MATHEMATICS OF

A HIERARCHY OF

BROUWERIAN OPERATIONS

TED BASTIN*

Original paper written in September 1964

Cambridge Language Research Unit

Re-typed in TEX, with some editing, 1998-FEB-27,
by John Amson**.

* Contact Address :-
The Mill House, Penbryn, Sarnau, Ceredigion, SA44 6QL, Wales
'Phone : 01-239-810-404

** Contact Address :-
The Old Manse, 5 Shore, Anstruther, Fife KY10 3DY Scotland
'Phone & Fax : 01-333-310-087, e-mail : john.amson@which.net

Editorial comment by John Amson

A typescript draft of this 1964 paper was found amongst old files in my attic. The original had been lost in the fire at Ted Bastin's home that destroyed so much of his archive material.

I believe the paper to be significant on both historic and practical grounds. Especially because it helps to recall what I remember to be the importance attached by Ted Bastin, when introducing me to this area of thinking, of the vital rôle of ideas of Brouwerian Intuitionism on the early development of what later became known as the Combinatorial Hierarchy — a rôle which, in my considered opinion, should once again be accorded primacy of place in current studies in this area.

Only the minimum of purely editorial changes have been made, occasional supplementary remarks and comments have been added, and the bibliography updated.

....

Section 1. Introduction.

The purpose of this paper (which is partly based upon work by A.F.Parker-Rhodes [see Refs.(1) & (2)] is to develop an algebraic system in which a set of operations upon a second set of quantities can be treated as operands for another set. The required algebraic system thus essentially has a hierarchical structure in respect of the operator/operand relationship.

The general reason for being interested in such an algebraic system is best explained by referring to the accounts given by Brouwer of his understanding of the nature of mathematics and of mathematical *creation*. This understanding was the original motivation — as far as Brouwer himself was concerned — for the technical developments of the Intuitionist School of mathematicians and logicians. I shall not be concerned much with these developments in the present paper, but my starting point is the same[1].

Brouwer describes two "acts of intuitionism" and writes [see Ref.(3), p.2] :-

> The first act of intuitionism *completely separates mathematics from mathematical language, in particular from the phenomena of language which are described by theoretical logic. It recognises that mathematics is a languageless activity of the mind having its origin in the basic phenomenon of perception of a* move of time*, which is the falling apart of a life moment into two distinct things, one of which gives way to the other but is retained in memory.*[2] *If the two-ity thus born is divested of all quality, there remains the common substratum of all two-ities, the mental creation of the* empty two-ity. *This empty two-ity and the two unities of which it is composed, constitute the* basic mathematical systems. *And the basic operation of mathematical construction is the* mental creation of the two-ity of two mathematical systems previously acquired*, and the consideration of this two-ity as a new mathematical system.*

The basic thought I wish to adopt and develop from this statement is that the juxtaposition of two mathematical entities generates a new mathematical entity, as a direct consequence of the fact that they are entertained in the mind in temporal succession. The mathematical system developed in this paper is intended to describe the generation of mathematical entities of this sort where complete freedom is allowed to the generating process. The mathematical entities considered are simple — being restricted to arrays of binary quantities. This approach has to be contrasted with intuitionist mathematics as that term is usually understood. In conventional intuitionist mathematics infinite choice sequences (in which some freedom is allowed at each stage) are shown to necessitate certain logical distinctions which are then taken as a sufficient representation of the creation process in mathematics. The creation process itself is then forgotten. In this way, indeed, intuitionist mathematics as developed by Heyting [see Ref.(4)] has ignored its own origins and concentrated on the systematization of these logical distinctions in abstraction from the

---

[1] *Ed.* For a comprehensive account of Brouwer's ideas, and a bibliography, see Ref.(9).

[2] *Ed.* The following two sentences in this quotation were omitted in Ted Bastin's original paper, but are here included, for completeness.

study of the Urphenomenon that they were introduced to illustrate. I take this view that this development has restricted the possible field of application of Brouwerian philosophy, and I shall attempt a more positive application of it in the present paper by systematizing the time concept directly.

The difference between the present approach and current intuitionist mathematics becomes important as soon as any operation is defined in which members of a set are selected and combined with other elements in any way. any such operation would be equivalent to a method of retrieving the elements in question, and any retrieval process must consist (by the original definition of the mathematical creation process) of a series of temporal juxtapositions each of which is a mathematical operation which must be represented within the system.

In the present approach the act of remembering the elements has to be part of the mathematics, and from this position our development of the formal properties of a mathematical system will begin. Originally [see Ref.(1)] such a mathematical system was proposed as a description of a hierarchically organized automaton, and Parker-Rhodes has applied the techniques of matrix algebra over the Boolean ring of two elements ($\mathbf{J}_2$) to describe such an automaton. In § 2 matrix multiplication which will be formally introduced later in this paper is used to illustrate the principles laid down in this present introductory section, by establishing a model of a Brouwer *spread* of infinitely proceeding sequences. These sequences are chosen subject to restrictions which make them appropriate to illustrate the principles in question, but in § 2 I do not aim at a deductive presentation, and therefore I shall not attempt to justify these restrictions in that section. In the subsequent sections I shall give a systematic development of the binary matrix algebra from a set of simple principles. The finiteness theorem of Parker-Rhodes [Ref.(2)] will be shown to be deducible from these principles without the arbitrary features of the form in which the theorem was originally proved. Finally, the finiteness theorem of Parker-Rhodes will be compared with the 'spread theorem' of Brouwer [Ref.(3)], and by that comparison further light will be thrown onto the relationship between the mathematics of this present paper and technical intuitionism.

Section 2. A Mathematical Model to Illustrate the Hierarchical Principle.

Suppose that data are presented to a system capable of manipulating them (in ways I shall presently describe) in the form of a time sequence of ordered pairs of quantities :-

$$\text{- - - - - - - - - -}\longrightarrow \textit{time}$$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \dots , \begin{bmatrix} x_j \\ y_j \end{bmatrix}, \dots \tag{2.1}$$

where x_j, y_j stand for either 0 or 1. suppose further that the passage of time is from left to right as is shown by the arrow in diagram (2.1) so that at any given point in the sequence the generation of the next term on the right may be discussed before it is given. Moreover, in accordance with the principles of § 1, the sequence is considered two terms at a time in the first instance.

The act of considering two successive terms will be represented in the following way :-

if the pairs are written V_i, V_j, then we seek a 2×2 matrix A_{ij} such that V_j is the transform of V_i under A_{ij}.

Following Parker-Rhodes [Refs. (1) & (2)] the defining operations used in these matrix transforms are

`Multiplication`

$\times$	0	1
0	0	0
1	0	1

`Symmetric Difference`

$+$	0	1
0	0	1
1	1	0

The explanation of the reason for adopting the symmetric difference operation is deferred to the next section. In general there will be more than one matrix A_{ij} such that $A_{ij}(V_i) = V_j$. We choose any one of these <u>at will</u>. Having chosen it we see if it will fit the next pair of terms (*i.e.* V_j and V_k where V_k follows V_j). If it does (*i.e.* if V_k is the transform of V_j under A_{ij}) then we proceed to the next pair, and so on. If it does not fit, then A_{ij} has at this stage to be rejected as the specification of the <u>development</u> of the sequence. An example will make this clear.

Level $\mathbf{J}_2$

$$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, ---\rightarrow \tag{2.2}$$

Level $(\mathbf{J}_2)^2$

$$\begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}, \quad \begin{pmatrix}1 & 1\\0 & 0\end{pmatrix}, \quad \begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}, \quad ---\rightarrow$$

The notation " $(\mathbf{J}_2)^2$ " with an exponent '2' to represent the changed level, is a convenient one, but it should not be taken to mean that in any ordinary sense that $(\mathbf{J}_2)^2$ is the square of $\mathbf{J}_2$. However, the <u>number</u> of <u>elements</u> in the matrices in $(\mathbf{J}_2)^2$ is in fact the square of the number of elements in the vectors in $\mathbf{J}_2$, and this relationship persists at higher levels. The sequence of vectors labelled " Level $\mathbf{J}_2$ " is the original data sequence, and the sequence of matrices labelled " Level $(\mathbf{J}_2)^2$ " specifies limited portions of the development of the original data sequence $\mathbf{J}_2$.

The level $(\mathbf{J}_2)^2$ in diagram (2.2) is now regarded as a second order sequence of data for repetition of the process of analysis at a new level $(\mathbf{J}_2)^4$ in which 2×2 matrices have been replaced by 4×4 matrices. For this purpose each 2×2 matrix is to be treated as an array and replaced by a 4×1 array which can then be regarded as a vector for the new matrix transform. Provided that the rewriting of the arrays is always performed in the same manner, it is obvious that there is an order preserving 1 : 1 correspondence between the elements of the set of 2×2 arrays and elements of the set of 4×1 column arrays. A simple way to rewrite the 2×2 arrays is to put the second column beneath the first column and in general to 'string out' the columns one beneath another. This method is obviously applicable to all other levels. Thus the second level in Diagram (2.2) is rewritten as

$$\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad - - - \rightarrow$$

I shall now compare the properties of this model with the requirements of the general principles of § 1. In the first place I recall that matrices were described in the model as being chosen "at will". This phrase needs further explanation. We may imagine that the choice is made by a random process[3] so far as any knowledge is concerned that can be described at level $(\mathbf{J}_2)^2$. However, if we are to consider a more complex level, then a specification may exist of the choice which can only be regarded as random in the absence of the more complex level. In terms of this model, we imagine a system containing a series of levels of increasing complexity which correspond to the possibility of surveying — in stages — larger segments of the original sequence. In § 6, below, it will be shown that this sequence cannot be continued indefinitely, and that an infinite regress is avoided.

It is now necessary to make an important distinction between the cases where the original data sequence is considered as having an objective existence apart from the process of analysis provided by the model that I am describing, and cases where no such independent existence is granted to the data sequence. The present paper will be entirely concerned with the second of these cases (an essentially simpler situation). Thus in the case I shall consider, the length of the data sequence that can be considered at one time is determined by the number of levels of the model that have been developed, and elements of the data sequence that lie outside the reach of the system in this way are inaccessible and cannot be recalled, since to recall them would require us to specify the individual characteristics of each. If the power of recalling or retrieving an item when presented with another — related — item (as in memory association experiments) can be taken as a definition of memory, then the systems described in the present paper could be described by saying that within them no meaning can be attached to entities lying outside the memory defined in the system.

The model given in this section was constructed by Bastin and Kilmister [see Ref.(5)].

[3] By "random process" I mean a process (a) whose causation we are not prepared to look into further for the purpose in question, and (b) whose causation we have no reason to expect to be connected in any way with the purpose we have in mind for its application.

Section 3. The Matrix Model and Intuitionist Logic.

Brouwer [see Ref.(6), p.113] has defined a number, S_f, which enables him to demonstrate that the general philosophical principles of intuitionism make necessary an intuitionistic logic[4]. He has shown this, for example, by considering the assertion :

$$S_f = 0 .$$

This demonstration consists in observing that though this assertion will be proved to be true by any proof that proves it to be non-contradictory, yet it does not satisfy the principle of the excluded third. It is easy to show that an analogous proof can be carried out for the model of § 2 — in which case all the conclusions regarding intuitionist logic adduced by Brouwer would follow also for that model — provided that we can assume that integers in natural order are automatically assigned to the elements of the sequences in the model, and that this assigning of integers does not have itself to be represented in the model.

Before I can consider this vitally important *proviso* I have just mentioned however, I shall construct S_f.

The definition of S_f requires first the definition of a *fleeing property*[5], which Brouwer defines to satisfy these conditions :-

(1) for each natural number n, it can be decided whether or not n possesses the property f; (the adaptation of the model just described was necessary to satisfy this condition);

(2) no way is known to calculate a natural number possessing f;

(3) the assumption that at least one natural number possesses f is not known to be contradictory.

To relate these 'fleeing properties' to the rules of § 2 — expressed by matrix transformations — that were chosen at each stage of a sequence to specify the development of the sequence, we take f to be the property that a given matrix A ceases to specify the development of a sequence V_0, V_1, $V_2, \ldots$ at the elements of the sequence under consideration. Clearly this 'fleeing property' then has a sense for each natural number defining a term in the sequence, (condition (1) above); clearly also conditions (2) and (3) are satisfied.

Let us now return to consider the adaptation of the model of § 2 that made it possible to construct a number that had the properties of Brouwer's number $S{-}F$, within the model.

---

[4] *Ed.* In the later Ref.(9), Brouwer uses the symbol $\kappa\kappa_f$ here, and reserves S_f for a real number which is a limit of a sequence of rational numbers that depend on $\kappa\kappa_f$.

[5] *Ed.* Brouwer [see Ref.(9), p.6] gives this example of what would be a fleeing property:- "There exists a natural number n such that in the decimal expansion of π the nth, $(n+1)$th, $\ldots$, $(n+8)$th and $(n+8)$th digits form a sequence 0123456789." The question as to whether this is true "relating as it does to a so far not judgeable assertion, can be answered neither affirmatively nor negatively. But then, from the intuitionist point of view, because outside human thought there are *no* mathematical truths, the assertion that in the decimal expansion of π a sequence 0123456789 either does or does not occur is devoid of sense."

We find at once that no provision has been made within the model for the adaptation in question — namely the assumption that a natural number can be assigned to each element of a sequence, without including the assigning process within the hierarchy of levels. So far as the model has been developed at present, the process of deciding that a sequence has fitted a given law of development has meaning only in so far as it is possible to compare more than two terms, any such comparisons have to be conducted in stages — built up from comparisons of pairs. Hence when we "decide, for a given natural number n, that n does possess or does not possess f," this decision can be taken at any one of a sequence of levels depending on how much of the sequence we are prepared to consider, and all the decisions of this sort that can be contemplated within the model have already been incorporated within the levels. As moreover the mathematical model of § 2 is intended to be applied to the understanding of the process of observation, though I do not develop this application in the present paper, the following point needs making. The mathematical model of § 2 is not yet a description of the process of observation; it is at a more primitive stage — being a formulation of the reference frame within which all observations have to be made. We have not a scheme rich enough to distinguish observing system from the system that is observed so long as further mathematical properties cannot be discussed in detail in future work and cannot be short-circuited by having an independent assignment of numbers to the elements of a sequence.

let us now turn to consider Brouwer's own position. It seems on the face of it that if the model of § 2 correctly represents Brouwer's time philosophy, and if as well we have to adapt the model in a way quite foreign to the spirit of this philosophy to define the number S_f within it, then Brouwer must have been wrong in thinking that an intuitionist logic would adequately exemplify the philosophical principles from which he started. I think this conclusion is correct. I can see no justification for Brouwer's implicit assumption that we have — as it were — an automatic sense of the order of the natural numbers (which would provide names for the terms in the data sequence, thus distinguishing it from the sequence of names for the terms provided by the hierarchy of levels). If such an explicit sense exists, then it needs to be introduced explicitly as a modification or extension of the account given by Brouwer of the faculties of mathematical creation — an account which we have seen to exclude the possibility of it.

If I am correct in insisting on a separate first stage in the development of a mathematical system from Brouwer's time philosophy,[6] then certain conclusions follow regarding the

[6] It was first pointed out to me by Margaret Masterman that in Brouwer's calculus of infinitely proceeding sequences as presented by Heyting [Ref.(4)], a logical distinction ought to be drawn between the numbers in the sequences used a "grid" or "reference system", on the one hand, and the use of the numbers to represent stages in the development of mathematical entities, on the other hand. In the algebraic hierarchy of the present paper there is a distinction which is analogous to this though not identical with it, for in the hierarchy (later called the ground hierarchy, § 7) in which all the mathematical possibilities are included, it is possible to represent only as it were a skeleton of mathematical meaning. More subtle concepts would have to be represented by imposing restrictions upon the possibilities allowed by the ground hierarchy. In this way the ground hierarchy plays the part of Masterman's "grid".

conventional link between Brouwerian mathematics and logic. My own summary of the situation is that the importance customarily attached by intuitionists to denying the law of the excluded third resulted in an unfortunate side-track in the development of brouwerian mathematics; successful conceptual innovations issue in special techniques, and it was felt that the Brouwerian innovation had already found its proper issue in intuitionist logic, and development of the Brouwerian mathematics into a general discipline applicable in empirical fields of an unfamiliar sort was thereby inhibited.

Before I leave the subject of the connexion of the model of s 2 with intuitionism as usually understood, I shall describe the counterpart of my model to the 'spread theorem' of Brouwer. The task of the present paper is to consider the ramifications of the total set of levels of entities constructed by specifying relations between entities existing at previously specified levels. In carrying out this task, great importance will naturally attach to circumstances in which there is a mathematically determined or automatic end to the process of the developing levels, and I shall devote the later sections of this paper to a discussion on a combinatorial basis that will show that when the greatest possible generality is given to the generation of new elements, the process does naturally come to an end — (Parker-Rhodes Theorem in Ref.(2)). This work has clearly a connexion with Brouwer's *Spread Theorem* [Ref.(3)][7]. Brouwer proves that if it is known that a number can be assigned to every sequence, within a spread of infinitely proceeding sequences, then that assignment will have to be made before any sequence can be stated. The proof of this theorem depends essentially upon the possibility of taking the *arguments* [Ref.(3), p.14] which lead to the conclusions that given sequences have numbers assigned under given conditions, themselves to have the structure of sequences in the spread.

Brouwer's proof applies to a hierarchical structure in the sense of the present paper, since Brouwer allows the possibility that the decision to consider a given sequence may generate a new sequence: indeed Brouwer deduces consequences from the fact that a connexion between statements that is established by any argument must itself be subject to the laws governing the spread calculus. Parker-Rhode's proof, also, describes a hierarchical situation, but by contrast with Brouwer's, its scope is restricted to the logically more primitive situation in which data sequences are not independent of the hierarchy that forms the subject of the present paper. Within this simpler context Parker-Rhodes' proof is stronger than that of Brouwer. It shows that a 'stop-rule' for the hierarchical generation process exists even when all possible arguments are included in it as elements, with no *protasis* concerning the assigning of numbers to such arguments.

Section 4. The Algebra of the Hierarchy : Binary Representation of Operators.

In preceding sections use has been made of an algebraic system to represent the structure provided by the total set of decisions that can be made about the progress of any Brouwer time-sequence. The concepts of level, element of level, and the hierarchy of

[7] *Ed.* Ref.3 has a 'Fan' Theorem, but does it have a 'Spread' Theorem ?

levels have been introduced. In addition it has been shown than an element of a given level determines a decision about the progress of a sequence by specifying two contiguous terms at the adjacent simpler level of the hierarchy. Now it is possible to consider each element of a given level as an operator upon the elements at the adjacent simpler level, whilst being in the relation of operand to the elements at the adjacent more complex level. With the aid of this idea I shall deduce some properties of a hierarchy from the principles that have been established earlier in this paper, and this deductive development will take the rest of the paper.

Let us consider operators p_μ in a finite family[8] $p_1, \ldots, p_n$ which define a level in a hierarchy. Without any knowledge of other levels of the hierarchy there is only one thing we can know about any element p_μ — namely whether at a given time the operation it describes is taking place or not. We represent a change by associating two operators, p_i, p_j, by a new operator P_{ij}. We call P_{ij} the *discriminator*[9] because we wish it to discriminate between the cases

(a) the operator has changed, and
(b) the operator has not changed.

We can write the two possibilities for the discriminator :-

$$P_{ij} \qquad \begin{cases} \textit{exists} & p_i = p_j \\ \textit{does not exist} & p_j \neq p_j \end{cases} \tag{4.1}$$

The discriminator relation is really the basic Brouwer "two-ity" relation [Ref.(3)] in the primitive form in which we can distinguish the present thing we are considering from the past thing, and nothing else. It cannot tell us whether p_j is equal to any other element of the family at the same time. The choice of " 0 " to represent the null effect of the discriminator, is a suitable notation because it conveys the property of an operator that when it has no effect it might equally be considered not to exist. We shall also write the symbol " 1 " in Diagram (4.1) instead of "*exists*", but at this stage 'multiplication' is not defined. hence we write :-

$$P_{ij} = \begin{cases} 0\,, & p_i = p_j \\ 1\,, & p_j \neq p_j \end{cases} \tag{4.1a}$$

let us now define a composite structure to represent the situation of more than one discrimination process taking place at the same time. We form a family[10] of discriminators,

[8] *Ed.* Here and elsewhere, where need be, the original term 'set' has been replaced by 'family' since the members of a family need not all be distinct as they must were the collection a 'set'.

[9] *Ed.* This may historically be the first occurrence of the term 'discriminator' in this context.

[10] *Ed.* See the earlier footnote about 'set' and 'family'.

and I shall write such a family as a column with single square brackets on the left[11] :-

$$\left[\begin{array}{l} (\,P_{ij}\,)_1 \\ (\,P_{ij}\,)_2 \\ \quad\vdots \\ (\,P_{ij}\,)_n \end{array}\right. \quad = \quad \left[\begin{array}{l} a \\ b \\ \vdots \\ n \end{array}\right. \tag{4.2}$$

where $a, b, \ldots, n$ on the right take values 0 or 1. Such a *column* will be referred to henceforward as a *vector*.

It is clearly possible to write such a structure formally, provided we do not allow the notation to lead us into doing things that have not been defined for constituent discriminators. In particular, in writing such a composite structure as a column we assume that the elements of the structure are <u>ordered</u> in the sense that we know which element has been put into which place (so that — for example — it is possible to compare one structure with another, element by element)[12] Now to assume ordering (in this sense and not in the sense in which the integers are 'ordered' or 'simply ordered') is to beg the whole question that is to be discussed — namely the question of what mathematical operations are involved in recalling the elements of a segment of a time-sequence[13] nevertheless we choose deliberately to use a notation that does beg this question, because it will be convenient for our later demonstration that the order in fact exists, which will shall provide by defining a sequence of recursive steps between the levels of the hierarchy.

The question of order, in the sense above defined, will accordingly play a large part in the rest of this paper.

We now wish to extend the concept of a discriminator to apply to a vector, and we define the operation $\{P_{ij}\}$ on $[P_{ij}$, by :-

$$\{P_{ij}\} \quad = \quad \begin{cases} 0 & \textit{if} \;\; a = b = c = \ldots = 0, \\ 1 & \textit{otherwise}^{\star}. \end{cases} \tag{4.3}$$

(* *i.e.* if any $a, b, c, \ldots,$ is $\neq 0$.)

This definition is appropriate to the idea of an operator, since if and only if each of the constituent operators produces no change can the combinations be said to produce no change.

We shall call the discriminator operations applied to pairs of vectors the *generalized discriminators*, and say that the discriminator has been *generalized*[14].

To determine the form of the discriminator in terms of ordinary algebraic operations we first prove

Theorem **1.** *The discriminator operation must be the symmetric difference operation.*
`proof:` Let A and B be distinct vectors, and write

$$(\,P_{AB}\,) \;\equiv\; A \,\star\, B$$

[11] *Ed.* This notation was introduced by Parker-Rhodes in his 1962 paper (see Ref.(8))
[12] *Ed.* It is this requirement that prompted the change from 'set' to 'family' as noted in an earlier footnote.
[13] *Ed.* The original paper here has *"the elements of a set in a time-sequence"*.
[14] *Ed.* These notions do not appear in the sequel.

where " $\star$ " is the operation we wish to specify. We require

$$\begin{aligned} A \star B &= 1 \quad \text{if } A \text{ and } B \text{ are different, so that} \\ A \star B &= B \star A, \quad \text{and also} \\ A \star A &= B \star B = 0 \end{aligned} \tag{4.4}$$

We divide the problem into two parts :-
(i) We first define a binary operation $\wedge$, which makes a vector correspond to any pair of vectors A and B. This vector (V) will be said to have a certain form *designated* if and only if A and B are identical.
(ii) Then we define a function, f, in such a way that $f(V) = 0$ if and only if V is designated.

For part (i) we have to consider a particularly simple binary operation (a 'separable' one[15]) such that if $A = [\,a_n$, and $B = [\mathsf{b_n}$, then $A \wedge B = [\,c_n$ where $c_n = a_n \wedge b_n$.

The possible tables for $\wedge$ are of the form

$\wedge$	0	1
0	x	y
1	w	z

From our assumptions, $x = z$, $y = w$, and $x \neq y$. Hence possible tables are

$\wedge_{(a)}$	0	1
0	0	1
1	1	0

$\wedge_{(b)}$	0	1
0	1	0
1	0	1

For part (ii) we consider first (a) (the symmetric difference, or simply " + " in the field $\mathbf{J}_2$. the designated vector is then the null vector $(0, 0, \ldots, 0)$, and if $V = [\,V_n$ we can define

$$f(V) \quad = \quad \begin{cases} 0, & \text{if every } V_n = 0, \\ 1, & \text{otherwise.} \end{cases}$$

If we choose (b), which is related to " + " in $\mathbf{J}_2$ by

$$x \wedge_{(b)} y \;=\; x + y + 1\,,$$

then the designated vector is $(1, 1, \ldots, 1)$, and so

$$f(V) \quad = \quad \begin{cases} 0, & \text{if every } V_n = 1, \\ 1, & \text{otherwise.} \end{cases}$$

15 *Ed. I.e.* one that acts 'component-wise' on the components of its vector operands.

The operation required is therefore either the symmetric difference or its complement. Between these two possibilities we are in a strict sense free to choose[16] without loss of generality (though we must stick to the choice we make). But if we choose the complement we do so at the expense of intuitive meanings of 0 and 1, since we should expect the effect of discriminating between two vectors to be zero if the vectors differ. Accordingly we choose symmetric difference. *This establishes the theorem.*

Section 5. Mappings.

I shall now establish the recursive relation between levels that will enable us to order the elements at a given level in the hierarchy (in the sense of "order" defined in § 4).

By means of the discriminator relation we can derive new vectors from ones which have already been found. But any pair of vectors corresponds to a third vector, and in a family of n vectors repeated use of the discriminator operation will produce in general

$$n + \tfrac{1}{2}n(n-1) + \cdots \;=\; 2^n - 1 \tag{5.1}$$

vectors, unless it happens that the process throws up vectors which have already occurred before. If the full number of vectors is produced by this process, we can call the original family *linearly independent* (this definition agrees with the usual definition for the field of two elements when multiplication has been defined, but is appropriate for the present stage of development of the theory before a second operation has been introduced). If the two vectors have m elements, a necessary condition for linear independence is clearly

$$n \leq m\,.$$

To establish recursion we prove

Theorem **2.** *When a family of vectors is ordered in the sense used in § 4, the larger family derived from it by repeated use of the discriminator operation can also be regarded as ordered.* We establish some preliminary results before giving the proof. The discriminator relation corresponds roughly to the idea of "next" : if a family of vectors is ordered, then there must also be a mathematical representation of a "memory" which is capable of deciding whether the association of vectors defined by the discriminator operation is, or is not, in accordance with the sequence as it exists in the memory. This idea requires a store of pairwise associations of vectors. Such an association can be defined by an expression

$$V \;=\; f(\,U\,)\,,$$

where a functional operator f defines a unique vector V for each vector U. The elements of U and V belong to $\{\,0,1\,\}$.

Lemma *The functional operation f is a matrix transformation.*
`proof:` We know from a result of Boole that if we define a new binary operation between[17] the $0,1-$elements u, v of the vectors U, V, which it is convenient to denote by multiplication as uv, and which is defined by the table

---

[16] *Ed.* This concept of choice led eventually to other notions of what Pierre Noyes in 1987 called 'Amson Invariance'.

[17] *Ed.* This sentence is a clearer re-wording of the original.

uv	0	1
0	0	0
1	0	1

(5.2)

then, with this operation and the original one, any function can be expressed. For example, if U has two elements u_1, u_2, then a general function of U will have elements of the form

$$a + bu_1 + cu_2 + du_1u_2 \tag{5.3}$$

Thus Boole's argument shows that (5.3) is the most general form fro the mapping. For the kind of algebra we are constructing, however, it is still too general, for we introduced the two kinds of terms — $a, b, \ldots$ and $u, v, \ldots$ — so that the two basic type of operation could be represented (discrimination and mapping), and it is therefore unnecessary to construct the quantity du_1u_2 : the mapping relation is represented between terms of the different type.

At this stage almost any other binary operation would do for the mapping; the only exception being the one which is complementary to addition (+), since this gives us essentially nothing more than we have already assumed when we take addition as the first operation. I shall discuss the choice of operation.

The next problem is to ensure that the operation that defined the mapping is consistent with the discriminator operation already considered in the sense that the "non-linear" term in (5.3) will not reappear even if the vectors to be mapped are considered as composite arrays. (An obvious necessary condition since vectors are meant to be considered as composite arrays.) For this purpose it is convenient to change the notation for a moment, and to write

$$D(U, V) \tag{5.4}$$

for the discriminator of two vectors U, V, and to write $\bar{U}$ for the vector corresponding to U in the mapping association. The condition of consistency can then be written as

$$\overline{D(U, V)} \;=\; D(\overline{U}, \overline{V}),$$

and when we translate this back into the notation with which we are familiar we get

$$f(U + V) \;=\; f(U) \;+\; f(V). \tag{5.5}$$

Equation (5.5) is simply the familiar condition that f should be a linear[18] function, and we may notice that our development gets us into conventional algebra by developing the concept of linearity from the more primitive notion of an operation that is capable of representing a mapping (of levels).

[18] *Ed.* Strictly speaking, an 'additive' function, but over the field of two elements, 'additive' and 'linear' are identical notions.

It is now necessary to consider the form of the second binary operation. When we have rejected the choices that would give different results depending on which vector in the mapping of a pair of vectors we chose first, we are left with only one operation as a serious contender competing with "multiplication", (5.2). This is the set union

$u \cup v$	0	1
0	0	1
1	1	1

.

We reject the possibility because it gives us an algebra of mapping operations in which (as can easily be shown) there exists a zero element in the set of mappings but no unit element. Such an algebra could not be reduced to the simple form in which all the vectors represent binary situations.

Given the choice of the second operation, the condition of linearity confines us in fact to a matrix formalism, as we can see at once in the following way. If we choose a basis for the vectors so that any vector can be written in term of its components by

$$U \;=\; \sum_i u_i\, e_i\,, \tag{5.6}$$

where the basis elements are

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \qquad etc.$$

then we have at once that

$$V \;=\; f(\,U\,) \;=\; f\left(\sum_i u_i\, e_i\right) \;=\; \sum_i u_i\, f(e_i). \tag{5.7}$$

However, in this equation the functional operations on the basis elements give us vectors which must be expressible in terms of the basis (since any vector is so expressible). Accordingly we write

$$f(\,e_i\,) \;=\; \sum_j f_{ji}\, e_j\,, \tag{5.8}$$

and this at once gives us

$$v_i \;=\; (\,f(u_i)\,) \;=\; \sum_j f_{ij}\, u_j\,,$$

which is simply an expression of the matrix algebra. *This proves our lemma.*

Theorem 2 follows directly from the lemma : each matrix transform of vectors X_1 and X_2 orders the pair (X_1, X_2), and if there are enough matrices to order all the pairs (X_i, X_j), then a 1 : 1 relation is established between the elements of the larger family and the matrices. Since the smaller family is ordered, the order of the larger family follows provided that there not more pairs than there are matrices — *i.e.* provided that

$$2^n \leq m^2 ,$$

where m is the dimension of the space. *This proves Theorem 2.*

In the following table these results are applied over and over again in successive stages of a hierarchy which begins with two linearly independent vectors each with two elements.

m	n	dim(f)	$2^n - 1$
2	2	4	3
4	3	16	7
16	7	256	127
256	127	$(256)^2$	$2^{127} - 1$

(5.10)

Section 6. Level Recursion.

It is now possible to specify some of the mathematical properties of a hierarchy of levels from the results established in Theorem 2. This we do by a recursive method, since to order the operations at a given level it is necessary to construct the contiguous simpler level, and to assume ordering of its elements. This in turn requires us to consider the level simpler than that, and so on.

We can see on general grounds and in an intuitive way that at a given level in a hierarchy the effective number of operators is not the number in the level in question; rather it is the number at that level together with the numbers at all simpler levels that is to be taken as the effective number. This we shall call the *multiplicity* of the hierarchy at a given level. The intuitive reason is that the hierarchy is a structure which occupies time, and therefore operators at different levels (which would be kept strictly incomparable by type distinctions in the class logic) can be taken together. One cannot say that if there are a horse and a cow in a field then there are three things in the field : the horse, the cow, and the (horse and cow). In a hierarchy one could (by analogy) say, however, that there was the perceiving cow event, the perceiving horse event, and the perceiving (horse and cow) event. More formally, the summation of numbers of operations can be understood if we consider that every individual process represented in the hierarchy is a binary one in which a mapping of two operations is represented by an operation at a different level. If many operations at one level are to be considered then they have to be brought under consideration by these binary processes which will in general involve a tree of operations at all the new simpler levels.

In applications of the theory of the present paper the multiplicities of hierarchies will be of considerable importance.

A recursive definition of the multiplicity, $\Lambda\,\lambda$, of level λ of a hierarchy (*i.e.* the total number of operations up to level λ) can be expressed by the relation :-

$$\Lambda\,\lambda \;=\; (\,2^n - 1\,) + \Lambda(\,\lambda - 1\,)\,,$$

where n is the increase in the total provided at level $\lambda - 1$, in accordance with Theorem 2 and on the assumption that the largest family of operators that can be ordered is also the largest family of operators that can be enumerated by any means whatever at that level.

A hierarchy has now been defined in terms of a recursive rule. We now wish to apply the rule successively to find the form that emerges for hierarchies. There is clearly a hierarchy of specially simple form[19] in which the lowest level is based on a vector of two components; successive levels then have basis vectors of

$$2^2, \qquad \left(2^2\right)^2, \quad \ldots$$

components. We shall call this the *ground hierarchy.* This name derives from the fact that more special hierarchies with distinguishing features can all be obtained by imposing restrictions on the ground hierarchy. to put it another way, the ground hierarchy is the hierarchy from which all distinguishing features have been removed. In the present paper we restrict our attention to the ground hierarchy. These properties of the ground hierarchy follow from a theorem to be proved below, but first we illustrate the first stages in the construction of the ground hierarchy.

We take a bottom level consisting of two operators. According to § 1 these are to be written 0, 1. And

$$\lambda_1 \;=\; \lambda_1(\,0,\,1\,)\,.$$

The level λ_2 of operators upon λ_1 is now constructed, according to our recursive procedure. From these we construct a further level λ_3 of $2^3 - 1$ operations and obtain the total number of operations Λ_2 at this level by adding the increments at the levels hitherto generated. This gives

$$\Lambda_2 \;=\; 3 + 7\,,$$

By continuing in this way we generate the following table :-

Level	Number of elements of array	Increment in number of operators at level	Total number of operators
1	$2^1 = 2$	3	3
2	$(\,2\,)^2 = 4$	7	10
3	$(\,(\,2\,)^2\,)^2 = 16$	127	137
4	$(\,(\,(\,2\,)^2\,)^2\,)^2 = 256$	$\sim 10^{38.2}$	$\sim 10^{38.2} + 137$

[19] If the lowest level has one element, no non-trivial hierarchy results. The level of operations upon this also contains one operation, and is formally indistinguishable from the first. And so on.

The table immediately provides the following important

Theorem **3.** **Parker-Rhodes' Theorem**

The ground hierarchy has only 4 levels.

`Remark:` This result follows at once from the fact that at level 4 there are more operations than the dimension number of the vector space at the new level. Therefore the construction terminates.

Parker-Rhodes' Theorem shows that the *protasis* of Brouwer's Spread Theorem is satisfied by the hierarchical algebra that has been developed in this paper. That is to say, in Brouwer's proof that we can "assign a number" to each sequence of operations defined in the hierarchical algebra. Brouwer's theorem, with its implications for ordering the continuum, follows, and we have shown that the program suggested by Bastin and Kilmister [see Ref.(7)] of developing a continuum suitable for physical theory from a Brouwer space is logically feasible. The original program lacked a suitable form fro the Brouwer spread, and this has been provided by the hierarchy algebra of the present paper. As, however, the hierarchy algebra is (in ways fully discussed in the opening sections of this paper) more general than the spreads contemplated by Brouwer, it has been necessary to show that sequences in the hierarchy can be treated as forming a spread for the purposes of the spread theorem.

In an earlier form of discussion of the hierarchy, Parker-Rhodes [Refs. (1) & (2)] used the device of mapping vectors at one level onto points in a vector space at a higher level, as the first method of hierarchy construction. This device led to a somewhat involved calculation of the hierarchy multiplicities which we have calculated in the present paper without any reference to a special choice of mapping since we have made no assumptions about the relation of the m-vectors to the vector space at the next most complex level.

Naturally, further applications of the hierarchy mathematics will make use of special forms of mapping, and in particular it will become important to consider mappings defined by eigenvalues as vectors at one level that are mapped onto the vector space at the higher level. this was the form originally used by Parker-Rhodes to define the hierarchy, but in fact the multiplicities that have been calculated do not depend on any such special forms.

ooOOooOOooOOooOOooOOooOOooOOooOOooOOooOOooOOooOOooOOooOO

REFERENCES

1. U.S. Air Force Office of Scientific Research, Annual Reports, Information Structures Project, 1960,61,62,63, Technical Reports I,II,III;
in particular, Report III–part II (by Ted Bastin and K.Sparck Jones)[20], published as a paper read (on May 15th, 1961) by E.W.Bastin to the British Society for the Philosophy of Science, entitled *Self-Organising Mechanisms as Models for Scientific Theories.*

2. Parker-Rhodes, A.F. [1964], An Algebraic Hierarchy.
Unpublished[21]

3. Brouwer, L.E.J. [1954], Points and spaces.
Canadian Journal of Mathematics, **6**, p.1–17.

4. Heyting, A. [1956], Intuitionism. An Introduction.
North-Holland Publishing Company, Amsterdam.

5. Bastin, E.W. and Kilmister, C.W. [1964],
U.S. Air Force Office of Scientific Research, Annual Report, Information Structures Project, 1964, p.28 *et seq.*

6. Brouwer, L.E.J. [1955], The Effect of Intuitionism on Classical Algebra of Logic.
Proceeding, Royal Irish Academy, Section A, **57**, p.113-116.

7. Bastin, E.W. and Kilmister, C.W. [1955], The Concept of Order, II.Measurements
Proceedings, Cambridge Philosophical Society, **51**, p.454–468.

ADDITIONAL REFERENCES (Supplied by John Amson **1998**)

8. Parker-Rhodes, A.F. [1964], Hierarchies of Descriptive Levels in Physical Theories.
Unpublished[22]

9. Van Dalen, D. (Editor) [1981], Brouwer's Cambridge Lectures on Intuitionism.
C.U.P. Cambridge.

---

[20] *Ed.* John Amson has a copy, which he can make available.
[21] *Ed.* This paper, completely rewritten in collaboration with John Amson at the time under the title *Essentially Finite Chains*, eventually appeared in *Int.J.General Systems* 1998 **27**, Nos.1-3, p.81-92.
[22] *Ed.* This paper, resurrected, edited and annotated by John Amson, appeared, posthumously, under the same title, in *Int.J.General Systems* 1998 **27**, Nos.1-3, p.57-80.