THE FINE STRUCTURE OF HYDROGEN *

REVISED and EXPANDED DRAFT
CIRCULATED FOR COMMENTS AND CORRECTIONS

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

* Presented at the 10th Annual Meeting of the Alternative Natural Philosophy Association, Cambridge University, England on Thursday, August 24, 1988 and the 5th Annual ANPA West Meeting, Stanford University, January 29, 1989.

I. INTRODUCTION

The equivalent formulae of Sommerfeld and Dirac for the fine structure of hydrogen were crowning achievements of their day. The methods for achieving this result were apparently quite different and yet have recently been shown to be equivalent (The "Sommerfeld Puzzle" Revisited and Resolved by L. Biedenharn). In this paper we introduce a new and quite distinct derivation of the formula based upon the Combinatorial Hierarchy (Bastin, T., Studia Philosophica Gandensia, 4, 77, 1966) and the ordering operator calculus (McGoveran, D.O., Foundations of a Discrete Physics).

The Combinatorial Hierarchy led to the definition of the fine structure constant as 1/137 as a first order approximation. Our conceptual model of the hydrogen atom leads us to compute a second order correction factor for α. Estimates of the third order corrections (including $(1/137)^2$ and $(m_{e/m_p})2$) are beyond current experimental precision, so that the correctness of the second order approximation must be correct to about four decimal places <u>given our current conceptual model</u>. [Note that these statements apply to low-energy values of α. Further relativistic refinements are required for higher energy measurements (e.g., at the mass of the W.]

This theory is very different from that of quantum mechanics, not only in its conceptual but in its mathematical foundations. It leads to different results than do standard theories when they are extended beyond ordinary quantum mechanics to include relativity. Indeed, our theory is intrinsically relativistic and discrete, so that our model of what the Sommerfeld formula represents is different than the Sommerfeld semi-classical, the non-relativistic, or the Dirac models of the hydrogen atom.

II. PRELIMINARY CONCEPTS

<u>Terminology</u>

Throughout this paper, terms should be read as being mathematical rather than physical, unless explicitly given as physical. Thus, the terms event, cycle, period, and frequency do not have the familiar temporal and physical connotations.

Value and probability measures become indistinguishable in the ordering operator calculus. The probability measure is taken as counts of occurrences of some abstract event over an unnormalized (we do not mean unnormable) space - we only divide counts by the maximum cardinality N of the space when we want the relative probabilitistic frequency of some combinatoric attribute. So the space (e.g., 1-dimensional) is not on the [0,1] interval but on the [0,N] interval.

For our purposes, all objects of investigation will be treated as occurrences of what we call **combinatorial events** and the relationships between combinatorial events. When representing a given occurrence of some combinatorial event (using a bit string representation) we look for a bit string pattern (not substring) as output from the ordering operator (which could be the PU algorithm). If the combinatorial event recurs in a regular way, we say it is **periodic** and the number of bits generated between occurrences is the **cyclicity**. Then any generation of that many bits constitutes a **cycle**. These are mathematical, not physical terms. Whenever we use the term event, we shall mean a combinatorial event unless the term is otherwise modified.

A string of cyclicity n means that the string can be seen as the concatenation of similar substrings of length n. Under a physical interpretation (meaning the combinatorial event is interpreted as a physical event - see next paragraph), this cyclicity would be called a **period**. However, we are reluctant to use this term since it introduces connotations of physical time, which is not strictly implied by the mathematics. We will refer to a bit string representation of an event having a cyclicity as a **periodicity** for purposes of the discussion in sections IV and V.

By a **Program Universe event**, we shall mean a vertex in the sense defined by Noyes in the interpretation of Program Universe. It is not clear that all these vertices have a physically observable interpretation however, and so we distinguish between Program Universe events and physical events. Thus, a **physical event** as used here is a Program Universe event which has a physically observable interpretation. A physical event which participates in a Coulomb interaction will be referred to as a **Coulomb event**.

Ordering Operators

An **ordering operator** is a finite computational object which can be thought of as generating a **finite directed graph**, a mathematical object consisting of nodes and directed (i.e., ordered) arcs between nodes. Such a graph defines the ordering operator and vice-versa. Thus, the ordering operator calculus is process-oriented. When a node of the graph is connected by directed arcs leading either toward or from (but not both) two or more other nodes, a partial ordering is represented and the nodes are indistinguishable. Two nodes with a single directed arc between them are totally ordered. The possible walks of a graph generated by an ordering operator include all the possible "jumps" between indistinguishable nodes. One may think of such jumps as "virtual arcs", although no direction is implied since indistinguishables are not intrinsically ordered. Note that any node may be thought of as an abstract representation of a graph that has been previously constructed.

Ordering operators are quite general and powerful, but are always well-behaved. As a result, there is often no need to impose constraints of linearity on the mathematical objects constructed within the ordering operator calculus. Operators may act on graphs to generate an ordering of the indistinguishable nodes of the graph. Two or more operators may be thought of as operating concurrently on the same graph and operators may order two or more graphs to create more complex objects. Under a particular ordering operator, particular nodes of two otherwise distinct graphs may be taken as indistinguishable (e.g. when they are partially ordered under the given operator).

A sub-graph can be understood as representing a particular property or attribute of the graph, called a combinatorial attribute. It is often convenient to represent a walk of a particular graph by a binary sequence called a bit string. The representation of the walk is always in terms of some combinatorial attribute. One writes down a '1' if the attribute is encountered in the walk and a '0' if it is not after having traversed a number of nodes equivalent to the number of nodes required to traverse the sub-graph which represents the attribute.

If one decomposes the graph into several sub-graphs, these form a representation of the graph, so long as we know how to re-connect the sub-graphs. For a given attribute, the sub-graphs can be walked to generate bit strings representations. Given the context, the bit strings may be used to reconstruct the graph. It is often convenient then to think of ordering operators as generating bit strings instead of directed graphs. In this paper, we will use this concept of bit string generation when referring to the output of an ordering operator.

Probabilities

The ordering operator calculus treats all quantities as being combinatoric. This has the effect of (1) ensuring that all values and equations are algebraic, and more important (2) unifying the way in which both values (numbers) and probabilities are treated. The only difference is that values are typically given as integer counts (of the occurrences of some combinatoric attribute) and probabilities as that number normalized by the largest value then possible, i.e. the frequency interpretation of probabilities, so that the rational result lies between 0 and 1. If the probability space remains unnormalized, then there is no difference in the two spaces. Thus, in this paper, we will use the terms number, probability, and frequency interchangeably, context making it clear that normalization may be implied in order to obtain a probability or frequency from a number (count of occurrences).

Independence

There is an additional consideration. Ordinarily, there is the concept in probability that two events are either mutually independent or they are dependent. In the ordering operator calculus, due to the treatment of indistinguishables, there exists a smooth, though discrete progression from mutual independence to complete dependence. (Whether or not it is always appropriate to measure the degree of dependence via correlation coefficients is a matter of current investigation.)

This "spectrum" forms a multi-dimensional discrete and finite vector space which we have called a combinatorial attribute d-space. A d-set of completely independent combinatorial attributes forms a d-basis for the d-space. Two or more mutually independent attributes add as orthogonal vectors in the space.

It is also important to understand that the process notion of independence differs somewhat from the "static" definition. Because the size of the space is always changing in a process view, constraints and relationships come into play at various stages.

For example, although we might have a final relationship between x and y (whose sequences of values is counted by t), this relationship could not be said to be constructive until the space was large enough (say of cardinality N) to represent the computation of the relationship. Until the space has been generated to cardinality N, x and y are independent. Thereafter, they are dependent.

As a second example, and a simple extension of the first, suppose that x and y must be satisfy some relationship whenever t satisfies some condition. Then, it is not the case that x and y are correlated except when this condition on t is also satisfied. Between these pairs of values of x and y, no correlation can be stated.

An important point here: since this is a process system, independence can be transient. Clearly two periodic events are not independent if they are taken to have commensurate periods (just as the minor and major generators of an elliptic oscillator are not independent over time intervals larger than period of the system). However, they must be independent outside of this to 'lift the degeneracy' (i.e. - to get away from the equivalent of a circular orbit).

Objects

"An object is a conceptual carrier of [possibly complex] combinatorial attributes between [possibly complex] combinatorial events [which may correspond to physical events under a particular interpretation]."

Note that this definition reduces to the definition of particle given by Noyes (Noyes, H.P., and McGoveran, D.O., An Essay on Discrete Foundations for Physics, Physics Essays, 2, No. 1, (1989)) in presenting the interpretation of Program Universe:

"A particle is a conceptual carrier of conserved 3-momentum and quantum numbers between events."

Classical physics actually imposes many constraints on the nature of experiments as does quantum mechanics and relativity. A correct statement of a correspondence principle must take into account all the differences between these various domains. Between quantum mechanics and classical physics the primary differences have to do with how properties combine, which is only approximately correlated with object size. Part of the reason for this lack of correlation is due to the differences between how objects are viewed. In quantum mechanics, properties can be exchanged between objects and objects can be created and destroyed. In classical physics, object integrity is always maintain homogeneously - objects may be broken as in an elastic collision, but they do not change properties.

The classical notion of an object as having well-defined boundaries is contrary to normal experience. In addition, a similar problem exists with regard to what properties define an object. This problem has been analyzed in a manner different from that one used here by such researchers as Lofti Zadeh using "fuzzy logic" and, by application, "fuzzy pattern recognition". Those involved in other approaches pattern recognition (such as the clustering analysis methods of Nils Nilson) are closer to our approach and have long recognized the detailed problems.

While is it easy to see that one might argue that an upholstered chair is still a chair if the upholstery is removed, it is less easy to see that the question of chairness might or might not include the "warmth" of the object or its sentimental history or the airflow around the object. Indeed, the physical extent of an object into its environs is always in question. Anyone involved in the legal issues of property rights is well aware of this on-going and difficult debate. The point is that the classical physics notion of object does not meet the needs of everyday experience.

Part of the problem is due to the insistence that objects with well-defined boundaries and mix of properties PERSIST between observations. This is not, of course, altogether unreasonable. We are able to "intercept" the object along the classical trajectory and observe a confluence of roughly the same set of properties. But clearly this is not always the case: we can not always backtrack to a prior observation nor does an object evolve along all possible classical trajectories. We give reasons for this, but these are reasons which are valid only if one constrains experiment according to the classical physics world view. The notion of what constitutes a clean experiment

not only serves to define the conditions, but to make them repeatable. And in so doing, the procedure eliminates numerous possible interactions with the environment, eliminates the complexities of impure objects and operations, and the complexities of what happens when the experimental prediction must be based on partial information.

In addition, the classical physics paradigm assumes that all objects are distinguishable. But this is no more true in the macroscopic domain than in the microscopic. Consider the "poor identification" of criminals based on eye-witness accounts. The lack of information, in principle available, which might distinguish the accused from the guilty leads to exchange of individuals as equals. A rose is a rose is a rose... The only difference between exchanging eggs in two omelets and electrons in two preparations is an ontological statement in the latter that the distinguishing information is, even in principle, not obtainable (see Bohm, Causality and Chance in Modern Physics for an extensive discussion of the lack of justification for or against this commitment).

The manner in which individuals decompose a complex object into supposedly less complex objects (and this is a major assumption) is heavily biased by culture. While it is certainly the case that decomposition of the environment is promoted along the lines of what can be easily manipulated. An object is perceived to be distinct if it can be separated physically from the rest of its environment, if it can be moved. But what can be and can not be moved is dependent on the technological devices available to the culture. And even if available, the perception will not be realized unless this is motivated by the needs of the culture.

If we give up the absolute character of the object (i.e. its Platonic reality) then we are left with the confluence of its properties. This gives us a characterization of an object at a particular time and in a particular context. We use pattern recognition techniques to correlate this confluence of properties with others. If there exists a strong correlation and a causal structure connecting the two confluences, we are ordinarily inclined to say that there is an object which carries the properties. (This is exactly what the magician depends upon to work his deception!) But this is a mistake. The interaction between and confluence of properties is an evolving aspect of systems. Treated in this way, the macroscopic world begins to look more compatible with the microscopic world.

There are other simplifications about the macroscopic world which classical physics promotes. For example, the notion that all classical operations commute. The notion that pure states do not exist. The notion that superposition is invalid. There are macroscopic counterexamples to each of these concepts.

A quantum mechanical interaction is more closely akin to a classical chemical reaction than to a classical object interaction: the properties may be exchanged between the reactants and products and apparent new properties may arise. However, the fundamental properties are always conserved. With a quantum mechanical interaction, many of the reaction parameters are only defined statistically, so that reactants and products are defined probabilistically.

Distance Functions and Independence

The adding of independent values arises from a straightforward consideration of (a) the requirements given in Foundations for a distance function and a norm, (b) the assumption of a locally flat space, and (c) the independence (which becomes orthogonality under a value measure) of the generations of the two probabilistic frequencies.

The ordering operator calculus defines an inner product as a recursive function n() which satisfies the following where a is any rational number and x is a d-vector on a vector d-space:

P8: $n(ax) = |a| * n(x)$ (1)

It is a distance function if it satisfies the following postulates:

P9: $n(x) >= 0$ and $n(x) = 0$ iff $x = 0$ (2)

P10: $n(x + y) <= n(x) + n(y)$ (3)

It is bilinear and symmetric (as it must be if the space is homogenous) if it also satisfies:

P11: $n(x + y)^2 + n(x - y)^2 = 2[n(x)]^2 + 2[n(y)]^2$ (4)

In terms of '0's and '1's, the distance function can be defined as

$$(I-D)/N \quad (5)$$

where I is the number of increments, D is the number of decrements, and N is the number of possible increments and decrements. Note that the maximum value of I+D is not necessarily N in the event that indistinguishables are encountered. This is just a modified Hamming distance. The counts, I and D are, of course, not pure numbers - they depend upon the walk of the particular graph and also on the particular attribute whose occurrences are being counted. Thus, a count represents a particular sequence: it has a structure. The manner in which the count was obtained is important as to how we may manipulate the value.

Such a distance function is used to ensure that magnitudes are always used consistently. Clearly, the counts resulting from two partial walks of the same graph should add linearly - they are not independent if they must add to the length of the total walk. Independent walks, however, imply that no nodes may be walked twice. Adding squares implies that there exists a graph the walk of which could replace the two independent walks, and that there are decompositions of this graph which can represent each of the independent graphs. In other words, the replacement graph preserves the basis for the original graph. Then we must generalize the distance function given above to include d-vector resultants.

As in a Euclidean continuum, the resultant magnitude c of two orthogonal d-vectors in a locally flat (discrete but otherwise Euclidean) d-space of magnitude a and b is given by

$$a2 + b2 = c2 \qquad (6)$$

In a locally flat discretum however, the value of c^2 may be computed, but the value of c as a square root may not exist. Instead, we look for the symmetric factors, so that the structure by which the count was obtained is taken into account. By symmetric factors we mean a pair of values c'+e and c'-e where both c' and e are rational and

$$c^2 = (c'+e)2(c'-e)2 = c'2 - e2 \qquad (7)$$

Note that (7) assumes the commutativity of c' and e, but this can be proven from the construction if a and b are truly independent. In a non-flat space, this commutativity will fail. Since we deal only with algebraic quantities (i.e. with c either rational or irrational but not transcendental), there is always a solution set for (7). Thus independent quantities add via their squares if the space if locally flat (by local here we mean simply over the extent of the walk by which the quantities are obtained).

If the discretum is not locally flat, then there is a correction due to the "curvature". This fact will have a consequence when we discuss possible corrections to the results presented below.

Comparing New Theoretical Results to Old

The computational results of analysis of any particular experiment is dependent upon the interpretation according to a given theory. Of course, some of the analysis is dependent upon more obvious constraints given by the experimental situation. However, the determination of which numbers correspond to physical quantities is dependent upon the interpretation. For example, the relative importance of values of the factors of a polynomial versus the evaluation of the polynomial itself is key.

If the polynomial is incorrect (in that, like the wave equation, it does correspond to any structure in the system being modeled but is only a computational procedure), then the relationship of the factors to the evaluation of the polynomial is also incorrect.

Suppose that, in the correct theory, there are two factors - (1-e) and (1+e). Since these factors are symmetrical, it is easy to confuse them and measure simply +e and -e so that only one value is perceived experimentally. It is then natural to assume that the polynomial is K(1-e)(1-e) rather than K(1-e)(1+e), K being some third factor, if the extant theory ascribes only a simple structure (a value squared) to the polynomial. In order to compute the accepted experimental value, one must make this same mistake. However, the value of e is the theoretical value given by the new theory.

Within a methodology which provides for iterative improvement, it is considered to be a task of a new theory to explain how the theory it displaces computes the values it does, why the old theory computes correct answers when it does and why it computes incorrect answers when it does. It is important in all cases to understand why the two theories differ and why they agree. To the degree that this is possible, the theory can be purposefully improved to provide a more encompassing explanation of the known facts.

Combinatorial Hierarchy

Many of the results of the ordering operator calculus depend on how the space is generated, that is, on what ordering operators are used to construct the space and on how they interrelate. For applications to physics, we have (since 1981) attempted to exploit the point of view that the ordering operators must also generate what is called the Combinatorial Hierarchy. The reasons for this selection have been discussed in other papers (e.g., McGoveran, D.O, Foundations of a Discrete Physics, in DISCRETE AND COMBINATORIAL PHYSICS, H.P.Noyes, ed.; ANPA WEST, 25 Buena Vista Way, Mill Valley, CA 94941).

Roughly speaking, one expects a hierarchy in physical processes which are decomposable into self-similar sub-systems. If a discrete, recursive process is an appropriate description of the elementary processes at work in the physical world, then clearly such a structure is called for in order to explain the rapid changes in degree of complexity as one changes scale.

Furthermore, at each level of the hierarchy, it must be possible to treat the elements as a complete and independent set of operators which operate on objects of less complexity and as a complete independent basis for representing more complex objects. If the representational basis is to be understood in a traditional (physical) sense, then independence (orthogonality) and linearity are appropriate properties. Simply put, these

notions of completeness and independence mean that the hierarchy will contain its own means for abstract representation of all possible operations and all possible objects. The abstraction will be minimal and yet no conceptual power will be lost.

In addition, the algebraic structure for a sufficiently large numbers of elements (whether because the number of elements within a given level of the hierarchy is large or the hierarchy can be repeated) should be practically indistinguishable from a classical continuum structure. We are motivated to use a system which emphasizes the importance of discrimination (exclusive OR) over the field Z_2, since the ability to distinguish two objects is clearly of fundamental importance. That the Combinatorial Hierarchy satisfies this mathematical "correspondence principle" was pointed out to the author by Herbert Doughty, III (private correspondence).

The Combinatorial Hierarchy is generated from two recursively generated sequences. The first is governed by the recursion formula

$$n_{(i+1)} = 2^{n_i} - 1, \qquad (8)$$

(a formula familiar to those who have studied the Mersenne primes), and begins with the term $n = 2$ leading to the sequence $3, 7, 127, 2^{127} - 1, \ldots$ The cumulative cardinals of this series (ignoring the initial term) also form a series which has interpretive significance, namely $2, 3, 10, 137, \sim 1.7016\ldots * 10e38 + 137, \ldots$

The second recursively generated sequence is governed by the formula

$$m_{(i+1)} = m_i^{\,2} \qquad (9)$$

also beginning with the term $m = 2$ and leading to the sequence $2, 4, 16, 256, 65536, \ldots$

These two sequences have various justifications. Perhaps the clearest presentation has been given by Clive Kilmister (correspondence to H.P.Noyes, date Oct.16, 1978), paraphrased here as follows:

Definition: By a **combinatorial hierarchy** is meant a collection of levels related as follows:

a) the elements at level L are a basis of a vector space V/Z_2

b) the elements at level L+1 are non-singular (i.e. invertible) linear operators mapping V/Z_2 into V/Z_2

c) each element A at level L+1 are mapped to a subset S of the elements at level L by the correspondence: the proper eigenvalues of A [i.e., Av = v] are exactly the linear subspace generated by S.

d) each element at level L+1 is chosen independent, allowing the process to be repeated for level L+2, L+3, L+4, ...

Theorem 1: There exists a unique hierarchy (up to isomorphism) with more than 3-levels and it has the following successive numbers of elements:

$$2, 3, 7, 127, 2^{127} - 1$$

and terminates at Level 4 due to the fact that the operators have m^2 elements if the vectors are m-fold, and 2n (required for V/Z_2 increases too fast.

One can think of the number of elements at each level of a combinatorial hierarchy as being the number of subsets of a set of n things. Given the definition of a hierarchy over the field Z_2, this generates the first sequence mentioned above. Over a finite field V, all operators which map V into V may be thought of as permutations in that they map elements of V into other elements of V and this map can be given a pairwise representation. Note then that is it possible to think of the number of independent operators as simply the number of permutations of n things taken two at a time (i.e. exchange of a and b corresponds to mapping a to b) or m^2 total at any level of the hierarchy, given m at the previous level. Clearly, this forms a complete set of independent operators - all other permutations of n things can be created by successive application of this set of permutation operators. This generates the second sequence mentioned above which governs a combinatorial hierarchy.

Mapping the first sequence onto the second and treating the second sequence as independent basis strings, one finds that the mapping is not uniquely possible beyond the fourth term or level. Thus the first sequence can not be a coordinate basis beyond level four of the hierarchy. The inter-relationship forms a **stop rule** which signals the end of global novel structure generation, although not necessarily the end of generation and redundant structure. In other words, if one codes the algorithm as a program, there will be a point at which no further novelty will result, although the program need not halt altogether. The cardinals of elements and operators at each level will be determined by the two sequences given above. As noted in Theorem 1, any two runs of the program will produce hierarchies which are isomorphic, even though the particular evolution may differ and the particular objects used to represent the elements may differ. For this reason, and since the details of evolution are generally not significant for us, we refer to **The Combinatorial Hierarchy.**

This way of thinking about the Combinatorial Hierarchy allows one to understand the conceptual nature of the increases in complexity. In a sense, the subsets of a set are objects created through a higher level of abstraction than is the set or its elements. Indeed, it would not be proper to say that the subset of a set S consisting of the element x is conceptually identical to the element x. To do so ignores the context of x.

Because the Combinatorial Hierarchy is defined over the field Z_2, it is now considered appropriate (and convenient) to think of the cardinals as counting binary strings, leading to the so-called "bit-string" representation, with the basic operations of discrimination (exclusive OR) and concatenation of arbitrary bits onto extant bit strings. This representation allows the hierarchy generation to be given in algorithmic terms, an exercise which, along with a particular interpretation of the physical meanings of the bit strings, forms the conceptual basis of the Program Universe (PU) interpretation and representation.

Labels

It is important for our purposes to understand the purpose of a basis representation such as is provided by the second sequence above. In particular, it provides a unique label or name for every distinguishable object to be represented in the space. This is true even for ordinary vector spaces: the basis representation allows one to uniquely represent a given point in the vector space as a unique n-tuple.

We were convinced long ago that the kind of labeling scheme suggested here is completely equivalent to Noyes' labeling scheme (interpretation of Program Universe), but noted that Noyes fails to make this relationship to the mapping space (i.e. the space of n^2 basis elements) explicit. Furthermore, Manthey does not provide a particular mechanism (code) within Program Universe (as written) for it. Indeed, this is one of the reasons for wanting to rewrite the program into "UP", being explicit in the constructive generation of Combinatorial Hierarchy everywhere rather than generating a structure which was "rich" enough to support the Combinatorial Hierarchy and simply showing that a model of the Combinatorial Hierarchy exists in the output of Program Universe.

We will refer to the generation of a bit string under various constraints (see the discussion on 3- and 4-events in Noyes, H.P. and McGoveran, D.O., An Essay on Discrete Foundations for Physics, SLAC-PUB-4528, pp. 35-43) as an event. This is simply a combinatorial event, provided no physical interpretation is implied. If the constraints are properly defined, the event can be interpreted as representing a physical event. Otherwise, it is simply an occurrence of a particular bit string having a computable probability of occurrence given the mechanism by which the background population of extant bit strings has been generated.

Keep in mind that there are two independent constructions in the Combinatorial Hierarchy (PU generated) space of strings which are distinct from the representation or mapping space. The first are given in a space which models the system, whereas the latter provide a common representation which appropriately labels the first strings. When strings generated by two (or more) independent constructions are mapped to the same label in the representation space, this does not mean they are identical but only that they are taken as representing the same information in the constructive context. It is incorrect to conclude that "if A maps to L and B maps to L, then A is the same as B." This allows the multiple runs of PU (or any other Combinatorial Hierarchy algorithm) generated strings to be truly independent even in the sense of interpretation and this must be so.

III. AN IMAGINATIVE PICTURE

The imaginative picture which we describe here is not meant to be accurate. Rather it is intended to motivate the reader conceptually. As we go, we will try to point out inaccuracies which might otherwise be misleading.

We begin by considering an experiment which we describe operationally as follows. The equipment consists of a digital sampling oscilloscope. Each count in a time period is based upon the occurrence of a particular digital pattern. There are two inputs. Each trace may be adjusted both horizontally and vertically. In addition, the magnification may be adjusted.

There are two ways in which a trace may be considered to be periodic. First, we may have a periodic structure in which the counts in each time period are periodic. In this way the amplitudes form a stable trace.

The second is an abstraction of the first to which we will appeal in the derivation which is to follow. Each trace occurs periodically. As a result, we can represent the entire trace as a pattern and set the input to detect this pattern. Then the trace consists solely of one or more points.

Each point on the trace represents what we call an event. Whether or not an event is elementary as in the first description or complex in the second does not matter. What is important is whether or not the events are periodic.

Suppose that we wish to compare two traces. Clearly, with the adjustments available, we can adjust the two traces so that the first points are coincident. However, there are not sufficient degrees of freedom to adjust the position of the beginning of the next period so that the corresponding initial points are guaranteed coincident as well. In particular, this will not be possible when either (1) the periods are relatively prime or (2) the minimal sampling period will not resolve the

differences or (3) the resolution allowed by the adjustments is too course.

Suppose that we select one of the two inputs as the reference and adjust the second trace to be as nearly coincident with it as possible. If exact coincidence is not possible, then the best we can infer is that the first point of the second trace will derive from an event which occurs within one time increment from the corresponding point from the second trace, even if it is coincident on the trace. It will be convenient to think of the degree of coincidence between the two traces as the relative "phase" of the two traces. However, one must be careful not to let this classical language intrude on the analysis and remember the discrete nature of the system.

Now suppose that due to the experimental arrangement we are led to believe that the two traces derive from two independent aspects of the same source. Suppose further that the two inputs are known to be digitally produced; if we know exactly how they are generated, we can inquire as to the probability that the two inputs are distinguishable and under what circumstances the initial points of the beginning of each period of each trace remain distinct.

Again, one must not forget that the entirety is digital - the sources of the signals as well as the measurement apparatus and the adjustments it provides are digital and finite. Not only does this mean that the signals might not be perfectly synchronized in the measurement process, but also that the degree to which they are not synchronized is, in principle, statistically computable in a combinatorial fashion. In a sense, more information is available than would be the case if the signals were continuous - there are only a finite number of values by which the two signals may differ.

Suppose that the measurement apparatus with its intelligent inputs is connected to the source via a single channel - namely that we depend on computational pattern recognition to distinguish (i.e., separate or unmix) the two input signals. Suppose further that display consists of a single point per signal per period. Then pure synchronization of the two signals corresponds to indistinguishability. On the other hand, this "degeneracy is lifted" if subsequent periods beyond the first may not be synchronized on the display - they are thus shown to be distinct within the context of the experimental sources and the measurement apparatus. Given knowledge of the manner in which the two sources are generated and assuming perfect pattern recognition, we may compute the probability that the degeneracy is lifted in some way so that the display of the two traces differs by one or more divisions.

The difficulty with this picture is that it suggests that the problem of perfect synchronization can be solved by a measurement apparatus of higher resolution or other improvements. It also suggests that the two signals might be differentiated a

priori. In our theory, neither of these is possible. The discreteness and the finiteness of the system is absolute and can not be circumvented. What follows will demonstrate the manner in which this situation explains a number of otherwise unexplainable phenomena. We will refer to the patterns being detected as bit strings and to their detection as bit string events.

IV. DERIVATION OF THE FINE STRUCTURE CONSTANT

Let us examine how the ordering operator calculus tells us to treat the problem of multiple bit strings events where the bit strings have some cyclicity n. We begin with a review of how a single periodic structure of bit string events can be used to derive the Bohr atom, modified from Noyes argument (McGoveran, D.O. and Noyes, H.P., On the Fine Structure of Hydrogen, SLAC-PUB-4730(rev. 3/15/89), submitted to Physical Review Letters).

The Bohr Atom

Part of a Program Universe bit string is interpreted as representing the quantum numbers generated by the Combinatorial Hierarchy and the remainder of the string of length n represents the increments, k 1's, and decrements, n - k 0's, in pre-space between events. Each step has an attribute velocity which we interpret to be the limiting velocity c and a length of h/mc; hence the attribute velocity between events is

$$\beta c = [(2k/n) - 1]c.$$

If we wish to model a "constant velocity" β, this restricts a string length of nN to have kN increments, defining the "deBroglie wavelength periodicity" N as the "positions" where events could (but need not) occur. Because the step-wise change in position h/mc implies a change in momentum mc, both of which can be reversed at the next step enclosing an area h in a discrete phase space (i.e., a discrete space having momentum and position as bases), or more generally enclosing a area nNh when we return to a cyclic starting point after nN steps, we have derived relativistic Bohr-Sommerfeld quantization from our model, including the zitterbewegung associated with the string mass specified by the system label.

We attribute the binding of m_e to m_p in the hydrogen atom to Coulomb events, that is, only to those events which involve one of the 137 labels at level 3, and hence occur with probability 1/137. (Noyes has phrased this as "on the average one in 137"; we will not introduce in this paper the complexity of assumptions which speaking of averages implies. Rather we assume that it is sufficient to be more deterministic for now so that the occurrence of a Coulomb event is strictly periodic in our sense. This does not mean that the strict periodicity is observable; remember that we have not introduced observable time in our construction.) The changes due to other events are

indistinguishable in the absence of additional information. We can have any periodicity of the form 137j where j is any (positive) integer.

So long as j is the only periodicity, we can say that

$$137j \text{ steps} = 1 \text{ Coulomb event}$$

or

$$137j \text{ steps/Coulomb event} = 1 \qquad (10)$$

Thus, (10) is a units conversion factor or, in our terminology, a rule of correspondence. Now characterize the system by some attribute which we identify with the energy. The value of the total system rest energy in terms of the system mass u is just μc^2 (where $\mu = (m_e m_p)/(m_e + m_p)$. We meet the bound state requirement that the attribute associated with the energy E be less than that associated with the system rest energy. Since the internal frequency 1/137j is generated independently form the zitterbewegung frequency which specifies the mass scale, we can express the system rest energy in terms of a portion which is due to 1/137j proportional to that part which independent of this frequency (which we will call E). There are two ways in which these two portions may combine (to represent probabilities in the frequency sense) combinatorially, namely

$$E(1 + 1/137j) = \mu c^2 \text{ and } E(1 - 1/137j) = \mu c^2$$

These behave as two factors in the observed system, combining by multiplication, yielding

$$(E/\mu c^2)^2 \, [1+(1/137j)^2] = 1 \qquad (11)$$

Identifying $e^2/\hbar c$ as 1/137 to first order as implied by our interpretation of the Combinatorial Hierarchy, (11) is just the relativistic Bohr formula (Bohr, N., Phil.Mag, 332-335, Feb. 1915).

The Bohr-Sommerfeld Atom

In a system of multiple combinatoric constructions, the constraints are always sufficient to define the possible combinatorial paths leading to a particular generation of an attribute instance or combinatorial event (or particular confluence of them). In the present case, the event is either a Coulomb event or confluence of the multiple periods.

The constraints are just a combination of theoretical and operationally justifiable polynomial conditions. These must be understandable in terms of the rules of correspondence and also consistent with the algebraic operations which standard theory uses to compute the standard values of any structure constants of the system under investigation.

Now combine a second distinguishable period **s** with the first **j**, considering two independent periodic bit strings, which we represent by a count of the number of respective cycles as j and **s**. Let **j** have a cyclicity j_0 of one and therefore be of integral period. This just means that we represent an entire period with one bit. We will refer to this integral period requirement later as constraint C(1). Let **s** similarly be of integral periodicity. Now we explore how to combine the period counts for these periodic bit strings. Note that, in the system so far constructed, there is no absolute scale. If the cyclicities of j and **s** were identical, they would be indistinguishable and no question of when **s** begins with respect to j is meaningful. However, it has not been assumed that the periodic bit strings need have the same cyclicity, so that if only the number of periods are counted, it is possible for the beginning (or end) of a period of j not to match the beginning (or end) of a period of **s** regardless of how a cycle of **s** is represented.

The best that can be done is to represent **s** in such a way that the magnitude of each period is integral and can be mapped locally to a single period of **j**. We must then take into account the relative origins of the two representations. Furthermore, we must also consider the possibility that the particular period of **s** actually occurs after multiple periods of s_0. If the lowest of these s_0 is given in the same representation as that of j, then s_0 is not necessarily an integer, because the lowest period of j need not correspond to a minimum value for **s**. (Physically, this is similar to two waves which are slightly out of phase.) However, without further constraints being imposed the difference between successive values of **s** (due to counting lowest periods by twos, threes, fours, etc.) must still be integral, i.e.

$$s_n = n + s_0. \qquad (12)$$

We refer to this as constraint C(2).

The combinatorial paths are related to Feynman paths (Feynman and Hibbs, Quantum Mechanics and Path Integrals), which represent information (contributions to the kernel). Each path represents a distinct factor for a polynomial constraint on the , all of which are required to produce the 'observed' constraining condition (in this case, the relationship between j^2 and s^2; since j and **s** are independently constructed. As a result, whereas information is summed, the independent probabilities represented by combinatorial paths are multiplied (Nature, January, 1989).

There is not a non-integer starting point here for s_0 at all but a mapping between to independent integer counts that results in a non-integer representation of one of j or **s**.

From the ordering operator calculus, distances are only comparable if they add according to a distance function. Given the conditions which a distance function must satisfy, and

assuming a locally flat space, independent quantities add according to the familiar Pythagorean formula (i.e., in quadrature). This fact has special meaning within the ordering operator calculus; it is the squared values of independent quantities that are related to each other in an observable manner. Only completely dependent quantities add linearly.

Since j and s are independently constructed, they must add according to a single norm or distance function if they are to be treated as comparable, i.e. they add in quadrature. We can think of constructing this relationship by considering the two discretely generated paths by which j and s can combine. The use of j or s has no independent meaning so long as only one periodicity occurs in the system under consideration. If, however, both j and s occur concurrently and it were possible to state that both j and s simultaneously begin and end synchronously and are still distinguishable, then we could give (11) as in (13) below.

$$137j\ \frac{\text{steps}}{\text{Coulomb event}} + 137s_0\ \frac{\text{steps}}{\text{Coulomb event}} = 1 \qquad (13)$$

There is no way in the system as constructed that j and s could be distinguished under these conditions. Rather, we must assume that some asynchrony in the periods of j and s (i.e., a non-zero "relative phase") occurs and that this leads to two possible conditions depending on whether s begins before or after j in whatever common representation we choose for comparisons of the two.

$$137j\ \frac{\text{steps}}{\text{Coulomb event}} + 137s_0\ \frac{\text{steps}}{\text{Coulomb event}} = 1 + e \qquad (14)$$

and

$$137j\ \frac{\text{steps}}{\text{Coulomb event}} - 137s_0\ \frac{\text{steps}}{\text{Coulomb event}} = 1 - e \qquad (15)$$

From this, and ignoring the physical units for now, we have:

$$137j + 137s = b_+, \quad \text{or} \quad j + s = (b_+/137) \qquad (16)$$

or

$$137j - 137s = b_- \quad \text{or} \quad j - s = (b_-/137) \qquad (17)$$

Note the subscripts on b, representing two distinct values. We either add or subtract values of 137s depending on whether the relative beginning of s with respect to j is positive or negative value in order to "synchronize" the periodicities. The b_- and b_+ are rational values which are less than or greater than 1 by some small amount e in accordance with our understanding of using

symmetric factors as discussed above:

$$b_+ = 1 + e \tag{18}$$

and

$$b_- = 1 - e. \tag{19}$$

In a continuum theory, the values of **b** are not differentiated. Instead, a symmetric average is used which for us is just the assumption of a unique root rather than the existence of multiple polynomial factors, and this assumption is not valid in a discrete theory. Remember, however, that in the discretum both factors must be used to obtain the "squared" value. The usual value is then just the square root of the product of $(b_-)(b_+)$, or

$$(b_-/137)(b_+/137) = (j + s_0)(j - s_0) \tag{20}$$

$$= j^2 - s_0^2 \ . \tag{21}$$

Let this last quantity be represented by a^2 :

$$a^2 = j^2 - s_0^2. \tag{22}$$

Thus, from (18), (19), (20), and (22),

$$a^2 = (1 - e^2)/137^2 \tag{23}$$

Note that a^2 is a <u>theoretical</u> quantity, whereas the multiple values of 137b are in principle observable. Thus we must be careful when comparisons to empirical values are made below.

Solving for s_0 in terms of the above equation, and putting in the constraint C(2) that higher values must differ from lower values by integers **n**, and that s_0 must be positive, we have that

$$s_n = n + \sqrt{j^2 - a^2} \tag{24}$$

although we must remember not to solve for the square root in the usual fashion if we are to compare our results with experiment.

<u>Computing the Probability that j and s Coincide</u>

We must now compute the probability that **j** and **s** are mapped to the same label in the representation space. To do so, we impose further constraints on **j** and **s** by noting that all events must be mapped to a single basis representation, as discussed in the subsection entitled <u>Labels</u> above. This space is constructed within the Combinatorial Hierarchy in the usual manner.

We can understand the quantity a as an event probability corresponding to an event **A** generated by a global ordering operator which generates the entire structure under consideration. Note again that these are probabilities in the frequency interpretation sense of probabilities. Then a is given by a combinatorial analysis of the relative frequency of a particular attribute instance (an event) within the given construction (corresponding in probability parlance to the sample population of possible events). We can now demand that additional constraints on the sampling be met as well, based on knowledge of the construction of the population. Each of the two events having probability counts j and s are derived by sampling from the same population.

For us, the population which defines a is defined via some ordering operator which generates the combinatorial hierarchy. **Footnote:** In the presentation given at ANPA 10, I attempted to justify the number of labels required differently. That means of computing the probability that j and s coincide in the representation came after the method used in the current version of the paper. It was motivated by (1) a misunderstanding of a physical constraint on the allowed values of k for valid velocities in Noyes' interpretation of Program Universe, and (2) a desire to make the mathematical and conceptual presentation less difficult to follow.

In order to represent the 127 basis strings at level 3 of the Combinatorial Hierarchy, the canonical basis strings must be of length 7. Just prior to the completion of level 3 via the generation of s, no more than 6 bits are required. (No relationship to total angular momentum is implied here by the use of the symbol L). For lengths L_a of a in terms of bit counts and L_{s0} of s_0, $L_a = L_{s0} + 1$.

The combinatorial ways of generating only labels which have even numbers of 1 bits, and eliminating those which under discrimination map any string either to the null string or to itself, the number of labels available for s_0 is 30. These can be divided into two sets of 15, each the dual of the other set.

There appears to be some connection between the two methods of computation. Under the requirement for constant velocity, the partitions of a representation string of length $n = 6 = pN$ must have $k = qN$ 1's for integer p, q, $N < n$. Except for $N = 1$, this condition reduces to the condition that k is even. How this relates to the representation is not clear, although I suspect the situation is more than coincidence. Also, if one uses Feynman's method of computing the amplitude of a discrete one-dimensional path integral (Feynman, R., and Hibbs, A., Quantum Mechanics and Path Integrals, problem 2-6, pp.34-36, c.1965), one sees that this constraint is equivalent to demanding that the amplitude is positive. This likely has some Program Universe interpretation in terms of spin.

This operator could be specified by the Program Universe algorithm, for example. In particular, let the population consist of the 137 basis strings defined at the first three levels of the hierarchy. Let a be the probability that a Program Universe event completes a set of 137 basis strings. In Noyes interpretation via the Noyes argument, this means that 1 out of every 137 vertices is a Coulomb event.

Now if the event S_0 is to be independent of **A**, it can not have a string representation that is greater than or equal to that of **A** (i.e. it must have shorter period in the terminology of the interpretation). For simplicity, we will use the canonical bit strings basis representation, in which the bit strings which are used to represent (i.e., label) the generated bit strings are just their binary enumeration (see Footnote below).

We must generate a population of 127 basis strings in order to represent level 3 of the Combinatorial Hierarchy since, under the interpretation, we require level 3 to have the Coulomb event probability (as in the Bohr atom model discussed above). These bit strings do not represent the occurrences (counts) of the steps and Coulomb events; they serve as labels for them. To combine the two independent constructions for j and s, they must have a common representation in this space of labels subject to the constraints of independence, namely, they should not require the same label at the same point in the generation.

The next constraint C(3) on the values of a, s and j is given by understanding what counts as a string that participates in a physical event.

If one considers the number of discriminately independent strings required to generate level three (again excluding the trivial initial strings of 0 and 1), we have 137. Indeed one may consider these strings as being discriminately independent with respect to the strings 0 and 1 of the initial level as a basis. Until the full basis is generated however, we must use as the basis the sixteen strings at the previous level.

Any elementary Program Universe event corresponding to level 3 of the Combinatorial Hierarchy occurs once in every 137 strings. For a particular representation, such an event maps to a specific basis string at level 2. This mapping can be understood as providing a distinguishable label for the string. We can not then include the ten discriminately independent strings from levels 1 and 2 as these are already represented by the basis strings in the mapping. Thus, one maps 127 discriminately independent strings at level 3 (as used for j) to the 16 basis strings (labels) available at level 2. However, the null string does not meet the requirements of the interpretation, reducing the number of available labels to 15.

Any other periodicity generated in the same period as j must then map to one of 15 basis strings if it is to have a distinguishable representation and must be one of the 127 strings

that can be generated at level 3. There are then 127*15 possible distinguishable representations that differ for any given event **s**. Thus an indistinguishable string representation will occur one out of 127*15 times. Due to our construction, we do not know if s_0 completes its cycle later than or earlier than **j**. This corresponds to twice as many possible ways in which the 127 strings can be labeled, so that we have 127*30 possible ways j and s can jointly occur.

Now comes the final constraint **C(4)**, the independence of the events **S** and **A**. Thus, given a level 3 Combinatorial Hierarchy event **A**, it is decomposable into independent sequences of events j and s subject to the constraint that j and s are indistinguishable with a probability of

$$e = \frac{1}{30 * 127} \qquad (25)$$

The event which completes s_0 may occur anywhere within an interval **2e** around the Coulomb event for **j**. The expectation of this event is just **1/(127*30)** according to (25) above. However, as noted in the subsection on Comparing New Theoretical Results to Old, we must take care in comparing our theoretical results to experimental values. Note that the value given by (25) can either subtract from or add to the probability of a Coulomb event. These correspond to two different combinatorial paths by which the independently generated sequences of events may close (the "relative phase" may be either positive or negative). In the particular physical situation, we are only interested in the portion that occurs within the period for j, namely 1-e. The reason for this is strictly operational. If the event which complete the period of s is later than that which completes the period of j, it will be counted operationally with the next period of j. Therefore the empirical value which corresponds to the probability that all s_0 vertices are generated distinguishably within one period of j is:

$$1 - \{ \frac{1}{(127*30)} \} \qquad (26)$$

The difference d^2 between j^2 and s^2 is then just computed as the square of this "root":

$$d^2 = (1-e)^2 \quad \text{or} \quad 1 - 2e + e^2 \ . \qquad (27)$$

It is now legitimate to work only with the single factor since we claim that d^2 corresponds to the square of α as the actual empirical value). Substituting, we obtain the following for **d**:

$$= \quad 1 - \frac{1}{127*30} = \frac{3809}{3810} \qquad (28)$$

Performing the substitution into **(1/137)**α for **d** we obtain

$$\frac{\alpha}{(137)} \simeq \frac{3809}{3810} \qquad (29)$$

Multiplying through we obtain the value
1/137.0359674... **(30)**

in comparison to the accepted empirical value of

$\alpha =$ 1/137.035964(15).

recently replaced by

$$1/137.0359895(61) \text{ at } Q^2 = m_e$$

(M.Aguiar-Benitz, et.al., Particle Properties Data Booklet, April 1988, p.2, p.4.)

V. THE FINE STRUCTURE OF THE HYDROGEN ATOM ENERGY LEVELS

Now that we have the relationship between **s**, **j**, and **a**, we can consider a quantity **H'** interpreted as the value(s) of the "Hamiltonian" energy attribute **H** expressed in dynamical variables at the **137j** value of the system. We seek to obtain a formula which governs the values of **H** in this system of two periods. According to our previous derivation of the relativistic energy (Foundations of a Discrete Physics), we represent this energy in units of the "mass" for the bit string defining j as

$$\mu c^2 \qquad (31)$$

where μ is a bit string invariant which transforms as mass under the Lorentz transformations, and **c** is the minimum limiting velocity for combinatorial system. The independent amount of additional energy due to the shift of **s** relative to **j** for a

period can be given then as a fraction of this energy by

$$(a/s_n)(H') \qquad (32)$$

One may wish to think of this quantity as a correction term to the quantity H' due to the fact that two independent periodic bit strings serve to define the system, s being the second such string. These independent values may then be either added or subtracted to yield the total system energy over one period, subject only to the constraint C(5) that the resultant may not exceed the total energy μc^2. This gives two factors:

$$1 - (a/s_n)(H'/\mu c^2) = H'/\mu c^2 \qquad (33)$$

and

$$1 + (a/s_n)(H'/\mu c^2) = H'/\mu c^2 \qquad (34)$$

These factors (33) and (34) then multiply just as the factors for a multiplied above. We then obtain the (elliptical) equation

$$(H'2/\mu 2c4) + (a2/s_n\,2)(H'2/\mu 2c4) = 1 \qquad (35)$$

Note that both (35) and (22) - rearranged with j as the dependent variable - constitute expressions of legitimate distance functions as demanded by the definitions given above.

On rearrangement of the terms of (35) and substitution of the equation for s_n (22), we have the desired Sommerfeld result for the fine structure of hydrogen:

$$H'/\mu c^2 = \left[1 + \frac{a^2}{\{n + \sqrt{(j^2 - a^2)}\}^2} \right]^{-(1/2)} \qquad (36)$$

Footnote: As Kilmister has pointed out, an inequality is appropriate for the general form of the Hamiltonian energy (Not only due to the physical bound state requirement that the energy be less than or equal to the system energy, but also due to the mathematical conditions for adding combining independent combinatorial quantities.) This is the case if one allows for other than a flat space with totally independent generators of the two portions of the energy attribute. As soon as the degree of independence and curvature is known, the inequality goes away and a corresponding correction is asserted.