Philosophica 50 (1992, 2) pp. 55-71

SPACE, TIME, DISCRETENESS

C. W. Kilmister

To the memory of A.F. Parker-Rhodes

1. *Introduction*

This paper reports my personal view of a fragment of a continuing programme in which a number of workers including myself are involved. I stress a personal view because the programme is still in active development and ideas are changing. Historically my starting point for investigating the theory was a very curious algebraic construction by Frederick Parker-Rhodes in the fifties, which is described by Ted Bastin in (1). The result of this construction became known as the Combinatorial Hierarchy and it has led to the study of the whole of physics on a combinatorial basis. A rather inadequate presentation of it is to be found in (2), (pp. 445-488 give the relevant algebra). Here I am concerned only with the fragment of the theory on the space-time continuum. (In what follows, very similar considerations apply to space, time and relativistic space-time. I shall not try to differentiate between them). The present paper owes much to conversations with Ted Bastin, David McGoveran, Pierre Noyes and Alison Watson but in expressing my gratitude I am more than usually concerned to accept the sole responsibility for this form of the ideas.

2. *Discrete space and time*

The interest in a discrete space and time arises because it has proved extraordinarity difficult to reconcile discrete aspects of the world with an objective space-time continuum. There at least two indications of this.

Most immediately, in physics the discrete aspects are dealt with in quantum theory, which has no theoretical connection with the description of gravity in terms of space and time by general relativity. This lack of connection shows up as a lack of explanation of limits to the ranges which continuum variables have, the quantum limitations embodied in Planck's constant and the event horizon caused by the maximal and invariant character of the speed of light. But, secondly, various mathematical attempts have been made to replace the real number field in terms of which the space-time continuum is defined by the rationals or by the integers modulo some large number N. The nature of failure here is more complex and more instructive:

(i) What value should be chosen for N? Such a theory will be complete only if it also provides an argument fixing N.

(ii) Any such attempt wil be sabotaged by Zeno-like difficulties. Grünbaum (3) draws attention to the most immediate of these: that a line of length L, from the origin to the point L, cannot be considered as a union of its points if, as in rational geometry, there is only an enumerable infinity of points and each point has length zero. To my mind a more serious difficulty of the same kind is the one exhibited in measure theory: assume that the length of the segment from a to b is b-a. Since the rationals are enumerable, enumerate the points between 0 and L. Situate the first point of this enumeration at the middle of a segment of length of length L/4, the second point similarly in one of length L/8, and so on. In this way all the points are potentially covered by line-segments. Some segments overlap, so that a more subtle covering could be provided by shortening some, but this is not necessary for my argument. All the points of the line are covered by a set of line-segments of length (L/4) $(1+\frac{1}{2}+\frac{1}{4}+...)$ = L/2 and this contradicts the original assumption that the line was of length L. This is a more serious difficulty because it does not depend on possibly bizarre attempts to constitute a line from its points, only from its line-segments.

(iii) In any such attempt, an acute difficulty arises over dimensionality. If space is described as a manifold over the reals, its three-dimensionality is a well-defined property. The existence of monsters like Cantor's space-filling curve does not upset this, for it represents a transformation which, though continuous, is not bi-continuous. But the whole notion of three-dimensionality slips away if the manifold is defined only over the rationals, or *a fortiori* over a finite field. For the enumerable character of the rationals means that there is a one-to-one mapping between the set of triples of rationals and the set of rationals themselves. Continuity will no longer serve to prohibit such mappings for it plays no important part in a rational theory.

(iv) It is sometimes argued that there is even a geometrical difficulty. In a plane tiled with minute equal tiles the diagonal for the unit square has as many tiles along it as the sides do. Weyl (4) traces this idea back to Plato and quotes Riemann in his "Über die Hypothesen..." of 1854: "that for a discrete manifold the principle of measurement is already contained in the concept of this manifold, but that for a continuous one it must come from elsewhere".

(v) Finally, such attempts have rarely tried very hard to overcome the difficulties and the reason for this is that no new philosophical or physical insight seems to come from the mathematical trick.

In this paper I give up the goal of reconciliation in favour of starting with the discrete and gradually arriving at concepts that have a recognisable similarity to those of continuum space-time. The formulation of a three-dimensionality will serve as an initial test of success; I shall also be able to make tentative proposals to avoid the Zeno-like difficulties, and the number N, in so far as it occurs, is fixed by the theory.

I propose to construct space and time from a more primitive beginning. What is this beginning? When Kant characterises space, not as a conception derived from outward experience (since such experience needs space already as a form of representation) but as a representation *a priori* which is necessary for external experience to be possible, and similarly, with "internal" replacing "external", for time, he is drawing attention to an important aspect which cannot be denied even if one rejects his example of geometry as synthetic *a priori*. I conclude from Kant's analysis that a construction of personal space and time must be in terms of the process of experience. The critical word here is "process", not "experience" although it is not used by Kant. It is clear from his use of plurals that he understood that a single experience will not serve to give space, only a continuing process of experiences can do that. The notion of process is the foundation of the new construction of space and time, and so it will be necessary to build it into the theory from the beginning. One consequence of this, perhaps the most profound, is that the nature of the mathematics has to be changed to take account of the inclusion of change. This may conjure up the name of L.E.J. Brouwer, but although Brouwer has been an influence on some of the workers on the programme in the past, the present theory is not intuitionism.

Those considerations came from Kant's analysis of personal space and time. I want to argue that, correspondingly, a construction of objective space-time must be by means of an abstract process. I cannot now say "of experience" because that would imply that someone had the experience and so the objectivity would be lost. In elaborating the idea I want to avoid any anthropomorphic temptation; the theory must work by itself, without the outside interference of the mathematician. What remains of the notion of personal experience in the abstraction is that of an increase of information. The picture of the universe is of one divided into a known part and an unknown part. Entities may change from being unknown to the known. Such an entity must then be labelled. The label shows up in the mathematics (which describes only the known part) as a newly generated element. All change, whether it be generation of a new element or incorporation of new information, takes place in discrete steps and the sequence of such steps is the process.

4. *The Principle of Choice*

My first step in seeing the structure of this process is to recognise that the system will not always give a new label when an entity arises. Sometimes an entity is given the same label as an earlier one because the process identifies them as two copies of the same kind of thing. This restriction on the process is necessary if arguments involving probability are to be possible. But I can add a further prescription: that the labelling should be systematic, constructing new labels as strings from a fixed label-alphabet, (instead of, for example, employing girl's names in the manner of labelling hurricanes). My justification for this prescription is not that it is necessary, for it is not, but that a process for which it holds and one for which it fails to hold will work in exactly the same way. I therefore choose to analyse further the one which is easier to treat. This form of argument will recur several times below; I call it the *Principle of Choice*. It is unimportant what label-alphabet is used. I find it convenient to use the infinite alphabet of symbols 1, 2, 3, ... These are not the cardinal numbers but I shall use the Principle of Choice to use them as ordinal numbers in a sense, the symbol m marks the m th step in the labelling, with a suitable definition of step.

At any stage in the process, let S be the set of entities already labelled. When a new entity arises, the relation between S and the as yet unlabelled entity is made explicit by some signal; so it is necessary for there to be some determinate set Z of signal-strings such that, if any is produced by the process, this signals that the entity should be given the same label as one of the elements of S. If such a signal results, the process will have to go on to determine which one of the earlier labels is to be used, so the process must be able to test an unlabelled entity a number of times before it is labelled. If a signal not in Z results, the process continues by labelling the entity and adding the label (and the signal, if it is new) to S. So S is not a fixed set but continually growing; it is in order to cope with this changing nature of S that the determinate set Z is needed in the process. I use the *Principle of Choice* again to rule that the new label to given at any stage is the least label-string not previously used, according to some conventional ordering of label-strings (which it will be convenient to state later).

5. *Discrimination*

In the situation just discussed, in which a signal in Z is produced, how can the process ascertain this? It seems as if it will have to test the signal-label in turn by a repitition of the test just used, with Z replacing S, and that in turn requires another signal to be tested and so on in an infinite regress. The determinate character of Z saves the day; there is just one way in which the process can avoid being trapped in a regress and that is to have as the members of Z symbols that are definitely not labels, so that it is at once determined whether a signal element is in Z or not. In order that Z be determinate, it must be finite or recursive. I shall confine myself to the finite case; I do not know whether the recursive case brings in anything new; I think not. In that case I can, by the Principle of Choice, restrict attention to a process in which Z has only one element. A useful notation for this element is 0, and in the ordering of strings 0 is to counted as less than any string.

I do not wish to distinguish between the developing set S and the process of determining its elements, so the testing can be symbolised by one of the two forms:

$S \rightarrow a, \qquad S \rightarrow 0$

according as the unlabelled entity is new or not. But in the course of further testing it may be necessary for the process to use the same test on an already labelled element, b say. The notation can accomodate this in the form:

$S, b \rightarrow a, \qquad S, b \rightarrow 0$

A mathematical reader will see a similarity to function notation, in which one might write

$S(b) = a, \qquad S(b) = 0$

But the mathematician's notation may mislead. He thinks of $S(b) = a$ as specifying a precise rule; the present process is not precise in this way. If the unlabelled entity is new, S can give rise to any signal except 0. I shall utilise this fact to apply the Principle of Choice several more times and it will then transpire that it is possible to fix attention on a particular process in which $S(b) = a$ is indeed a precise rule and the function S is determined by a recursive process. The details are in the notes. In the first stage the investigation is of the special case where S has one element only, say s and so for the signal $S(x)$ I write (s, x).[1]. The specialisation of (s, x) is then to one that satisfies:

$(s, s) = 0 \qquad (s, x) = (x, s)$

$(s, x) = (s, y)$ only if $x = y$

$(s, x) \neq x$ for any x.

To explain the recursive process, I need to specify the ordering of label-strings at this point. Because of the equality of (s, x) and (x, s), it is convenient to use the label-alphabet so that each label-string does not depend on the order in which the elements of the alphabet occur in the string. The string 12 and 21, for example, are to be treated as the same label and 12 may be taken as the canonical form. The ordering is defined in this way:

(i) Order one-element strings in the way that their labels suggest.

(ii) Order longer strings by the order of their largest element.

(iii) In each group of strings at the same point in the ordering under (ii) order them by the order of the next to largest element.

Continuing in this way orders the strings as follows:

1 2 12 3 13 23 123 4 14 24 124 34 134 234 1234 5...

With this ordering the recursive process is this: Firstly, $(1, 1) = 0$. Then $(1, 2)$ cannot be 1 or 2 so is set equal to 12. Next $(1, 12)$ cannot be 1 or 12 so is 2 and similarly $(2, 12)$ is 1. Thus the set [1, 2, 12] is closed under the bracket operation in the restricted sense that (p, q) belongs to the set if p, q are two *different* members of it. Next (1, 3) cannot be 1 or 3 or (1, 2) = 12 or (1, 12) = 2, so it must be 13, and so on. The labelling has been adapted to the particular bracket process selected by the Principle of Choice so as to give a simple rule that is easy to justify:

> If p, q are label-strings, then (p, q) is the *reduced form* of the string $p \cup q$, where the reduced form of a string s is defined as got by deleting every repeated pair of labels in s.

Thus, for example, (135, 1256) is the reduced form of 1123556 which is 236.

This function is called *discrimination*, for obvious reasons and since it is easy to prove that it is commutative and associative, the notation $p+q$ is used instead of the bracket. Sets like [1, 2, 12] and [1, 2, 12, 3, 13, 23, 123] are called *discriminately closed subsets* (dcss), will prove to be important below and have sizes 3, 7, 15, $\ldots 2^r-1 = r^*$. A set of entities with a discrimination operation (commutative, associative operation + such that $x+y = 0$ if and only if $x = y$) is called a discrimination system.[2] The *order* of a label-string is defined as the largest element of the label-alphabet in it, so that the dcss above with 7 elements has one of order 1, two of order 2 and 4 of order 3.

6. *Characteristic Functions*

Discrimination has thus been shown to be always a possible specialisation of the process in which a putative new element is tested against a single one. In general however the process will be concerned with testing against a growing set $S = [u_1, u_2, \ldots]$ of already labelled ones. It cannot do this by simply testing each u_i in turn, for then a later stage in the process would have to determine whether this u_i had been used before or not and this could only be answered by a further test and so on in an infinite regress. Instead the process must treat the set S as a whole and give a signal, (S, x) say, which (using some of the simplifications above) is 0 if and only if x belongs to S. If (S, x) = 0 further testing against subsets of S will be needed before the process succeeds in labelling x by determining which member of S is the same as x. The various acts of testing may occur in various orders and it is not possible here to use the Principle of Choice to select a definite order, because of the infinite regress just mentioned. This causes an unexpected limitation on the set S against which x may be tested, in order that the labelling may be unambiguous.

To see what this limitation is, consider an actual process. The first element is labelled 1, the next different one is labelled by the next label in the ordering, 2, and the signal that it is different, 12, is also added to the set S giving [1, 2, 12]. The next new element, however this is determined, will be labelled by the next least string, 3 and the signal which showed it to be new may perhaps be 13. The difficulty arises at the next stage, with a further new element; should it be labelled 23? Only if the process has not already used 23 in one of the subsidiary tests and we do not know that. A clue to the cause of this ambiguity is its non-occurrence at the earlier stage. Because [1, 2, 12] is a dcss, there is no ambiguity in giving the label 3 to the next element which does not belong to it. The closure prevents 3 from having come up earlier. For unambiguous labelling, the next set is the dcss [1, 2, 12, 3, 13, 23, 123]; a set cannot be a candidate for the testing process until it has been "filled up" that is, is discriminately closed.

I can now use the Principle of Choice on the same lines as in section 5 to show that the testing process is equivalent to (works in the same way as) a *characteristic function* S, so that (S, x) may be written S(x).[3] The nature of the recursive rule detailed in the notes can be seen from a detailed example. Take as the dcss S = [1, 23, 123]. Then:

(i) Set S(1) = S(23) = S(123) = 0

(ii) Consider the least element outside S, that is, 2. Set S(2) as the least element of S, that is, 1.

(iii) Set S(u+2) = 1 if u is in S. This then defines the function S for the dcss S_1 = [1, 2, 12, 3, 13, 23, 123].

(iv) Consider the least element outside S_1, 4 and set S(4) as the least element after 1, that is, 2.

(v) Set S(u+4) = S(u)+2 if u is in S_1. This defines the function S for the dcss $S_2 = S_1 \cup [4] \cup [S_1+4]$.

And so on, giving S(5) = 3, S(6) = 4 ... The construction is of a form of S(u) defined for all u but with an infinite tail of values which carries no information, since the values are determined by a primitive recursive function. The only important part of the function S is the first seven values, that is, the values for the initial dcss T which contains S and is the least to do so. (By an initial dcss I mean one such that, if u is a member and v is less than u, then v is a member.)[4] These values are specified by giving the values of S for the elements of the label-alphabet 1, 2, 3 from which T is generated by discrimination.

7. *The Combinatorial Hierarchy*

The importance of this functional form is that, if S, T are two dcss with their corresponding characteristic functions, then a binary operation is defined between these functions by the usual rule:

$$(S+T)(x) = S(x)+T(x) \text{ for all } x.$$

Subject to defining also that $S = T$ if and only if $S(x) = T(x)$ for all x, this $S+T$ is a discrimination. So the set of characteristic functions is itself a new discrimination system, the process can be repeated and the system is hierarchical.[5] I return to the notion of change which the process describes. Entities change from the unknown to the known. I do not wish to draw a sharp distinction between a physical theory and the manipulations with bits of the physical world which derive a meaning from it. So the process will itself lead to examples of such change and the most obvious of these will be the emergence of discriminately closed subsets. When the process is dealing with these subsets, for which purpose their characteristic functions may be employed, it is operating at a higher level of complexity than when it is dealing with the individual members of the subsets.

It is then possible for the process to move to a still higher level of complexity, since the characteristic functions are themselves a discrimination system, and so on. Whether the process moves to the more complex level is a question of how much self-organising it is carrying out. There is, however, a limit to the extent to which this self-organisation can proceed. In the general programme I spoke of in section 1 this limit is identified with the source of the limits on the ranges of continuum variables. In the particular process concerned with experience in space and time which I am dealing with here, the limit provided the three-dimensionality of space. To determine this limit, I construct the most highly organised behaviour for the process. The core of this construction was put by Parker-Rhodes in essentially this form:

> Start with a basis B of r elements. Every subset of B generates a dcss which has a characteristic function. These characteristic functions

then form a basis for the next level and the construction begins again.

Parker Rhodes does not carry out the construction in terms of characteristic functions but in terms of eigenvectors. But the two are fully equivalent, because if S(u) = 0, and T = S+I, where I is the identity function, I(u) = u for all u, then T(u) = u. Where I would write S, Parker-Rhodes has T. He shows that the maximum self-organisation comes from beginning with 2 elements. Successive bases then have 3, 7, 127 and $2^{127}-1 = 10^{38}$ elements and then the construction breaks down.

My aim has been to set this construction in a context in which it could be understood, with some necessary modifications. The original version made use of linear vector spaces, matrix operators and the choice of these as linearly independent; these mathematical devices seem *ad hoc* because they are being used to incorporate ideas that have not been fully discussed. To see what these ideas are, consider first a set of elements [1, 2, 3, ...r]. These generate a dcss with $r^* = 2^r-1$ members so that the result of arbitrary discriminations will be to give one of these r* elements or 0. In terms of the signal-processing concept of information, to specify one of these $r^*+1 = 2^r$ possibilities is to give r bits of information. For shortness, I say that each element carries r bits. If S is an arbitrary set of strings, r in number, then each element may carry r bits or less. For, if S = [1, 2, 12, 3] so that r = 4, the dcss has only 7 members and so each string carries only $\log_2 8 = 3$ bits.

Consider next, for a set of r strings, each carrying r bits, the r* characteristic functions, one for each dcss generated by strings of the set. Any characteristic function is specified by listing the r strings into which it respectively carries 1, 2, 3, ...r. Hence such a characteristic function carries r^2 bits. The initial stage of the construction is described as *possible* by Parker-Rhodes so long as $r^* < r^2$, which limits r to 2, 3 or 4. The possibility referred to here is really that of repeating the construction at the next higher level. At this next level the r* elements carry r^2 bits each and this is more than r*. The easiest way of dealing with this is to consider the r* elements as a subset of r^2 independent ones, each carrying r^2 bits in the usual way. Then the next lot of characteristic functions at this level will carry r^4 bits. What is not yet determined is how many such characteristic functions there are. There may be (r*)* or there may be fewer. If the most highly organised situation of (r*)* results, the possibility of further extension involves $< r^4$ which requires that r should be 2.

To begin again with 2 elements, then, each carrying 2 bits, the first higher level involves $2^* = 3$ characteristic functions each carrying $2^2 = 4$ bits. At the next stage there are then 7 characteristic functions, each carrying 16 bits and this gives rise to a further stage. Presumably there will then be 127 characteristic functions at this level, each carrying 256 bits. But at the next stage the situation is of 10^{38} characteristic functions, each carrying only $256^2 = 65536$ bits. Since the elements carry too few bits, the situation is like the example quoted above and the construction cannot be carried on. This stoppage draws attention to the possibility that similar diffficulties might have arisen unnoticed at an earlier stage. In fact the 3 characteristic functions at the first level are uniquely determined and do give rise to 7 dcss. The 7 characteristic functions at this level are not uniquely determined but it is easy to see that some choices of them (about 90% of the possibilities) give rise to 127 dcss. It is more difficult, but possible to prove that some choices of the 127 characteristic functions do indeed give rise to 10^{38} dcss. The argument for this is in the notes.[6]

8. *Dimensionality*

I begin by summarising the argument up to here. I assumed that information about the world increases by means of a discrete, self-organising process. Such a process leads to algebraic structures on several levels. The mathematicians' language is to speak of a graded algebra, whose elements are of the form:

$$u = u_1 \oplus u_2 \oplus \ldots$$

the u_i being the elements at diffferent levels. Operations between elements, especially discrimination, take place individually:

$$u+v = (u_1+ v_1)\oplus(u_2+v_2)\oplus..$$

The extent to which the higher levels come into play is a measure of the process's self-organisation but there is a limit to how much self-organisation can take place, specified by the combinatorial hierarchy construction. The result shows that the graded algebra is of finite type, as it is called; it has only four levels. The elements at the first level form an algebra of dimension 3, then those at the next level come in as well so that a graded algebra of dimension $3+7 = 10$ results. Those at the third level adjoined to this give a graded algebra of dimension 137; and the fourth level is then exhibited as different in nature from the other three. My contention is that, because this abstract process (i) has this unique $3+1$ bounding structure and (ii) represents the process of increasing information about the world, this provides an explanation of why we describe all our experience as taking place in a particular framework, which we call space and time. The nature of the construction overcomes the dimensional objection (iii) in section 2 because the hierarchy levels play the same role that bi-continuity plays in the theory with the real field. Indeed, it does much more, for it offers an explanation of why the dimensionality of space is three and not any other number.

Because of the ambitious nature of my claim, I must expect objections. One of these I wish to disarm at once because I earlier held it myself. It is that it appears bizarre to identify the very different three initial levels of the hierarchy with the three dimensions of an isotropic space. But to urge this is to ignore the process aspect of the theory. The first level is pursued until it is no longer possible to fit in the new information coming in. It is natural for the complexity to increase at each such step.

I conclude by referring briefly to the other objections listed in section 2. If the theory is to be one with a large finite number, then the number N will be $2^{127}-1$, so that objection (i) is overcome. But I do not believe that the completely elaborated theory, of which this paper is only a preliminary sketch, will prove to have only a finite number of elements. The objects of the theory (sets of space-time points) will not be single elements of the algebra, but *complexes*: a finite set of elements, a set of functions for the dcss generated by then, another set of functions for all the dcss generated by the first set of functions and so on. This elaboration is not cut off by the Parker-Rhodes theorem, for that is engendered by seeking the most highly self-organised system. Such an infinite complex can be coded in bit-string notation[2] but with the strings now potentially infinite there is no finite limit on the number of elements.

I do not regard objection (iv) as a serious one; its weight is no more than is already contained in (ii), to which I return below. As to (v), my claim is that there is good reason to pursue this theory because of the unforced appearance of the well-known and highly significant physical scale constants, 3, 10, 137 and 10^{38}. These four numbers, with a stop at that, are a good indication that further elaboration will lead to increased philosophical insight into physics. This programme is under way but has not yet been fully carried out; it is, in any case, beyond my present brief.

I am left with (ii) as a serious problem and one to which I have not yet a convincing reply. I am hopeful that a way out exists on these lines: the basic notion of the theory is that of a dcss. Any other set has a unique *discriminate closure*. It seems possible to define a "discriminate topology" on the set of complexes in terms of a closure operation, in the manner of Kuratowski, with the closure being, of course, discriminate closure. But this is not quite straightforward. A closure operation is defined as having the property that the closures of the intersection and union of two sets are the intersection and union of their closures. Only the first of these two properties holds for discriminate closure; the other has to be taken in the weaker form that the closure of the union of the closures of two sets is the closure of their union. So the actual setting up of a topology will be significantly different from the usual one and the exact nature of these differences still remains to be prised out. But if a clear topological picture emerges, it will then be possible to define a metrical picture from it – as an artefact, as it were – which will satisfactorily deal with (ii).

NOTES

1. The process of determining whether x belongs to S is now that of determining whether x and s are equivalent. I rephrase the three conditions of an equivalence relation as:

   (i) $(s, s) = 0$ for any s,

   (ii) if $(s, x) = 0$, then $(x, s) = 0$,

   (iii) if $(s, x) = 0$ and $(s, y) = 0$, then $x = y$.

   Consider (ii): the Principle of Choice allows the selection of an equivalent process, with $(s,x)^*$ replacing (s, x) where

   $(s, x)^* = \min[(s, x), (x, s)]$

   for $(s, x)^* = 0$ if and only if $(s, x) = 0$. Then (ii) takes the form

   (ii)' $(s, x)^* = (x, s)^*$

   and (iii) can then be rewritten with $(\ )^*$ for $(\ )$ without change of content. To put it differently, we can simply drop the stars and assume the process to fulfil (ii)'.

   By another application of the Principle of Choice it is possible to strengthen (iii) as well, to

(iii)' (a) If $(s, x) = (s, y)$, then $x = y$
(b) $(s, x) \neq x$ for any x.

The demonstration must now be recursive, for it must deal with all values of x. I call the method "Conway's trick" as it is modelled on a definition used by him in a different context in (5). Consider the set B of all possible "brackets", that is, all (s, x) satisfying (i) and (ii)' without the star. Now select from this set those, called ( )* again, satisfying

$$(s, x)^* = \min[(s, x):(s, x) \neq (s, x') \text{ for any } x > x';(s, x) \neq x].$$

Here, as above the ordering refers to that given in the text. Since ( )* is equivalent to the members of B, it lies in B and it obviously satisfies (iii)'. Dropping the star again, one has a bracket that satisfies the conditions given in the text.

2. The labelling can be coded in a different way, which makes comparison with the Parker-Rhodes construction easier. The label-alphabet element r is represented by the (potentially infinite) *bit-string* 0...010..., with a 1 in the rth place but nowhere else. A label-string is then represented by addition of the corresponding bit-strings and the discrimination operation becomes addition modulo 2. The label-string equation in the text reads:

$$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ \cdot \\ \cdot \end{bmatrix}$$

The only significant way in which this bit-string representation differs from Parker-Rhodes' is that the bit-strings have no definite length.

3. In note 1 the starting point was the three conditions satisfied by any equivalence relation. The first of these becomes in the more general case:

(i) $(S, x) = 0$ if and only if x is in S.

The second, symmetry, condition has no simple analogue but the third corresponds to the statement that I have argued for, that S is a dcss:

(ii) If (S, x) = 0 and (S, y) = 0 and $x \neq y$, then (S, x+y) = 0.

I follow the same general lines of the earlier argument to show that (ii) can be replaced by

(ii)' (a) If (S, x) = (S, y) and $x \neq y$, then (S, x+y) = 0;
(b) (S, x) = x for any x.

The means of doing this is essentially the same recursive process as before; (S, x) is taken as the least element in play up to that point which satisfies (i) and (ii)'. The phrase "in play up to that point" needs qualification. The set of elements in play up to a certain point means the discriminate closure of all those which have been mentioned in the construction up to that point. This recursive construction then gives a function, which I write as S(x). The working out of this rule is exemplified in the text.

4. The function S in the text can therefore be represented by an *array* (0, 0, 1). In general an array R is defined by k elements, either label-strings or zeros, r^1, r^2, ...r^k, none of which is of order greater than k. Such an array represents a linear operator by these rules:
   (i) If s is in the label-alphabet, $R(s) = r^s$;
   (ii) if u is a label-string, $u = s_1+s_2+\ldots s_i$, where the s_j are in the label-alphabet, then $R(u) = R(s_1)+R(s_2)+\ldots R(s_i)$.

   In the bit-string notation an array is very like a square matrix, but is of indefinite size. In the Parker-Rhodes construction (in section 7) the bit-strings are artificially restricted in length and the arrays are matrices of definite size.
5. The fact that the characteristic functions form a discrimination system implies the existence of a corresponding bit-string notation for them. A good deal of discussion was wasted in the early years of the construction in justifying particular notations, as, for instance, writing (0, 1, 1) as a square matrix

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

   and then setting the columns one above the other, giving a bit-string which in terms of the label-alphabet is 47. But this discussion was beside the point; any bit-string notation will serve.

6. Instead of proving the theorem directly, it is easier to deduce it from a simpler, though more general, result:

   Consider r elements giving rise to r* dcss. The r* characteristic functions can be chosen to give (r*)* dcss at the next level if and only if they carry more than r* bits.

   The necessity of the result is clear and I prove the sufficiency by showing how to construct the sets of functions. The method is a fairly obvious algorithm:

   1. Order the dcss (a) in decreasing order of cardinality, (b) for all those of the same cardinal, order by the order of the largest element, (c) if two sets have the same largest element, order by the next largest, and so on.

   2. Beginning with the first dcss, construct an array for it in exactly the way specified in notes 3 and 4 and exemplified in the text. (The first set will usually be all the elements and so the array will be zero).

   3. Repeat with the next dcss and check whether the result is linearly independent of the first array.

   4. If linear independence has been retained, repeat on the next dcss. If not, modify the array to regain linear independence by increasing the first element of the array which will do this by the amount necessary to do so. This last step casts doubt over whether the method is truly algorithmic or whether the construction might fail. The algorithmic character is equivalent to the truth of the theorem. The proof takes the form of first exhibiting the process in the initial stages and then showing that the later ones will in fact be easier. It will be sufficient to detail stages 1 and 2. At stage 1 we have:

	1	2
1,2	0	0
1	0	1
2	2	0

   where the top row labels the columns of the arrays, which are listed below it. The left-hand column lists the dcss by giving the elements of which they are the discriminate closure. The same procedure at stage 2 gives the table on the next page. At the sixth line no new element enters in any of the three columns, which is a sign that linear independence is in danger. We try to avoid the danger by increasing the 2.

	1	2	3
1,2,3	0	0	0
1,2	0	0	1
1,3	0	1	0
2,3	2	0	0
1	0	1	2
2	2	0	1
3			

The value 12 is ruled out because then the effect of the array on 12 is to give 12 and 3 is ruled out for similar reasons but 13 is permitted and so the scheme is completed:

2	13	0	1
3	3	1	1

It would be tedious to give the details at the next stage r = 4; it can be carried out. But at stage r = 5 things get simpler because the conditions of the theorem are such that the arrays have six entries, not five and this means that as well as filling up the first five places in accordance with the algorithm, the sixth is available to help over the linear independence. The same is true at every higher stage.

REFERENCES

1. Ted Bastin, *Studia Philosophica Gandensia 4*, 1966, 77-101.
2. Ted Bastin, H. Pierre Noyes, John Amson & Clive W. Kilmister. *International Jour. Theor. Phys. 18*, 1979, 445-488.
3. A. Grünbaum, *Philosophical problems of space and time*, Routledge and Kegan Paul, 1964, (pp. 172, 175).
4. H. Weyl, *Philosophy of Mathematics and Natural Science*, Princeton, 1949, (p. 43).
5. J. Conway, *On Numbers and Games*, Academic Press, 1976 (p. 53).