic problem in concurrent programming – is viewed as at least analagous to the non-determinism encountered in quantum mechanics[11,12] . Thus the basic coding for the routine which returns either a zero or a one with "equal prior probability" (and whose output is symbolized below by "R") as given by Mike Manthey(in Appendix IV) starts from two memory locations which flip a bit backward and forward on one time interval; one bit is read whenever the (asynchronous) operation of the main program calls up this routine. However, when Manthey and I had occasion to need this routine for an EPR computer simulation we are working on, he fell back on the standard (pseudo)-"random number generator" available on his local computer. I turn to another expert for discussion of this issue.

The term McGoveran prefers to use when talking about what is often called "randomness" is "arbitrary choice". By this he means "not due to any finite, locally specifiable algorithm"; of course in standard practice, the *local* operation of the computer is deterministic (if it is working properly), so this means calling on some number generated in a larger system not under local control. In the same termininology, he would "define" *random* as "not due to any finite, local *or* global algorithm". Since we have no operational way to meet this requirement, the concept of "random" is, strictly speaking, meaningless in our context;

we must content ourselves with the currently available pseudo-random number generators for our simulations. McGoveran's basic thinking on this was spelled out recently[13] in response to a query from Kilmister. I quote:

"I think that we must insist on computability to the detriment of randomness for a number of reasons, each of which I have previously discussed. This position does no harm to the power of the model since, as proved by Shannon (1965), – an infinite state machine with a random element can be replaced by an infinite state machine (*infinite* being "constructively infinite"), and as I demonstrated (1984 ANPA West Proceedings), there will always be a method for constructing certain repeating binary inputs which a given finite state machine with finite memory can not distinguish from 'random' binary inputs.

"In some sense, randomness is a local phenomenon. So long as a 'generator' is available which has significantly more states than the 'detector', there will be a possibility of generating strings which are random from the detector's point-of-view. Similarly, given a string which passes all computable tests for randomness of a fixed complexity (i.e. by a finite state machine with m possible states), it will be possible to construct a finite state machine with $n >> m$ possible states which produces that string. In algebraic terms: there exists a computable function g (pseudo-random number generator) for each finite class of computable functions f_i such that, whenever the range R_g of g is sufficiently large compared to the union of the domains of f_i (call this D), it is impossible to prove that G is computable based on the f_i. In pictures:



[C sees some outputs of G as random; G sees no outputs of C as random. The computational cost of detecting the orderliness of all G's output is too great for C.]

"Because we have neither the means for specifying what we mean by randomness, nor for detecting it in a finite system, *and* we can be certain that a means exists for constructing *ANY* sequence, I have insisted that we have no need for

the concept of randomness, replacing it with 'arbitrary'. We implicitly recognize an as-yet-unspecified 'finite computable function' as the source of 'arbitrary' strings.

"So long as we are blind to the nature of a deterministic system, the effect is the same *locally* as having "random choice". Furthermore, true randomness *implies* infinities (an infinite state machine is required for generating random output – i.e. undecidable output). I think we need consistency here and so deny randomness in favor of parsimony."

Now that we have spelled out how, in practice we can select either of our two symbols 0, 1 with what is close enough in practice to "equal prior probability" in the frequency theory sense of probability, we can understand the basic arbitrary choice from which the algorithm called PROGRAM UNIVERSE starts. But the generation of the strings in this universe and the discrimination between them require considerably more background if the algorithm is to be followed.

The basic elements with which the hierarchy work started are ordered strings of the symbols 0, 1 of the form $S^a(N_U) = (....., b_n^a, ...)_{N_U}^a$ where N_U is the integer specifying the number of symbols in ("length" of) the string, n is the usual integer ordering parameter $n \in 1, 2, 3, ..., N_U$, and $b_i^a \in 0, 1$. We will discover below that when our construction has proceeded far enough we can specify the label a in terms of the sequence of symbols in the N_L positions $n \in 1, 2, 3, ..., N_L < N_U$. For those who wish the integers themselves to be constructed, one can follow Gefwert's approach in terms of primitive recursive functions[14,15], or follow Kilmister's foundational discussion in Appendix I where in a sense they come to us along with the hierarchy itself. If we define the null string $0_N = (0, 0, ..., 0)_N$, the operation $\oplus$ which tells us whether two strings are the same or different (and hence *discriminates* between them) gives $S^a \oplus S^a = 0_{N_U}$ when they are the same and has two equivalent definitions:

$$(..., b_n^a +_2 b_n^b, ...)_{N_U} \equiv S^a \oplus S^b \equiv (..., (b_n^a - b_n^b)^2, ...)_{N_U} \tag{2.1}$$

whether they are the same *or* different. For the first definition the operation $+_2$ is addition (mod 2), or symmetric difference, or exclusive "or", the symbols are bits and the operation is the standard XOR of computer practice. For the second definition the symbols are integers and we can define operations such as

$k^a(N) = \Sigma_{n=1}^{N} b_n^a$ which gives us the number of "1" 's in the string. This fruitful ambiguity was first noted by Kilmister and myself; we refer to *either* operation as *discrimination* in order to preserve the generalisation that goes beyond XOR. The anti-null string is symbolised by $1_N \equiv (1, 1, ..., 1)_N$, allowing us to define the "bar operation" $\bar{S}^a(N) \equiv 1_N \oplus S^a(N)$ which interchanges "0"'s and "1" 's in a string.

By 1980 Kilmister[16] realised that the discrimination operation by itself would not suffice for the theory, since it gives us no clue as to how the strings arise in the first place. He therefore introduced a generation operation by modifying a construction of the integers used by Conway (originally due to von Neumann), and found out how to go on to arrive at the discrimination operation using this approach. The final(?) version of his approach (which was sketched out at ANPA 7 and completed since) is given as Kilmister's Appendix I. In this way, or by using discriminate closure and the matrix mapping due to Parker-Rhodes[9][17] or the set-theoretic derivation due to John Amson[18] one arrives at the *unique*, 4-level *combinatorial hierarchy* with the cumulative cardinals $3, 10, 137, 2^{127}+136 \simeq 1.7 \times 10^{38}$.

As Kilmister and I soon realised, once one has introduced the generation operation, there is nothing to stop it from generating additional elements even when the full hierarchy has closed off. In terms of the bit-string representations used in my work, this means that the first bits in the string can be put into 1-1 correspondence with *any* representation of the hierarchy, and that as we go on cranking out new elements of still greater length there will come to be many strings with the same label. Kilmister and I called the portion of the string beyond the *label* the *address* and thus arrived at the idea of *labeled address ensembles*. These come to play the role in our theory of *quantum state vectors*, but there are subtle differences from the conventional quantum mechanics which we will discuss at the appropriate point.

When Christoffer Gefwert heard of our work, he saw that constructive mathematics could offer the appropriate philosophical framework in which to achieve consistency, and suggested to me that if we were indeed trying to create a "constructive physics", it would have to be expressible as a computer program. This encouraged me to re-establish contact with Michael Manthey and led to the first version of PROGRAM UNIVERSE[19] . Since we did not see any simple way to

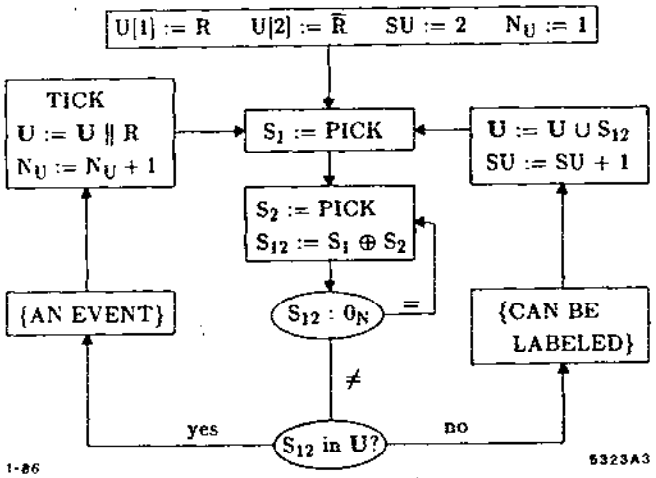
code up Kilmister's generation-discrimination construction, we decided to generate strings in the simplest way we could think of. What we now have is simply described. If there are SU strings in a universe of length N_U, it is allowed to evolve in only two ways. Two strings are picked arbitrarily and discriminated. If the resulting string is not already in the universe, it is adjoined; the number of strings goes up by one. If the string is already in the universe, an arbitrary bit is selected for each string and concatenated with that string; the length of the strings goes up by one. This second operation is called TICK.

We generate the strings according to the flow chart:

Figure 1.The flow chart for Program Universe.

PROGRAM UNIVERSE

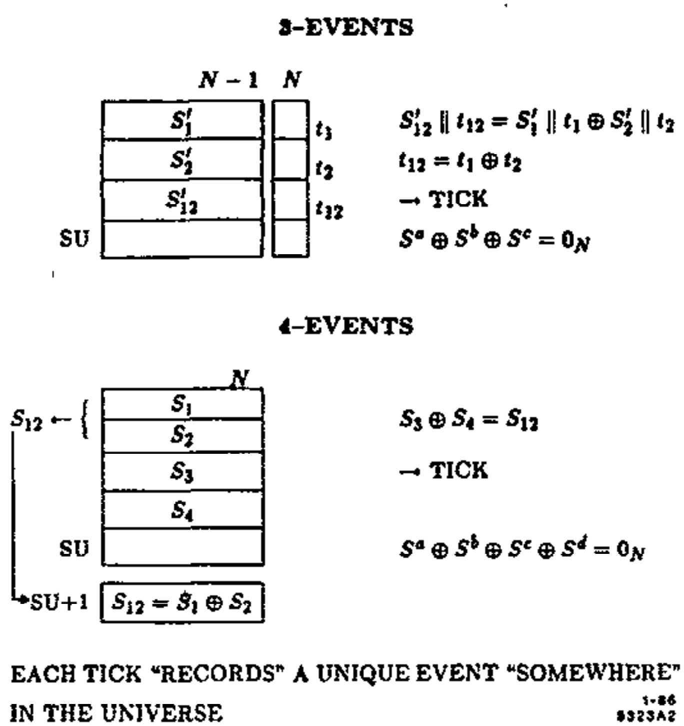
NO. STRINGS = SU   R ⇒ 0,1 (FLIP BIT)
LENGTH = N_U PICK := SOME U[i]   p = 1/SU
ELEMENT U[i]   TICK U := U ∥ R
$i \in 1,2,\ldots,SU$ $\bar{S} = 1_N \oplus S$



For those who prefer explicit coding, this is provided by Manthey in Appendix IV. The program is initiated by the arbitrary choice of two distinct bits: $R :=$ 0 *or* 1, $\bar{R} = 1 \oplus R$. Entering at $PICK$, we take $S_1 := PICK; S_2 := PICK; S_{12} := S_1 \oplus S_2$. If $S_{12} = 0_{N_U}$, we recurse to picking S_2 until we pass this test. [A still simpler alternative, which occurred to me in writing this paper, would be to allow the null string to occur as one of the elements in the universe. So far as I can see, this would not affect the running of the main program, this change might require a little care, and perhaps some change in the coding when we turn below to the extraction of the hierarchy from the results of a run of the program.] The program then asks if S_{12} is already in the universe. If it is not it is adjoined, $\mathbf{U} := \mathbf{U} \cup S_{12}, SU := SU + 1$, and the program returns to $PICK$. If S_{12} is already in the universe, we go to our third, and last, arbitrary operation called $TICK$. This simply adjoins a bit (via R), arbitrarily chosen for each string, to the growing end of each string, $\mathbf{U} := \mathbf{U} \| R$, $N := N + 1$, and the program returns to $PICK$; here "∥" denotes string concatenation.

What may not be obvious is that TICK results either from a *3-event* which guarantees that at string length N_U the universe contains three strings constrained by $S^a \oplus S^b \oplus S^c = 0_{N_U}$ or a *4-event* constrained by $S^a \oplus S^b \oplus S^c \oplus S^d = 0_{N_U}$ that these are the only ways events happen in the bit string universe is illustrated in Figure 2.

Figure 2. How events happen in *Program Universe.*



In the case of 3-events the universe just before $(N-1)^{st}$ TICK contained three strings constrained by $S'_1 \oplus S'_2 \oplus S'_{12} = 0_{N-1}$ which were replaced by $S_1 = S'_1 \| t_1$, $S'_2 \| t_2$ and $S'_{12} \| t_{12}$ respectively as a consequence of that TICK. Before the N^{th} TICK S_1 and S_2 are picked and S_{12} generated by discrimination. Clearly, if $t_{12} = t_1 \oplus t_2$ then S_{12} is *already* in the universe, and the program will proceed to carry through the N^{th} tick. However it can also happen that when two strings are picked the S_{12} generated by discrimination is *not* already in the universe, and hence will be adjoined to it. Eventually however (if the program does not encounter some circumstance that produces a 3-event first) it will pick two strings S_3 and S_4 (which could even be a second pick of S_1 and S_2) such that $S_3 \oplus S_4 = S_1 \oplus S_2$; clearly this will then lead to the N^{th} TICK as a 4-event.

In the original version of PROGRAM UNIVERSE, I was hung up with the idea that only 4-events should occur, because energy *and* 3-momentum cannot in general be conserved in 3-events (a fact familiar to particle scattering data analysts). I therefore went to some elaboration to insure this, and only later stripped down the program to the present form. Once I had done this, James Lindesay then saw that 3-events could also occur by the mechanism just described. As we will see below both are needed in the scattering theory, so this fact turned out to be extremely fortunate. This is only one of many instances in the course of this research where the attempt to arrive at simpler formulations has had profound consequences.

It is important to keep in mind both here and in what follows that the actual structure of the memory and the specific strings in it generated by our computer simulation are *not* to be thought of as modeling "real" elements in the world. We are not allowed to access them directly, even conceptually, when it comes to interpretation. The string length, whether a specific event is a 3-event or a 4-event and how many other combinations of strings satisfy the event constraint at that TICK are hidden from us. We can only talk about them as structural constraints and in terms of statistical arguments. Contact with experiment can only be made *indirectly* via the counter paradigm. This "simulates" in another sense what actually goes on in the laboratory. We can never know "what it is" that initiates the chain of happenings which leads to the firing of a counter. All we can do is to use the connections provided by theory and experiment by means of some more or less successful type of analogical thinking to refine and improve the *statistical* behavior of our counters, or more sophisticated detectors.

In order to see that this program *also* leads to some representation of the combinatorial hierarchy and to the label-address schema, we must first discuss the idea of discriminate closure, originally due to John Amson. Given two distinct (*linearly independent* or l.i.) non-null strings a, b, the set $\{a, b, a \oplus b\}$ *closes* under discrimination. Observing that the singleton sets $\{a\}$, $\{b\}$ are closed, we see that two l.i. strings generate three *discriminately closed subsets* (DCsS's). Given a third l.i. string c, we can generate $\{c\}$, $\{b, c, b \oplus c\}$, $\{c, a, c \oplus a\}$, and $\{a, b, c, a \oplus b, b \oplus c, c \oplus a, a \oplus b \oplus c\}$ as well. In fact, given j l.i. strings, we can generate $2^j - 1$ DCsS's because this is the number of ways we can choose j distinct things one, two,... up to j at a time. This allows us to construct the combinatorial hierarchy[9] by generating the sequence $(2 \Rightarrow 2^2 - 1 = 3), (3 \Rightarrow 2^3 - 1 = 7), (7 \Rightarrow 2^7 - 1 = 127), (127 \Rightarrow 2^{127} - 1 \simeq 1.7 \times 10^{38})$ provided that we can find some "stop rule" that terminates the construction.

The original stop rule was due to Parker-Rhodes. He saw that if the DCsS's at one level, treated as sets of vectors, could be mapped by non-singular (so as not to map onto zero) square matrices having uniquely those vectors as eigenvectors, and if these mapping matrices were themselves linearly independent, they could be rearranged as vectors and used as a basis for the next level. In this way the first sequence is mapped by the second sequence $(2 \Rightarrow 2^2 = 4)$, $(4 \Rightarrow 4^2 = 16)$, $(16 \Rightarrow 16^2 = 256)$, $(256 \Rightarrow 256^2)$. The process terminates because there are only $256^2 = 65,536 = 6.5536 \times 10^4$ l.i. matrices available to map the fourth level, which are many too few to map the $2^{127} - 1 = 1.7016... \times 10^{38}$ DCsS's of that level. This (unique) hierarchy is exhibited in Table 1.

Table 1

The combinatorial hierarchy

hierarchy level ℓ	$B(\ell+1) = H(\ell)$	$H(\ell) = 2^{B(\ell)} - 1$	$M(\ell+1) = [M(\ell)]^2$	$C(\ell) = \Sigma_{j=1}^{\ell} H(j)$
(0)	-	2	(2)	-
1	2	3	4	3
2	3	7	16	10
3	7	127	256	137
4	127	$2^{127} - 1$	$(256)^2$	$2^{127} - 1 + 137$

Level 5 cannot be constructed because $M(4) < H(4)$

Although this argument proves the *necessity* of the termination (which is no mystery in the sense that an exponential sequence must cross a power sequence at some finite term), it did not establish the existence of the hierarchy. This was first done by me by creating explicit constructions of the mapping matrices[17] and later more elegantly by Kilmister[20] . That the termination, and indeed the combinatorial hierarchy itself, is a much more fundamental object that the apparently *ad hoc* mapping procedure which first led to it can be seen either by Kilmister's latest derivation as included here or by the very different way Parker-Rhodes now gets it out of his *Theory of Indistinguishables*[21] ; a useful discussion of that theory is provided by him IN Appendix III.

The method Manthey and I use to construct the hierarchy is much simpler; in fact some might call it "simple-minded". We claim that all we have to do is to demonstrate explicitly (i.e. by providing the coding) that any run of PROGRAM UNIVERSE contains (if we enter the program at appropriate points during the sequence) all we need to extract some representation of the hierarchy and the label address scheme from the computer memory *without* affecting the running of the program. The obvious intervention point exists where a new string is generated, as indicated on the flow chart (Figure 1) by the box [CAN BE LABELED]. The subtlety here is that if we assign the label i to the string $U[i]$ as a *pointer* to the spot in memory where that string is stored, this pointer can be left unaltered from then on. It is of course simply the integer value of $SU + 1$ at the "time" in the simulation [sequential step in the execution of that run of the program] when that memory slot was first needed. Of course we must take care in setting up the memory that *all* memory slots are of length $N_{max} > N_U$, i.e. can accommodate the longest string we can encounter during the (necessarily finite) time our budget will allow us to run the program. Then, each time we TICK, the bits which were present at that point in the sequential execution of the program when the slot $[i]$ was first assigned will remain unaltered; only the growing head of the string will change. Thus if the strings i, j, k.... labeled by these slots are linearly independent at the time when the latest one is assigned, they will remain linearly independent from then on.

Once this is understood the coding Manthey and I give for our labeling routine should be easy to follow. We take the first two linearly independent strings and call these the basis vectors for *level 1*. The next vector which is linearly independent of these two starts the basis array for *level 2*, which closes when we have 3 bases vectors linearly independent of each other and of the basis for level 1, and so on until we have found exactly $2+3+7+127$ linearly independent strings. The string length when this happens is then the *label length* N_L; it remains fixed from then on. During this part of the construction we may have encountered strings which were *not* linearly independent of the others, which up to now we could safely ignore. Now we make one *mamouth* search through the memory and assign each of these strings to one of the four levels of the hierarchy; it is easy to see that this assignment (if made sequentially passing through level 1 to level 4) has to be unique. From now on when we generate a new string, we look at the first N_L bits and see if they correspond to any label already in memory. If so (since the address part of the string *must* differ) we assign the address to the *address ensemble* carrying that label. If the new string also has a new label, we simply find (by upward sequential search as before) what level of the hierarchy it belongs to and start a new labeled address ensemble. Because of discriminate closure, we must eventually generate $2^{127} + 136$ distinct labels, organised in the four levels of the hierarchy. Once this happens, the label set cannot change, and the parameters i for these labels will retain an *invariant* significance no matter how long we continue to TICK. We emphasize once more that *what* specific representation of the hierarchy we generate in this way is irrelevant.

Each event results in a TICK, which increases the complexity of the universe in an irreversible way. Our theory has an ordering parameter (N_U) which is conceptually closer to the "time" in general relativistic cosmologies than to the

"reversible" time of special relativity. The arbitrary elements in the algorithm that generates events preclude *unique* "retrodiction", while the finite complexity parameters (SU, N_U) prevent a combinatorial explosion in *statistical* retrodiction. In this sense we have a *fixed* – though only partially retrodictable – *past* and a necessarily *unknown future* of finite, but arbitrarily increasing, complexity. Only structural characteristics of the system, rather than the bit strings used in computer simulations of pieces of our theory, are available for epistemological correlations with experience.

What was *not* realised when this program was created was that this simple algorithm provides us with precisely the minimal elements needed to construct a finite particle number scattering theory. The increase in the number of strings in the universe by the creation of novel strings from discrimination is our replacement for the "particle creation" of quantum field theory. It is not the same, because it is both finite and irreversible; it also changes the "state space". The creation of novel strings by increasing the string length (TICK) implies an "exclusion principle"; if a string (state) already exists, the attempt to fill it leads to an "event", and a universe of increased complexity. Note that the string length N_U is simply the number of events that have occurred since the start up of the universe; this order parameter is irreversible and monotonically increasing like the cosmological "time" of conventional theories. Our events are unique, indivisible and global, in the computer sense; consequently events cannot be localized, and will be "supraluminally" correlated.

3. THE COUNTER PARADIGM; THE COSMIC FRAME

To make contact with physics we must now relate our bit string universe to the laboratory measurement of mass, length and time or three independent dimensional standards which can be related to these measurements. Laboratory practice in elementary particle physics is to use "counter" experiments or their equivalent for velocity measurement, momentum conservation for mass ratio measurement, and to find some phenomenon that brings in Planck's constant for the third connection (charge via $e^2/\hbar c$, Compton scattering, deBroglie wave interference, black body radiation, photo-effect,...). The inter-relationships between these measurements provide tight standards of self-consistency, and numerical values for the fundamental constants which in the end are more important than the comparisons with the standard meter, kilogram and second. Thus all we need do in principle is to make contact with three aspects of our theory in such a way that these connections follow.

As Heisenberg realised long ago, one of the easiest ways to make contact with macroscopic laboratory physics is through particulate momentum measurement, for example the firing of two counters a distance L apart with a time interval T, and identifying the "particle" which naively speaking "fires the counters" by measuring its mass (eg by momentum conservation in a scattering from a particle of known mass). Since the counters can, in principle, be placed as far apart as we like the velocity $V = L/T$ can be measured to as high precision as our budget allows. Empirically all such velocities are less than or indistinguishable from the limiting velocity c, and the momentum P and energy E are related to the mass m by $P = m\beta c/\sqrt{1-\beta^2}$ and $E = mc^2/\sqrt{1-\beta^2}$ (or $E^2 - p^2c^2 = m^2c^4$) where $\beta = V/c$. Thus if the basic quantum mechanics used is written in momentum space, and all physical quantities can be computed from the momentum space scattering theory, then contact with laboratory measurement is about as direct as possible. This is sometimes called the S-Matrix philosophy, and is adopted here. From this point of view, the representation of quantum mechanics in space-time is then obtained by Fourier transformation, and has only a formal significance, particularly for short distances where direct measurement with rods and clocks is impossible. Hence if we can show that our bit string universe supports a momentum space scattering theory of the same structure as conventional relativistic quantum mechanics (or at least in close enough correspondence to that structure

so as not to be in conflict with experiment), our interpretive job has been done for us by Heisenberg and Chew. We develop this scattering theory in the next chapter, but still find it instructive to go as far as we can in interpretation without invoking that formal apparatus.

The means used to connect the bit string universe to the practice of particle physics is to assume that

any **elementary event,** *under circumstances which it is the task of the experimental physicist to investigate, can lead to the firing of a counter.*

The typical laboratory situation we envisage is that in which one of a beam of particles of some known type (which eventually we will have to connect to some label a in the bit string universe) enters and fires a macroscopic counter, and at time T later a counter a distance L from the first which is sensitive to the same type of particles also fires. Ignoring the practical details which will occur to the experimentalist, and the many sophisticated steps he will have to take to convince his colleagues that neither firing was "spurious", we follow conventional practice and say that this sequence of happenings means that a particle of type a has been shown to have a velocity $v = L/T$, and until something else happens will (if it carries a conserved quantum number such as charge or baryon number) continue to have that velocity in the geometrical direction defined by the first two counters. This assumption can be checked by adding counters down stream and checking that indeed (within uncertainties of measurement and corrected for energy loss in the counters) the expected velocity is again measured. We call this the "counter paradigm".

The first step in connecting the counter paradigm to the bit string universe is to assume that the first firing is connected to some unique event involving label a and that N TICK's later there was a second event involving the same label connected to the second firing. Further we assume that for some relevant portion (to be spelled out in detail later) of the address *ensemble* with this label the average number of ones added by these TICK's was $< k^a(N) > = < \Sigma_{i=1}^N b_i^a >$ allowing us to define a parameter $\beta^a = \frac{2<k^a(N)>-N}{N}$. Since $-1 \leq \beta^a \leq +1$ we identify it with a velocity measured in units of the limiting velocity c, and connect it to the experiment by requiring that $\beta^a = v/c = L/cT$. Following Stein[22] we interpret this ensemble of strings of length N as a biased random walk in which a 1 represents a step in the positive and a 0 a step in the negative direction.

Since we now know how to relate sub-ensembles of bit strings to velocities in laboratory events, the question naturally arises as to what coordinate system the full ensembles generated by PROGRAM UNIVERSE refer to. Fortunately this is an easy question to answer. We now know that the solar system is moving at approximately 600 km/sec with respect to the coordinate system in which the 2.7^oK background radiation is at rest; we also have measured the direction of this motion with respect to the distant galaxies. But the statistical method by which the strings are generated guarantees that on average they will have as many zeros as ones, defining uniquely a zero velocity frame with respect to which non-null velocities have significance. Clearly this must be identified with the empirically known "cosmic" zero velocity frame. Further there are two strings, 1_{N_U} and 0_{N_U}, which describe two states in which corresponding labels, 1_{N_L} and 0_{N_L}, have had the limiting velocity in opposite directions from the start. Thus we have an event horizon, to which we cannot assign any further content even after we have constructed our version of 3+1 "space"; the event horizon must be *isotropic*. Of course within that event horizon we could still be receiving signals from the remnants of collections of events which can be expected to be isotropic only in a statistical sense. We find it very satisfactory that these observed cosmological features emerge so readily from our interpretation of the model.

Now that we have confidence that the address strings do indeed specify discrete velocity states in general and not just in the laboratory, we next note that once the hierarchy has closed off at level 4, the set of available labels is *fixed* and simply keeps on reproducing itself in subsequent events. Thus labels have an *invariant* significance no matter how many subsequent TICK's occur, and can be used to identify both quantum numbers and elementary particle masses. Of course it will then become the task of the theory to compute the ratios of these masses to m_p (or to M_{Planck}). The problem is to make this assignment in such a way as to guarantee both quantum number conservation and 3-momentum conservation between *connected* events. Just how to do this is not obvious, and I have made several false starts on the problem, from each of which I learned something. The key turned out to lie in the parallel development of a finite particle number relativistic quantum scattering theory[23-37] which I hope will one day be considered as a candidate to replace both quantum field theory and S-Matrix theory as the theory of choice for practical problems in relativistic quantum mechanics. That, of course, lies in some *very* uncertain future. Fortunately the development

has proceeded far enough to give the essential clues as to how to connect the bit string universe to at least one version of relativistic quantum mechanics.

We now spell out in more detail precisely how the counter paradigm is used to connect the firing of two laboratory counters as described above to two events in the bit string universe. These two events will involve some label L^a of length N_L. We assume that the address string A^a is of length $N_A = N_U - N_L$ when the first firing occurs and of length $N_A + N$ when the second firing occurs. The laboratory velocity $V = \beta c$ is then to be computed from the bit string model by $\beta^a = (2k^a/N) - 1$ where for a single string $k^a = \Sigma_{n=N_A+1}^{N_A+N} b_n^a$. As we have already discussed, we are not allowed to access the computer memory directly, so our knowledge is not this precise. The macroscopic size of the counters ΔL and finite time resolution ΔT necessarily require us to consider discuss all strings in the bit string universe in some range $\beta \pm \Delta\beta/2$ where $\Delta\beta = (L+\Delta L)/c(T-\Delta T) - L/cT$. We will see in the next chapter that this "wave packet" description is essential for the calculation of the "propagator" in the scattering theory.

Before I fastened on the counter paradigm as the correct point of contact between the theory and experiment, I tried to make use of Stein's[22] "derivation" of the Lorentz transformation and the uncertainty principle. He assumed that the basic "objects" underling what we call particles are ensembles of biased random walks of N steps of length ℓ with a probability p of taking a step in the positive direction and $q = 1 - p$ in the negative direction, and hence the probability distribution $N!/p!q!$ for the most probable position of the peak. To relate this to the velocity of the "particle" take $p = \frac{1}{2}(1+\beta)$ and $q = \frac{1}{2}(1-\beta)$, where βc is indeed the velocity of the most probable position. From the fact that the standard deviation from the peak is $\sqrt{Npq} = \sqrt{\frac{N}{4}(1-\beta^2)}$ Stein then arrives at the Lorentz transformation, and by taking $\ell = h/mc$ gets the uncertainty principle as well.

Once I had the counter paradigm in mind, I took over Stein's "random walk" idea by assuming that the 1's in the $N_A + 1 \rightarrow N_A + N$ portion of the address strings represented steps in the positive (first firing to second) direction between the counters and the 0's steps in the opposite direction. The definition of β remains the same, and by taking the step length as $\ell = hc/E$ the velocity of the most probable position and the momentum are correctly related to the energy. Further the velocity of each step is the deBroglie "phase velocity". If we make up "wave packets" from these discrete "velocity states", it is easy to show that the most probable position still moves with the mean velocity and that the "coherence length" which determines interference phenomena based on these periodicities is indeed[38,39] the deBroglie wavelength $\lambda = h/p$. Our discrete theory therefore relates momentum measurement to interference phenomena and the "wave-particle dualism" in much the same way that it is done by following the S-Matrix philosophy.

We now have $\hbar, c$, and m/m_p related to measurement in a precise way. In the next chapter we complete this part of the argument by showing that we can indeed construct a scattering theory with 3-momentum and quantum number conservation in events using the strings of program universe. But our identification of address strings with velocity states already allows a number of cosmological connections between our theory and experimental fact to be made independent of the technical details of the scattering theory. As was spelled out above, we have the cosmological event horizon and its isotropy, and the identification of the coordinate system in which the theory is constructed with that coordinate system in which the 2.7^oK cosmic background radiation is at rest.

4. SCATTERING THEORY; CONSERVATION LAWS

We must now proceed to show that the events discussed above can be interpreted as supporting conservation laws that will be preserved by all relevant TICK-connected happenings. This will be done by invoking a new multi-particle relativistic quantum mechanical scattering theory[23−37]. The basic idea in this scattering theory is to use Faddeev-Yakubovsky equations for the dynamics rather than a Hamiltonian, or Lagrangian or analytic S-Matrix formulation. The basic input to the *linear* integral equations is then a two-particle scattering amplitude with one or more spectators. Because neither particle from the scattering pair is allowed to scatter again with its partner until something else has happened, there can be no "self-energy-loops" or infinities such as occur in field theory. Because the equations are linear, the solutions are unique, in contrast to the non-linear ambiguities that occur in the analytic S-Matrix theory. Because of the algebraic structure of the equations probability flux is conserved for those degrees of freedom which are included.

The basic theory allows any finite number of distinguishable particles. Fortunately we will not have to explore the combinatorial explosion that results in the standard Faddeev-Yakubovsky theory when one tries to go from N to $N+1$ with $N > 4$ because elementary events in PROGRAM UNIVERSE can involve at most four distinct strings. The 4-process has two cases: (3,1) three particles can coalesce to one (or one dissociate to three) with the fourth particle as a spectator; (2,2) two particles can scatter, and the scattering of the second pair can be the spectator. The 3-process allows one pair to scatter with the third particle as a spectator; adding a spectator to this process will lead to one of the two previous possibilities on the first iteration.

The Faddeev (3-particle) theory has three input processes: $a + b \leftrightarrow a + b$, c spectator; $b + c \leftrightarrow b + c$, a spectator; $c + a \leftrightarrow c + a$, b spectator. But when "crossing" is considered[40] the dynamics have to describe as well the anti-particles $\bar a, \bar b, \bar c$ with no change in the dynamical degrees of freedom. In quantum field theory or S-Matrix theory, any particulate state with velocity $\underline{v}$ and quantum number(s) $Q_{(s)}$ must enter the theory in such a way that no prediction of the theory is altered by changing the (conventional and arbitrary) choice of sign of the quantum numbers and choice of reference direction for velocities to $-\underline{v}, -Q_{(s)}$ and inverting the coordinates (parity operation); the *relative* sign between velocities and quantum numbers *is* significant. Since for any labeled address A^a, $A^{\bar a} = 1_{N_U-N_L} \oplus A^{a}$, $\beta^{\bar a} = -\beta^a$, all we need do to insure this rule is to require that for any quantum number we define using the (unique) label string for label a ($a \in 1, 2, ..., 2^{127} + 136$)) $Q^{\bar a} = -Q^a$; all rules used below meet this requirement. Then any 3-event can be viewed as a two particle amplitude

$$\begin{matrix} a \to \\ b \to \end{matrix} \leftarrow \bar a \oplus \bar b \equiv a \oplus b \to \begin{matrix} \to a \\ \to b \end{matrix};\ c = a \oplus b\ spectator \tag{4.1}$$

or the velocity reversed equivalent; note that we cannot distinguish this locally from any cyclic or anti-cyclic permutation on a, b, c. In terms of the scattering theory we have developed[23−37] the basic scattering process starts from a collision between a particle and an anti-particle with opposite velocities, which is isomorphic to the bit-string 3-event described by Eqn. (4.1) if we look at is as $a + b \to\leftarrow \bar a + \bar b$. Because the distinction between the symbols 0 and 1 does not depend on which is which (a point brought home forcefully by John Amson's discussion of the *Bi-Orobourous* included here, the masses of particle and antiparticle must be the same. Consequently the basic process has zero total momentum, which is consistent with the assumption made above that the construction refers to the zero momentum frame. The extension to 4 events, taking proper account of the two cases, is immediate:

$$(3,1): \; a \oplus b \oplus c \rightarrow d \equiv \bar{d} \leftarrow \bar{a} \oplus \bar{b} \oplus \bar{c}; \; d = a \oplus b \oplus c \; spectator \tag{4.2}$$

$$(2,2): \; a \oplus b \rightarrow (ab) \equiv (\bar{a}\bar{b}) \leftarrow \bar{a} \oplus \bar{b}; \; (cd) = a \oplus b \; spectator \tag{4.3}$$

Since our basic process is what is called in high energy physics "anelastic" (2 in, 2 out but not necessarily the same two), it would appear that there are only two degrees of freedom – energy and scattering angle or the manifestly covariant Mandelstam variables. Actually this is true so far as the coupled integral equations go, but the coupling between the three Faddeev channels necessarily brings in a third dynamical degree of freedom. If we take these three degrees of freedom to be the magnitudes of the three momenta p_a, p_b, p_c, 3-momentum conservation guarantees that the vector triangle formed by them closes, so the magnitudes fix the internal angles. One vector in the plane of the triangle then can be used to relate the scattering triangle to space-fixed (laboratory) axes, providing 3 kinematic degrees of freedom. Since 3-momentum is conserved the plane of the triangle is fixed, as is the *total* 3-momentum in any arbitrary laboratory frame; total 3-momentum provides 3 more kinematic degrees of freedom. Since the particles are "on mass shell" ($E^2 - p^2 = m^2$ with $c = 1$), 9 of the 12 degrees of freedom are needed only to relate the fundamental dynamics to the manifestly covariant description in terms of the 4-vectors $\vec{k}_a, \vec{k}_b, \vec{k}_c$. A similar analysis shows how the Faddeev-Yakubovsky dynamics used in the 4-particle equations in the zero momentum frame, again under the assumption of 3-momentum conservation, suffices to provide all that is needed for the interpretation of the results in terms of standard relativistic kinematics.

Now that we know where we are headed, we can try to connect this theory up to the events in the bit string universe. The scattering theory uses single particle basis states with energy, momentum, mass and velocity connected (with $\beta^2 = (V/c)^2$ and $c = 1$) by

$$E^2 - p^2 = m^2 \geq 0; \; 0 \leq \beta^2 = p^2/E^2 \leq 1 \tag{4.4}$$

Calling these states $|m_a, \beta_a > \equiv |a>$, the single particle mass $\mathbf{M}^a$, velocity $\mathbf{B}^a$,

momentum $\mathbf{P}^a$ and energy $\mathbf{E}^a$ operators have these states as eigenvectors:

$$\mathbf{M}^a|a>= m_a|a>; \mathbf{B}^a|a>= \beta_a|a>= \frac{p_a}{\sqrt{m_a^2 + p_a^2}}|a> \tag{4.5}$$

$$\mathbf{P}^a|a>= \frac{m_a \beta_a}{\sqrt{1-\beta_a^2}}|a>; \; \mathbf{E}^a|a>= \frac{m_a}{\sqrt{1-\beta_a^2}}|a> \tag{4.6}$$

So far the connection to the bit strings of specified address length thought of as states is immediate if we make the identification $S^a(N_U) = L^a(N_L)\|A^a(N_U - N_L) = |a>$. All values of the parameters compatible with the constraints expressed in Eqn. (4.4) are allowed in the conventional scattering theory. Bit string dynamics is more specific. Only the discrete velocity eigenvalues $\beta_a = \frac{2\Sigma_{n=N_1+1}^{N_2} b_n^a}{N_2 - N_1} - 1$ are allowed. Here N_1 and N_2 are the string lengths of the universe when the two events of interest in defining the velocity state space occurred; of course $N_L \leq N_1 < N_2 \leq N_U$.

There is an interesting convergence between the basis states the bit string universe generates automatically and the "light cone quantization" states which Pauli and Brodsky[41] find peculiarly appropriate to simplify the quantum field theory problem. They introduce a finite momentum cutoff Λ and discretize the problem by using a finite quantization length L (the old trick of periodic boundary conditions). In the Introduction to their second paper Pauli and Brodsky call the parameters L and Λ "artificial", which indeed they are in their context of trying to discretize a "continuum" theory; for us two related finite parameters are *necessary.* Our theory has a finite momentum cutoff: the smallest finite mass particle recoiling from the rest of the universe. The maximum invariant energy we can discuss in our theory is, so far as we can see now, $M_U c^2 = (2^{127}+136)M_{Planck}c^2 = (2^{127}+136)^2 m_p c^2$. Working out the connection to the maximum finite values for N we can discuss consistently in our framework would get us into a discussion the issues raised by Amson's *Bi-Ourobourous*, so we defer it to ANPA 8 or later. We see no likelihood of finding *direct* experimental confirmation of our finite philosophy by exploring that limit experimentally. Both for the Pauli-Brodsky approach and for ours the momentum cutoff is set by the computing budget rather than more fundamental considerations.

Fortunately the minimum resolution we can achieve sets practical limits that are simpler to discuss, and which are directly related to the states Pauli and Brodsky use. They relate their invariant 4-momentum M to their "harmonic resolution" K by requiring, as we do, zero center-of-mass momentum (cf. p. 1999, Ref 41). Their unit of length is $\lambda_C = h/Mc$ which is the same as the step length ℓ in our random walk. Hence their harmonic resolution $K = L/\lambda_C = L/\ell = N$, that is the number of steps taken in the random walk, or the number of bits in the relevant portion of the address strings. As they say "One must conclude, that the wave function of a particle in one space and one time dimension depends on the .... value of the harmonic resolution K". This should make it clear that we can map our results onto theirs or visa versa, and find out the equivalent of their Lagrangian, creation and destruction operators, etc. in our context - or visa versa. The details remain a problem for future research.

There is a difficulty in their approach in going to 3+1 space since one needs two basic operators in addition to the invariant four momentum and the harmonic resolution. But, as Pauli and Brodsky assure me the obvious high energy particle physics choice of $M, P_{\|}, P_{\perp}, L_z$ works very well. Our problem is different in that there is nothing in the definition of "event" which insures that 3-momentum will be conserved, a fact which Kilmister pointed out rather forcefully at ANPA 7. Hence it is not obvious how to put these single particle states together to describe a 3-event or a 4-event. Actually this is a difficulty in *any* quantum theory, not just ours. The quantum framework is in fact more general than the 3-momentum conservation which (so far) is always observed. Non-relativistic quantum mechanics meets this problem by requiring that any interaction used either conserve 3-momentum, or be an approximation in a system where some large mass is allowed to take up arbitrary amounts of momentum. In quantum field theory the problem is met by assuming certain symmetries in the space of description *and* the allowed interactions, which lead to 3-momentum conservation for *observable processes*. "Vacuum fluctuations" (or disconnected graphs) which violate various conservation laws can still occur; they correspond to the events in our theory which we also wish to exclude. If one takes the symmetries as more fundamental, then momentum conservation can be "derived"; however, I would claim that the symmetries were introduced in the first place in order to insure this result. From a logical point of view momentum conservation is an added postulate.

S-Matrix theory starts from physically observable processes, and hence imposes momentum conservation from the start. The finite particle number scattering theory I am modeling simply requires 3-momentum conservation for all driving terms in the integral equations, and the structure of the equations guarantees that this propagates through the solutions. I claim we have at least as much right to restrict the interpretation of the bit string theory to those connected events which conserve 3-momentum when we discuss physical predictions as does any other quantum theory.

Actually the recent work by McGoveran and Etter[4] puts us on still firmer ground in making this restriction. The basic fact about a discrete topology is that distance cannot be defined until ordering relations, which define attributes of the resulting partially ordered sets, are imposed on the initially *indistinguishable* finite elements. Once this is done, the "distance" depends on the number discrete ways in which the information content of two different collections differs with respect to each attribute. Thus the metric, and the rate of information transfer, is attribute-dependent. Consequently there will be various "limiting velocities", the one which refers to *all* attributes being the minimum of these maximum allowed velocities. In the physical case, this is clearly the velocity of light and is the maximum rate at which information (i.e. anything with physical efficacy in producing change) can be transferred. However correlations (or in computer terminology *synchronization*) can occur supraluminally. This is our basic explanation of the EPR effect. With regard to the point under discussion, since 3-momentum conservation is one of the known attributes of physical effects, we are clearly justified in requiring this of the events that enter our scattering theory. Our bit string universe is then richer that the physical portion we discussion this paper, – a point worth pursuing in the future.

A second difficulty which emerges is that even though we restrict ourselves to (eg for 3-events) those strings for which

$$|p_a - p_b| \leq p_c \leq p_a + p_b; \;\; a, b, c \; cyclic \tag{4.8}$$

we will not have all the richness of Euclidean geometry. We can of course define our angles in the triangle implied by 3-momentum conservation [which will close if Eqn. (4.8) is imposed] by $p_c^2 = p_a^2 + p_b^2 + 2p_a p_b cos\theta_{ab}$, but the digitization of the momenta (via the digitization of the velocities) will allow only certain angles to occur. This, of course, is familiar in the old "vector model" for quantum mechanical angular momentum; it is sometimes still called "space quantization". One loose end that still needs to be tied up is the connection of angular momentum quantization in units of $\hbar$ to the $\hbar$ we have already introduced via our random walk. We obviously cannot introduce Planck's constant twice. In a metric space restricted to commensurable lengths Phythagoras' Theorem does not always hold. McGoveran has pointed out that when we try to close triangles in a discrete space the restriction to integer values is one way that non-commutativity can enter a discrete topology. So all of this should work out in the long run. If it doesn't we are in serious trouble.

We nail down the 3-momentum conservation law by allowing only those labeled address ensembles for which it holds to provide dynamical connection between TICK - separated events The next step is to show that there are conservation laws arising from the labels which can stay in step with the kinematics. This is considerably easier. A 3-event requires that $L^a \oplus L^b \oplus L^c = 0_{N_L} \equiv 0_L$ and hence that $L^a \oplus L^b \oplus c = 1_{N_L} \equiv 1_L$ cyclic on a, b, c, where $L^{\bar{a}} = 1_L \oplus L^a$. If we define the quantum number operators for some attribute x by $\mathbf{Q}_x^a|a>= q_x^a|a>$ and require that

$$\mathbf{Q}_x|0_L>= 0 = \mathbf{Q}_x|1_L> \tag{4.9}$$

and that $q_{\bar{a}} = -q_a$, quantum number conservation in 3-events follows immediately. Further, velocities and particle-antiparticle status reverse together, as in usual in the Feynman rules. We defer the discussion of "spin" to the next chapter.

Probably the most significant step taken since ANPA 7 is the derivation of the "propagator" for the scattering theory directly from the bit string universe via the counter paradigm. The breakthrough was achieved last fall in collaboration with Mike Manthey, who got me out of the rut of a "binomial theorem" connection between TICK-separated events I had failed to make work. In the scattering theory, the connection between events is provided by the "propagator" $\frac{1}{E'-E-i0^{\pm}}$. Here +(-) refer to "incoming" ("outgoing") boundary conditions, and are all that remains in the "stationary state" scattering formalism to record the "time dependence" of the wave function in the Schroedinger representation. The unitarity of the S-Matrix $\mathbf{S} = \mathbf{1} + i\mathbf{T}$, that is $\mathbf{S}^\dagger \mathbf{S} = \mathbf{1}$ or the corresponding

restriction on the scattering amplitude $\mathbf{T}$, is then all that is needed to insure flux conservation, detailed balance and time reversal invariance in the conventional formalism. This is the formal expression of the Wick-Yukawa mechanism, which attributes quantum dynamics to the "off-shell" scatterings at short distance which conserve 3-momentum but allow the energy fluctuations consistent with the Heisenberg energy-time uncertainty principle and the Einstein mass-energy relation. In words, the propagator is the probability amplitude for having the energy E' in an intermediate state in the scattering process when one starts from energy E for the incoming state. Since only the value at the singularity survives in the end (i.e. "asymptotically", or to use more physical language, in connections between numbers that can be measured in the laboratory), the normalization of this singularity can be fixed by the unitarity (flux conservation) requirement, and need not concern us. The scattering equations are simply the sum over all the possibilities allowed by the conservation laws with this weighting.

To obtain the statistical connection between events, we start from our counter paradigm, and note that because of the macroscopic size of laboratory counters, there will always be some uncertainty $\Delta\beta$ in measured velocities, reflected in our integers k_a by $\Delta k = \frac{1}{2}N\Delta\beta$. A measurement which gives a value of β outside this interval will have to be interpreted as a result of some scattering that occurred among the TICK's that separate the event (firing of the exit counter in the counter telescope that measures the initial value of $\beta = \beta_0$ to accuracy $\Delta\beta$) which defines the problem and the event which terminates the "free particle propagation"; we must exclude such *observable* scatterings from consideration. What we are interested in is the probability distribution of finding two values k, k' within this allowed interval, and how this correlated probability changes as we tick away. If $k = k'$ it is clear that when we start both lie in the interval of integral length $2\Delta k$ about the central value $k_0 = \frac{N}{2}(1+\beta_0)$. When $k \neq k'$ the interval in which both can lie will be smaller, and will be given by

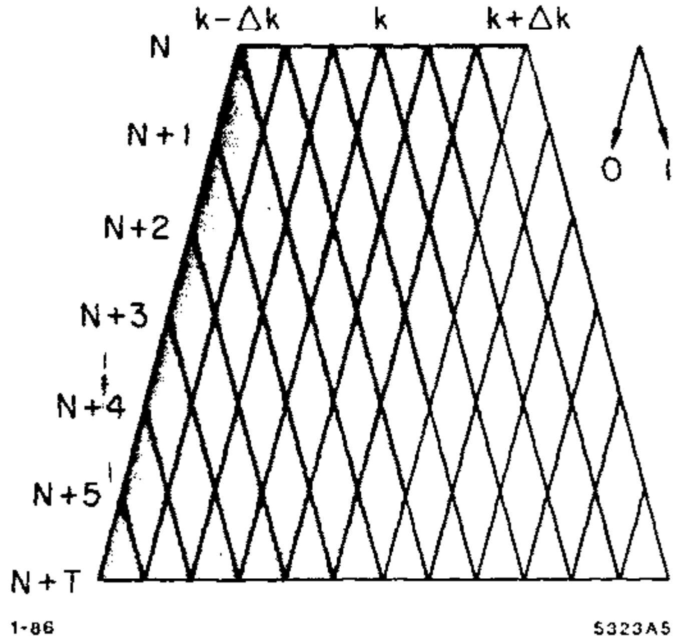
$$[(k + \Delta k) - (k' - \Delta k)] = 2\Delta k - (k' - k) \tag{4.10}$$

when $k' > k$ or by $2\Delta k + (k' - k)$ in the other case. Consequently the correlated probability of encountering both k and k' in the "window" defined by the velocity resolution, normalized to unity when they are the same, is $f(k, k') = \frac{2\Delta k \mp (k'-k)}{2\Delta k \pm (k'-k)}$, where the positive sign corresponds to $k' > k$. The correlated probability of finding two values k_T, k'_T after T ticks in an event with the same labels and same normalization is $\frac{f(k_T,k'_T)}{f(k,k')}$. This is 1 if $k' = k$ *and* $k'_T = k_T$. However, when $k' \neq k$, a little algebra allows us to write this ratio as

$$\frac{1 \pm \frac{2(\Delta k - \Delta k_T)}{(k'-k)} + \frac{4\Delta k \Delta k_T}{(k'-k)^2}}{1 \mp \frac{2(\Delta k - \Delta k_T)}{(k'-k)} + \frac{4\Delta k \Delta k_T}{(k'-k)^2}} \tag{4.11}$$

If the second measurement has the same velocity resolution $\Delta\beta$ as the first, since $T > 0$ we have that $\Delta k_T < \Delta k$. Thus, if we start with some specified spread of events corresponding to laboratory boundary conditions, and tick away, the fraction of connected events we need consider diminishes in the manner illustrated in Figure 3.

Figure 3. The connection between the address strings in tick-separated events resulting from an initial uncertainty in velocity measurement.



Consequently if we ask for the correlated probability of finding the value β' starting from the value β we have proved that in the sharp resolution limit this is 1 if $\beta = \beta'$ and ± 0 otherwise. That is we have shown that in our theory a free particle propagates with constant velocity with overwhelming probability – our version of Newton's first law.

Were it not for the $\pm$, the propagator would simply be a δ-function, and since we are requiring 3-momentum conservation the theory would reduce to relativistic "point particle" scattering kinematics. But the limit we have derived approaches 0 with a sign that depends on which velocity is greater, which in turn depends on the choice of positive direction in our laboratory coordinate system, and hence in terms of the general description on whether the state is incoming or outgoing. In order to preserve this critical distinction in the limit, instead of something proportional to a δ-function we must write the propagator as

$$P(\beta,\beta') = \frac{lim}{\eta \to 0^{\pm}}\left[\frac{-i\eta}{\beta' - \beta - i\eta}\right] \tag{4.12}$$

where the limit is to be taken *after* summing over the allowed possibilities. Thus we find that the complexity of the wave function, and the propagator needed for scattering theory, can actually be derived from our interpretation of the bit strings. As already noted, the actual normalization of the propagator depends on the normalization of states, so we can use the conventional choice $\frac{1}{(E'-E\mp i0^+)}$ just as well.

What I like best about this derivation is that the macroscopic dimensions of the counters enter explicitly into the structure we need, just as "wave packets" have to be brought in for careful discussion of fundamental problems in standard quantum theory. It is also very satisfactory that the dichotomic choices at the lattice connections arising from TICK are strongly reminiscent of Finkelstein's "space-time-code" checkerboard. Scattering theory is one way to connect the imaginary time dependence in the Schroedinger equation of the conventional treatment to the discrete time scale we have to use to describe our time evolution. For consistency, this must also connect to the complex representation of angular momentum, and non-commutativity. As mentioned above McGoveran has some profound ideas here that are crying out to be explored.

Now that we have the propagator for a free particle, it is easy to write down the basic two-particle scattering operators as poles in the invariant two-particle 4-momentum which occur when the two particles coalesce to form a "bound state" of the mass appropriate to the resulting label and clothe this with 3-momentum conservation. We now have derived all the ingredients needed for the scattering theory. Since we have on hand a preliminary description of the theory[37] we repeat here the portion relevant to this paper.

Fortunately the "zero range scattering theory" developed in a non-relativistic context[28] allows scattering amplitudes to be inserted in Faddeev equations without specifying their relation to the non-invariant concept of "potential energy distribution". The model then reduces to the *kinematic* requirement that the "elementary" (or input) two-particle amplitude for meson-nucleon scattering have a pole when the invariant four-momentum of this pair is equal to the nucleon mass. As has been noted many times[25] the use of Faddeev dynamics guarantees unitarity without ever producing the self-energy infinities caused by the quantum field theory formalism. Clearly our general philosophical framework is that of S-Matrix theory, although we part company from the usual approaches to that theory by restricting ourselves to finite particle sectors. The second critical physical input is that 3-momentum be conserved in each elementary scattering. All particles are "on-shell"; only the energy of the system as a whole is allowed to fluctuate within the limits provided by the uncertainty principle. Again this is hardly new; Wick used this idea long ago[2] to provide physical insight into Yukawa's[3] meson theory. Putting this together with the requirement that observable probabilities be conserved specifies a minimal theory, as we now show.

Although the two-nucleon one meson system described by four-vectors has twelve degrees of freedom, our mass shell requirement $(\vec{k})^2 = \vec{k}\cdot\vec{k} = \epsilon_m^2 - \underset{\sim}{k}\cdot\underset{\sim}{k} = m^2$ reduces these to 9, and total 3-momentum conservation to 6. We restrict the Faddeev treatment (which would include the kinematic equivalent of particle "creation" and "destruction") by assuming that we start and end with a "bound pair" plus a free particle, and hence need only consider the residues of the double poles in the Faddeev amplitudes. Under these circumstances, 3-momentum conservation fixes the scattering plane in the *external* (and then laboratory) frame and reduces the *dynamical* (internal) degrees of freedom to 3. The remaining 3 simply allow the result of solving our dynamical equations to be related to external, and via the total 3-momentum to laboratory, coordinates. In general there will be nine "elastic and rearrangement" amplitudes (for example if we have a nucleon and an anti-nucleon, there will be a pole at the mass of the meson), but our "confined quantum" assumption[30,31] reduces these to four. Finally, the δ-function on spectator momentum reduces the 3 degrees of freedom to two dynamical degrees of freedom for each Faddeev channel (of course care must be exercised because the Faddeev description is "overcomplete"); we take these to be the magnitude of the momentum and the scattering angle, as in nonrelativistic potential scattering, or a single vector variable $\underset{\sim}{p}$ with the understanding that the azimuthal angle (or magnetic quantum number) is an "ignorable coordinate".

There is a further non-trivial kinematic fact which simplifies our result. We use the Goldberger-Watson[42] propagator $R_0^{-1}(z) = \epsilon_1 + \epsilon_2 + \epsilon_\mu - z$ where $\epsilon_i = \sqrt{p_i^2 + m_i^2}$, $i \in 1, 2$ and $\epsilon_\mu = \sqrt{q^2 + \mu^2}$. Since we are in the zero momentum frame, this is related to the invariant $S = (\vec{k}_1 + \vec{k}_2 + \vec{k}_\mu)^2 = (\epsilon_1 + \epsilon_2 + \epsilon_\mu)^2$ by $R_0(z) = (\sqrt{S} - z)^{-1}$ Here $\underset{\sim}{p_1}, \underset{\sim}{p_2}, \underset{\sim}{q}$ refer to the "internal" coordinates where all three particles are "free". But the "external" coordinates refer to a particle of mass m_a and "bound state" of mass μ_a, with the invariant $s_a = (\epsilon_a + \epsilon_{\mu_a})^2$ or $\epsilon_a = \frac{1}{2}\sqrt{s_a} + \frac{m_a^2 - \mu_a^2}{2\sqrt{s_a}}$ because $p_a^2 = \epsilon_a^2 - m_a^2 = \epsilon_{\mu_a}^2 - \mu_a^2$. The model requires the driving terms to have a pole at $S_{i\mu} = (\vec{k}_i + \vec{k}_\mu)^2 = m_i^2 = (\epsilon_i + \epsilon_\mu)^2 - p_j^2$ where we have used the fact that $\underset{\sim}{p_1} + \underset{\sim}{p_2} + \underset{\sim}{q} = 0$. Hence (for equal mass nucleons) $S_{i\mu} - m_i^2 = (\sqrt{S} - \epsilon_j)^2 - \epsilon_j^2 = \sqrt{S}(\sqrt{S} - 2\epsilon_j)$, and the pole also occurs at $S = 4\epsilon_j^2$. Finally, we note that on shell, $S = s_i = s_j = 4\epsilon_j^2$ and $p^2 = (p^0)^2$, so the pole also occurs when the two momenta are equal. This allows us to write the driving terms as

$$\frac{g^2 \delta^3(p - p_0)}{p^2 - (p^0)^2 - i\eta} \tag{4.13}$$

[In this treatment we are using the continuum approximation]

Now that our space and the operators in it are defined, we can start from the Faddeev decomposition of the three body transition operator

$$\mathbf{T}^{(3)} = \Sigma_{i,j \in \alpha,\beta,\gamma} \mathbf{M}_{ij} \tag{4.14}$$

where the Faddeev operators $\mathbf{M}_{ij}$ are defined by the operator equations

$$-[\mathbf{M}_{\alpha\beta} - \mathbf{t}_\alpha \delta_{\alpha\beta}] = \mathbf{t}_\alpha \mathbf{R}_0 [\mathbf{M}_{\beta\beta} + \mathbf{M}_{\gamma\beta}] = [\mathbf{M}_{\alpha\alpha} + \mathbf{M}_{\alpha\gamma}] \mathbf{R}_0 \mathbf{t}_\beta \tag{4.15}$$

The δ-function in the driving terms reduces the corresponding integral equations immediately to coupled equations in two variables. Further, since we are concerned here only with the (2,2) sector, and hence with the residues of the double poles, which in a non-relativistic context would be called "elastic and rearrangement amplitudes", we can define

$$M_{\alpha\beta} - t_\alpha \delta_{\alpha\beta} = \frac{g_a}{s_a - \mu_{bc}^2} H_{\alpha\beta}(\underset{\sim}{p_a}, \underset{\sim}{p_b^0}; z) \frac{g_b}{s_b^0 - \mu_{ca}^2} \tag{4.16}$$

For the 3-nucleon paper we are relying on here, we assumed two nucleons and one meson with no direct nucleon-nucleon scattering, and called the four surviving amplitudes K_{ij}.

The final result for the nucleon-nucleon amplitude in this (scalar) model is then that

$$T(\underset{\sim}{p}, \underset{\sim}{p'}) = K_{11}(\underset{\sim}{p}, \underset{\sim}{p'}) + K_{12}(\underset{\sim}{p}, -\underset{\sim}{p'}) + K_{21}(-\underset{\sim}{p}, \underset{\sim}{p'}) + K_{22}(-\underset{\sim}{p}, -\underset{\sim}{p'}) \tag{4.17}$$

where

$$K_{ij}(\underset{\sim}{p_i}, \underset{\sim}{p_j}) - V_{ik}(\underset{\sim}{p_i}, \underset{\sim}{p_j}) = \int d^3 p_k' \frac{V_{ik}(p_i, p_k') K_{kj}(p_k', p_j)}{(p_k')^2 - p_j^2 - i\eta} \tag{4.18}$$

and

$$V_{ij} = -(1 - \delta_{ij}) \frac{g^2}{\epsilon_\mu^{ij'} (\epsilon_\mu^{ij'} - \epsilon_{\mu_i} + \epsilon_j')} \tag{4.19}$$

with $\epsilon_\mu^{ij'} = \sqrt{(\underset{\sim}{p_i} + \underset{\sim}{p_j'})^2 + \mu^2}$.

If the "bound state" is required to contain exactly one particle and one meson, three particle unitarity fixes a unique constant value for the coupling constant[27]. However, as has been discussed in connection with the "reduced width" (also the residue of a "bound state" pole) in the non-relativistic theory[43] it is possible to treat the residue as a measure of how much of the state is "composite" and how much "elementary"; the density matrix derivation given in the reference is due to Lindesay. In the case at hand, since the K_{ij} satisfy coupled channel Lippmann-Schwinger equations, their unitarity and that of the T constructed from them is immediate, and is independent of the value of g^2, making this, as well as the meson mass available as adjustable parameters for use in low energy phenomenology. In fact the equations in the non-relativistic region correspond to an ordinary and exchange "Yukawa potential" or for negligible meson mass and $g^2 = e^2$ to the usual coulomb potential. Thus we finally have made contact with both Rutherford Scattering and the Schroedinger equation for the hydrogen atom starting from bit strings!

5. THE STANDARD MODEL OF QUARKS AND LEPTONS; COSMOLOGY

We saw in the last section that our quantum numbers are to be defined in such a way that they reverse sign under the "bar" operation $\bar{S}^a = 1_U \oplus S^a = 1_L \| 1_A \oplus L^a \| A^a$, as do the velocities in the address part of the string. Hence for each string we can single out one quantum number which defines the relative sign between velocities and quantum numbers, and hence defines a "direction" in the space of quantum numbers which is correlated with the directions in ordinary space. This obviously is "helicity" which can be directed either along or against the direction of particle motion. Putting this together with the 3-momentum conservation we have already assured, the fact that this does not reverse sign when the coordinates are reflected makes this a "pseudo-vector" or "spin", and we must assume that it is to be measured in units of $\frac{1}{2}\hbar$ if we are to make contact with well known experimental facts. As already noted, one remaining foundational problem is to connect up the unit with the "orbital angular momentum" from our definitions of 3-momentum and the lengths that occur in our random walks (deBroglie phase and group wavelengths using h rather than $\hbar$ as the unit with these dimensions). In what follows we will assume that this can be done without encountering difficult problems.

Once we have identified the necessity for one quantum number in each label being interpretable as "spin", or more precisely "helicity", (including of course the possibility of the value 0 for some strings), the interpretation of Level 1 is essentially forced on us. The dichotomous spin state with no other structure is the "two component neutrino" familiar since the parity non-conserving theory of weak interactions was created by Lee and Yang, and demonstrated experimentally by Wu. A simple way to represent this is, for a two-bit representation (b_1, b_2), is to take $h_z = (b_1 - b_2)\frac{1}{2}\hbar$, as is shown in Table 2a.

Table 2
Conserved Quantum Numbers

2a. Level 1.

String (b_1, b_2)	$q_0 = b_1 - b_2$
(1 0)	+1
(0 1)	−1
(1 1)	0
(0 0)	0

2b. Levels 2 and 3.

String $(b_1 b_2 b_3 b_4)$	q_1 $b_1 - b_2 + b_3 - b_4$	q_2 $b_1 + b_2 - b_3 - b_4$	q_3 $b_1 - b_2 - b_3 + b_4$
(1 1 1 0)	+1	+1	−1
(0 0 0 1)	−1	−1	+1
(1 1 0 1)	−1	+1	−1
(0 0 1 0)	+1	−1	+1
(1 1 0 0)	0	+2	0
(0 0 1 1)	0	−2	0
(1 1 1 1)	0	0	0
(0 0 0 0)	0	0	0
(0 1 1 1)	−1	−1	+1
(1 0 0 0)	+1	+1	−1
(1 0 1 1)	+1	−1	−1
(0 1 0 0)	−1	+1	+1
(1 0 1 0)	+2	0	0
(0 1 0 1)	−2	0	0
(1 0 0 1)	0	0	+2
(0 1 1 0)	0	0	−2

Then, if we adopt the usual convention that the electron neutrino is "left handed" and has negative helicity relative to the positive direction of motion, we have (for massless neutrinos) $\underset{\sim}{\nu}_e = (01)\|1_A$ and $\underset{\sim}{\bar{\nu}}_e = (10)\|1_A$. There are only two states because (thanks to invariance under the bar operation) for intermediate states a neutrino moving the positive direction is indistinguishable from an anti-neutrino moving in the negative direction. Only in the laboratory, where we can use macroscopic "rigid bodies" to establish directions, can we measure both parameters.

Having an obvious interpretation of the basis states for Level 1, the interpretation of Level 2 is almost as straightforward. We use the representation (b_1, b_2, b_3, b_4), which allows three quantum numbers which meet our restrictions to be defined: $q_1 = b_1 - b_2 + b_3 - b_4$, $q_2 = b_1 + b_2 - b_3 - b_4$, $q_3 = b_1 - b_2 + b_3 - b_4$. These are exhibited explicitly in Table 2b. Since Level 2 only has three linearly independent basis vectors, we require $b_1 = b_2$, which arises naturally from the mapping matrix construction of the hierarchy, as we have discussed in detail in previous work. Under this restriction $q_1 = -q_3$, so there are only two independent quantum numbers. The obvious choice is to identify q_1 as lepton number (or electric charge) and q_2 as helicity in units of $\frac{1}{2}\hbar$, which leads to the particle identifications in Table 3. The graphical representation of these numbers given in Figure 4. may be more informative. We defer discussion of the "Coulomb interaction" called C until we have made the Level 3 assignments.

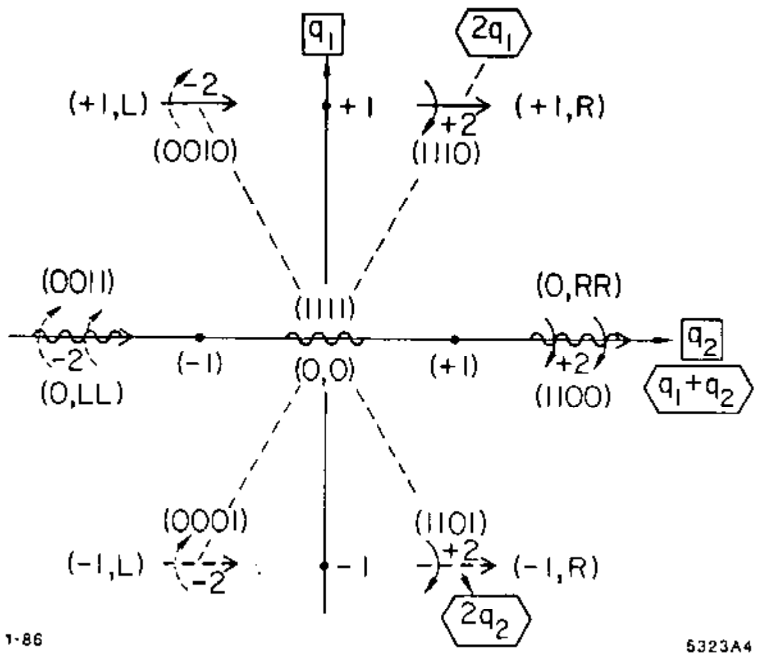
For level 3 we use one four-bit string allowing all 16 possibilities concatenated with a second for-bit string resembling level 2 and hence having only 8 possibilities. The first has four basis vectors and the second three, making up the required 7. Together we have 128 possibilities, or if we subtract the null string, the usual hierarchy 127. Assuming for the moment that the second 4-bit string is (1111) or (0000) – which we will see shortly is a QCD (quantum chromodynamics) color singlet – we have in fact the 16 states which can be formed from two distinguishable fermions and antifermions. These are clearly the nucleons with the associated pseudoscalar and vector (since a fermion-antifermion pair has odd parity) mesons. The identifications are spelled out in Table 3.

Table 3

Particle identifications for Levels 2 and 3

String	*Level 2*	*Level 3 (color singlet)*
(1 1 1 0)	$e^+_{+\frac{1}{2}}$	$n_{+\frac{1}{2}}$
(0 0 0 1)	$e^-_{-\frac{1}{2}}$	$\bar{n}_{-\frac{1}{2}}$
(1 1 0 1)	$e^-_{+\frac{1}{2}}$	$\bar{n}_{+\frac{1}{2}}$
(0 0 1 0)	$e^+_{-\frac{1}{2}}$	$n_{-\frac{1}{2}}$
(1 1 0 0)	γ_{+1}	ρ^0_{+1}, ω_{+1}
(0 0 1 1)	γ_{-1}	ρ^0_{-1}, ω_{-1}
(1 1 1 1)	C	$\pi^0, \rho^0_0, \omega_0$
(0 0 0 0)	(C)	$(\pi^0, \rho^0_0, \omega_0)$
...		
(0 1 1 1)		$\bar{p}_{-\frac{1}{2}}$
(1 0 0 0)		$p_{+\frac{1}{2}}$
(1 0 1 1)		$p_{-\frac{1}{2}}$
(0 1 0 0)		$\bar{p}_{+\frac{1}{2}}$
(1 0 1 0)		d_0
(0 1 0 1)		$\bar{d}_0$
(1 0 0 1)		π^+, ρ^+_0
(0 1 1 0)		π^-, ρ^-_0

Figure 4. Level 2 quantum numbers represented as strings.



To make this into quantum chromodynamics, we need only note that the level 2 quantum numbers also define an SU_3 octet, as is shown in Table 4 in terms of I, U and V -spin; again Figure 4 illustrates the relationships.

Table 4

The SU3 octet for "I,U,V spin"

	$(b_{11}b_{12}b_{13}b_{14})$	$2I_z$	$2U_z$	$2V_z = 2(I_z + U_z)]$
STRING:	1110	+1	+1	+2
	0010	-1	+2	+1
	1100	+2	-1	+1
	1111	0	0	0
	0000	0	0	0
	0011	-2	+1	-1
	1101	+1	-2	-1
	0001	-1	-1	-2

$$2I_z = b_{11} + b_{12} - b_{13} - b_{14}$$
$$2U_z = -2b_{11} + b_{12} + 2b_{13} - b_{14}$$
$$2V_z = -b_{11} + 2b_{12} + b_{13} - 2b_{14}$$

For color we could take red = (0001), anti-red = (1110); yellow = (0010), anti-yellow =(1101); blue = (1100), anti-blue = (0011). Then three colors or three anti-colors give the color singlet (1111), as do the appropriate combinations of color and anti-color. The three basis strings so constructed concatenated with the four already discussed give us two distinguishable colored quark and the associated gluons. Since $a \oplus a \oplus a \equiv a$, three colored quarks (or anti-quarks) add to give a color singlet and yield the spin and helicity states of a nucleons and anti-nucleons as we have just shown. Speculatively, since the scattering theory employed allows three states of the same mass to combine to single state of that mass, we can take both the quark and the nucleon mass to be the same; this would mean that quark structure would only appear at the 3 Gev level, which is desirable if nuclear physics is to continue to use mesons and nucleons as a first approximation.

Clearly we now have the quantum numbers for the first generation of quarks and leptons familiar from the standard model. Because of the closure properties of the hierarchy it is obvious that we will get higher generations simply by duplicating the structure we already have as many times as we need to get to $2^{127}+136$ quantum states. We see that level 4 gives us a combinatorial explosion of higher generations with the same structure, but only weakly coupled because of the large number of combinatorial possibilities.

This is all very satisfactory until we ask (a) how to interpret the level one closure $(11)0_A$ (*or* $(00)1_A$) and (b) how to extend this interpretation to label strings of length L, which our PROGRAM UNIVERSE construction forces us to do. This problem has not been faced in previous discussions, and the conclusions reached for purposed of this report are frankly speculative. The problem is to get the coupling *between*levels and generations right. The speculative idea starts with the conjecture that the label $1_L = (11)_{L/2}$ is simply the universal Newtonian gravitational interaction which couples to any pair of labels with probability $[2^{127}+136]^{-1}$ But then, so far as level 1 labels go, it is indistinguishable from (11). To go on, the analagous level 1 - level 2 cross level coupling would be the unit helicity Z^0 with unit helicity extensions to the $W\pm$. The 1-2-3 cross level coupling would be (as before) the coulomb interaction, with care taken so that the neutrinos carry no charge. As we have noted before, our theory is analagous to doing QED in the "coulomb gauge", so the spin-flip $\gamma_{\pm|}$ which come along are down in probability by 1/137. Including these must correct our first approximation $\alpha = 1/137$ toward the *observed* value, but all strong as well as weak interactions will enter the calculation of the correction. Thus the mixing between generations cannot be ignored *a priori*. Conventional theories are now struggling with the problem of how best to combine weak-electromagnetic unification with the standard model, the generation structure, and gravitation in some sort of "super-unification" scheme. So the problem we hit in our own theory lies close to the cutting edge of conventional physics, as promised.

The cosmological implications of our theory are also interesting. We have already noted that our first approximation gives us Newtonian gravitation, so a "flat space" cosmology can also be anticipated. Our "big bang" - like that in an early version of Parker-Rhodes' *Theory of Indistinguishables* starts out "cold" in that we have to generate the labels first and only begin to develop "heat" after the basis vectors close and we begin to accumulate addressed label ensembles. Since the initial scatterings can take place in $\simeq (1.7\times 10^{38})^2$ ways, and baryon number and lepton number appear to be very well conserved in our scheme, this initial condition gives approximately the baryon number and lepton number of the universe within the (rather broad - but all "flat space") observational limits. Since the initial address strings are short, they correspond to very high velocities and the resulting temperature will be extremely high. Even though we start "cold" we get a cosmic fireball early on. Once the average temperature falls (due to the expanding event horizon) down to the Tev range now being explored by particle accelerators our cosmology will develop much like others. The question which lies open is whether our rather unusual boundary condition will have consequences at variance with more conventional models in such a way as to lead to feasible observational tests. Only the uncertain future can decide.

6. THE MASS SCALE

[This section is quoted from SLAC-PUB-3566, "A discrete foundation for Physics and Experience (January, 1985).]

What is still missing in our fundamental theory are the mass ratios of the particles relative to our standard m_p identified by $\hbar c/Gm_p^2 = 2^{127} + 136$. Here we adapt a calculation of Parker-Rhodes[44] based on his alternative, but closely related, approach to the problem of constructing a fundamental theory. He confronts the problem of *indistinguishability*, which in modern science goes back at least to Gibbs, but poses the problem in the logical (static) framework of how we can make sense of the idea that there are *two* (or more) things which are indistinguishable other than by the *cardinal* number for the assemblage *without* introducing either "space" or "time" as primitive notions. Clearly his starting point is distinct from the constructive program, and the "fixed past - uncertain future" implicit in our growing universe with randomly selected bit strings.

We have seen above that, for a system at rest in the coordinate system defined internally by $<\beta>=0$ or externally by zero velocity with respect to the background radiation, the minimal fundamental length is h/m_pc, inside which length we have no way of giving experimental meaning to the concept of length without external coupling [2]. We have also seen that our scattering theory has, for zero mass coulomb photons, a macroscopic limit in Rutherford scattering, a non-relativistic limit in Bohr's theory of the Hydrogen atom, a continuum approximation in deBroglie's wave theory provided by continuum interpolation using Fourier analysis, and hence the usual formalism for the macroscopic e^2/r "potential" up to $O(1/137)$ spin-dependent corrections or relativistic corrections of the same order (either of which corrections — relativistic spin(Dirac) or relativistic motion (Sommerfeld) — account quantitatively for the empirical hydrogen fine-structure to that order). We have also seen that our momentum-space S-matrix theory has (within our digital restrictions) the usual properties of rotational and Lorentz invariance in 3 + 1 momentum-energy space, and hence by our interpretive paradigms in 3-space.

We therefore can assert that outside a radius of $h/2m_pc$, the energy associated with the (minimally three) partons connected to an electron, the electrostatic energy of an electron can be calculated statistically from $<e^2/r>$ with three

degrees of freedom and $r \geq (h/2m_pc)y, y \geq 1$. Since the conservation laws we have already established require charge conservation, the electrostatic energy must be calculated from the charge separation outside this radius with charges ex and $e(1-x)$, so $<e^2>=e^2<x(1-x)>$. At first glance x can have any value, but in any statistical calculation the charge conservation we have already established requires that these cancel outside of the interval $0 \leq x \leq 1$. We have seen that the leptons are massless until they are coupled to hadrons at level 3 of the hierarchy (with, as the first approximation, $e^2/\hbar c = 1/137$). Hence, in this approximation, we can equate m_ec^2 with $<e^2/r>$, and arrive at the first Parker-Rhodes formula

$$m_p/m_e = \frac{137\pi}{<x(1-x)><1/y>};\ 0 \leq x \leq 1;\ 0 \leq (1/y) \leq 1$$

From here on in, the only point to discuss is the weighting factors used in calculating the expectation values, since we now have from our S-matrix theory the same number of degrees of freedom (three) as Parker-Rhodes arrives at by a different argument based on the *Theory of Indistinguishables*. For the $(1/y)$ weighting factor this is almost trivial; our carefully constructed derivation of the Coulomb law and the symmetries of 3-space imply that $P(1/y) = 1/y$. For $x(1-x)$ the two-vertex structure of our S-matrix theory requires one such factor at each vertex in any statistical calculation: $P(x(1-x)) = x^2(1-x)^2$. The calculation for three degrees of freedom is then straightforward, and has been published several times[5,6,10,17,38,39],[45-47]. The result is $<1/y>=4/5$, $<x(1-x)>=(3/14)[1+(2/7)+(2/7)^2]$, leading immediately to the second Parker-Rhodes formula

$$m_p/m_e = 137\pi/[(3/14)[1+(2/7)+(2/7)^2](4/5)] = 1836.151497...$$

in comparison with the experimental value of 1836.1515 ± 0.0005. Although this result has been published and presented many times, we know of no published challenge to the calculation.

The success of this calculation encourages us to believe that the seven basis vectors of level 3 will lead to a first approximation for $m_p/m_\pi \approx 7$ with corrections of order 1/7, but this has yet to be demonstrated.

7. CONCLUSIONS

We have tried to show in this paper that a complete reconstruction of relativistic quantum mechanics, elementary particle physics and cosmology can be based on a simple computer algorithm organized to exploit the closure properties of the combinatorial hierarchy and known basic principles of modern physics. Some people found the initial success in 1966 already impressive; others called it numerology. By now there are a number of quantitative and qualitative successes to our credit and no known failures. In the language of high energy physics, it may not yet be convincing as an "experiment", but it is beginning to look like the basis for a reasonable "research proposal".

The question remains – first asked me at Joensuu – what it will mean if the results are close to experiment, but can be shown to be in *quantitative* disagreement in a way that no fundamental remedy seems likely to fix. We will assume that a reasonable amount of effort has gone into minor tinkering with the fundamentals – they are so rigid that this gives very little scope. I expect this to be the case some day, hopefully soon. That I anticipate *close* agreement with experiment once the scattering calculations are carried out rests on the overdetermination Chew has already shown to exist between the structure of scattering theory, unitarity and "crossing". The theory has "bootstrap" properties, a finite and convergent equivalent of "renormalization", and once we get the quantum number assignments nailed down, will*have to* (if Chew is correct come close to experimental results. Our advantage over his theory is that our equations are *linear* allowing unique answers for any finite number of particulate degrees of freedom. So what if we are wrong? I would claim that *any* theory must contain the minimal elements we have used, and so must *also* fail unless additional basic principles are added. So a "failure", if we have done our job properly will either point to an are in which to look for *new* physics (eg. extension of the present three (MLT) dimensional concepts [to 7??]) or even a way to look *beyond* physics. This is, as I see it, the most one should ask of any model or theory or philosophy.

ACKNOWLEDGEMENTS

In this paper I have reported on a program aimed at reconstructing quantum theory "from the ground up" which started with work by Bastin and Kilmister in the 1950's[7,8], reached an intermediate success with the discovery of the combinatorial hierarchy by Amson, Bastin, Kilmister and Parker-Rhodes in the 1960's[9], and to which I started contributing in the 1970's[17]. For a survey of the historical development one must follow "preprint" literature and conference proceedings [5,6,10,38,39,46–48],[48] . My own recent work on this problem, carried out primarily in collaboration with David McGoveran, Michael Manthey, and Christoffer Gefwert, including the developments that have occurred since ANPA 7, forms the body of this paper. None of it would have taken place without those collaborators, and all of the ANPA members have contributed in one way or another, but I have decided on this occasion to present my own views without qualification, and make no attempt at achieving consensus. I trust this will lead to lively discussion and criticism at ANPA 8.

REFERENCES

1. A.Aspect and P. Grangier, "Tests of Bell's Inequalities...", in *Symposium on the Foundations of Modern Physics*, P. Lahti and P. Mittelstaedt, eds., World Scientific, Singapore, 1985, pp.51-71, and references therein.
2. G.C.Wick, *Nature, 142*, 993 (1938).
3. H.Yukawa, *Proc.Phys.Math.Soc., Japan, 17*, 48 (1935).
4. H.Pierre Noyes, David McGoveran, Tom Etter, Michael J. Manthey and Christoffer Gefwert, "Minimal Postulates for a Discrete Relativistic Quantum Physics"; ABSTRACT: "Starting from a minimum number of postulates, which do not refer to supposed physical properties of space and/or time, we show that any parametrized change in the distinguishable number of states for a specific attribute satisfying the postulates has the properties of the relativistic doppler shift and the relativistic composition law for velocities; additionally, such changes satisfy the Lorentz transformation. Each of these parameterized changes has an upper limit which may be characterized as a velocity. For a finite set of attributes, we show that the minimum of these maximum limiting velocities dominates relationships between the attributes. This yields a theory which is a discrete, finite version of the accepted properties of relationships in physical space-time. However, this space is multiply-connected and allows for correlations exceeding the minimal limiting velocity. Hence this construction is relevant to the supraluminal correlations demonstrated in Aspect's and other EPR-Bohm type experiments." Submitted to the 2nd *INTERNATIONAL SYMPOSIUM ON FOUNDATIONS OF QUANTUM MECHANICS -In the Light of New Technology*, Tokyo, September 1-4, 1986.

5. For a summary of the early history cf. H. Pierre Noyes, Christoffer Gefwert, and Michael J. Manthey, "Toward a Constructive Physics" SLAC-PUB-3116 (rev. September, 1983); a short version of the paper based on this material was presented at the 7[th] *Congress of Logic, Methodology and Philosophy of Science*, Salzburg, Austria, July 11-16, 1983 (see next reference).
6. H.P.Noyes, Christoffer Gefwert, and Michael J. Manthey, "Program for a Constructive Physics", in *Foundations of Physics*, P. Weingartner and G. Dorn,eds, Hölder-Pichler-Tempsky, Vienna, Austria, 1986.
7. E.W.Bastin and C.W.Kilmister, *Proc. Roy. Soc.*, **A 212**, 559 (1952).
8. –*Proc.Camb.Phil.Soc.*, **50**, 254, 278 (1954), **51**, 454 (1955);**53**, 462 (1957); **55**, 66 (1959).
9. T. Bastin, *Studia Philosophica Gandensia*, **4**: 77, (1966).
10. H.Pierre Noyes, Christoffer Gefwert and Michael J. Manthey, "A Research Program with No Measurement Problem", *Proc. New York Academy of Sciences*, "New Techniques and Ideas in Quantum Measurement Theory", D. Greenberger, ed., New York City, 21-24 January, 1986; available as SLAC-PUB-3734
11. M.J.Manthey and B.M.E.Moret, "The Computational Metaphor and Quantum Physics", *Communications of the ACM*, **26**, 137-145, 1983.
12. M.J.Manthey, "Non-Determinism Can Be Causal", Tech. Report No. CS(83)-7, Univ. of New Mexico, Albuquerque, NM 87131.
13. David McGoveran, *ANPA Newsletter* **6**, No.1, 6-7 (1986).
14. Christoffer Gefwert, *On The Logical Form of Primitive Recursive Functions*, SLAC-PUB-3334 (May 1984), to be submitted to the *Journal of Philosophical Logic*.
15. –, *On The Logical Form of Mathematical Language*, SLAC-PUB-3344 (May 1984), to be submitted to the *Journal of Philosophical Logic*.
16. C.W.Kilmister, Appendix II.1 "Brouwerian Foundations for the Hierarchy" (based on a paper presented at ANPA 2, King's College, Cambridge, 1980) and Appendix II.2 "On Generation and Discrimination" (based on a paper presented at ANPA 4, King's College Cambridge, 1982) in H. Pierre Noyes, Christoffer Gefwert, and Michael J. Manthey, "Toward a Constructive Physics" SLAC-PUB-3116 (rev. September, 1983).
17. T.Bastin, H.P.Noyes, J.Amson and C.W.Kilmister, "On the Physical Interpretation and Mathematical Structure of the Combinatorial Hierarchy", published in *The International Journal of Theoretical Physics*, **18**, 455-488 (1979).
18. J.Amson, Appendix to Ref. 17.
19. M.J.Manthey, Appendix IV. in SLAC-PUB-3116 (rev. September, 1983).

20. C.W.Kilmister, Appendix II.3 "HIERARCHY CONSTRUCTION(second version)" in SLAC-PUB-3116 (Rev. September, 1983).
21. A.F.Parker-Rhodes, *The Theory of Indistinguishables*, Synthese Library, *150*, Reidel, Dordrecht, 1981.
22. Irving Stein, seminar at Stanford 1978, and papers at ANPA 2,3, King's College, Cambridge, 1980,'81.
23. H.P.Noyes, "Three Body Forces", in *Few Body Problems*, I.Slaus, et.al. eds, North Holland, Amsterdam, 1972, p.122.
24. P.Noyes, J.Bell, and T.E.Phipps, Jr., *Fixed Past and Uncertain Future; an exchange of correspondence*, SLAC-PUB-1351, December 1973 (unpublished),p.30.
25. H.P.Noyes, *Czech.J.Phys.*,**B 24**, 1205 (1974).
26. H.P.Noyes, *Foundations of Physics*, *5*, 37 (1975) [Erratum *6*, 125]; *6*, 83 (1976).
27. J.V.Lindesay, PhD Thesis, Stanford, 1981, available as SLAC Report No. 243.
28. H.P.Noyes, Phys.Rev *C 26*, 1858 (1982).
29. H.P.Noyes and J.V.Lindesay,*Australian J. Physics*, *36*, 601 (1983).
30. H.P.Noyes and G.Pastrana, "A Covariant Theory with a Confined Quantum", in *Few Body Problems in Physics*, B. Zeitnitz, ed., North Holland, Amsterdam, 1983, p. 655.
31. J.V.Lindesay and A.Markevich, "A Minimal Unitary(Covariant) Scattering Theory", in *Few Body Problems in Physics*, B. Zeitnitz, ed., North Holland, Amsterdam, 1983, p. 321.
32. H.P.Noyes and M.K.Orlowski, *Nuovo Cimento* **81 A**, 617 (1984).
33. J.V.Lindesay, A.J.Markevich, H.P.Noyes and G.Pastrana, "A Self- Consistent, Poincaré-Invariant, and Unitary Three-Particle Scattering Theory", *Phys.Rev.***D 33**, 2339-49 (1986).
34. A.J.Markevich, "Angular Momentum and Spin within a Self-Consistent, Poincaré-Invariant, and Unitary Three-Particle Scattering Theory", *Phys.Rev.*I **33**, 2350-56 (1986).
35. H.P.Noyes and G.Pastrana, "A Covariant Approach to Mesonic Degrees of Freedom in Light Nuclei", presented at the *International Symposium on Mesons and Light Nuclei*, Bechyně Castle, Czechoslovakia, May 27- June 1, 1985.
36. H.P.Noyes, "Comment in Reply", SLAC-PUB-3916, April 1986 (unpublished).
37. H.P.Noyes "A Minimal Relativistic Model for the Three Nucleon System" in *Three Body Forces in the Three Nucleon System*, B.Berman, ed, Springer-Verlag (in press), available as SLAC-PUB-3973, June, 1986.

38. H.P.Noyes, "A Finite Particle Number Approach to Quantum Physics", SLAC-PUB-2906 (April, 1982, unpublished).

39. H.P.Noyes, "A Finite Particle Number Approach to Physics", in *The Wave-Particle Dualism* S.Diner, et.al., (eds), Reidel, Dordrecht, 1984, pp. 537-556.

40. G.Pastrana, PhD thesis, Stanford (in preparation).

41. H-C. Pauli and S.J.Brodsky, *Phys. Rev.* **D 32**, 1993, 2001 (1985)

42. M.L.Goldberger and G.M.Watson, *Collision Theory*, Wiley, New York, 1964.

43. Ref.43, pp. 1869-70.

44. Ref. 21, pp 182-185.

45. H.P.Noyes, M.J.Manthey and C.Gefwert, "Constructing a Bit String Universe: a Progress Report" in *Quantum Concepts in Space and Time*, R. Penrose and C.J.Isham, eds, Oxford University Press, pp. 260-273 (in press).

46. H.P.Noyes, "A Discrete Foundation for Physics and Experience", SLAC-PUB-3566, January 1985 (presented at the Esalen Mind-Physics Conference, 4-8 February 1985).

47. H.P.Noyes, C.Gefwert and M.J. Manthey, "The Measurement Problem in Program Universe", *Symposium on the Foundations of Modern Physics*, P. Lahti and P. Mittelstaedt, eds., World Scientific, Singapore, 1985, pp.511-525.

48. H.P.Noyes, "Non-Locality in Particle Physics", SLAC-PUB-1405 (1974).

8. Appendix I. A FINAL FOUNDATIONAL DISCUSSION?

Clive W. Kilmister
Chelsea-Kings, University of London

8.1 INTRODUCTION

My aim is to derive the Parker-Rhodes construction described in[49] from the acknowledged properties of quantum mechanics. Wheeler has said[50] "Quantum theory presents us with the strangest object in all of physics, an entity which has no localization in space and time, the elementary quantum phenomenon... all the information we acquire we get, directly or indirectly, from elementary quantum phenomena." This is my starting point. It characterizes quantum mechanics in this way:

(i) there must be elementary quantum events, our only source of knowledge, not localized in space or time,

(ii) we cannot get outside this sequence of quantum events, no matter how much this is at variance with the classical picture of things.

But since the theory is not to be simply meaningless sequences of such events, the sequences must organize themselves in some way to produce higher order events, and these in turn likewise. This means that the system must have such levels and so the mathematical features are

(a) the system must be one with levels,

(b) new features arise and so the mathematical formulation must allow the entry of new symbols.

Classical mathematics does not fit very well with either feature but especially not with (b). Indeed such a creative feature suggests intuitionism and I have flirted with that in the past, but it turns out not to be necessary to go so far. Rather to adopt an intuitionistic stance has traces of simply taking what is to hand instead of looking for exactly what is needed.

8.2 GENERATION AND DISCRIMINATION

Let us begin with (b) above. New entities *arise*. How can this be? There must be some generation process, G, and as each new entity is created it must receive a label. We may use as a set of labels the symbols $0, 1, 2, 3, \ldots$ but these are not, of course, the cardinal numbers. On the other hand they can be read as ordinal numbers, and if, as we shall see is appropriate, we consider 0 as in a special category, we can read the symbol m as labeling the m^{th} element produced, either by the generation operation or internally inside the mathematical system. This notation emphasizes, what is clear from the generation idea, that the resultant system will be either finite of enumerable.

An important question is begged in the preceding paragraph. When an entity is produced, by whatever means, how can we be sure that it is a new one? In one sense it must be, because it differs from all the other entities in its stage of production; but to regard the entities in that way is to produce a system with no structure at all. Rather we must have a notion of equivalence between two examples of an entity, so that each potentially new element must be checked to see if it is equivalent to one which has occurred before. At this point it is again necessary to take note of the continuing generation; for it means that we cannot simply ask whether the new element is or is not the member of some given set. This set would be that of the already generated elements and so is not "given" – whatever that means for the classical mathematician– but is continually changing.

A device which avoids this (it may not be the only possible way to proceed) is to use the nineteenth century idea of a function, that is, a rule which gives a value for each value of the argument. It must not be assumed that any ideas of domain or range are implied by the use of the word "function"; there is simply a rule. This rule is then constrained to be such that the values of the function for elements equivalent to already generated elements lie in a fixed set which is disjoint from the set of values for truly new elements. This condition goes some way to determining the rule, as will become clear later. Such an operation will be called discrimination, D. A "run" of the system will consist of a sequence of G's and D's. The sequence is arbitrary, but if an ensemble of runs is considered it is possible to introduce the ideas of probability. This is not considered in this paper.

8.3 THE SIMPLE CASE

The simple case in which one compares the "new" element with a single known one determines much of the structure. Let u be the known element, and x a potentially new one. Denote the function by $f(u,x)$ so as to include reference to the known element. Denote the fixed set of elements which are values of the function when u, x are equivalent by Z. Since we are specifying an equivalence relation, we must have:

(i) $f(x,x) \in Z$ for all x,

(ii) If $f(x,y) \in Z$, then $f(y,x) \in Z$;

(iii) If $f(x,y) \in Z$ and $f(y,z) \in Z$ then also $f(x,z) \in Z$.

There is an obvious equivalence relation between such functions f, which holds when they do exactly the same job. That is, one defines f, f^* to be equivalent if, whenever $f(x,y) \in Z$, then also $f^*(x,y) \in Z^*$. It is then straightforward to prove:

THEOREM 1.

Each equivalence class of f's contains an element, g say, which is such that

(a) The Z for g has one element, say 0.

(b) $g(x,x) = 0$ *for all* x.

(c) $g(x,y) = g(y,x)$.

(d) $g(g(x,y),z) = g(x,g(y,z))$.

It will be noticed that each of (b), (c), (d) are specializations of the conditions (i), (ii), (iii). But since these conditions on g are exactly those of commutativity and associativity in a field of characteristic two, it is convenient to use the notation $x+y$ for $g(x,y)$. The proof of the theorem may easily be worked out by the reader, as far as (b) and (c) are concerned. The final step is accomplished by what I call "Conway's trick" because it is used (in a somewhat different context) in his book.[51] This is to take as the value of $x+y$:

$$x+y = \min z \ (f(z,x'+y) \neq 0, f(z,x+y') \neq 0)$$

$$\textit{for all } x' < x \textit{ and all } y' < y$$

Here f is a member of the equivalence class which has already been chosen to satisfy (b) and (c). It is then easy to see that $0+0=0$, and indeed $0+x=0$. So $1+1$ cannot be the same as $1+0=1$, but can be 0. Then $1+2$ cannot be 1 or 0 or 2 so must be a new element 3. Of course, $1+3 = 1+1+2 = 2$ and so the set $(1,2,3)$ is "discriminately closed" (that is, any two different elements x, y are such that $x+y$ is in the set). If now the generation process throws up a new element, it will be labeled 4, and $4+1=5$, $4+2=6$, $4+3=7$ and the set $(1,2,3,4,5,6,7)$ is again evidently closed. It is clear that in general every discriminately closed set is of order $2^r - 1$, for some integer r. Correspondingly, if the set is not yet closed:

THEOREM 2.

At every stage the system can be embedded in a discriminately closed one of finite size $2^r - 1$.

Thus the numbers $3, 7, 15, 31, 63, \ldots$ are going to be of importance in describing quantum mechanical systems.

A useful change of notation is provided by:

THEOREM 3.

The system of theorem 2 has an injection into $V_r/\mathbb{Z}_2$, where V_r is the vector space of r dimensions, and $\mathbb{Z}_2$ is the field of two elements.

The proof of this is easy as soon as one realizes that the injection in question can be taken in the form that the element k, where $k = pqr \ldots s$ in the scale of 2, is mapped onto the vector $(s, \ldots, r, q, p)^t$. Thus $5 = 101$ and so is mapped onto $(1,0,1)^t$, whilst $11 = 1011$ and so is mapped onto $(1,1,0,1)^t$. It is of course important to bear in mind that the dimension of the space V_r is not fixed but increases.

8.4 THE GENERAL CASE

The case considered in Chapter 3 gives a guide to the general case. Here our concern is whether a new element x belongs to a set $S = (u_1, u_2, u_3, ...)$ of already generated ones. We use the same trick of defining a function $f(S, x)$ which vanishes if and only if one of the terms $u_i + x$ vanishes. Because of the asymmetry it is more convenient to use the notation of characteristic functions and write

$$f(S,x) = F_S(x)$$

If we define an addition operation between characteristic functions by the obvious induction:

$$F(x) + G(x) = (F + G)(x)$$

for all x (in play up to the point at which the definition is being made), then it is easy to see that this operation has exactly similar properties to the discrimination operation, so that the set of characteristic functions is itself a discrimination system.

We next make a similar definition of equivalence between characteristic functions to the one made in the last section. That is, we define F, F^* to be equivalent if, whenever $F(x) = 0$, then also $F^*(x) = 0$ and visa versa. Amongst a set of equivalent F's it is obvious that there will be some of the special form

$$F(u_1, u_2, ..., x) = h(u_1 + x, u_2 + x, ...)$$

where h is a function which vanishes if and only if one of its arguments vanishes. (Such F's will in fact be those with the property that the non-zero values of $F(u_1, u_2, ..., x)$, when x is not one of the u_i, depend only on the n variables $u_1 + x, u_2 + x, ..., u_n + x$.) Of course one defines two such functions h, h^* as equivalent just when the corresponding F, F^* are equivalent. Then, by an argument that closely parallels that for theorem 1, we can prove

THEOREM 4.

Every equivalence class of h's contains a particular one with the properties:

(a) $h(x_1, x_2, ...x_n)$ *depends only on the set* $(x_1, x_2, ...x_n)$, *so is invariant under permutations of the arguments;*

(b) $h(x_1, h(x_2, x_3)) = h(h(x_1, x_2), x_3)$; *with obvious generalizations to more arguments;*

(c) $h(x_1, x_2 + x_3) = h(x_1, x_2) + h(x_1, x_3)$.

Here again the proof relies on Conways's trick, but the proof is more intricate than before. Conway uses a least number principle applied to $h(x', y) + h(x, y') + h(x', y')$ to define $h(x, y)$, where as before x' and $y'$are any smaller elements than x, y. But since (b) , (c) are simply the associative rule and distribution over addition, it is intuitively simpler (though more complicated to work in practice) to define $h(x, y) = xy$ as the least element not forbidden, taking account the commutative and associative rules and distributivity over addition. For example, $0x = 0$ and $1x = x$ are straightforward. Then 2^2 cannot be 0 or 2; could it be 1? No, because if it were, then $2 \cdot 2 = 1 = 2(3 + 1) = 2 \cdot 3 + 2$, so that $2 \cdot 3 = 3$ which will in due course be forbidden since $1 \cdot 3 = 3$. So $2 \cdot 2$ is in fact 3 (and it is easy to see that this is not forbidden).

The set $(1, 2, 3)$ is now closed again $(2 \cdot 3 = 1)$; when 4 is introduced, the distributive law alone allows the completion of the table up to 7×7 in the following way:

$4 \cdot 1 = 4$; let $4 \cdot 2 = a, 4 \cdot 4 = b$, so that $4 \cdot 3 = a + 4$, $4 \cdot 5 = b + 4$, $4 \cdot 6 = a + b$, $4 \cdot 7 = a + b + 4$ and so on for $5 + 2$ etc. From these results it is easy to see that the values $0, 1, 2, 3, 4, 5, 6, 7$ are forbidden for a, the values $0, 1, 2, 3, 4, 5$ for b and the values $4, 6$ for $a + b$. the least solution of these conditions is $a = 8, b = 6$,

but the table does not give a closed set. Now the associative law allows the rest of the table up to $15 \cdot 15$ to be filled in which gives a closed set of size 15. In the same way the next closed set of size 255 arises; and the general conclusion is that amongst the numbers $2^r - 1$ which have already been noticed as being of importance in quantum mechanics those particular ones with $r = 2^s$ will be of particular importance.

I call the sequence of discrimination systems isolated here the *discrimination fields*, $\Phi_1, \Phi_2, \Phi_3, ...$ where the number of elements of Φ_i is 2^s, where $s = 2^i$. Because they are fields it is convenient to include the zero element when speaking of them; strictly speaking, the discrimination system corresponding to a discrimination field has one element fewer. It is appropriate to note here that the addition and multiplication operations have a different status here. The addition is discrimination, and the operation at one level induces a corresponding one at the next higher one. But the multiplication is then imposed on a discrimination system of appropriate size; If the system is not of the size of one of the discrimination fields, it can be embedded in a larger system which is. The exact way in which the numbers mentioned here are important will become clearer in the next section.

8.5 THE LADDER CONSTRUCTION

The set of all characteristic functions on a set S, as remarked above, forms a discrimination system. If S is itself a discrimination system together with its zero and so has $s = 2^r$ elements, the number of characteristic functions is evidently the number of ways of specifying s independent elements of S, i.e. $s^s - 1 = 2^{rs} - 1$ (subtracting one for the omission of the zero) and so has the dimension $rs = r \cdot 2^r$ Thus for $r = 1, 2, 3, 4, ..., rs = 2, 8, 24, 64, ...$. But this discrimination system is redundant, since it includes a number of representatives of each equivalence class which corresponds to the set of objects at the level of S. There are obviously $2^s - 1$ equivalence classes (a sequence of s choices of zero or non-zero) so that, if there were single representatives of each of these which formed a discriminately closed set, then it would have maximal dimension s by theorem 3. In fact, this is so:

THEOREM 5.

One can choose one representative from each equivalence class to form a discriminately closed set of dimension S.

It is easy to prove this by considering (for example) the discriminate closure of the set of s functions F_i where $F_i(x) = \Pi_{j \neq i}(a_j + x)$, where $S = \{a_j\}$. In particular, if S is a discrimination field, so that $r = 2, 4, 8, \ldots$ the numbers $4, 16, 256, \ldots$ are the dimensionality of certain discrimination systems, and it is this which gives them their importance.

Returning to the considerations of Chapter 1, it is now clear that levels exist in the present construction. Any set of elements can be specified in two ways: either in terms of the individuals in the set or as the functions characterizing the set at the next level. So we now extend the notion of generation and discrimination by including under the generation process **G** the formation of characteristic functions, and under **D** the multiple discriminations performed by these functions. A run of the system will now allow the self-organization of the sequences of elementary quantum events into higher order events, as mentioned in Chapter 1. The following ladder construction exhibits one way in which this self-organization proceeds; this way is a kind of envelope as it were, of possible runs of the system.

A sequence of levels is called a *ladder* in the following case:

(a) The *foot* consists of k_1 elements;

(b) the *first rung* consists of the $k_2 = 2^{k_1} - 1$ characteristic functions, a discrimination system of dimension k_1, and the discriminate closure of these can be embedded in the discrimination field Φ_i.

(c) The next rung consists of the $k_3 = 2^{k_2} - 1$ characteristic functions of this new system, a discrimination system of dimension k_2, and the discriminate closure of these can be embedded in the field Φ_{i+1}, and so on. Then one can prove:

THEOREM 6.

There is only one non-trivial ladder (i.e. having more than two steps), that in which $k_1 = 2$ and $i = 1$, and this terminates at Φ_4, giving rise to sets of functions of size $3, 7, 127, 2^{127} - 1$ ($\simeq 1.7 \times 10^{38}$).

The proof that this is the only non-trivial one is straightforward, by exhibiting the shortness of other candidates. But the proof that the longer ladder exists is less easy because of the need to show the embedding. However, there is no need to give the proof here, as there is a different version of the construction, the original one due to Parker-Rhodes, which is equivalent and for which the proof has been given.

To introduce this, notice that, for a particular set S it can never be the case that

$$F_S x + F_S y = F_S(x + y),$$

but

THEOREM 7.

If and only if S is discriminately closed, the equivalence class of characteristic functions for S contains a member F for which $F(x) + F(y) = F(x + y)$.

In the corresponding vector space picture such an F is represented by a square matrix, and the addition defined between functions becomes ordinary matrix addition. In the original construction[1] by Parker-Rhodes the sets considered are derived from those described by consistently replacing each set by its discriminate closure. It is evident from the ladder construction that the Parker-Rhodes version (called the Hierarchy construction) is isomorphic to it. Then for Parker-Rhodes the matrix specifying a discriminately closed subset is taken as that leaving unchanged all the elements of the subset, i.e. having them as eigenvectors. These Parker-Rhodes matrices simply differ from those constructed here by the addition of the unit matrix. In the Parker-Rhodes construction one chooses linearly independent functions which then in their turn have as additive closure another discriminately closed subset. The final stages of this version of this construction are then to find $2^7 - 1 = 127$ linearly independent operators in 256 dimensions and finally 1.7×10^{38} in 65536. The last case cannot give rise to linearly independent operators so that this construction stops there and suggests the importance of the numbers $3, 10, 137, 1.7 \times 10^{38}$ (which arise as the cumulative totals of elements) in quantum mechanics. In fact a well known argument identifies 137 as the maximum number of charged particles (electrons) which can be packed in a region determined by Compton wavelength without producing so much energy as to allow pair creation which would render the number indeterminate. A similar argument identifies 1.7×10^{38} as a maximum number of gravitating particles.

The Hierarchy construction is, in many ways, a more convenient one to handle mathematically. The advantage of the ladder construction is that it removes the dependence on linearity of functions, a condition which is very difficult to understand physically.

REFERENCES

49. T.Bastin, H.P.Noyes, J.Amson and C.W.Kilmister, *On the Physical Interpretation and Mathematical Structure of the Combinatorial Hierarchy,* published in *The International Journal of Theoretical Physics, 18,* 455-488 (1979).
50. J.A.Wheeler in *General Relativity and Gravitation,* ed. B.Bertotti, F.de Filece, A.Pascolini, Reidel, Dordrecht, Holland, 1983 (pp.507-8).
51. J.H.Conway, *On Numbers and Games,* Academic Press (for the London Mathematical Society) , 1976.

9. Appendix II. A SYSTEMS-THEORETIC CRITIQUE of PROCESS-HIERARCHY MODELS

Ted Bastin
Pond Meadow, West Wickham, Cambridgeshire, CB1 6RY

Abstract

The possible range of 'process' or sequential or computing models which use the combinatorial hierarchy is considered from the point of view of general systems theory. General principles which must govern any interpretation are discussed, and it is recognized as inevitable that there will be confusion between those who start from the principles which sequential models have to conform to whether or not they ever get to physics as we are familiar with it, and those who make short cuts to real physical problems because it is necessary to think at that level too. The most notorious sources of confusion are listed and offered as appropriate objects of philosophical clarification. Some thoughts on where the wave-function might come from if we are taking the safe but vague road conclude the paper.

Of hierarchies, why some be abolished and some retained.
(Edward VI Prayer Book)

9.1 INTRODUCTION

All of the models which have adopted a process, or sequential, view of the interpretation of the hierarchy algorithms (roughly everything except that of Parker-Rhodes) fall into the domain of what has been called general systems theory. It is true that our thinking has one fundamental difference from what usually goes under the name of systems theory, for in the latter it is usually assumed that one has an existing substratum of material which can be pushed and pulled about by the forces which are under investigation, whereas for our enterprise it is of the essence that whatever the rules of the system turn out to be it is they which provide the substratum. In spite of this major difference I believe it to be useful to bring to bear on our problems the most general thinking from the area of systems theory that we can find.

At a recent meeting on self-organization at Cumberland Lodge[52] (May 1985) I discussed the hierarchy theory with the intention of showing that a lot of it was at least very natural, and at best actually entailed by the general principles of systems theory. I hoped to get the views of the conference. In particular, I started from the position that in any system there must be what we call the 'primal division' between two interacting but quasi-independent subsystems. Please imagine the Venn diagrams. I then developed this position slightly by pointing out that for the case in which I was interested where the total system was the constructive potential of the universe itself, it would always be the case that the most natural interpretation of the primal division was between the part of the universe known to us and the part about which we could only make inferences of different degrees of certainty from what happened on the boundary between the two systems. This much has always seemed to be the point from which any systems approach has to set out, and it follows that if we think in systems terms then we are taking a view of the universe in which the way we get our information about the universe has to be part of our basic theory. I will say in anticipation of what I shall argue in detail later that the computer models that have been formulated by us recently have often been difficult to follow because they have not made it clear where they stand on the primal division in its epistemological interpretation – even to the extent that one gets the feeling that the proponents of the models would prefer not to have to think that way at all. However I have found that the more one tries to make the division go away the more it comes insistently back through a variety of back doors in the manner that logically basic constraints have.

I requested the audience at the self-organization conference for comment on my starting position – and got it. E.R.Zimmermann[1] maintained that the position from which I wished to start was inadequate to define separable subsystems, and that to give meaning to the separation of subsystems one had to define an environment for two or more subsystems to be put into. I think that Zimmermann conceded my main point – possibly thinking it too obvious to be worth making (namely that there had to be separation of subsystems). I replied to him by entirely agreeing on the need for an explicitly provided environment, but

contended that our aim of providing the material substratum for the systems forced us to make logical provision in a way that had not formerly been seen to be necessary for the earlier stage at which one could not make a distinction between other systems and an environment. There was just the system of which we are a part in some way, and the unknown beyond.

It will be obvious that my starting point presupposes that our basic theory must include an account of how our knowledge is obtained. I contend that so much is entailed by the adoption of any process model. It would follow that the intuitions of the founders of the quantum theory about the great change that was implicit in the new epistemology appropriate to the quantum theory, were right by our reckoning, though the use made of that intuition in the mathematics remains inadmissible (and, needless to say because it has been said in these meetings so many times) cannot be saved by what is usually called measurement theory.

9.2 FOUR PHILOSOPHICAL PROBLEMS

Profound theories create philosophical puzzles. It seems that the ageless metaphysical or ontological categorizations can only be lived with in the crystallized forms which seem always to be appropriate to the scientific framework which is just on the way out. In this situation my recommendation – at any rate for most of us – is not a lot of *detached* philosophizing, but a lot of attention as a matter of amendment of the theory itself to those points in which there is philosophical perplexity. The history of conventional quantum theory provides an excellent example. There has always remained philosophical perplexity over the part played by measurement, and this meant that the theory was still not satisfactory.

I find four points at which there has been philosophical perplexity in the process forms of the hierarchy theory:

1.*Levels.* In any way one may try to imagine the hierarchy algebra as governing process, one has at every point in the sequence to decide (or accept) that one is operating at some level, and it follows that this decision (or acceptance) must have physical interpretation. It is true that some like presentations of the algebra which apply whatever level is chosen, but as soon as one has to algorithmetize probabilistic choices one has in effect to make level decisions. Noyes took the decisive way of insisting that algorithms representing process define levels. He correctly saw it a necessary concommitent of the sequential interpretation. I had myself produced a completely sequential model using computing realization which was published in 1974[53], and later on I shall be talking about the possible variety of computing models at which point I shall refer again to that work. However I was rather tentative about it – lacking the courage of my conviction in the face of the then tacitly held common position that the hierarchy algorithm must somehow describe the static structure of space and time and then derivatively base dynamics as had then been the pattern in general relativity. Anyway it was left to Noyes to take the bull by the horns, and everyone is now a processist EPR (except Parker-Rhodes).

2. *Construction.* The main thing that has perplexed us all, EPR, arises directly out of the problem of giving meaning to working at a given level. You recall that we had to go processist because no sense at all could be given to working at a given level in a static general-relativity-like picture. So far so good, but then how did all the darting about between levels happen? Who was responsible for it; or did it happen automatically though perhaps at random? In particular, did all the interesting level construction happen at some big bang after which things just ticked over, or does it keep happening at each new particle process? Worse, is it all anything to do with conscious activity – is there a subjective element? Indeed, does the construction – the generation of the successive levels whether once-for-all or repeated with reiteration and change – represent the process of a conscious investigation? Kilmister supposed it did in his work on the logic of construction. Of course he was concerned with logical necessity in the first place, but insofar as he thought about interpretation he thought that, I think. It is now accepted, EPR, that the hierarchy algebra is logically incomplete without a constructive operation distinct from discrimination, and my belief is that all possible such processes will be as good as isomorphic with the device from Conway which Kilmister originally used to exemplify it. However what interpretation we are to give it I regard as still very much a live issue.

3. *Measurement* (or observation). As far as this paper is concerned, measurement or observation are merely the names we give to our recognition that we have to be on one side of the primal division. We think of ourselves as just part of the subsystem, and of any consciousness we may have of what is going on as being irrelevant to the framework of physical law we discover. Of course this assumption is not incompatible with our directing things so as to get information in an efficient way, but it differs radically from conventional theory in not having to make the structure of theory depend upon a distinction between measured and non-measured events. This position is consistent with what I have been taking to be implied in the phrase 'participant observer'. It does however pre-empt a certain range of views about measurement at the expense of some traditional positions. In particular it is sometimes argued that the very possibility of measuring things in a real world requires that there be three basic units or kinds of dimensional magnitude. The classical ones are mass, length and time, but there could be equivalent quantum-specified ones.

Now we know that in the hierarchy model we interpret coupling constants in terms of ratios of atomic and cosmological constants, and that there are sufficient of these identifications to determine the number of independent units at three. It has to be the case that the classical account of the dimensionality necessary for measurement is a deduction from the hierarchy model, whether or not we have a satisfactory alternative to the correspondence principle. Any other view would require some sort of pre-arranged harmony which would be unacceptable. In any case the derivation of the dimensionality is a major success of the theory.

It is true that I claim to get this consequence about the units from general

considerations about systems, whereas I am using a result from the hierarchy which is a very special system. However I would say that I am really only taking a short cut. The coupling constants have got to be our way into physics if we adopt any process model whatever, as became increasingly evident from a long scrutiny of the logical place occupied by those constants in modern physics which Noyes and I conducted some time back, and I am going to assume the results of that work without further comment.

4. *Completeness.* To my knowledge the term 'completeness' first appeared in physics with Einstein's criticism of the new quantum theory, and it has been in evidence in ANPA discussions recently. All I think I can say about its use is that it indicates a desire to have everything in the way of common sense interpretation that the classical physicists assumed – that there is a stage that has to be reached, in fact, at which one is saying one is doing real physics. These issues are so complex that in what I shall have to say I shall be taking a socratic stance and asking what people's demands are over this completeness, and then – particularly in the context of the appraisal of the range of possible models allowed by general systems considerations which I shall be undertaking – consider how far they can be met (if they can be met at all). In the discussions that have led us, corporately to where we now stand, one very important position has been established, and that is that the world described by the hierarchy algebra is not – in the first place at any rate– a classical world but the different world of high energies from which, as is acknowledged on all sides, we have some hard thinking to do to find the correct replacement for the correspondence principle. I see a dialogue which would be very valuable taking place in which Noyes offers for scrutiny what he sees as the demands of completeness in that new sphere, and where others of us consider (a) whether the ways he has chosen to satisfy the demands are the only possible ones and (b) if not, what arguments would we give for alternatives we might suggest.

9.3 THE MINIMAL MODEL

It seems from systems-theoretic considerations that the following set of characteristics are necessary for all models. I shall consider whether they are sufficient separately. I shall be very grateful for comment on this crucial stage in my argument.

1. An initial set of elements is *scanned* and some sense is given for the result of this scan to be consistent or inconsistent with the empirical situation which the model is put forward to describe or explain. It is usual for the elements to be treated indifferently by the operation of the model since otherwise implicit assumptions have to be made and the aim is usually to have all such assumptions following from the properties of the model and not imposed upon it. However this requirement is not logically necessary.

2. The scan is capable of being repeated with different results, and a numerical measure derived. This measure is interpreted probabalistically. There are great advantages in building up a world picture in which the probabilistic values

are not far from integral values assigned by the model. This last requirement again is not logically necessary, but it does seem necessary to employ a probabilistic mode of getting a manifold with the right properties to represent physical measurement. I should be grateful for comment and discussion on this point also.

3. There must be a change from frequency representation to occupancy representation. This requirement seems to be the most basic way to express the fact that the structure of the model has to change to represent what has been learned from the empirical background. The deductive potential of this requirement has not been realised in our discussions. It would be too much to say that from it one may deduce our particular combinatorial model with its unique hierarchical structure. However it is not too much to say that my argument isolates it as the natural place to start.

4. Occupancy requires the distinction between label strings and operational strings. Without the hierarchy principle of increasing complexity this provision would be artificial. With labeling there is scope for representation of a history, memory or part of the model universe.

5. All sets are finite. This is automatically ensured in the hierarchy model, but it is a more general requirement which follows from the assumption that one is investigating an unknown background. If the sets are not known to be finite one cannot assign frequencies.

6. Any model will have to proceed in a sequence of scans at each of which the occupancies and the labels are updated. It seems inevitable to interpret the individual scan as some sort of interaction at the primal division. In some sense it has to be an observation if we are interpreting the general system as having to do with the world of physics, but of course our special approach imposes on us great specificity at this point, and I am unable to avoid saying that each scan is a scattering process, or perhaps part of a scattering process.

7. Finally I recur to my discussion with Zimmermann and to his principles about the environment having to be specified before one can introduce subsystems separate from the part of the primal division from which one starts. I actually do acknowledge the correctness of what he says, but I do not think we have yet got to the stage at which we can satisfy him in our model, though I do think we are now finding ourselves up against the difficulties which result from our not having a complete grasp of this next step. A great deal of what rest I have to say will be related to this matter. The problem in the context of current physics has been the subject of current discussions at our meetings – particularly with Aerts and his discussion of 'the *one* and the *many*' and how *quantum* theory has not got its own method of defining multiplicity except by appeal – explicit or implicit – to classical physics.

9.4 MODELS WITH MINIMAL INTERPRETATION EXEMPLIFYING THE PRINCIPLES OF THE MINIMAL MODEL

9.5 THE BARE MODEL.

In the 1974 paper to which I have already referred[2], I produced a model to be implemented on a computer which had all the properties listed in the last section for the minimal model and with very little physical interpretation except that the coupling constants had to emerge in something like the correct relation to idealised scattering processes. As with Noyes' models strings were obtained from the vasty deep by random processes subject to weighting by what had happened in past scans. Labels were used in a way consistent with the hierarchy treatment of occupancy. Only strings at the right level were selected by the random selection procedure, and therefore there was no problem about what level one was working at. There was no universal time. The model was not complete in any of the senses that anyone is likely to give to that term and I certainly regarded it at that time as being desirable to show as clearly as possible that completeness is not logically necessary. I would have regarded (and still regard) many of the forms of completeness which people do require – though probably not all – as in fact unobtainable. The reason why I did not pursue the construction of models at that time was partly because no one else took it up, but more because I was worried by the looming and immense problem of taking the first steps in dynamics without any solution to the one and the many problem. I used at that and later times to say that I was only able to give a picture of a universe under the aspect of a *single particle process*.

9.6 NOYES' MODEL.

Noyes again has all the properties of the minimal model, and I hope that my argument will be of some use to him in giving reasons for his choice of computer procedures. When *it* comes to *interpretation* Noyes is *in the same position* as I to claim the advantages of the definite point of contact with experiment of the specific hierarchy model via the coupling constants, and can claim the same theoretical superiority over other models because of the natural inherence of the use of labels in the way the hierarchy describes increases in complexity. Noyes would certainly wish to go further and claim that he has done much to meet the requirement of Zimmermann that there be an environment and the possibility of independent subsystems. (Actually I do not know quote what he would say about this since what he defines is the 'off-shell kinematics' of quantum field theory and not a fully classical situation). I shall come back to this question since it is clearly a vital one.

Noyes' technique to reach a universality which transcends particular interactions between the known structure and the unknown background is to stipulate 'tick' which increases string length universally. I do not think this is objectionable but I find difficulty in understanding how to imagine the construction process which it effects. I have to stick to one statement of Noyes on the subject, and for this I choose his abstracts for the Abbreviated Summer School (held the day before ANPA 7). He says: "To construct *a new theory* we start from arbitrary choice between two symbols..." and so on. I should have thought that if he wished to avoid an uncontrollable sort of subjectivism he would have thought it irrelevant what theories have been or were being constructed to the actual properties of the physical world, and that the construction had to be of some thing. Of course the activity of the person constructing theories has to be represented in the model, but then we should naturally expect to see that actually done.

The 'things' that offer themselves for construction are the universe, and the scattering processes respectively. There may be other possibilities, but I don't know how they would be handled. The hierarchy model really entails that both be taken together since the contact with physics via the coupling constants imposes constraints on cosmological as well as atomic quantities. However, one would have to be very careful in expounding the topic to avoid suggesting that one could identify the construction with the big bang in the crudely realist sense that the term usually attracts. Noyes gets out of the problem by insisting that he is dealing with a specifically quantum mechanical set of concepts all the time, and that the progress toward a fully intuitive dynamics identifiable with classical physics is a long road. I certainly want to understand this position since the negative aspect of it is certainly right at least. However, though I have sweated blood to hold the parts of the argument together in my mind I have not been able to, and it is always the obscurity about what the construction represents that is the sticking point. What I myself think is that the model as we have it simply has not yet got the logical facilities which are needed for the task. Both Pask and Manthey have said that the ideas of concurrent computing are the way through but we have no model before us which incorporates this solution

I make two observations: (1) There is no actual necessity that any given theory must be capable of succeeding in the representation of classical multiplicity. Quantum theory fails at this point as we, collectively, spend a good deal of effort pointing out. (2) We get enormous benefit and profundity from the way that the hierarchy increases complexity of description by changing levels (and we get the advantage of a real representation of how scattering processes are theoretically linked) but this makes it the more important not to bulldoze the barrier into classical pluralism. I think this means – among other things – that we have to make the scattering processes really primitive and cannot construct them from a pre-existing dynamics – at least one that has any tendency to slither into classical dynamics while you are not looking. This goes particularly for velocity. I say all this tentatively: I may get reassuring answers, but I cannot know until what I will describe as my mental block over the construction is removed.

9.7 A MODEL BASED ON SIMILARITY

I have always wanted to exploit to the full the fact that the most secure connection we have with experiment is the values of some dimensionless constants of nature. I have taken this estimate of the right approach to indicate that whatever theoretical constructions have been used to arrive at the values of those constants are the best guides in getting towards a way of representing and – ideally – calculating a wider range of physical magnitudes. It is vital for us to remember at all time that whereas in the usual physical theories there is an infinity of numbers any of which could, in principle, be the result of a measurement and the problems only begin with the question of which – if any – we calculate, our situation is very different. We actually have to propose a meaning for measurements which are not those of the basic scale constants. This is why I have taken the line that the operationalist device of giving experimental meaning to the concepts of theory by identifying them with a particular measurement process is not automatically admissible. This is because the primary identification has set the type of measurement procedure, and all subsequent must be consistent. In particular, it is important to be sure about the 'counter paradigm'. It follows from my argument that since measurement is counting the most important measurements will employ counting techniques, and this position is fortunately consistent with the way we see modern physics going in its universal reference back all the time to scattering processes. However it would be a fundamental error to think that we could wave a wand over scattering processes and that suddenly at this point in our development we were doing physics in the proper 'complete' sense for the first time. The reasons: (a) it is impossible because we can't use our operational criterion in two ways which we have not demonstrated to be equivalent and we have used it for counting and cannot get supplementary support from the Bridgman form, (b) it shortcuts the real work and the insights to be gained therefrom (we have had a major advance from McGoveran in his recursive analysis of dimensionality recently, and this is very much part of the real work), (c) when the 'real work' has been done nothing remains to be done in the way of identification.

The vital step in hierarchy models is in one form or another the assumption of equal prior probabilities for the background processes which generate the hierarchical structure. Together with some sort of ergodic hypothesis to the effect that given long enough the universe will run through all the possibilities that are allowed by the structure, we can get to our beloved constants. In this last sentence I have ridden roughshod over several sensitive points that have been much in discussion. However I have no desire to be philistine, nor in fact any need. I can say simply that the right way to proceed in the absence of knowledge of the background is to treat the finite sets of theoretically allowed possibilities *indifferently*. (Other people have different ways of saying it.) I believe, too, that the same principle has to be invoked even when the logical underpinning provided by Kilmister is used. Now as I understand the situation absolutely everything depends upon this idea of indifference. It is not that other distributions may not occur; probably they do all the time; but by definition they are not part of physics. At least they do not become part of physics in any automatic way without out explaining how we can extend the principle of indifference to include them.

Kilmister, Amson and I tried not long ago to develop an idea which we called 'inexact matching' to break out of our prison. The effort produced no immediate success, but it may be useful to recall it. First, the standard way suggested by systems theory is that of concurrent computing. Pask suggested this years back, and it has been extensively discussed in our present ambience by Manthey. We thought that concurrent computing was too like a counsel of despair. It was safe in the sense that one could create independent starting points of the sort demanded by our ordinary intuition of the physical world (the separable subsystems in an environment of my earlier discussion) but how as one to get in any dynamics except by arbitrary fiat? We thought that if we could 'fuzz out' the values of the constants which had an interpretation over a finite domain in the neighborhood of their discrete values so as to have a set of the right measure we might be able to apply the dynamical ideas that went with the interpretation over a range of values that 'inexactly matched' the discrete value. This suggestion requires one to have a definition of degree of similarity of strings, and Kilmister and Amson explored some of the combinatorial possibilities for obtaining such a definition – obtaining a distance relation in this way. However the distance relation has not so far shown itself adequate to the needs of scattering theory.

The model I now tentatively propose makes a change from the ideas I have just been referring back to, while keeping the notion of inexact matching – or at any rate the idea of estimating the similarity of strings – with a numerical measure. The change depends upon a different view of the possibilities of interpretation of strings, and in this change I have of course been affected by the Californian work on string universe models. It is convenient to say that the earlier work on similarity of strings allowed an interpretation only for the ends of the strings. That is to say for the event horizon corresponding to each string. This way of speaking needs clarification, I know, but I will hope for the moment that it conveys some meaning in view of the fact that the boundedness of the length of strings is what determines the scale constants, and that it is these which receive a conventional cladding in terms of event horizons.

Consider a number j (j finite) of strings which are related by being similar. I call this a *yealm* (pronounced to rhyme with 'realm' or in some parts of the country yellum). It is an array of parallel-aligned straws whose ends are further apart than the ends of the longest constituent straw (and therefore the name is better than 'sheaf'). The yealm is the working unit of the thatcher. A yealm can be subdivided and the parts are also yealms. A yealm can be transformed by parallel sliding of the straws.

Two strings p, q will be said to have similarity $S_{pq} = l$ if they have a segment of length l in common. It will obviously be necessary to place a lower bound l_0 on the admissible range of values that l can take, and when the value of l_0 is specified, the number of yealms is determined.

The next question obviously is how the segments are located, and information about them stored. I think the answer must lie in a process of running

along the strings to their ends and being 'reflected' back at the event horizon. Of course we already have the necessary theoretical apparatus for the expansion and subsequent contraction of the string segment under scrutiny in the hierarchy construction itself, and it is clearly at this point that we fine the general principle of successive scanning which we saw earlier to be universally necessary in systems design coming in. However there is a more novel aspect. The event horizon which is essential for the search procedure has to be what I will call a *local event horizon*. We see all the complexity of the world as having as its paradigm case the global event horizon – every piece of complexity, be it local field, local particle distribution, chair or table. The global event horizon arises out of our presumption of indifferent treatment. If that is not what is actually in the background then we find something else, and it shows up as multitudinous complexity. (I note in passing that we have no reason a priori to think that the complexity will be law abiding. In fact the law-abiding complexity is by observation a special case, and we have a further task in finding out in what the law-abidingness of physics resides.)

The model as described so far could easily be made a complete algorithm in the computing sense. I have just stopped at the point at which it is no longer clear which of many algorithmic courses is the right one, and I hope for comments at this early stage. The clarity of having the algorithm filled out is offset by the reader's feeling that arbitrary choices are being made. The Californian models seem to me to be dictated too much by the analogy of classical particle dynamics at the stage at which they meet this arbitrariness, but that may only be a personal problem, and in any case as I have already made clear I already recognize the importance of getting to something which enables us to think and talk physics.

The model I have described has not yet got to the point of being able to speak of independent systems. We are still stuck with our own system and the unknown. However the introduction of similarity measures was meant to make the bridge. We start with two yealms characterized by two similarity measures (sets of pairwise similarities) and with numbers specifying their respective event horizons. This gives us two nodal points in a primordial spatial relation. Since the specification is not in the hierarchy structure it provides independence in the classical sense (as is required if 'spatial' is to have its usual meaning). The two nodes define a wave function, and this introduction of the wave function concept is appropriately fundamental: it is the bridge into the physical world as ordinarily understood.

What I have been saying has overtones which remind us of Bohm's implicate order and its connection with explicitly ordered things. Nearer to home in the sense of having a lot of structure is Alison Watson's metaphor of the meaning of a 'vertical' description of the world with the ordinary 'horizontal' spatio-temporal description.

It may be that all this is nearer to Noyes' model than at first sight appears. However at the moment I feel his choice of detailed algorithmic form owes too much to the analogy with particle dynamics of a classical sort particularly over the introduction of independence of subsystems, and I associate this trend in some sort of way with the difficulty I have over the interpretation which might allow the independence to be slipped in unnoticed.

A further remark: I would expect to be able to introduce an Ur-energy very early on at the point at which one is forced to introduce a parameter to limit the length of string segments to define the similarity function. It would make sense to have this length correspond to the available energy in whatever scattering process was imagined to generate the yealm. With a lot of energy one could sustain very vague similarities. Again. there would be a lowest available energy corresponding to the universal or bounding event horizon (zero point energy of course). I also notice that in this model one would find the origin of dimensionality in the 'vertical' description if one accepted the vertical account of dimensionality. This would go along with Noyes too; as also would the primitive allocation of quantum numbers.

REFERENCES

52. *Third Meeting of the International Study Group on Self-organization and Dissipative Structures*, Cumberland Lodge, May 1985; selected papers from the three meetings are to be published by Reidel.

53. T.Bastin, "Probability in a discrete model of particles and observations", contribution to *Logic and Probability in Quantum Mechanics*, P.Suppes, ed., *Synthese* **29** (1974) pp. 203-227.

10. Appendix III. AGNOSIA

A Philosophical Apologia for INDISTINGUISHABLES

by A.F.Parker-Rhodes, Ph.D
33 Apthorpe St, Fulborn, Cambridgeshire CB1 5EY

As an introduction to my topic, I propose to offer a brief historical prelude. Somewhere around 1962 I hit upon a series of numbers of which Ted Bastin noticed that the last two (the generating procedure could not produce more than four) were close to two well-known physical constants , the reciprocals of the fine-structure constant and of the gravitation coupling constant. He drew the attention of Clive Kilmister at King's College London, and John Amson at St Andrews, to this series, and these two began to work on the algebraic formulation of the series, whose self-terminating property intrigued the mathematicians as much as the contents of the series did the physicists. I also worked on the problem, albeit divergently, having noticed that it might be based also, and perhaps more profoundly, on "indistinguishables". At a meeting where I expounded the germs of this idea, there was a student who drew attention to the lack of mathematical rigour in my exposition; this set me to try and correct the deficiency. The work took some years, and led me much further from orthodox mathematics than I had expected. Nevertheless, it eventually reached a form in which I could hope to publish it. Pierre Noyes of Stanford University, who had meanwhile initiated the setting up of the Alternative Natural Philosophy Association to further the work, gave valuable assistance in the final stages, and was instrumental in getting it accepted by Jaako Hintikka, general editor of the Synthese series published by Reidel of Dordrecht in 1981.[54]

The approach through indistinguishables has been viewed with suspicion, on philosophical grounds, by physicists in ANPA; for these entities are not physical objects, and it has been customary to look only to physical things to explain physical phenomena. I challenged this position; but it must be done at length if it is to persuade the opposition.

10.1 PHILOSOPHY

Heraclitus and Quantum Mechanics

Heraclitus, whose work survives only in fragmentary quotations by other writers, was odd-man-out among the early Greek philosophers. For a start, he was the only one who was also a king - not that the title in the Ephesus of his day carried either power or honour - but it might account for his misanthropic contempt for almost everyone else (except Bias of Priene) which shows up repeatedly in his sayings. His opinion on the nature of things is summed up in the well-known two-word sentence -*panta rhei* everything is in flux. His choice for the *arche*, the original material from which all else was evolved, was the ever-flickering fire; but he had no intelligible account of how the staider constituents of the world came from it.

Although all others among the Greeks looked for a more stable foundation, current fashion now tends to favour the Ephesian king rather than his opponents. Quantum mechanics for example finds nothing amiss in having some 200 "elementary particles" perpetually changing into each other, some of them so fleeting as to persist for barely seconds, though occasionally detectable because their immense speed carries them a measurable distance in that time . Heraclitus would surely have been delighted by the idea of so rapid a "flux". But there are a few stable, or nearly-stable particles; protons, if not actually everlasting, outlive the present age of the universe by an enormous factor; electrons, neutrinos, photons, and gravitons seem to be indestructible so long as they are left alone. And there are some laws of nature and 'universal constants' which most physicists assume to be immutable. But they offer explanations for the changes which are the general norm, and feel no urge towards — or no hope of — explaining the immutables. The explanations of change are so successful, that the non-explanation of the changeless doesn't worry them.

Heraclitus considered ceaseless change to be so obvious a property of things that it didn't need to be explained. And as there was nothing else, in his philosophy, there was nothing to be explained - only morals to be drawn. We still need the morals, certainly, but we also need explanations and quantitatively reliable ones too.

Parmenides - an Alternative Natural Philosophy

As I have said, Heraclitus was almost alone in his enthusiasm for ceaseless change. The most way-out of the opposite view was Parmenides. He has too often been treated, among the ancient philosophers, in parody, as holding that all motion is illusory. But we have to remember that the Greeks generally regarded explanatory principles as more 'real' than what they explained — like the Indians, whom some think they may have got it from — and that 'motion' was often used metaphorically for 'change' in general. So when Parmenides said that nothing ever really moves, he perhaps really meant that explanations (if correct) never change; or that changeable notions explain nothing. Sounds like common sense — but Parmenides, or his expositors, did go beyond present-day

tolerance in interpreting their metaphors literally. Anyway, one suspects that he was provocatively outspoken, if only with the aim of making others think.

Stephen Toulmin, in a recent article[55] , has made a case for the thesis that the Parmenidean era is now over. Newton was a Parmenidean, (as indeed was Einstein, but Toulmin doesn't say so), inasmuch as he believed in immutable laws of nature, but their purpose was precisely to describe and explain change, or at least motion (though in Newton's case, not without occasional help from God). Nowadays, however, we insist on explaining physical phenomena in strictly physical terms, whose unchangeableness is never more than an unsupported guess. This is to rely entirely on causal explanations , a position I have argued against elsewhere. To explain some effect by pointing to its cause, can never be more than a step along an infinite regression, or one more shot at a vicious circle. What is needed is to jump out of the system concerned altogether and find explanations by understanding its ground-plan. We may then see without argument that the only way from A to B is via P, or whatever.

The well-known principle associated with the name of le Chatelier is instructive in this context. It states that in chemical reactions whatever changes occur tend to counteract or reverse the situation which conditions them. Thus dilution of a solution tends to promote dissociation of the solute; endothermic reactions are promoted by heating; and so on. It is an example of an explanation that serves to help control the phenomena explained, and that stems from an attitude of mind which was foreign to the ancient philosophers. What it explains is one aspect of Heraclitus' ceaseless flux, and it is to that extent anti-Parmenidean.

The principle however is also arguing against causal explanation. It says in effect that chemical causes tend to self-destruct, – what we now call negative feedback. It is, in the grammar of explanation, what self-contradiction is in mathematics - just as positive feedback parallels tautology. Both can be useful in technology, and those who use technology don't need to know why it works or how to explain it. But those whose aim is to develop and expand technology need deeper understanding of a scientific kind; and at least some of us should be thinking towards ultimate explanations. This is a meaningless term for those who deal in vicious circles and infinite regressions - and not those alone. What sort of explanation could be called 'ultimate' ?

The Whole Problem

Anyway, the use of the term 'ultimate' inevitably collides with a host of prejudices. It smacks of intellectual hubris, from which we should not expect anything other than humiliating discredit. It further implies, if narrowly interpreted, a reliance on the strategy of reductionism, which is coming increasingly into disfavour. But worst of all, an ultimate explanation ought surely to be a complete explanation, and as such invite a philosophical account of the whole content of actual or potential experience.

Kant was, they say, the first to call in question the philosophical validity of statements concerning the totality of existence. His doubts, although expressed in philosophical arguments, may have been founded on the mere feeling that such statements overpassed the limits of human competence. Although not subservient, as many of his contemporaries were, to theology, he may not have wished to contradict the opinion that God could not be fully known. Whatever the motives or merits of the argument, it was followed by a long slow decline in the breadth of philosophic vision. In the following two centuries there took place an accelerating abdication, on the part of the more respectable philosophers, of the one topic in which their discipline might aspire to a legitimate leadership role. Not looking beyond mere parts or aspects of totality, they were in competition with scientists, who were increasingly better equipped and better informed about their specialities, and often disregarded their admonitions — but still made progress.

The time has come for this depressing abdication of leadership, through which narrow specialists win all the top prizes, to be reversed, and a return made to priorities which have guided philosophical speculation, with few remissions in all the world's civilizations. There is urgent need today to face the totality of things with courage and, if possible, with understanding. Admittedly this is to espouse an unlimited programme, towards which we can as yet make only strictly limited advances. Though I shall propound an 'explanation' with a plausible appearance of 'ultimacy', it will not explain very much. One step up — more like half a step — is all that I can promise. Beyond there remain several frontiers opposing reductionist explanation, where there is genuine novelty to be accounted for within any successful description adequate to the whole of experience. All I do is to lift the carpet of causality, and show you the floor beneath.

Information and Causality

Physics, in its usual and proper sense, is concerned with causal relationships. To reduce causality to anything else may therefore be seen as misplaced reductionism. But if it is merely a question of translating causal language into another idiom and back again ('bijectively , as the mathematicians say), there is no real reductionism involved . I maintain therefore that any causal relation can be accurately restated in information-language, notwithstanding that philosophically information originates in biology , not physics. Like many organisms it is capable of adaptation. If an event A is said to be among the causes (strictly , necessary causes) of another event B, it is implied that to know that A has occurred is at least relevant to the probability that B will also occur (or have occurred). Thus "A is a cause of B" is equivalent to "(A has occurred) increases the probability of (B will occur)". This translation into information-language can be made for any statement involving causality, and vice versa. In fact, although it acquires a new non-biological meaning in the process, information is a valid way of expressing causality.

Two corollaries follow from this idea. One is that we should expect, and not be astonished at, the relevance of probability to causal relations. And the other, distinct though related, is that whereas causal language tends to suggest determinism, information-language carries no such associations. The high incidence of 'randomness' as a physical concept translates quite simply into the non-existence of what might have been relevant information on a system.

We may further take into account the well-known principle that "information" also has a reflection in physics in the shape of "negentropy" and indeed its relation to causality bears this out. Le Chatelier's principle, for example, can be seen as an application to chemistry of the Second Law of Thermodynamics, that in a closed system entropy must increase if anything at all happens, and information therefore decreases . In other words, "causes self-destruct". This connection (if not the whole cause-information-entropy triangle) is of course well-known; less so is the notion that the universal expansion against the pull of gravity prevents the treatment of the universe, in its earlier stages, as a 'closed' system. There may be a time when the total information actually increases. Might it then be that, at the big bang, the total information action available was simply zero? The postulation of any other value would itself constitute information, thereby contradicting itself before anything followed, unless the value given were (vacuously) its own.

More precisely, let $\mathbf{T_0}$ be the initial information postulated, and let $I(I_0 = t)$ be the information involved in postulating it , which cannot be less than t. But $I(I_0 = t) = I(t)$ which is the number of steps in the optimum search strategy - logarithmic ranging - needed to reach the target t (assumed integral) from a presumptive value p, (normally $p = 1$, but here we are assuming $p = 0$); if $|t - p + 1| = 2^x$ (with x integral), $I(t) = log_2|t - p + 1|$, equal to t if $t = 0$ or 1 and $p = 0$, but greater for all larger values of t. For these, then, $I(I_0 = t) = I(t) > t$, so that one needs initially m more bits to specify the initial quantity of information than the quantity being specified, which *includes* the specification. We then have a contradiction, unless $I_0 = t \leq 1$.

The Inchoative

The science of physics rests upon a foundation including several kinds of statements , which have hitherto seemed to most physicists logically independent. There are, for example, (1) the 'law of mass action', or , more precisely, statistical thermodynamics; (2) a battery of laws (perhaps soon to be unified in a GUT) involving as parameters both dimensionless ratios and dimensional constants; (3) a long list of 'elementary' particles, to which the elegant theory of SU-groups has begun to lend a vague analogy to the Periodic Table; (4) the Uncertainty Principle; and no doubt others. Of these (1) is an immediate consequence of zero information, and (4) at least involves the concept of information. But (2), postulating things eternally so, is indelibly Parmenidean, apparently entailing certain quantity of information at all times and places. As for (3), the less said the better.

If we therefore follow up the result demonstrated in the previous section, we shall have to either explain (2) away, or else reduce it to a tautology which even Parmenides would have jibbed at. I shall take the case where $I_0 = 0$ as the simpler to start with; that is, the universe starts with no information at all, Now any hypothesis if it is to have any consequences, must contain information; if not, it is either a tautology or a contradiction, and in either case leads nowhere. But the hypothesis, that there exists an entity X such that X contains no information does carry information; for it asserts a non-vacuous attribute of X - and a distinctly unusual one. My hypothesis is therefore that there exists something that I call the Inchoative, whose only property is that it contains no information. It is not that we know nothing about it, but there is nothing to know. It is tempting to think of its existing in the BBB area, before the Big Bang, which is nonsense; it is timeless, like God — and devoid of all other attributes. That being so, it may be futile to start asking specific question about it. But if one answer implies information and the other doesn't, the latter must be accepted; and if both are informed (or both not) we must keep both branches alive.

So, is the Inchoative a One or a Many? in other words, does it have parts, or not? Either would tell us something, but their disjunction is suitably vacuous. I shall start with the "many". But how many ? Any specified finite number, or range of numbers, carries information, so it must be infinite; not a specified order of infinity, of course, but haven't we already committed information by excluding finite numbers? No, because all these are allowed for through the existence of finite subsets. "At least a denumerable infinity" is an acceptable answer. Is there, now, any difference between one of these parts and another? If, as in more usual contexts, there are, we would contradict the alternative description of the Inchoative as one whole, and would in any case involve information in identifying each 'part'; therefore all the parts must be indistinguishable from each other.

We thus arrive at a characterization of the Inchoative, defined merely by its "agnosia" — its want of all knowable particularities — as a single whole consisting of an infinity, at least denumerable, of indistinguishable parts. Does that sound like a claim to "knowledge"? Well, I haven't defined 'knowledge' up to now, so I can still assert that this sentence doesn't count as knowledge. It is how one disclaims knowledge, a prolix interpretation of 'agnosia'. But there remains a more important question:

Is it True?

And, if it is, what does it mean? As so often happens in this work, care at each step provides answers for questions one has not yet thought of. In this case, we found two possible answers to the quantity of information initially available: zero, assumed hitherto, but also one bit. The latter is not enough to discredit any part of the preceding 'apophatic theorem', except the one-many dichotomy, but it is enough for an answer to the question, "Does the Inchoative exist ?" So , having dodged the one-many question, we are free to assume that the Inchoative — that is, an entity defined by agnosia — does exist. This is important, because it answers those who might object, with good reason, that nothing follows from nothing. What follows is logically entailed by the thesis that the Inchoative exists - a foundation as exiguous as that of the columns of Coventry cathedral, and as sufficient. But still, what does it mean, the 'existence' of something like this? It certainly doesn't mean that one can say "look , there it is !" It is a strictly non-physical entity, a mentefact, not in the now accepted sense 'real'. In the quantum mechanical jargon, it doesn't "feel" either space or time, and is therefore unchangeable, and 'real' in the ancient's sense. It is , some will say, a piece of rank metaphysics, and the charge, whatever it is worth, is true. But physics, I maintain, cannot be justified by forever chasing its own tail; it must

be seated with dignity upon a foundation other than itself, - a foundation which cannot be denied. Of course, one can deny that the Inchoative exists (as I suspect most members of ANPA still do) —- a No is as good as a Yes — but then nothing follows, and your physical theory has to start a bit higher up. You have to ignore the difference between indistinguishability and distinction, and assume that whatever things are not identical are distinguishable , which in a few but irrefutable instances is demonstrably not the case (I refer to the vacancies of arrayed measurements) and probably many things which would have been theorems will have the status of axioms. The losses are perhaps aesthetic rather than practical; but practicality is not what is wrong with conventional physics. For these reasons I think I am justified in calling for more attention in ANPA to the metaphysical underpinnings of physical theory. On these grounds I recommend the thesis that the Inchoative exists, and the slender, obscure and consequential theory which follows from it . The Inchoative is , one might say as near to nothing at all as you can get and still keep talking. Its existence, if it explains anything , is surely a prime candidate for being an 'ultimate' explanation. In fact, it appears, it explains quite a lot, and even predicts a few things hitherto unobserved (or observed if at all by the wrong people). It might thus be the occasion for a considerable rethink.

10.2 THEORY

Triparitous Mathematics

In the phrase, "The Inchoative is an infinite collocation of indistinguishable parts", the term 'Indistinguishable' is used in contrast with both 'identical' and 'distinct'. A number of authors have entertained a notion of such a relationship and given various definitions under various names for it. The possible definitions are probably finite in number, and though it is important that the definition I use should exactly suit its usage in the above phrase, I cannot be sure that it (or its equivalent) has not been used before, and I apologize to any successful claimant for not acknowledging priority.

My definition is then that two things are indistinguishable if they contribute separately to the cardinality of any class to which they belong, but any proposition referring to either retains its value as true or as false if the other is substituted at every (unprotected) occurrence. The second clause implies that though one can tell them together one can't tell them apart, though the first "tell" doesn't "count" since counting involves labeling each item serially implying that when you come across a twin j of an i you have already counted,you say "has been counted" — which will lead to a wrong result. One can however tell how many items are in such a class by finding the smallest set into which it can be mapped.

I have used the term "triparitous mathematics" to include all mathematical theories which allow for the intermediate parity-relation of "twinship" between "identity" and "distinction", defined as above. Actually, there are six parity-relations, not three, since simple negation of any of them is ambiguous. "distinct" means either twins or identicals, "bipar" means either identical or distinct, and "cardinant" (= non-identical) is either distinct or twins Whereas biparitous mathematics is often presented as founded on Set theory, triparitous theories are derived from analogous entities called Sorts. In a "perfect Sort", any two members are twins (I use "twins" for the mathematical relation between entities which are "indistinguishable" or "ibs") and whatever its cardinal its ordinal is 1 because twins can't be ordered; in an "ordinable" Sort, all of whose members are distinct, the cardinal and ordinal are equal; in a "mixed" Sort of cardinal n, the ordinal o satisfies $l < o < n$, and equals the largest of all the ordinable Subsorts.

Ordinable Sorts have all the properties of Sets, except for the company they keep. Mixed and perfect Sorts clearly don't; this is why I have not used "Set" with an adjective for them, but made a new term out of a common noun (after the usual but rather deplorable custom in mathematics). Sort theory turns out to be very different from Set theory (which of course it contains). The difference that is most obtrusive comes from the need to maintain the distinction between "together" and "apart" rigorously and, so far as possible, perspicuously. The effect of this is inevitably that the grammar of statements in a triparitous theory is in general context-sensitive. The referent of a given symbol need not always be the same through all its occurrences in a single statement: they may be twins. There are ways round this, but they are 'cosmetic' only; if you use any diacritics to distinguish twins, you have to have rules for putting them in (and taking them out). Such rules (I have found) make things even worse, certainly slower.

Triparitous Degeneracy

As one develops the theory of Sorts, one keeps coming across instances where the presence of twins reduces constructions which are good enough for Sets to triviality. If f is a mapping from an ordinable into a perfect Sort, all the images by f are identical to a free choice among all the members of the perfect Sort. Suppose a, b to be two distinct members of an ordinable Sort, and h, i, j three twin members of a perfect Sort. Then $fa = h$ implies $fa = i$ or j (they are all twins). Not so the inverse mapping; considered as a Set of ordered pairs it is in the same case as before, but if we look at it as defining a functor f, that is taking all the pairs simultaneously, context-dependence, and the meaninglessness of a functor without its argument, means that the "h" in fh may not be the twin of i in fi; it is the formula fh we look at and if $fh = a$ and $fi = b$ because of the mapping, they are in fact distinct. Although we can't count twins one by one, we can map a whole Sort of them onto an ordinable Sort in one operation, symbolized by the "functor" f, even if this does not distinguish them all (having used up both members of the ordinable Sort, fj must map the same way as one of the other twins). If we had a third, distinct c we could distinguish all three, and with a further d, one of the four would go unused. That is how we can define the cardinal of a perfect Sort.

Insofar as any function implies a corresponding mapping, the degeneracy of mappings into perfect Sorts means a great reduction in the variety of functions definable cver such Sorts, and a considerable simplification of the whole theory in comparison with the biparitous situation, where functions can be arbitrarily

defined, if necessary by writing out operational tables in full. This kind of degeneracy turns out to be the source of a most significant patterning of structures unparalleled in biparitous theory.

Empirical Existence

This 'pattern' has to do with what class of triparitous constructions can be said to "exist" in a fully empirical sense. I take this to mean that one can find, or expect to find, among things observed or rigorously inferred from observations, something sufficiently isomorphic with the triparitous construction said to exist. This notion is not interesting in the case of biparitous constructions because they all satisfy the conditions. Anything, from a simple Set to an elaborate theory, may have an empirically verifiable manifestation. But the more complex the structure the more likely it is that such manifestation will not by found; but even if not, we shall know what it is we're looking for : a structured Set of some kind mapable onto the structure in hand. Given a perfect Sort, however, we get nowhere this way, since mappings onto such a Sort are necessarily trivial. What we can do is to construct a biparitous system from the triparitous one, using the whole of it and nothing more. Obviously, if one or more elements are left out from the perfect Sort, what follows refers to a proper Subsort only; if we use all the elements but leave out some function definable over them, we misrepresent the system, though to have functions in the biparitous representation which have no triparitous original is harmless for they can be merely disregarded when looking for manifestations; finally, if elements need to be added to the triparitous construction to find a biparitous model, we are way out.

If precisely the elements of the triparitous system and all but only the functions definable over them, can be used to construct a truly biparitous homomorph of the given Sort, we can argue as follows: if the triparitous T 'exists', and if a biparitous homomorph B can be constructed to the above specifications from T, then B also 'exists' in the same sense as T; but, being biparitous, we can look for an observable manifestation of B which' if found, will be 'sufficiently isomorphic' also with T to justify the claim of 'empirical existence' for the triparitous Sort T. Every perfect Sort has as many Subsorts, distinct from each other by differing cardinalities, to represent all its twin members; but the representation of functions may fail , as also may the symmetry relations among the elements.

The great majority of perfect and mixed Sorts do not meet these requirements; those which do , I call "rational" Sorts. Rational mixed Sorts arise from perfect ones by adjoining to the latter either a perfect Sort composed of functors under which different Subsorts are invariant (of which there are more than one kind) or a singular Sort whose sole member is the "Smudge", that is, any one unsuccessful member of the initial Sort.

Arrayed Measurements

If any empirical object is to be acceptable as a manifestation of a rational Sort, it must in general contain parts, in the right number , which are indistinguishable. We can't insist that any components of such an object should be in a strict sense distinguishable though (for a mixed rational Sort) that would be an added confirmation. However, it is readily shown that indistinguishability can never be demonstrated of actual concrete objects. One may legitimately ask then whether this condition can ever by met.

One situation which clearly meets the requirements arises whenever a set of measurements are required to specify an object or event, or its location (relative to an arbitrary origin) or any other attendant property. The simplest case to consider is the location of an event in space-time; this requires four measurements, one of which (time) is determined by quite different techniques from the others. One thinks of these as represented by Cartesian coordinates, but we need not be so specific; the point is that before anything is done we know we have four vacancies to fill, one different from the rest. Of the spatial vacancies we can postulate strict indistinguishability: (1) they are not 'concrete objects', (2) to change the order of two of them (before any measurement) leaves everything as it was, (3) each one, considered on its own, is identical to either of the others.

On beginning to make the measurements, the first step is to map the vacancies onto a set of labels, either three of those provided by the Combinatorial Hierarchy (see sec. 12), or the x, y, z, t of a cartesian framework. This distinguishes them and at the same time determines what measurements are required. Finally, we make the measurements and write the results in the labelled vacancies. The 'time' vacancy doesn't actually need labeling, since we know beforehand that it is distinct form the others. A majority of the perfect and mixed rational Sorts yield perspicuously enough to this analysis (though perhaps trivially for perfect Sorts). The two largest, from which the whole enterprise started, do not show obvious traces of an arrayed measurement basis; but since if they did the labelling process would be very hard to envisage as a practical operation (127 indistinguishable vacancies? or 2^{127} ?) this is perhaps unsurprising. It is in general however convenient to denote the "form class" of a mixed Sort of ordinal o by $(c_1, c_2, ..., c_o)$ where the c_i are the cardinals of its perfect subsorts; the form class of a perfect Sort is $[c]$, of an ordinable one $[1, 1...]$, and of spacetime $[3, 1]$ or $[1, 3]$.

Overview of Rational Sorts

At this point I insert a brief summary of the rational Sorts. First of all, the empty Sort is rational, since it is identical with the empty Set, whose 'manifestations' are infinite in number and also in triviality. Nevertheless, it bears thinking about.

The singular Sort, of the form class $[1]$, is in much the same case, but it does have descendants. Adjoined to the class of Identity functors, it yields $[1, 1]$, and if we bring in further the identity functors over the class of identity functors, we get $[1, 1, 1]$. This can go on for ever, and generates the whole class of ordinable Sorts isomorphic with the integers. Note however that individual 'classes of identity functors' are not rational Sorts, nor is any sequence of them having gaps or missing the first member; each term is only defined as a function of the term immediately below it, which must be included to satisfy 'rationality', recursively till we reach a component of the Inchoative.

The perfect Sort of two members is rational, of form class $[2]$. So also is the result of adjoining its 'smudge' which we might represent as the form class $[2, 1'']$. The notation " $''$ " indicates that the final subSort is represented by the "smudge functor" $I\Phi$ as sole member. However the smudge of any Subsort of a perfect Sort can be shown to be twin to each single member, so that $[2, 1'']$ is in effect identical with $[3]$, though if the extra element represents not the smudge itself but the smudge-functor which like functors is distinct from its arguments, it is best to keep the notation $[2, 1'']$. "$[3]$" refers to the "closure" of $[2]$, including the smudges of all its Subsorts of which the only proper ones are the elements themselves. By adjoining endomorphism functors to $[3]$ we get $[3, 7]$ and then $[3, 7, 127]$, and finally ,using a different, non-repeatable, endomorphism formula, we get $[3, 7, 127, 2^{127}*]$, which terminates the series and also gives an otherwise inaccessible $[2, 3^*]$. The notation " $*$" marks the last subSort as represented by the functors ID' over all non-empty subSorts D' of the preceding term D. As already mentioned there is a perfect Sort $[3]$. Adjoining to it the smudge functor gives $[3, 1'']$. The closure of $[3]$ is $[7]$ which however is not rational; but by adjoining endomorphisms of the non-repeatable class we get $[3, 7^*]$, which is regarded as distinct from $[3, 7]$.

All finite perfect Sorts with cardinal 4 or more are non-rational but there is at last one infinite perfect Sort (whose cardinal is surprisingly $2^{\aleph_0}$, unless you an 'Intuitionist'), of which only subsorts not greater than $\aleph_0$ are strictly rational. The Inchoative as a whole is not demonstrably rational; but there may be an infinite number of rationals in the form class ∞ just as there are an infinity of cardinable Sorts, which are rational from birth as it were.

The Combinatorial Hierarchy

I have mentioned that the logic underling the concept of rational Sorts rests on the possibility of seeking empirical evidence for the existence of biparitous structures built out of *and* isomorphic to the Sort concerned. But isomorphism is a transitive relation, and once one has demonstrated that a given Sort is rational, one is no longer restricted to using the biparitous representation which first proves it; any biparitous construction isomorphic to the primary one, even such as are not constructible in the manner prescribed from the triparitous Sort, are available for what can be got out of them.

Now there is a very productive instance of this in the case of the rational Sort $[2]$. Its two elements can be represented by the elements 0,1 of the ring of modulo-2 arithmetic, notwithstanding that these are distinct and have no *direct* derivation from indistinguishables of any kind. The two two-argument functions definable for the Sort are represented by $+$ (for the "G" of *T of I*) and by free choice for "J". The latter is represented in various "random" operations, but the former which is the essential constructive device, by which the Combinatorial Hierarchy is derived[56,57] . Further levels of the Hierarchy, corresponding to the form classes $[3], [3, 7], [3, 7, 127]$, and $[3, 7, 127, 2^{127} - 1]$ are constructed of matrix-like structures built out of the elements 0,1; the rational Sort construction has analogous structures for the first three levels, but a different type for the last level. It also yields two rational Sorts, of form classes [3,1"] and [2,3"] missing from the CH model, which also lacks parallels for the ordinable and infinite Sorts. But the CH treatment more than compensates for this in the greater power of biparitous mathematics and the greater freedom in its physical interpretation. Neither the ordinable nor the infinite Sorts seem likely to yield anything new by way of alternative biparitous representations; such extensions are likely to be limited to the form classes [3,1"] (which is not mentioned in *T of I*) and [2,3"] associated with the parton structure of hadrons (and by implication of leptons also).

10.3 THE FINITE SORTS

Since the theory is relatively weak on genuine prediction, as opposed to retrodictions or explanations of what is already known, I shall leave detailed discussion of the infinite rational Sorts till later, and deal here with the finite ones, which offer comprehensive explanations for many things, few of which will surprise the well-informed reader. The basic idea is that since rational Sorts have been selected from Sorts in general so as to have 'empirical existence', it would be nice if physical correlates could be found for all of them. This seems in fact to be the case; a few even have more than one interpretation. It is not of course claimed that all physical phenomena have such an explanation.

Few words need to be said about the empty Sort. It is exemplified by the "fact" that Nothing exists. The evidence for this is that much of spacetime appears to be empty. But as we have not yet got any concept of spacetime, we can't begin to explain how "nothing" and the various contrasting "things" are interrelated. All we can say is that the rationality of the empty Sort implies that the structure of the universe is basically discrete. In some sense or other, there are gaps, but that's as far as we can yet go.

The Sequence of Events

I have explained the ordinable Sorts are all 'birthright' rationals. That means that there is an infinite simply-ordered Set which has a unique status as indefeasibly empirical. Since it can be argued that all physical observations are of elementary events, or of congeries of events, each of which is a scattering encounter between elementary particles (or a particle's disintegration or 'decay') the obvious interpretation of these ordinable Sorts is that they represent in a simple-ordering the successive events in the history of the universe. The singular Sort [1] will then represent the Big Bang, and subsequent members of the series events occurring later and in due course elsewhere. Of course they won't be *much* later but they will be, as we have seen, discrete – which some cosmologists deny.

As for the 'elsewhere', if we consider a "time-slice" (for relativity theory presumably a light cone, but we need not be too specific) of the universe, the events in it must be finite in number (else the totality would not be simply-ordered as the ordinable-Sort series insists) we can prove that, if there is a decidable

parity-relation between any to orderings of a finite collection, it must be that of twinship. Since there are no rational perfect Sorts with more than three and less than an infinity of members, the events in any such 'time-slice' must have at most three decidable twin orderings, and any other must be expressible in terms of these three . An analogous situation (a bit more than an analogy actually) is presented by orthogonal axes which are mutually congruent but functionally unrelated, whereas any non-orthogonal companion can be expressed as a function of the orthogonals. But the up-to-three indistinguishable 'orderings' — the terminology is none too apt, but it's an on-off affair — are not any kind of 'axes', and their only use is to demonstrate that to locate any event in spacetime one needs only four ordinal numbers, specifying its position in each of the orderings chosen and the time-slice in which it occurs.

This is sufficient to show that the universe of 'events' is sufficiently mapped by a system of arrayed measurements of the form-class [3,1] — which has a rational Sort to justify it, that is, of which it is the empirically-observable manifestation. When we come to consider how the vacancies in this array may be filled, we have to go over to the combinatorial hierarchy to define the identification of 'axes' (other than time, which depends on the kind of 'time-slices' we use) in such a way as to suggest how foot-rules ought to work. But we have many steps to go within the *T of I* before we can decide what we mean by observation.

Observation

Events communicate with one another through the particles they give rise to. An event occurs if and only if two or more particles meet within their mutual scattering-radius; (which particles may be described as moving backwards through time, when we have to account for the decay of a single particle). The trajectories of the particles 'carry' (albeit in a Pickwickian sense, as we shall see) information sufficient to specify the nature of the event in hand, and at least relevant to the information carried by the emergent particle trajectories; 'at least' because the latter are always in some measure indeterminate unless there is only one of them.

Although this process in general involves a loss of information between each event and its successors, it is possible to contrive experimental situations in which the effect is counteracted by preserving in temporally persistent form (as in the particle tracks in a cloud-chamber or photo-stack) information from many connected events. By a sequence of further events, nearly all strictly unobserved, much of this information is transmitted to a human brain trained to interpret it. The conclusion of this highly indirect chain of causation is called the 'observation' of one or more events; it involves the leakage of a fraction of the information output by the event observed by way of an artifactual diversion into the observer's consciousness. Despite what would appear to be the gross implausibility of the idea, many eminent authorities have laid great weight on the fact of observation as an essential ingredient — not merely of our knowledge , which all must agree, but also in the actuality — of what is observed. It is indeed laudable that the unity of the world and the interconnectedness of all its parts should not be lost sight of, but interconnections ought not to be asserted without at least a tentative explanation of how they operate, – which in this instance seems to have been lacking.

A particular problem is raised here by the law of increasing entropy, which entails decreasing information within any closed system. The existence of physics depends on the existence of physicists which is only possible because the universe, especially the surface of planets, is at present in a sate where such systems are rather exceptional . Although information keeps on getting lost, a lot of it still survives, often frozen into near-perpetuity; thanks to this we can carry out our observations of elementary events as well as matters on a more human scale. Our lives are surrounded by traces of the past, which we tend to explain as 'historical accidents'. The distribution of continents and islands on the Earth, and the state of its atmosphere, and much more, are consequences of large-scale events occurring over many aeons. On this scale, the discreteness of information-transfer from event to event at the elementary level is imperceptible. Though Europe and America are drifting apart by several centimeters a year, we have no reason to doubt that this is the effect of a vast number of elementary events. But it is much more usefully described as a continuous motion powered by the continuous upwelling of material in the mid-Atlantic Ridge. The bizarre details at the elementary level are simply irrelevant here.

Anyway, they are not observed and never will be. Thus, the question whether they *are* observed cannot be consequential. And if not in plate tectonics, why any more so in a linear accelerator?

The Disposition of Events

Nevertheless, one can from time to time observe, in the rather remote fashion indicated, an elementary event or a consortium of events. What can one then do with information thereby obtained? One can't really assign it to an ordinal vector, of the kind initially considered, since one can never know its position in any ordering , even if it were possible to define an ordering of all past events. But we can choose an origin and specify axes for a Cartesian [3,1]-vector, whose components for a given event can be at least approximately measured and reordered. In this way observed events, and theoretically also unobserved ones, may be disposed in geometrical space (GS). If we could accept the idea that the particles emitted in an event could be assigned specific trajectories, these could also be mapped onto the same GS; but for this to be possible we shall need to know the locations of the two events between which the trajectory extends, whereupon we can plot the path followed (which in the absence of relevant force-fields will be a straight line) and infer the position of the particle at any instant between the two events. This will be a [3]-vector referred to axes manifesting a perfect Sort.

In fact all perfect Sorts have such an interpretation so long as they are rational. A [2]- vector can represent the orientation of a spin axis, and a [1]-vector, which relates its referent to the Big Bang, can be interpreted as a measure of the time when the event occurs. But, as I have said, there is another perfect rational Sort, of infinite cardinality. Can this also describe some kind of 'space' ?

If we assume (to start with the 'worst' case) that each component of the

[infinite]-vectors involved can have a countable infinity of values, any such value can be defined by binary logarithmic ranging, answering a countable infinity of yes-no questions, which are not enough to outstrip the uncountable infinity of the vector components. Each component may therefore be treated as capable of two values only, a 'yes' and a 'no'. The infinite perfect rational Sort can thus be manifested as an infinite Boolean lattice (possibly with many uninterpreted points), isomorphic with an infinite "similarity-space". Finite similarity-spaces are a familiar tool in various branches of science, and their uses and properties are well-known; a familiar example is a "dichotomous key" for the identification of biological specimens. They convey qualitative descriptions of things. I have hitherto described the resulting kind of 'space' in physical contexts as 'disordinate' (because each point has an infinity of neighbours); it is not too happy a term, for there is *some* order there, but I propose to stick to it. Geometrical Space will be contrasted with Disordinate Space (DS) hereafter.

The objects, or events, which are described, however qualitatively, by DS are the same as those located in GS. This of course implies a mapping between the two spaces, but in view of the disparity in structure between them, and the consequent impossibility of specifying points of either in terms of the other, this can only be a random mapping. Furthermore, it is GS rather than DS which we directly experience in normal consciousness, and therefore this mapping can only be apprehended as a structure in GS, the nature of the 'structure' being that of a probability distribution function appropriate to the phenomenon in hand. A simple example of such a distribution is worked out, under the name of 'disordinate statistics', in *T of I* (p.150 ff).

There is no reason to assume (as is conventionally done) that GS is strictly continuous; there are arguments, which I have briefly alluded to , for thinking it may be discrete. However probabilities, being essentially ratios, introduce the need to accept general rational numbers associated with the GS coordinates, which are thus dense rather than discrete. Insofar as the objects we observe in GS are at least in part influenced by DS descriptions, we shall not perceive GS as observably discrete (not even if it were governed by a graininess coarse enough to be seen as such, which is certainly not the case). One inevitable effect of DS is to soften the outlines of things in GS.

There is also the important question of when DS and when GS is relevant for the description of phenomena. This is virtually answered by saying that two events must be located to determine the path of any particle passing from one to the other. In the presence of only one event, or of none at all, we have no alternative to DS; once two or more events are in question, we go over to GS. Because our minds have evolved to cope with actual events in large numbers, we are conscious (normally at least) only things "in GS". This is an ever-present cause for confusion, and has (I believe) tripped up some very eminent thinkers.

The Connectivity of Geometrical Space

There are various theoretical possibilities for the overall structure of GS as a whole. Briefly, an intrinsic uniform curvature, which could be positive, negative, or zero, but could not compatibly with the cosmological principle vary from region to large region, could obtain. The question is one which evokes surprisingly little concern. It is for example customary to assume, following Aristotle, in 4-D, that "space is (hyper)spherical", implying but not seriously believing that the overall curvature is positive; that the word 'sphere' is much more familiar than 'sella' may be why people don't more often say that it is 'hypersellar'.

Either sphere or sella would involve a radius (or other measure) of curvature as an additional universal constant, which the underlying metaphysics of my theory would require to be somehow computable *a priori*. Not seeing how to do this I naturally opt first for the third alternative, a hypertoroidal space, which has zero intrinsic curvature. This is supported by some and opposed by other observation. In theory, it could become possible to measure the curvature of space with the average density of matter, and data on this, though not very reliable, seem to most authorities to lean heavily towards low values implying negative curvature; but revisions seem mostly to be upward. That is the main source of opposition to my conjecture.

The main support is geometrical, in that it is known that some phenomena have an intrinsic chiral bias; that is, some events produce more (say) left-handed spinning particles. An even more ungainsayable fact, of really the same kind, is that anti-matter is virtually banned as a permanent constituent of our world. But with any non-zero curvature of space one would expect every chiral particle to be balanced somewhere by its antichire (which would be disastrously in conflict, one might say, with the 'anthropic principle'!). None of this need happen in a hypertoroidal space however, for this alone is fully compatible with any arbitrary chiral bias.

It is possible also to suggest more or less plausible ways of computing the average density of matter. One of these is set out in *T of I*, p. 193. The value it yields is within one percent of that required for a zero intrinsic curvature by relativistic gravitation theory (very accurate for the present state of the art). But there are many objections which might be made to this calculation and not too much confidence should be placed in the result.

Thus, while the facts remain in dispute, and long will, there is a good (but not overwhelming) case to be made for hypertoroidal connectivity. One doesn't have to visualize a four-dimensional American doughnut — any regular crystal lattice will do, if you don't object to an infinity of exact copies of your universe. There are no analogues however of a hypersphere, still less of a hypersella, in objective experience.

Quantization

There is a further point in connection with chirality which may be important. In Sec. 10 I explained how the making of an arrayed measurement involved three stages separated by two operations. The stage of strictly indistinguishable vacancies is terminated by 'labelling' them, and the labelled vacancies are then filled by appropriate measurement techniques. To say that a given quantity is

'quantized' is to say that it cannot be divided indefinitely, that there is an ascertainable value than which only zero is less. In the above scheme , a hypothesis that a measure is, or is not , quantized, must be introduced before the labelling of vacancies, for after that only empirical facts are taken account of.

There are two possible sources of quantization. If, as may well be the case, *all* measures are fundamentally discrete, all are in one sense reducible to integral numbers of quanta (though this might require a non-linear transformation to get the conventional measurements corresponding to them). A seemingly different property is the quantization of action, which is certainly linear, and on a much coarser scale than is plausible to postulate for say distance. The first type is a preexistent attribute of the measures affected, theoretically based (or refuted) and not a supplementary hypothesis like the second type. It is the latter I am concerned with here.

The hypothesis of non-quantization (which is easier to argue from than its contrary) assumes that we can find as many diminishments r, such that $0 < r < 1$, as we have vacancies in an array, by which to diminish whatever actual values x may turn up in a measurement, to get not truth but some plausible fiction. When the vacancies are labelled we can choose numerical values for each of the r's, which are in general all different and chosen at random. If a 3-dimensional figure, say a tetrahedron, has three edges of lengths $a < b < c$, it will be chiral; but on applying the aforesaid diminishment we shall in general not find that $r-2a > r-2b > r-3c$ because the r's are unconstrained. Thus chirality cannot be an *intrinsic* property of a figure defined by continuous variables. Chirality survives unchanged only in the absence of all 'hypotheses' of this kind, and is inconsistent with arbitrary 'diminishment'. Since the hypothesis of non-quantization thus fails, it follows that all measures involving chirality, actual or possible, are quantized. Well — are they?

The simplest such quantity is angular momentum or 'spin'. It is associated with the rational Sort [2,1] in which the [2] can specify the orientation of the axis, and the [1], as usual attached to time (or its reciprocal), assigns a frequency or rotational speed. This is of course well known as a 'quantum number' and moreover it has been shown to be communicable , as rotation, to more massive bodies bombarded by particles of monochiral spin. Any quantity of the same dimensionality as angular momentum, action in particular, is also quantized.

A possible exception might be seen in the case of charge , which is quantized but apparently not chiral. Now I have not proved that a non-chiral quantity can not be quantized, but one might expect that its quantization might reside in the discreteness which may be present universally. However the pattern of the so-called CPT conservation law shows that change must involve some kind of chirality, and therefore need not be regarded as exceptional in this respect.

The Uncertainty Principle

It is obvious that, in measuring any quantity subject to quantization, the least error that can be made, assuming that complete accuracy is impossible, amounts to one quantum, since fractions of a quantum are meaningless. For either dense or continuous measures, ' complete accuracy' requires infinite information, and is therefore impossible; but in the case of a discrete measure, the information required is finite, but likely in practice always to exceed what is obtainable.

There are a number of pairs of measures, whose product has the dimensions of action, which are also those of angular momentum (which is physically interconvertible with action); this as has just been explained is quantized, and its quantum is Planck's constant h. Thus the product of the errors in any two such measures cannot be less than h. Pairs of measures covered by this relation are energy (AT^{-1}) and time (T), angular momentum (A) and phase (I), linear momentum (AL^{-1}) and distance (L). Electric charge, which is also quantized, does not participate in such pairs, since its dimensionality is irrational $(A^{1/2}V^{1/2})$. Thus the minimum-error ('Uncertainty') principle is largely confined to the pairs of measurements cited.

Any measure which is quantized is subject to a conservation law, stating that in a closed system it can neither be increased nor diminished, since a quantum cannot be created or destroyed (though two antichiral quanta *can* spontaneously arise or mutually destroy each other). Each of the above complementary pairs consist of an 'active' and a 'locative' member; if the latter's values are a dense set, capable of infinitesimal change, the active complement must also be subject to a similar conservation law, since their product is so and the locative member is unconserved. Thus the existence of conservation of charge, energy-mass, angular and linear momentum, are fully in accordance with expectation. The argument does not necessarily hold up if space-time intervals are discrete — this is the only apparent evidence against discreteness which I have encountered; the want of empirical confirmation of discreteness is of course equally expected. We really don't know.

How Many Forces?

At various times during the last century or so different 'forces' or 'fields' have been distinguished by physicists. Once could name gravitation, electricity, magnetism, the weak and the strong interaction and the 'chromodynamic' field. Many of these have since been explained in terms of one another, first the electric and magnetic fields, then electromagnetism and the weak interaction, were unified; and the strong and chromodynamic fields likewise. Now there is a chase on for a Grand Unified Theory, which is having a good success except, so far, that it can't quite cope with gravity; but this is confidently expected to follow suit in the end.

Neither the *T of I* nor the Combinatorial Hierarchy approach has had unqualified success in finding an explanation for all this. Two 'coupling constants' have been identified, both on frankly numerological grounds; the two largest cardinals (in both systems), 137 and $^{17}10^{37}$, are reciprocals of the required constants for electromagnetism and gravitation respectively, the first being less than .03% out, and the second quite close. The weak and the strong interactions should have analogous coupling constants, but they have not been convincingly demonstrated

by either of our approaches. It may be that they will be explained by the GUT theorists.

If however our numerological identification of the gravitation coupling- constant is not a diabolical coincidence, the implication may be that gravity will not, in the end, succumb to GUT treatment. It will be interesting to see what success comes to this research.

Larks

It is a necessary consequence of the thesis that the Inchoative contains no information , that its ultimate indistinguishable parts should possess no attributes. It follows that they can never be observed as such; everything which is observed must be an aggregate, either of ultimate "agnomes" or of more complex constituents. All the known 'elementary particles' above the level of quarks are 'observable' in that they can sometimes be 'seen' in particular places or following particular trajectories; they can be said to 'feel' spacetime. They can be located, that is associated with distinct vectors, and therefore distinguished by attributes of location, even though they remain qualitatively indistinguishable.

These remarks apply equally to baryons and leptons, and run counter to the idea that there is a likeness between leptons and quarks which baryons do not share. I am led thus to claim that there is a case to be made for regarding leptons as having a structure of components analogous to the quarks which are regarded as the ultimate constituents of baryons. Whether these are truly ultimate, that is to say 'agnomes'. is debatable. Quarks conventionally, and their lepton analogues "larks" hypothetically, are assigned attributes of 'colour' and 'flavour'; but these could be seen as specifying a subspace of the *discrete* 'space' which manifests the [3,7] rational Sort involved here (there is another [3,7] Sort defining energy-momentum 'space' which is usually assumed to be 'continuous'). Each quark could share its time between different points of this subspace without contracting 'attributes', subject to appropriate statistical parameters belonging to the points not the quarks. This kind of time-sharing permits the fiction of dense values for quantized measures, applied for instance to charge in the proton-electron calculation (original version).

This 'fiction ' works only because the colour-flavour space is discrete, so that different quarks, or larks, can occupy the same point in it and not be distinguished at all, whereas in spacetime, whether it is dense or microquantized, two particles must occupy different points, or be observed as one, even if their intrinsic properties are identical. Quarks, and larks, are thus at least a step nearer to 'agnomes' than leptons and baryons are, even if thy are not actually there.

Of course , quarks have been 'observed' inside protons — or at least protons have been found to be internally lumpy – whereas no structure has ever been found in the much larger electron. But the larks in the electron are massless (as is assumed in the m_p/m_e calculation), whereas at least some of the proton's quarks are credited with mass. If so, this is an attribute which can't be argued away (might one devise a comparable experiment on say muons?), and quarks would have to be split up into (perhaps at last) ultimate 'agnomes'. Massless components confined within an electron would not give it any observable inhomogeneity. There is evidently a lot to be found out in this field.

10.4 ON TO INFINITY

I shall now pass over to the discussion of some applications of the 'disordinate space' which was earlier indicated as the principal source of more or less novel predictions, in which my theory is disappointingly poor. The main 'predictions' not involving DS are that the connectivity of space should be hypertoroidal, for which there is only rather indirect evidence, and may seem to some to be uncomfortably similar to the theory, now universally abandoned ,that the surface of the Earth is flat; and that leptons have a structure of 'larks' analogous to the quark structure of baryons, which is at present widely contested, though what the evidence against it is I confess to not knowing.

In contrast, the notion that disordinate space is an ever-available alternative to geometrical space leads us into a variety of new viewpoints and perhaps surprising predictions — some of which are however 'predictive' in the strict sense, only for the ill-informed.

Disordinate Space — Recapitulation

Most of the essential points about DS will be found in the later paragraphs of Sec.15. It will however be helpful to summarize this material here in the from of bald statements.

First, DS is a 'space' of infinite dimensionality, which is represented as allowing only two values, 0,1, in each 'dimension'. It is, that is to say, an infinite Boolean lattice, and has $2^{\aleph_0}$ points. Of course, the infinity of points can never be exhausted; but they are needed theoretically in proving there is a rational Sort of which it is a manifestation, and their actual infinitude may have real relevance in certain cases.

Whereas GS provides for the assigning of locations to events, and trajectories to particles, DS contains no specific reference to space or time; its function is essentially that of a similarity-space, which encodes qualitative descriptions. In view of its infinite dimensionality it has limitless tolerance of redundancy, and should therefore be capable of specifying any integral value of an variable . But it is doubtful how far this would be applicable to spacetime intervals, and the procedure involved would be so way-out for the purpose, that I think that little if any use should be made of it in this way.

Though GS can specify the location of a single event relative to any arbitrary system of coordinates, no real information is conveyed by this, and in effect any matters not reducible to at least two directly relevant events must be recorded in DS or not at all. But this still leaves us in an event-based world with the mental faculties adapted to living in such circumstances — faculties which make GS 'self-evident' and DS 'obscure'.

The relationship between the two 'spaces' is one of 'random mapping', but some relationship there has to be, since any set of events can be represented

in either of them. The difference is that they appear singly and unrelated in DS but lined together in a spatial framework but undescribed in GS. Of course, whether what I call 'random' mappings are in general unconstrained by boundary-conditions expressing the peculiarities of given problems, which can take a lot of the 'randomness' out of them; but the fact remains that no detailed point-to-point correspondence can be postulated, such as one usually expects of 'mappings' . One general rule is however worth noting:DS is extremely 'short-sighted' since it is hard to look far without having to take geometrical relations into account, whereas GS can't 'see' less than two events. Under these circumstances, the random mapping relationship between them results in a local 'softening of outlines' as mentioned earlier. The scale of this effect depends on the local event-density, but doesn't get near to any acceptable scale of intrinsic quantization (discreteness) of spacetime. It is therefore open to doubt whether such a quantization could ever be empirically observed.

10.5 EINSTEIN-PODOLSKY-ROSEN

Though I intend here to show how the notion of DS leads us into new areas, and indeed beyond the conventional concerns of physics, it is helpful to start with yet another instance where all we get is an alternative viewpoint on the well-known . The so-called 'paradox' of Einstein, Podolsky and Rosen provides a good illustration of the operation of DS at the simplest possible level.

The situation concerned in Bohm's version of EPR is where a single event gives rise to two particle differing only in having mutually antichiral spins. If, before any other interactions have occurred, one of the particles is established as say, dextral, the other, however far away it may now be, is known to be sinistral *simultaneously*, in apparent defiance of the limiting velocity which ought to have imposed a delay in transmitting this information. It is not only the experimenter who knows this — the particle itself can reveal its chirality 'long' before it could 'know'.

Chiral orientation is not directly representable in DS; it cannot be defined except by reference to an arbitrary standard, which reference involves several geometrical comparisons which DS can't cope with. It could be objected that elementary particles are a lot simpler than scalene triangles; but we may not attribute to DS any sort of model of them such as would be needed to exploit this simplicity — that would be 'information' not derived from the situation in hand. What can be held in DS is the fact, established with the initial event, that we have an antichiral pair.

As soon as the chirality of one particle is discovered, by a second event, frame of description changes and we go over to GS, where chiralities are straightforwardly representable, including that of the companion particle. In GS there is no geometrical uncertainty, and the 'paradox' is resolved.

The Collapse of the Wave Function

Let us now look at the celebrated 'two slit' experiment. An electron is fired at a screen containing two slits, appropriately spaced to give a regular interference pattern in the arrival points of a large number of electrons at a second detector screen. The passage across the first screen is the first relevant event, in which scattering occurs, deflecting the path of the electron. Considering this event only, we have to rely on DS. Nevertheless, we have to *think* in GS (else we would never recognize our mothers), and this calls for an exercise in random mapping.

The result of this mapping is conventionally, and I assume indisputably (though I am not able to present the matter as a random mapping in detail), a wave-function, determining the probability that the electron will arrive at each specified point-volume of the space beyond the first screen. The actual arrival of the electron constitutes a second event, and as in the EPR case, we must go over into GS, where a specific point of impact is coded.

While the electron is in transit it is represented not by a fixed trajectory but by a complex probability-distribution, the form taken by the 'random' mapping of DS, under the conditions of this experiment, onto GS. It contains 'information' only in what I earlier described as a 'Pickwickian sense'. On arrival, its position is stated because with the second relevant event we must translate into pure GS terms, which do not allow of an extended probability-function, derived from the no longer adequate DS information, but give a plain answer to the question "Where is it ?". The 'collapse of the wave function' is thus presented as a translation from a qualitative to a quantitative language, as we pass from considering a single event to looking at two together.

If we ask the classic conundrum, "Which slit did the electron pass through?", there is no objection to the answer that it passed through each slit with a certain probability. If the probabilities are p and q, and if we pass through not one but say N electrons, Np will pass one slit and Nq the other; the situation is precisely analogous to the case with light, where Np, Nq correspond to specified intensities at the two slits. The only trouble is that we think of Np as a number less than N, and therefore p as a fraction of an electron, which sounds bizarre. But what else *is* a probability, but that which, multiplied by a numerosity, yields a frequency, and what is a frequency, but that which divided by its numerosity gives the underlying probability?

What goes for this rather elaborate two-slit experiment goes also for every scattering -event. Such an event results in a complex probability-function, which is resolved for each resulting particle when it takes part in some other event, and thereby 'earns' its GS description. One might say that most of the life of most particles is spent in DS, with a brief touchdown in GS each time an event occurs. Those who like to speculate that things are ultimately quite other than they seem would perhaps say that the DS picture of the world is 'real', while GS is its illusory portrayal to our senses. But the opposite judgment would make exactly as good sense, and would not imply the causal efficacy of 'illusory' phenomena.

Mental Images

The role of DS is not limited to the realm of micro-physics, it is also required adequately to understand the formation of mental images. Babies start their life with all senses functional but not mature; most of the first year is spent learning

how to use them. Very few faculties do not have to be learned: it can recognize its mothers's face even on the first day (but, surprisingly, not her teats), and, at the other extreme, the dimensional structure of space may be innate, though how things fit into it seems to be initially puzzling. But the main business of 'translating' sensory percepts into usable images has to be learned; all of us do this on our own, by trial-and-error, and our varying experiences and the complexity of the task dictate that the results shall be essentially idiosyncratic — that is, the neural events underlying our imagery are largely peculiar to each individual, and have no predictable relation to the neural representation of the sensory inputs.

Learning the translation process thus involves crossing a gap, with no (or little) secure causal connection. Between the event-based sensory inputs, and the do-it-yourself image-making, there is an interval where DS (mapped, as ever, onto GS) is relevant. How the mapping onto GS is carried out is of course subject to strong constraints – the process could not have evolved unless a reasonably reliable representation was eventually obtained — but the interpretation is far from a simple analogue of the 'collapse of the wave-function' by the next 'event' after the one initiating the wave-function in microphysics. Our mental picture is not a simple copy of anything, but rather a carefully edited version of a mapping of what is still, ultimately, a DS construction.

Evidence for these assertions comes from various sources, from individuals born blind or deaf for remediable causes, treated at an age when they can report progress; from the study of illusions, especially optical ones; from the absence of neurologically detectable correlates of imagery subjectively reported; and several others. Anyone watching babies 'playing' (so-called because being babies what they do is so defined) can verify some of the points, such as the difficulty they seem to have with simple geometry, in putting one cube on top of another for example.

What comes of it all that we live in a sort of compromise world partly DS and partly GS, in which the latter, as the more practically 'reliable' , dominates our language, whereas DS tends to dominate in our dreaming (the constraints then being off). Our senses are in unimpaired contact with GS; our mental images are basically sited in DS (because the causal connection with sensory neural events is broken or incomplete), but from practical necessity re-mapped into GS. Introspection has repeatedly led philosophers in all civilizations to regard the GS aspect as a possibly illusory convention; while more recently others have come to think of the DS contribution as nonsense. It is unsurprising that, confronted by the stark separation of these two in the microphysical evidence, the idea of mental illusion should have caught on; for in our quotidian experience GS and DS work together.

New Frontiers

If the legitimacy of explanations referring to DS is once admitted there will inevitably follow a redrawing of philosophical frontiers. This will entail disputes as to where they ought to come. It is coming to be recognized already that there may be 'frontiers' across which reductionist arguments can't be extended, among which perhaps the nearest to being non-controversial is the living vs. non-living one. The argument of the last section suggests another between human and animal life – not based on the possession of 'language', already widely abandoned, but on 'idiosyncratic learning of mental representations', which may take in several of our 'dumb friends' not previously considered.

Another time-honored dichotomy is that between materialism and idealism. It could be conjectured that a truer definition of the difference would be that 'materialists' tend to think exclusively in terms of GS, whereas 'idealists' are thinking, unbeknownst to themselves, that DS has also to be taken into account. The theory here expounded requires however that GS and DS are both equally real and relevant to the behavior of matter, so that a consistent materialist ought to have no misgivings about admitting both. It is much less clear what a 'consistent idealism' might have to say; but we shall no doubt be told in due course. Unless – happy thought – this particular war will be over.

A much less well-defined, but possibly more important, difference of opinion concerns how much in the way of prior comprehensibility should stand as a qualification for a subject to receive scientific study. Most scientists would probably reject reports of precognition as not worth investigating; many would ban the whole of parapsychology. Yet there are a number of cautionary tales – that stones can't fall from the sky because there are no stones in the sky; that continent can't move because there is no force adequate to move them; and so on – evidenced by the fossils of extinct schools, which may give grounds for continuing unease. But some defense must be put up against the (ever increasing?) prevalence of old wives' tales; is there a way of doing this, except by periodically revised criteria of prior plausibility?

The implied parity of esteem between GS and DS of my theory is obviously relevant here. Our bodies, firmly defined in GS, do not overlap; our minds, at least in large part based in DS, can hardly be kept strictly isolated. This consideration diminishes rather than confirms the implausibility of effects like telepathy; it could be quite respectably argued that the time is ripe for a redrawing of this particular frontier.

Strange but True

Things strange – and even things well-nigh absurd – have turned out to be true. Fishermen, who saw ships seem to sink at sea, but come back safely at the close of day, may well have wondered whether the sea was flat as then supposed. Julius Caesar told how the ancient Britons claimed to know our planet's actual size – old druid's tales for him, of course, but we now know that the claim may have been well-founded.

Then there was Olbers, puzzled by the darkness of the night – as well he might be, even if he was the first. What sort of world would not show stars packed tight throughout the celestial sphere, burning us all up? We must be placed at the center of a finite cosmos, much as Aristotle had believed, but what a strange coincidence. Within my lifetime Olbers has been answered – the universe is finite, unbounded, and expanding, having been, umpteen billion years back, a

single point. Absurd? Yet this strange scenario has passed from absurdity to almost self-evident fact, within a few decades. Incredulity has had some nasty knocks.

One might draw a parallel between the origin of the physical world from a single point, and the origin of my metaphysical scenario from a single bit. If the first absurdity has faded into commonplace, maybe the second will. But it still has a long way to go.

The existence of that concerning which there is nothing further to be known implies much of the framework of the world; but far from all. The most obvious gap is the long list of particles with their masses and half-lives, hardly any of which have been in any sense 'explained'. Admittedly no other theory does any better at this, as yet; but none has had the expectancy of doing so either. It may be that some light on this may come from the combinatorial hierarchy work – which can legitimately be presented as an extension of the Theory of Indistinguishables– but apart from some work of Kari Enqvist there is so far little to show.

Truth, it has been said, is found at the bottom of a well. Well, the odd notion of 'agnosia' certainly looks like the bottom of something. Perhaps, then, it is the truth.

among which perhaps the nearest to being non-controversial is the living vs. non-living one. The argument of the last section suggests another between human and animal life – not based on the possession of 'language', already widely abandoned, but on 'idiosyncratic learning of mental representations', which may take in several of our 'dumb friends' not previously considered.

Another time-honored dichotomy is that between materialism and idealism. It could be conjectured that a truer definition of the difference would be that 'materialists' tend to think exclusively in terms of GS, whereas 'idealists' are thinking, unbeknownst to themselves, that DS has also to be taken into account. The theory here expounded requires however that GS and DS are both equally real and relevant to the behavior of matter, so that a consistent materialist ought to have no misgivings about admitting both. It is much less clear what a 'consistent idealism' might have to say; but we shall no doubt be told in due course. Unless – happy thought – this particular war will be over.

A much less well-defined, but possibly more important, difference of opinion concerns how much in the way of prior comprehensibility should stand as a qualification for a subject to receive scientific study. Most scientists would probably reject reports of precognition as not worth investigating; many would ban the whole of parapsychology. Yet there are a number of cautionary tales – that stones can't fall from the sky because there are no stones in the sky; that continent can't move because there is no force adequate to move them; and so on – evidenced by the fossils of extinct schools, which may give grounds for continuing unease. But some defense must be put up against the (ever increasing?) prevalence of old wives' tales; is there a way of doing this, except by periodically revised criteria of prior plausibility?

The implied parity of esteem between GS and DS of my theory is obviously relevant here. Our bodies, firmly defined in GS, do not overlap; our minds,at least in large part based in DS, can hardly be kept strictly isolated. This consideration diminishes rather than confirms the implausibility of effects like telepathy; it could be quite respectably argued that the time is ripe for a redrawing of this particular frontier.

Strange but True

Things strange – and even things well-nigh absurd – have turned out to be true. Fishermen, who saw ships seem to sink at sea, but come back safely at the close of day, may well have wondered whether the sea was flat as then supposed. Julius Caesar told how the ancient Britons claimed to know our planet's actual size – old druid's tales for him, of course, but we now know that the claim may have been well-founded.

Then there was Olbers, puzzled by the darkness of the night – as well he might be, even if he was the first. What sort of world would not show stars packed tight throughout the celestial sphere, burning us all up? We must be placed at the center of a finite cosmos, much as Aristotle had believed, but what a strange coincidence. Within my lifetime Olbers has been answered – the universe is finite, unbounded, and expanding, having been, umpteen billion years back, a

single point. Absurd? Yet this strange scenario has passed from absurdity to almost self-evident fact, within a few decades. Incredulity has had some nasty knocks.

One might draw a parallel between the origin of the physical world from a single point, and the origin of my metaphysical scenario from a single bit. If the first absurdity has faded into commonplace, maybe the second will. But it still has a long way to go.

The existence of that concerning which there is nothing further to be known implies much of the framework of the world; but far from all. The most obvious gap is the long list of particles with their masses and half-lives, hardly any of which have been in any sense 'explained'. Admittedly no other theory does any better at this, as yet; but none has had the expectancy of doing so either. It may be that some light on this may come from the combinatorial hierarchy work – which can legitimately be presented as an extension of the Theory of Indistinguishables– but apart from some work of Kari Enqvist there is so far little to show.

Truth, it has been said, is found at the bottom of a well. Well, the odd notion of 'agnosia' certainly looks like the bottom of something. Perhaps, then, it is the truth.

REFERENCES

54. A.F.Parker-Rhodes, *The Theory of Indistinguishables*, Synthese Library **150**, Reidel, Dordrecht, 1981; hereinafter referred to as *T of I*.
55. S.Toulmin, "The End of the Parmenidean Era", in *The Impact of Philosophy on Science*.
56. Bastin, T. 1966. *Studia Philosophica Gandensia*, **4**: 77.
57. T.Bastin, H.P.Noyes, J.Amson and C.W.Kilmister, *Int'l J.Theor.Phys.*, **18**, 455-488 (1979).

11. Appendix IV. PROGRAM UNIVERSE

Michael J. Manthey
Computer Sci. Dept., State Univ. NM, Las Cruces, NM 88003

```
program Universe;{A Constructive Bit-String Model of the Early Universe}

{This program specifies algorithmically how the bit string universe emanating
from the Combinatorial Hierarchy develops. The language is Pascal with
extensions for spawning asychronous concurrent processes and extendible
arrays.}

{————————Data Structure Definitions————————}

const doomsday = false;
     BasisSizes = array[1..4] of [2,3,7,127];
type   onebit = (0,1);
       Uptr = [1..Usize];    {index of a string in U.}
       Uptrs = array[1..*] of Uptr;  {for bases and closures, *=[2,3,7,127]
                                      and [1,4,120,2^127 - 126] resp.}
       ensemble = record
         last: [1..*];    {index of current last element of E.}
         E   : array[1..*] of Uptr;
       end;
       string = record
         bits: array[1..*] of onebit;
         bome  boolean             {true=>string is in a basis or closure. See StringEvolution.}
         Label Uptr,     {index in U of 1st string in ensemble; 1st is self-referencing }
       end;
       semaphore = (avail,busy); {used to guarantee mutual exclusion on updates to U}
```



```
var U: array[0..*] of string;  {U[0]=zero-string; all non-ensemble loops ignore it.}
    Usize: integer;    {initially zero == no strings in universe}
    Umutex: semaphore; {initially = avail}

{level  I        II              III                 IV

       1 2 (3) 4..6  (7..10)  11..17  (18..137)  138..255  (256..2^127-1)

basis | .2. |  .3.  |       .7.       |       .127.       |
size  |     |       |                 |                   |
      |     |.........|.........strings in closures............|
}

 Labels: record
          last: [1..*]; {index of current last element of L.}
          L   : array[1..*] of ^ensemble
        end;
 empty: string;  {an empty string, i.e. one whose length is zero.}
 zerostring: string=0;  {zero string, always of length slength.len}
 slength: record        {current length of strings in U}
      sem: semaphore;
      len: integer
   end;

  Bit: onebit;     {one random bit...see function Random below}

  B: 0..256;       {number of basis vectors found so far, over all levels.}
  LL: 0..256;      {fixed length of label field def'd when level 4 closes.}

  CurLvl: 1..4     {the level currently being "constructed"}

  Bases: record
        bvecs: array[1..4] of Uptrs;  {basis vectors for each level...completed sequentially}
        Uiptr: 0..*       {indices of vectors/strings in basis}
       end;

  Closures: array[1..4] of record
          cvecs: Uptrs,  {vectors in closure of corres Bases[i]}
          Uiptr: 0..*    {indices of vectors/strings in closures}
         end;
```

```
{-------- Synchronization -------------}

procedure wait(var s:semaphore); {poll s until it's avail}
   var t: semaphore;
begin  {presumes mutual exclusion on procedure swap, which is formally undefined
         (universal primitive) and which interchanges the values of two variables.}

   t:=busy;
   repeat  swap(s,t)  until t=avail;

end; {wait}

procedure signal(var s:semaphore);  {signal that s is available again}
begin
   swap(s,avail)
end; {signal}

{-------- Random 1/0 Generation -----------}


procedure RandomBit;  {Actual random bit generation. This runs as an independent process.}
begin
   repeat      {flip Bit forever}

      Bit := 1;
      Bit := 0;

   until doomsday
end; {RandomBit}

{The randomness of the value returned by function Random below depends on the fact that
        procedure RandomBit runs as an independent process asynchronous to everything, flipping the value
        of Bit constantly. This occurs even as U is locked during discrimination and scattering calculations.}



function Random:onebit; {Called whenever a random bit is needed.}
begin
   Random := Bit
end; {Random}
```



```
{---------Managing the Universe------}

function Generate:string;  {generates the first two strings in U}
   var g: string;
begin
   if Usize=0 then Generate:=Random
   else {Usize can only be 1}
   begin
        repeat g:=Random until g<>U[1];
        Generate := g
   end
end; {Generate}

{------------------------------}

procedure LockUniverse;
begin
   wait(slength.sem);
   wait(Umutex)
end;

procedure UnlockUniverse;
begin
   signal(Umutex);
   signal(slength.sem)
end;

{------------------------------}

procedure Tick;  {increments the universal string length by one bit.  This is done under
        mutual exclusion, so U grows, but no one ever sees it, and all bit strings are (for all
        practical purposes) always of equal length.}

   var i:integer;
begin
   LockUniverse;  {stop the world while we change it}

   slength.len := slength.len + 1;

   if Usize=0 then
   begin
        U[1] := Generate;
        Usize:=Usize+1
   end
   else  {increase the length of every string in U by 1 bit.}
     for i:=1 to Usize do U[i][slength.len] := Random;

   UnlockUniverse;   {Let the world breathe again.}
end; {Tick}
```

```
{- - - - - - - - - - - Bit-Picking Routines - - - - - - - - - -}

function ones(s:string):integer; {counts # of ones in s}
   var i,c: integer;
begin
   c:=0;
   for i:=1 to slength.len do
       if s.bits[i]=1 then c := c + 1;
   ones := c
end; {fcn ones}

function zeroes(s:string):integer; {counts # of zeroes in s}
begin
   zeroes := slength.len - ones(s)
end; {fcn zeroes}


{- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -}

function discrim(s,t:string, len:integer):string; {exclusize-or of 1st len bits of s and t}
begin
   for i:= 1 to len do
       if s[i] = t[i] then discrim[i] := 0
           else discrim[i] := 1;
end; {fcn discrim}

{- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -}

function Pick:string {picks a string at random from U}
   var i,index: integer; {index will be random in 1..Usize}
begin
   index := 0;
   repeat   {build random base 2 index.}
     for i:=0 to ceiling(lg(Usize)) do index := 2*index + Random
   until index in [1..Usize];

   Pick := U[index]  {assign random string to Pick}

end;  {function Pick}
```



```
{- - - - - - - - - - Hierarchy Construction- - - - - - - - - -}

            {Does not affect string generation or growth of U in any way}

 function InU(s:string):boolean; {true if s in U else false}
   var i,j: integer; found: boolean;
 begin

   for i := 1 to Usize do {search all of U}
   begin
     found := true;
     for j := 1 to slength.len do found := found and (s[j]=U[i][j]);
     if found then goto 1;
   end;

1: InU := found
end; {procedure InU}

function LinDepL(S:string; lvl:[1..4], len:[1..Usize]):boolean;
   {true if S is linearly independent of the strings in level lvl only.}
begin
   {--computation intensive-- generates n(b)= B(l)!/{b![B(l)-b]! discriminations upto len with S.}
end; {fcn LindepL}


function LinDep(S:string; lvl:[1..4], len:[1..Usize]):boolean;
   {true if S is linearly independent of all levels 1 to lvl.  NB: Assumes (correctly) that
              it is not called if there is no room in basis[lvl]...because of the value of CurLvl.}
 begin
   Lindep := false; {default value}
   if lvl<1 then Lindep := true  {base case}
   else {check previous levels, then current level.}
     if Lindep(S,lvl-1,len) then Lindep := LindepL(S,lvl,len)
 end; {fcn Lindep}


procedure PutIntoBasis(S: Uptr); {inserts U[S] into basis of CurLvl}
   { Assumes Bases.Uiptr=0 initially.}

 begin

   i := Bases.Uiptr + 1;     {point to next open slot}
   Bases.Uiptr := i;         {update Uiptr}
   if i=BasisSizes[CurLvl] then {start next level of Combinatorial Hierarchy}
   begin
     CurLvl := CurLvl + 1;  {CurLvl guaranteed <=4}
     Bases.Uiptr := 1        {reset ptrs}
     i := 1
   end;

   U[S].home := true;  {this vector is home}
   Bases.bvees[CurLvl][i] := S;  {point to S's string in U}
   B := B + 1;                    {inc global # of basis vectors}

end {procedure PutIntoBasis}
```

```
procedure EnsembleLabel(me:Uptr);  {assign me to some ensemble}
   var i:integer;
begin
   for i:= 0 to 2^127+136 do
      if discrim(me, U[i], LL)=zerostring then go to 1;
   error('no ensemble for me');  {can't happen}
  1:U[me].Label:=i;
   U[me].home:=true
end; {EnsembleLable}


procedure PutIntoClosure(S: Uptr; lvl:[1..4]); {inserts U[Si] into the closure of lvl}
begin
  i := Closures[lvl].Uiptr + 1;
  Closures[lvl].Uiptr := 1;
  Closures[lvl].cvees[i] := S;
  U[S].home := true
end; {PutIntoClosure}


procedure Label(me: Uptr); {categorize U[me] in terms of the hierarchy}
   var i,j : integer; found:boolean;
begin
  if Usize=2^127+136 then  {labels close now - closures full}
     for i:=1 to Usize do U[i].Label:=i; {assign each string as 1st in own ens.}
  else
  if Usize > 2^127+136 then {just find any ensemble. Theorem: When 139 L.I. strings have been generated,
                          Usize < 2^127+136.}
  begin
   EnsembleLabel(me)
  end
  else if B=0 then {guaranteed at this point that U=[01,10,11] or equiv}
  begin
   PutIntoBasis(1);
   PutIntoBasis(2);
  end
                              {continue to enlarge partial closure sets...}

  else if B > 139 then {bases and partial closures exist at this point}
  begin
    found := false;
    while not found and not Bases exhausted do
    begin
     for all Bi,Bj in Bases[l], l=1..4, found:=discrim(Bi,Bj,LL)=U[me upto LL];
     if found then PutIntoClosure(me, {lvl=}l)
    end;

    if not found then
    begin
     i:=0;
     repeat i:=i+1 until not LinDep(me,{lvl=}i,LL);  {guaranteed i<=4}
     PutIntoClosure(me,{lvl=}i)
    end;
  end
                                   {build level 2 basis.}
  else if B<5 then
```



```
  begin
   if LinDepL(me, {lvl=}1, slength.len) then PutIntoBasis(me, {lvl=}2)
  end
                   {build level 3 basis}
  else if B<12 then
  begin
   if LinDep(me, CurLvl, slength.len) then PutIntoBasis(me,{lvl=}3)
  end
                   {build level 4 basis}
  else {12 <= B <= 139}
  begin
   if LinDep(me, CurLvl, slength.len) then
   begin
     PutIntoBasis(me);
     if B=139 then {assign all strings in U to a specific level's (probably still incomplete) closure set, if it
                  is not already in some basis (i.e. home=true).}
     begin
       LL := slength.len;
       for i:=1 to Usize do
       begin
         if not U[i].home then {find right closure for U[i]}
         begin
           j:=1;
           while not U[i].home do
           begin
             U[i].home:= not LinDepL(i,{lvl=}j,LL); {not l.i.=>in j's closure}
             j:=j+1   {cannot exceed 4}
           end {while}
         end {if}
       end {for}
     end {if}
   end {else}
  end { major if-then-else stmnt }

end; {procedure Label}
```

```
{- - - - - - - - - The Life of a Bit String- - - - - - - - -}

procedure StringEvolution(var MeString:string; me:Uptr);
 {every string (except empty) becomes a separate incarnation of this
  procedure, i.e. a separate, independent asynchronous process.}

  var d,m:string;  {local working variables}
      home: boolean; {true ==> I am a member of a basis, closure, or
                      ensemble.}

begin
  repeat
    if Usize=0 then    {we need two strings to get started}
    begin
      Tick; {go from no strings in U to one.}
      MeString:=Pick; {we become this first string: the original empty-process becomes the U[1] process herewith.}
      m := Generate;        {generate a second string}
      Usize := Usize + 1;    {Universe now has two strings}
      U[2] := m;
      Label(me); Label(m);
      spawn stringevolution(U[2], Usize); {give U[2] life.}
    end
    else {universe is already rolling, so just discrim w/someone}
    begin
      LockUniverse;
      m := Pick;
      d := discrim(mestring,m, slength.len);
      if not InU(d) and d<>zerostring then  {add d to U}
      begin
        Usize := Usize + 1;       {increase size of Universe}
        U[Usize] := d;            {add d (literally) to U}
        Label(Usize);             {categorize d in comb hier}
        spawn StringEvolution(U[Usize], Usize) {— d leaves me here —}
      end
      else Tick;  {no novelty was generated}
      UnlockUniverse
    end
  until doomsday {strings never die}
end; {string evolution }
```



```
begin {————————Universe starts here————————}
  {Initialization}
      CurLvl := 1;
      Slength := 0;
      LL := 0;
      B := 0;             {number of basis vectors}
      U[0]:=zerostring;  {"invisible" zero string}
      .
      .
      .
   {end of initialization}

  spawn RandomBit;     {start random bit generator going}

  BigBang: StringEvolution(emptyset,0);

end. {Universe  (we never get here) }
```

12. Appendix V.

Some Fundamental Characteristics of a Discrete Geometry

David McGoveran
Alternative Technologies
150 Felker Street, Suite E
Santa Cruz, California 95060
408/425-1859

Among the fundamental principles which have been pursued by those researching the combinatorial hierarchy are: discreteness, finitism, and constructivism. When combined with primitive recursive operations for counting, labeling, and ordering, a rich variety of mathematical, logical, and physical structures can be obtained. In recent years, efforts to construct physical theories based upon the hierarchy have met some success. However, certain issues must, sooner or later, be addressed if such constructions are to be accepted by the scientific community-at-large.

Whenever a new theory is developed, its acceptance is predicated on the ability to communicate the spirit, the formal structure, and the utility of the theory. For the combinatorial hierarchy, that the theory be "communicatable" is an exceedingly difficult requirement to meet. Most of physics is understood in terms of centuries old geometric paradigms: the continuum, limits, infinities, distance functions (even distances), metrics, etc. How does one talk about distance in a finite world of discrete objects without the underlying continuum? What is a (physical) vector space in such a world? An event? A collision? Is it even possible to construct a discrete, finite model of the physical world without having such "old" language creep in, let alone be able to describe the model to a physicist steeped in a paradigm whose roots are at least as old as the Pythagorean irrational?

The most common form of constructivism is that which insists on the ability to provide a (algorithmic) construction of each object and theorem in the system. We refer to this as "weak constructivism". In the strong form of constructivism, each object and theorem of the system must be formulated along strict finitist and pure discretist lines. In this paper, we shall use "constructively", constructivist, and constructivism to refer to the principle of strong constructivism. The following principles (McGoveran's Principles) may be taken as guidelines in pursuing answers to the questions raised in the preceeding paragraph:

1) There is nothing in the knowable (or observable) Universe which can not be described constructively.

2) There is nothing which can be described constructively which (that known as) the physical Universe can not produce (in the combinatoric sense).

3) There is nothing which can be observed or known which can not be described constructively.

It is not enough to state such principles; one must demonstrate their utility. In particular, these principles deny the necessity, relevance, and even the meaningfulness of concepts such as infinity, infinitesimal, randomness, asynchrony, and continuity. At best, such concepts are useful only as placeholders in incomplete models, theories, specifications, or descriptions of a system. In most cases, randomness and asynchrony in particular are local descriptions of more global properties.

We believe it is possible to build a conceptual bridge which would allow the operational use of conventional terminology without implying (or assuming) the usual underlying geometric paradigm. First attempts at constructing such a bridge (and simultaneously bringing powerful tools to bear on the efforts at hand) have been made by developing a "discrete differential topology". This effort provides a means of refering to distances, functions, derivatives, etc. with most of the standard rules of use intact, but without violating strictures against appeal to non-constructivist entities such as the continuum, limits, or infinities.

In going from a discrete topology to a geometry, a number of interesting problems arise. So much that is taken for granted is not generally given meaning by construction: the geometric constants, the trigonometric functions, notions of direction, etc. It is interesting to note that most of the non-finite, non-discrete entities are embodied and related in one beautiful equation:

$$e^{i\pi} = -1$$

Within this single relationship lies the assumptions that give rise to coordinate systems, the structures of a circle and a square, translational and rotational invariance, irrational and complex numbers, and, last but not least, the base of the natural logarithm. Even the notion that one could in any way raise a number to an arbitrary (perhaps non-integral) power is rooted in the geometric paradigm. Certainly complex numbers can be understood as ordered pairs of real numbers and these real numbers can be restricted to the positive integers for our purposes. And negative integers can be understood as the "dual" set of symbols obtained by "counting down" rather than "counting up". What of the other entities?

In what follows we construct a square and a circle, and derive a ratio which plays the role of pi. We begin by constructing the equivalent of a square, two-dimensional coordinate patch. The only elements allowed for construction are a finite (perhaps large) number of discrete elements (essentially indistinguishable mathematical objects), ordering relation operators, the ability to count, and the ability to label the objects through an operator.

Without reference to a distance function, a "square" can be defined having the following properties:

a) two-dimensionality
b) parallel sides
c) fixed center under interchange of the dimensional parameters

The following definitions will serve to provide the objects necessary for the constructions (more precise definitions can be found in "Getting Into Paradox"):

Def: Two objects are said to be INDISTINGUISHABLE if they are unlabeled.

Def: A SORT is an ensemble of indistiguishables with cardinality and without ordinality.

Def: A SET is a sort with ordinality (an ensemble is not a set).

Def: An ENUMERATION is a total ordering of a sort.

Def: An ORDERING OPERATOR is an operator which generates a partially or totally ordered labeling of an ensemble (note that we do not mean set). The cardinality of the labeled ensemble is fixed in advance as part of the definition of the specific operator. Thus a particular finite partially or totally ordered labeled ensemble defines an ordering operator and vice-versa. Note that the labeled ensembles produced are sets if and only if none of the labeled indistinguishables are twins.

Def: The DIMENSION of a sort is expressed as the number of mutually disjoint ordering operators on the sort.

Def: Two ordering operators are said to be INDEPENDENT if they are mutually disjoint in the sense that no more than one element of a chain produced by the first operator will all be also be in a chain produced by the second operator.

Def: By NEAREST NEIGHBOR of an element e in a chain ordered by operator p is meant any element n such that for a:a=p(e) or b:e=p(b); then for operator q mutually disjoint from operator p, n=q(e), n=q(a), n=q(b), e=q(n), a=q(n), or b=q(n).

The criteria for 2-dimensionality is satisfied by requiring two mutually disjoint ordering operators. The algorithm is as follows:

1. Select an element e_0 --- Figure 1.

2. Establish a chain x_0^{-1} of length n with e_0 as the supremum, using the ordering operator p_x. --- Figure 2.

3. Establish a chain x_0^{1} of length n with e_0 as the infimum, using the ordering operator p_x. ---- Figure 3.

4. Call the union of x_0^{-1} and x_0^{1}; x_0. Require that x_0 is totally ordered. --- Figure 4.

5. For each element i of x_0, establish chains y_i^{-1} and y_i^{1} of length n under the ordering operator p_y with the selected element of x_0 as the supremum and infimum of the chain. Require that the y_i are disjoint as are the pairs (y_i^{-1}, y_i^{1}). This is a unique labeling or total ordering requirement on the entire construction (i.e. there must exist an ordering operator q such that the elements of the entire construction are totally ordered. --- Figure 5.

6. Require that the nth elements of the y_i form chains x_i ordered by ordering operator p_{x_i}. --- Figure 6.

7. The resulting object satisfies the requirements: it is the discretum version of a 2-dimensional (square) coordinate patch. In particular, the 2-dimensionality of the construction is satisfied by the definition of mutually disjoint ordering operators: at most one element in a chain resulting from one operator will be found in a chain resulting from the other. For the given construction, at most two operators can be used: a third would result in a partial ordering instead of a total ordering of the elements of the construction and this would then represent an object which is not connected or in an object for which "multiple" elements are doubly labeled. Thus, the ordering operators "parameterize" the object.

We can now proceed to construct an object which behaves as a discretum version of the 2-sphere. A 2-sphere (again without reference to distance functions) has the following properties:

a) 2-dimensionality
b) every perimeter (boundary) element is like every other
c) fixed center for all "orientations"

The constructive algorithm is as follows:

1. Select a (square) coordinate patch with with center e_0 and all elements uniquely labeled. Call this patch P_0. Figure 7.

2. Constrain the possible ordering operators to those operators which produce chains of length n and which select for e_0 a nearest neighbor of e_0, then a nearest neighbor of this element, and so on. We refer to the operators which represent these selections as "radial permutations" of the coordinate patch. Figure 8.

3. Starting from e_0 construct a coordinate patch with new ordering operators which are radial permutations of coordinate patch. Figure 9.

4. Map the elements of this patch P_i to patch P_0 and eliminate any elements which do not have at least i labels. Figure 10.

5. Repeat this process for all pairs of allowed radial permutations. Figure 11.

The result is a discretum version of the circle, in that it has a fixed center (e_0) with radial symmetry (isomorphic to its radial permutations with identified center e_0). It has built in bounds on "precision". The relation between the number of "sides" of the polygon formed by a set of cardinality n and the number of permutations is fixed: it gives a measure of the "size" of the circle.

Given these two geometric objects it is possible to define a ratio which plays the role of the ratio of the area of the circle to the area of the square patch from which it was formed. This number is obtained by counting the number of elements contained in the circle and the number of elements contained in the square and forming the ratio.

A second ratio is obtained by counting the elements in the perimeter of the circle and forming the ratio with the length n of the chain x_0.

In general, these ratios will be functions of the length n of the chain x_0. Furthermore, the values of the ratios will not in general be those obtained under Euclidean geometry. However, if one insists on isotropy, homogeneity, and "density" (i.e. large n), it is easy to see that these values must be those obtained by the standard polygonal approximation to the circle. In particular, these ratios will be approximations to pi/4 and pi with the appropriate precision. These constructions and the results are closely related to numerical and statistical "approximation" methods as seen from within the traditional geometric paradigm. In fact, Archimedes came close to the construction used here (Measurement of the Circle). However, the definitions are completely constructive and general; matching the continuum definitions as desired.

It is a central point of this paper that a measure of the discrete cardinality and of the discrete topology of our observed Universe is given by the precision with which the ratios pi(area) and pi(lengths) are identical in value. That pi should be of cosmological significance is not surprising. Indeed, if the cardinality of the Universe is changing, then the two values of pi should be changing also. Furthermore, if the relevant discrete cardinality is related to a spatial volume, then as this region becomes smaller calculations involving pi can not be treated in a naive manner. Specifically, the multiple meanings of pi must be dissociated if the values are different (i.e., pi(areas)/pi(circumference) will not be 1) and the usual value can no longer be taken as a constant independent of spatial volume. Even more important, if the world is discrete and finite, and if the values of pi are not related to the spatial volume via a cardinality of the volume, it follows that the values of pi used in calculations relate only to the cardinality of the Universe. In other words, pi becomes a true Universal discrete topological constant and local physical properties are then immediately dependent on the global properties.

As noted above, we need only interpret complex numbers in the Hamiltonian sense in order to be consistent: namely, complex numbers are taken to be ordered pairs of numbers. This forces us to recognize two orderings at work for pairs of equations where as the imaginary notation obscures it. Complex numbers are a special case in that the additional ordering contains only a supremum and an infimum for the y_i. Thus, the commutation relation holds since commutation simply gives the trivial dual of the chain. In general, however, this will not be true where the constraints of disjoint operators apply and the number of ordered numbers is more than two. This fact is well known, having been discovered by Hamilton in trying to work out an algebra of quaternions, and having been generalized by Grassman. (Note that there are only three algebras which have an invertible vector product: those over the reals, those over the complex, and those over the quaternions (ordered 4-tuple). Also, algebras of all n-tuples greater than two are non-commutative.)

All of this leads us to consider a general property of discrete, finite spaces (now that we have a geometry we can call it a space instead of a topology), their commutivity. Assume for the purposes of illustration that a square coordinate patch is embedded in a flat space with continuous distance function. Because of the finiteness and the discreteness of the space, there can be no discrete correlate to the irrational distance of the diagonal, since according to the continuous distance function this will be square root of two times the length of a side. Thus, at best only an parallelogram law works for discrete distance functions on this space. Translated into the discrete version of Lie dragging, this means that the space has a torsion or is non-commutative. Geometrically one would say that the space is locally non-Euclidean. Alternatively, one could insist that there is no continuous distance function mappable to such a discrete space or that there is no discrete geometry. But this would be tantamount to a claim that all objects are non-local - i.e. infinitely extended. We prefer to conclude that discrete, finite geometries are not torsion-free.

FIGURE 1. e_0

FIGURE 2. The sub-chain of length n=4, x_0^{-1} with e_0 as supremum.

FIGURE 3. The sub-chain of length n=4, x_0^{1} with e_0 as infimum added.

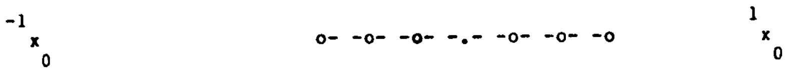


FIGURE 4. The chain of length n=7, x_0

FIGURE 5. The chains y_i of length n=7 added,

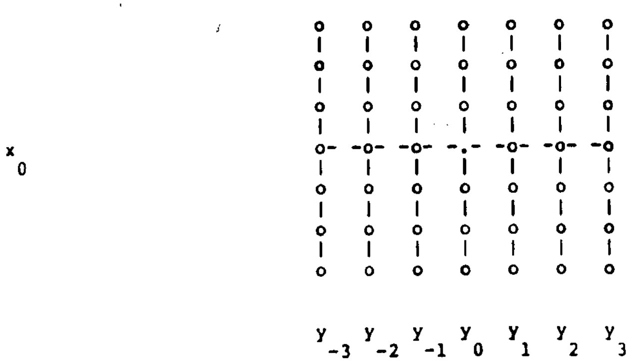


FIGURE 6. The chains x_i of length n=7 added.

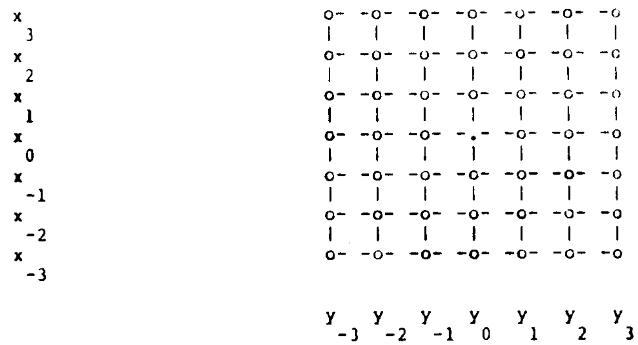


FIGURE 7. Select a patch, P_0.

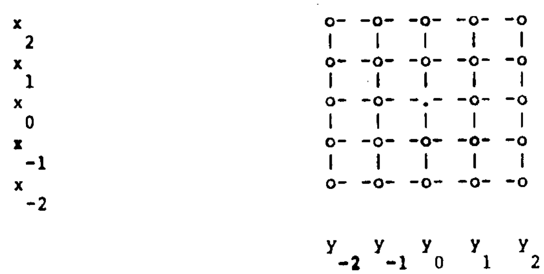


FIGURE 8. The nearest neighbors of e_0.

```
x           o- -o- -o- -o- -o
 2          |   |   |   |   |
x           o- -*- -*- -*- -o
 1          |   |   |   |   |
x           o- -*- -.- -*- -o
 0          |   |   |   |   |
x           o- -*- -*- -*- -o
 -1         |   |   |   |   |
x           o- -o- -o- -o- -o
 -2

            y   y   y   y   y
             -2  -1  0   1   2
```

FIGURE 9. A new patch, P_i.

```
x           o- -o- -o- -o- -o
 2          |   |   |   |   |
x           o- -o- -o- -o- -o
 1          |   |   |   |   |
x           o- -o- -.- -o- -o
 0          |   |   |   |   |
x           o- -o- -o- -o- -o
 -1         |   |   |   |   |
x           o- -o- -o- -o- -o
 -2

            y   y   y   y   y
             -2  -1  0   1   2
```


FIGURE 10. Mapping the new patch to the old.
* = new unmapped, # = old unmapped.

```
                                    *
                                   /
x             *   #- -o- -o- -o- -#
 2              \ | x | \ | \ | x |
x                 o- -o- -o- -o- -o
 1                | / |   |   | / |
x                 o- -o- -.- -o- -o
 0                | / |   |   | / |
x                 o- -o- -o- -o- -o
 -1               | x | \ | \ | x | \
x                 #- -o- -o- -o- -#   *
 -2                /
                  *
                  y   y   y   y   y
                   -2  -1  0   1   2
```

FIGURE 11. Elements remaining after all allowed radial permutations.

```
x                     o- -o- -o
 2                  / |   |   | \
x                 o- -o- -o- -o- -o
 1                |   |   |   |   |
x                 o- -o- -.- -o- -o
 0                |   |   |   |   |
x                 o- -o- -o- -o- -o
 -1                 \ |   |   | /
x                     o- -o- -o
 -2

                  y   y   y   y   y
                   -2  -1  0   1   2
```

CALCULATIONS:

A(square patch) = 16
A(polygon) = 12
ratio = pi(areas)/4 =
A(polygon)/A(square) =
12/16 = 0.75
or pi(area) = 3.00

C(polygon) = 12
d(polygon) = 5
ratio = pi(lengths) =
C(polygon)/d(polygon) =
12/5 = 2.4

ADDENDUM:

Note that in the example, all other radial permutations cause the same elements of the construction to be deleted or else do not map the coordinant patch. The reader may demonstrate this for himself. Also note that a central element is a matter of technical convenience for the algorithms and may be circumvented.

13. Appendix VI.

"BI-OROBOUROS" — A RECURSIVE HEIRARCHY CONSTRUCTION

John Amson
The Mathematical Institute, St Andrews University

(Work in progress ....)
(Prepared for ANPA VI, Cambridge, 1984)
(Revised, 1986)

ABSTRACT

First: The (Finite) Combinatorial Hierarchy is shown to possess a Dual Heirarchy; the "Top Levels" (i.e. the "Parker-Rhodes Stopping Levels") of both hierarchies are identified with the abstract tokens "0" and "1" out of which both hierarchies are constructed. This identification produces an infinitely recursive construction with a "two-headed orobouros" structure (the BI-OROBOUROS of this paper's title).
Second: The recursion in turn produces an infinite sequence of vector spaces of increasing dimension, which are then "glued" together (using a construction based on the idea of a Direct Limit) to form a single "background" space, an infinite dimensional space composed of vectors with infinitely many 0,1 coordinates (sequences). The "sequential act of construction" is interpretable as a primitive clock "timing" the "evolution" of the system. In this construction "time" has a beginning.
Third: Alternatively, this background space can be modified in a natural way to become a "two-sided sequence space" in which the original "starting" coordinate has lost its accidental privilege, This allows us to identify the location of the original Finite Hierarchy's successive four levels at an arbitrary place in the system (i.e. there is no longer any privileged location for this identification). This alternative construction corresponds (i) with the system's "timing" flowing from an infinitely remote "past" through an arbitrarily locatable "present", (ii) with any "present stage of construction" having reached an arbitrarily complex refinement.
Fourth: The passage from finite dimensional vector spaces to the infinite-dimensional background space allows the passage from (unavoidably) discrete topologies in the Finite Hierarchy to a "potentially reachable" non-discrete topology. This in turn allows the introduction of increasingly refined finite metrics at any stage of evolution in the system. These metrics provide increasingly precise approximations by rational numbers to a "potentially realisable" system of real numbers, for all measurement purposes.

1 VECTOR SPACES AND BOOLEAN DUAL SPACES

1.1 Suppose we are given Vector Spaces $E(n)$ ($n = 0,1,2,\ldots$) over the field $Z_2 = \{0,1\}$ of two elements, where the vectors in each $E(n)$ can be regarded as $\{0,1\}$-tuples of length n, and vector addition is coordinate-wise addition (mod 2). The neutral vector for vector addition is the vector $0 = (0,0,\ldots,0)$ of all 0s. Of course, $E(n)$ can also be regarded as a Boolean Algebra, in which the "dual" operation consists of replacing 0s by 1s and 1s by 0s in each vector – or, equivalently, by adding the "anti-neutral" vector $1 = (1,1,\ldots,1)$ to any vector. This leads us to the idea of a "dual" vector space $E'(n)$ in which vector addition is given by $x +' y = x + y + 1$, and the neutral vector is now 1.
(The dual "anti-neutral" for 1 is once again the original "neutral" 0.)

1.2 It is useful to write $E(n,0,+)$ and $E'(n,1,+')$ – or just $E(n)$ and $E'(n)$ for short – for these two ways of looking at the vector space $E(n)$ and its "boolean dual". Also, for any nonempty subset S in $E(n)$ we can construct its "dual" set S' by adding the anti-neutral vector 1 to each vector in S. A subset S is called "self-dual" if $S = S'$.

1.3 Of course, the two vector spaces $E(n)$ and $E'(n)$ are isomorphic, and to that extent are "indistinguishable" as algebraic structures. But it turns out to be remarkable useful to keep alive their distinctions.

1.4 I have described (at ANPA V, 1983) (and it is easy to prove) how a subset S in $E(n)$ can be a vector subspace in $E(n)$ if and only if its dual set S' is a vector subspace in $E'(n)$. It is also interesting (and should be potentially very useful) to identify which vector subspaces S are "self-dual" – so that each such subset S is simultaneously a vector subspace in $E(n)$ and in $E'(n)$; note that any self-dual vector subspace must contain both the neutral 0 and the anti-neutral 1 vectors. Keeping alive the distinction between a vector space and its dual allows us to see that the collection of dual vector subspaces is thus common to both vector spaces.

2 THE HIERARCHY AND ITS DUAL

2.1 We recall that the Combinatorial Heirarchy uses just four of these vector spaces (namely $E(2)$, $E(4)$, $E(16)$, $E(256)$) as its four LEVELS but ascribes no role to the field elements 0,1 other than their natural one in these abstract mathematical constructs. In point of fact it uses the "first" vector space E(1) as the base field (0,1) itself, from which the whole hierarchy is constructed. So in this sense we have a FIVE-LEVEL CONSTRUCTION - see Figure 1. I won't go into the actual hierarchy construction here; it is "well-known"[1] and the details are not actually important for what follows. What is important about it here is that E(2) can be embedded in a particularly significant way into E(4), and similarly E(4) into E(16) and E(16) into E(256), but that this process cannot continue beyond E(256) - the so-called "Parker-Rhodes Stopping" phenomenon.

E(1) →	E(2) →	E(4) →	E(16) →	E(256)	: Original Levels
0	I	II	III	IV	: Level Numbers
E'(1) →	E'(2) →	E'(4) →	E'(16) →	E'(256)	: Dual Levels

FIGURE 1. THE FIVE-LEVEL FINITE HIERARCHY

3 THE ROLE OF THE TOKENS 0 and 1

3.1 I have described (at ANPA V, 1983) how we can regard the special primitive quantities 0 and 1 as "tokens", that is, as representative of some unspecified things whose only properties - at this stage - are summarized in the fact that we can "discriminate" between them by some as yet unspecified procedure which allows us to say that if "the things being discriminated are the same" then the result of the discrimination is signaled by the production of a "0" , and if "the things are not the same" then the result is a "1". I also described how the lack of any other substance that could be accorded these tokens gave them an undesirable primacy in the scheme of things. That is to say, it gave them an essentially "primitive" exogenous status which was in conflict with the idea of the theory under development being in some way self-describable, self-organizing or self-generating.

3.2 I have always insisted (from my earliest talks with Ted Bastin twenty years ago) that whatever theory should ever arise from the development of a Combinatorial Hierarchy approach it should intrinsically contain the capacity (or at least the potential) reflexively to describe

the theory developed, together with its appropriate calculus. That is to say, the theory should possess an essentially self-organizing or boot-strapping feature. But to that end I always had to face the obstacle of how one could satisfactorily discharge these tokens of their accidentally special role. For me, the missing step was found when I recognized that one could identify the tokens 0 and 1 with LEVEL IV and its "dual", respectively. The identification also serendipitously removed another accidental specialization; namely that of the LEVEL IV (and its dual) as the Parker-Rhodes Stopping Level in the Hierarchy (and its dual). We can now say:
"0 stands for Level IV" or, interchangeably, "Level IV stands for 0", and similar statements about "1" and the dual of Level IV.

3.3 Such an identification, whilst immediately denying the tokens 0 and 1 and the Parker-Rhodes Stopping Levels of any undesirably special status, also provides us with an unexpectedly rich system, an infinitely recursive two-headed construction: THE BI-OROBOUROS — see Figure 2.

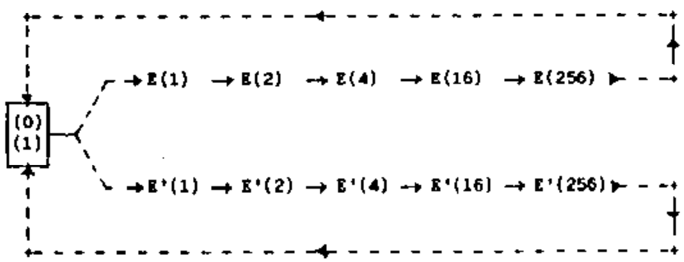


FIGURE 2. THE BI-OROBOUROS CONSTRUCTION

3.4 Thus this construction has allowed us to deflect the question 'What are 0 and 1 ?' with the sphinx-like answer: E(256) and E'(256), i.e. LEVEL IV and its dual, each of which is itself comprised of $2^{256} \approx 10^{78}$ vectors each represented by an n-tuple made up of 256 copies of the 0s or 1s, each of which in turn is a copy of E(256) or E'(256) each of which in turn is .... ad infinitum ......
From one point of view, the Bi-Orobouros has no beginning and no end. Yet from another point of view it "starts" with the recognition of the initial need to use "primitive tokens" 0 and 1 but immediately goes out from this

to support the entire infinitely recursive construction, endlessly recapitulating the Combinatorial Hierarchy and its Parker-Rhodes Stopping. In other words, any attempt (or feeling of obligation to attempt) to "define" the substance of the tokens 0 and 1 exogenously (itself a meaningless action) is removed infinitely far away by this recursion. The requisite endogenous nature of the Heirarchy is brought one step nearer actualization.

3.5 For later use, we notice that if we try to construct the formal steps in this recursion, we are led to an infinite sequence of vector spaces:- thus, beginning with E(256) we can - in each 256-tuple - systematically replace each 0 by E(1), that is the pair (0,1), and replace each 1 by the dual E'(1), that is the pair (1,0). In fact it's easier and more exact to achieve this by replacing 0 by the pair 00 and 01, and 1 by the pair 11 and 10; doing this helps to keep intact the "history" of the replacements[2]. It also allows us to see at once that E(256) is modified to E(256x2) = E(512). Interchangeably, we could of course just as well start with E(1) and replace 0 by E(256) and 1 by E'(256) using a similar kind of labeling scheme to keep track of the "history" of the replacements. (Note that each label here has to be written with 257 digits and there are 2^{512} distinct labels !) The result is again a larger vector space E(2x256) = E(512). Continue this process:- using E(256) and E'(256) in E(2) we get E(4x256) = E(1024); using E(256) and E'(256) in E(4) we get E(16x256) = E(4096); and so on.....[3]
Similarly, dual replacement steps are also to be carried out:- starting with E'(1) we get E'(2x256) = E'(512); and so on.....

3.6 It is clear that these dual recursion processes do not have to be carried out separately; they can (and must) be thought of as taking place simultaneously, and instantaneously. Moreover, they thoroughly and completely mix up both the 'original' primitive branch E(1)---E(256) and its dual (or 'anti') primitive branch E'(1)---E'(256). At each stage of the recursion, the Bi-Orobouros acquires an ever increasingly complex fine-grain structure comprising both 'original substance' and its dual 'anti-substance'.

4 THE RECURSIVE HEIRARCHY

4.1 The first practical problem now has to be faced:

How can we incorporate the idea of Bi-Orobouros in an algebraically and topologically meaningful way?

4.2 One interestingly practical way makes use of the idea of a Direct Limit (sometimes called an Injective Limit) of mathematical structures[4]. Very roughly speaking, we can imagine a new vector space E(oo) (it is a "super-space", of infinite dimension) consisting of vectors each of which is an infinite-tuple (sequence) of 0s and 1s, with only a finite number of 1s in any of them. (We shall also need a "dual super-space" E'(oo) in which vectors have only a finite number of 0s in them.) Vector addition is again coordinate-wise addition (mod 2). We can now recognize a sequence of vector subspaces in this super space E(oo):-
First pick out the two vectors which start (0,0,0,...) and (1,0,0,...) and which have all 0s after the first place; these clearly form a subspace isomorphic to E(1), i.e. to the base field Z_2 = (0,1). Second, pick out the four vectors which start (0,0,0,0,0,....), (0,1,0,0,0,...), (0,0,1,0,0,...), (0,1,1,0,0,....) and have all 0s after the third place; these form a subspace isomorphic to E(2). Do the same with vectors which have all 0s in 1st,2nd and 3rd places and after the 7th place, and at least one 1 in the 4th,5th,6th,7th places; these together with (0,0,0,....) make up a set of 16 vectors which is isomorphic to E(4). Then move on to the next 16 places, the next 256 places, the next 512 places ... and recognize E(16), E(256), E(512),... It's as if we were glueing together the infinite sequence of finite-dimensional vector spaces E(2), E(4), E(16), E(256), E(512), ... in a meaningful and organized way[5]; see Figure 3.

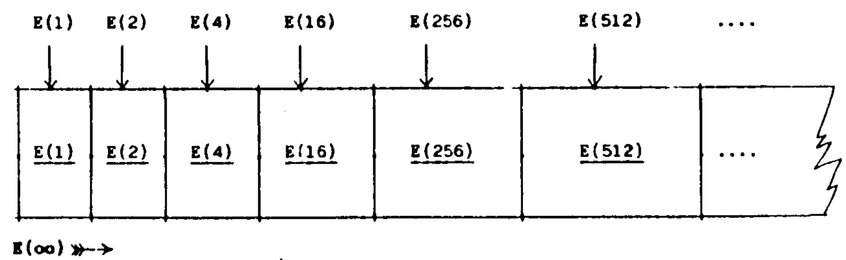


FIGURE 3. EMBEDDING THE FINITE HEIRARCHY (E(1)--E(256)) AND ITS OROBOUROS EXTENSIONS (E(512), E(1024),...) IN THE INFINITE BACKGROUND SPACE E(oo)

4.3 The degree of organization is larger than at first appears. For (see Note 5) there are maps between these vector subspaces which allow one to identify a lower level vector with a higher level one in the super-space E(oo). These maps are (or should turn out to be closely related to) the older level-matching maps of the Finite Heirarchy[6]. Moreover, since these maps are linear they are themselves representable by (infinite) matrices with entries made up of 0s and 1s (almost all of which are 0s, of course); and any such matrix can in turn be identified with a vector in the super-space E(oo), thus making the whole system very self-contained (and very Hierarchical in spirit..) !

4.4 Some other problems that have still to be tackled:

(A) Should we formally carry out the identification of E(256) with (0) and the dual identification of E'(256) with (1), in the super-space E(oo), and if so, how ?

(B) Is there a formal role for the dual super-space E'(oo) ?

(C) If E(oo) and its dual E'(oo), though isomorphic, are to be regarded as distinct in the same way as each E(n) and E'(n) were, can we not carry out yet another identification — this time of (0) with E(oo) and (1) with E'(oo) — and so repeat the entire Bi-Orobouros construction on this even grander scale by producing further super-spaces E(oo,oo) and E'(oo,oo), and so on and so on ..... ?

My own conjecture is that a satisfactory answer to (A) will resolve (B) and thence (C) in such a way that the first two super-spaces E(oo) and E'(oo) are shown to be mutually identifiable in the sense that neither can ultimately provide any further structure-based information. Moreover, it is very likely that E(oo) and the seemingly grander E(oo,oo) would turn out to be structurally indistinguishable[7], so that nothing new was to be gained by further creating an infinite succession of such grander identifications for (0) and (1). In this ultimate sense there would be no infinite regress of constructions for Bi-Orobouros. Bi-Orobouros would be formally identifiable with either E(oo) or E'(oo); the choice could be made without any more preference than is currently given the tokens "0" and "1" for whatever formal objects they are habitually required to represent.

5 TEMPORAL EVOLUTION IN THE RECURSIVE HEIRARCHY

5.1 There is another point which I ought to mention here, though this too is a very raw idea. The "extension" process by which I described the

setting up of the infinite recursion by using the spaces E(1),E(2),E(4),..., E(512),... could be regarded as a "ticking" universe, growing in complexity at each extension. The universe at any "time" is finite because only a finite number of places in any vector in E(oo) have had the possibility of being occupied by a 1 (all other places "further on" still being 0s). But if we wish to take on board an infinitely remote past, we have only to modify this construction and replace the super-space E(oo) by an isomorphic copy consisting of "two-sided" vectors (sequences of 0s and 1s which have no start and no end, e.g. (.......,0,0,0,0,1,0,1,1,0,0,0,0,.......). Formally these are described as sequences indexed by all the positive and negative integers and zero {...,-2,-1,0,1,2,...} rather than by the positive integers {1,2,3,....} alone.

5.2 The advantage in doing this, is that even the "start" of the construction loses any privilege or priority: any point can be taken as a convenient "start" for identifying a "local" E(1), and the system's structure then rolls away "locally" to the right as before; but any other point more to the left would have served equally well and when it comes down to a hard choice, there are no grounds for picking on any point.....; the "origin of time" is infinitely far to the left and "time" stretches infinitely far to the right, so to speak. But at each "tick" the universe becomes that much more rich and complicated, structurally speaking. Perhaps there is a sense in which this idea can be identified with the Noyes-Manthey-Gefwert ticking universe[8]?

6 INFINITE DIMENSIONAL SPACE TOPOLOGIES

6.1 The next positive advantage which I think will flow from these ideas is the potential removal of the limitation of having to work with a finite collection of finite vector spaces, as in the original Finite Hierarchy. As Clive Kilmister and I have described (at ANPA V, 1983), this unavoidably forces any topology on them to be discrete (and rather uninteresting as a result); we are left with only a "Hamming" metric (the distance between two vectors is just the number of places in which their coordinates differ). But now that we have infinite dimensional super-spaces E(oo) and E'(oo) to work with - even if only as background spaces, only "potentially realizable" - we may be able to appeal to a larger variety of topologies and related topics.

6.2 Some possible ideas here are these: (i) It is known[9] that E(oo), as defined above, is identifiable with all subsets of rational numbers, and hence (ii) that it contains a chain (i.e. a collection of vectors each one of which in some sense is an extension of all "previous" ones in the collection) which is identifiable with the real number system. This means we would have (most? all?) rational numbers available at any "time" and real numbers available at a potentially infinite "time". The metric "refinement" available to us in any finitely-realizable part of the universe would thus depend simply on the (theoretical?) extent to which we have followed the Bi-Orobouros' development starting at any given "origin". We are free to choose any such starting-origin (in the two-sided vector system described above) to suit the problem in hand, and so are free to discuss theoretical physical problems to any degree of topologic/metric "precision". But what effect this would have on any attempt to introduce the notion of quanta (in energy or spatial measurement, for example) is not at all clear.

========================

NOTES:

(1) See for example: Bastin, et al., 'On the Physical Interpretation and the Mathematical Structure of The Combinatorial Hierarchy'; Int.J.Theor.Phys.,Vol.18, No.7, 1979, pages 445-488.

(2) What in fact we are doing here is forming the Kronecker (or Tensor) Product of each vector in E(256) with the vector (0,1) and re-writing the resulting 256x2 matrix as a 512x1 vector. And similarly for other such constructions.

(3) Note that the sequence of dimensions grows with a "double rhythm":

(	2	4	16	256	)
(	512	1,024	4,096	65,536	)
(	131,072	262,144	1,048,516	16,777,216	)
(	33,554,432	67,108,864	268,435,456	4,294,967,296	)
(	8,589,934,592	17,179,869,184	68,719,476,736	1,099,511,627,776	)

....

The 1st entry on the (n+1)-th row has the value 2×256^{n};
the 2nd, 3rd, 4th entries are then 2x , 8x , 128x that value, respectively. Alternatively, note that the numbers in the FIRST ROW grow by squaring, then the numbers in each COLUMN grow by being mutiplied by 256.

(4) Useful accounts of this idea are given for example in
(a) Bourbaki, Elements of Mathematics, THEORY OF SETS, (Translation, from the original French, published by Addison-Wesley, 1968), (Chapter III, Section 5, (page 202));
(b) Dugundji, Topology, Allyn & Bacon, 1966; (Appendix Two, Section I, (page 420)).

(5) Strictly speaking, the Direct Limit involves a further construction. First there has to be an ordered set of mappings each of which reaches from one "embedded" space to another in the system, and they must composable in a sensible way, so that if spaces (a), (b), and (c) occur in that order, then if Fab maps you from space (a) to space (b), and Fbc from space (b) to space (c) then their composite Fbc.Fba is the same as the map Fac from space (a) to space (b). The extra construction now consists of identifying any vectors in the extended system which can be linked by a chain of these mappings stepping from one "embedded" subspace to another. (This idea reminds me of my earlier idea of "spines" of vectors reaching up through the Finite Heirarchy via chains of level-mappings — see paragraph A.4.7 in the reference in Note 1 above.) It is known that if two vectors are related by such a link, then the relationship so formed is an equivalence relation on the system. The final step comes in regarding any two such "equivalent vectors" as being one and the same object. (Formally, we construct the factor space E(oo)/R from E(oo) by this equivalence relation R.) This new "condensed system" is obviously very exciting:- any one of its "objects" consists of all vectors in the Bi-Orobouros which can reach each other by chains of level-mappings; the collective of vectors in such a chain have lost their "individuality", so to speak. Perhaps they correspond to "physical entities" in an "observable universe" ? A natural question to ask is: What is the cardinality of the system E(oo)/R ?

(6) The difficulty at present is in finding the correct way of mapping E(256), the LEVEL IV of the old Finite Heirarchy, into E(512), the first of the Extended Levels ("LEVEL 5"). Once that is agreed upon, the successive mappings can then be constructed iteratively up through the infinite sequence of all the Extended Levels in both the Recursive Hierarchy and its dual.

(7) In the same way that the vector space V(oo) of all infinite 0,1 sequences with finitely many 1s is isomorphic to the vector space V(oo,oo) of all infinite 0,1 matrices with finitely many 1s ("infinite sequences of infinite sequences"); and so on....

(8) See, for example: Noyes, Manthey, Gefwert,; 'Towards a Constructive Physics', SLAC-PUB-3116(rev. September 1983).

(9) See, for example: Dwinger, Introduction to Boolean Algebras, Physica-Verlag, Wurzburg, 1971, page 8, Problem 4.11.

****