On the Origin of the Scale Constants of Physics

Abstract

Four dimensionless constants are calculated by reducing dynamics in a particle assemblage to all-or-none interactions. The resulting structure contains no space or time continuum, and therefore determines the scale of microscopic in terms of cosmological phenomena independently of special assumptions about the space-time continuum. Definitions of mass and momentum are derived from the theory without appeal to the classical continuous concepts, and these definitions are shown to fit important existing developments in high energy physics.

1. Introduction

In this paper I put forward an approach to the problem of describing a particle in a field without assuming the space-time continuum. I deduce as much as possible from very simple assumptions concerning interactions between the elements of a « bootstrap » type assemblage in which each particle in the assemblage is built out of the interactions of all the others. Interaction either exists (a situation denoted by the digit 1) or else it does not exist (denoted by the digit 0) and there is no other possibility. No dynamical properties are assumed for the particles beyond the discrete, all-or-none interactions, and dynamics, therefore, — including the momentum concept — has to be built later. The theory being proposed differs vitally in this respect from the bootstrap theories that are based on the S-matrix technique.

It is now becoming increasingly widely recognised that fundamental difficulties exist in the application of the space and time concepts to high energy physics, and a great advantage of the very simplified approach that is being proposed is that it contains no continuous dynamics and therefore makes no appeal to conventional space and time. Accordingly there is no reason why its conclusions should apply at one particular magnitude rather than at any other. In fact we find numerical values which we identify as measures of the strengths of the main fields of physics that interact with particles. Because these values are calculated at a simpler stage in the theory than that in which continuous dynamical variables can be defined we are forced to suppose that they specify dimensionless ratios of the natural units (or fundamental constants) which are ultimately required to specify every measurement, and therefore every particular value of each continous dynamical variable, however that concept has later to be defined. This supposition is, of course, in accordance with the definition of the strengths of fields in terms of coupling constants which are currently written as dimensionless ratios of fundamental constants. In this paper the numbers which come to be identified with these dimensionless ratios will be called *scale-constants* because of their status in measurement.

I have therefore three tasks. Firstly to establish a mathematics to describe the discrete interactions ; secondly, to show that with such a theory results can be obtained which cannot be obtained without the simplification of current theory to a theory based on discrete interactions ; and thirdly, to show the lines on which dynamics can be developed from the theory of discrete interactions.

Because of its highly inter-disciplinary character the work described in this paper is actually the result of close collaboration. The vital mathematical method of relating levels in a discriminatory structure by using matrix transforms over the cyclic field of order 2 as the elements of a new level is due to A. F. Parker-Rhodes. The finiteness theorem (Theorem 3, § 3) is part of Parker-Rhodes's mathematics. J. C. Amson helped formalise the concept of discrimination and of discriminate closure as a joint author of an earlier draft of the paper. C. W. Kilmister extended this formalisation and proved theorem I § 3.
The Research reported in this paper has been sponsored by the Information Research Division : through the European Office of The United States Air Office of Scientific Research. Grant N°: AF EOAR 65-78.

2. General Physical Picture

The theory to be presented in this paper depends upon a physical picture which is described in the following set of principles.

1. There exists a set of elements with no spatial or temporal relationships defined initially among them.

2. Any element of the set can be chosen as the *initial* element from which interaction processes are considered.

3. The only kind of interaction (including signalling) that is defined between the elements of the set is *discrimination.* An element can be brought into consideration together with the initial element, and a process of discrimination takes place as a result of which the former element is recognised as dictinct from the initial element or else as not distinct from it.

4. The result of a discrimination process is itself a new element that can be added to the original set.

5. The total set can be expanded or contracted in order to incorporate the newly generated elements and so a structure results in which discrimination can take place in successive stages of the original simple type.

6. A total set of any given complexity that has been constructed by the discrimination process must be capable of acting as a set of individuals. For this to be possible a new *mapping* operation has to be introduced to replace the discrimination in successive stages.

7. It is logically necessary to postulate a *base-function* which causes elements to be brought into consideration together and which decides whether the original discrimination operation or the mapping operation is brought into action. This base-function will be responsible for the *contingent* features of distribution of matter in the world, while the *fundamental* features are those which can be deduced whatever the behaviour of the base-function. The base-function is like the ψ-function of quantum theory in that it is not directly observable.

8. For an observation to be possible there must be at least one localised particle (it must be remembered that the discrimination processes are not defined at a particular point in space or time).

9. When mappings (6. above) are invariant under choice of initial element (2. above), a *physically localised particle* results. Localised particles may be stable if the invariance of the mapping is independent of the base-field, or may have differing degrees of stability in cases where this independence is only partial. Degree of stability is thus a statistical concept having complete stability as a limiting case where a mathematically specified number is given whatever the behaviour of the base-function.

10. Fundamental numerical properties of particles arise as the cardinals of the sets of the total numbers of discrimination processes that are required physically to execute these mapping processes.

11. Mass, being the scalar quantity (other than the quantum numbers) that is associated with a particle, is the quantity we shall associate with the numbers described in 7. above. The physical idea is that the more activity a particale represents the more mass it has.

12. A number must be associated with a *ratio* of two masses, so that it is necessary to postulate a unit mass as well as the number itself. In fact

this unit mass would be an unobservable quantity, and we can work instead with the current conception of a coupling constant that specifies the strength of interaction of a particle with a given field, since any measurement of the mass of a particle must unltimately depend on an observation of acceleration of the particle (or some other particle) in a field. In this way a given field coupling can serve to compare masses of particles and hence serve as a unit of mass. In the present paper the scale constants of fields are the only actual calculations made.

13. Momentum — the first spatial or temporal concept to appear — is defined as an ordered sequence of measurements of particle masses.

3. Formalisation of a Discrimination System

Definition 1. The distinct (i.e. non-identical) elements of the discrimination system are written 0,1. These will be called *discriminators.*

Definition 2. Discriminators may be written in a *column* where d_1 — d_n

$$\begin{matrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{matrix}$$

are discriminators. Such a column will be written c_n. The d_j will be called the *components* of the column.

Definition 3. A binary operation B on columns acts on any two columns, and (a) gives different results when the columns are alike from what it gives when the columns are unlike, and (b) the result of the operation B is again a column.

If the columns are of different lengths, then the shorter is made equal in length to the longer by adding discriminators having value O. The operation B is *symmetric* : i.e., it is irrelevant which column is introduced first. If c_1, c_2, c_3 are columns, then we write $B(c_1, c_2) = c_3$.

Definition 4. A unary operation C on columns acts on any one column and the resultant of the operation C is a discriminator. Thus if c is a column, then we must write either $C(c) = O$ or $C(c) = 1$.

Definition 5. A column c such that $Buu = c$ for some column u is said to be *designated.*

Definition 6. If an operation B on n-columns has the property that each element of the resultant n-column is determined by the corresponding elements of the initial n-columns, we shall say that B treats the columns *elementwise.*

This definition of columns has to fit the physical picture of § 2 according to which the elements of an interaction system carry information which can either be used to distinguish them as elements of an existing set, or can be broken down into simpler information structures by a (stage by stage) process which explains how the information is built into the structure. It is as though each element can either be accepted as an individual straight away or else can be challenged to establish that individuality. None of the conclusions we reach in this paper depend upon the particular sequence in which these different operations take place (we postulated a base-function to settle this order) and the definition we give of column is an attempt to specify these requirements minimally. The idea of the columns having different lengths is one we avoid introducing because it requires a method of detecting sameness or difference that we are concerned to formalise. However we need the idea that two columns with different amounts of information can be compared, in order that we may break an element into sub-elements which shall be comparable with the originals. To do this we need an idea of pairing off columns in a well-defined way and the method of adding « neutral » zeroes does this. Of course we are putting something into the mathematics here which will later have to be brought to light and made explicit.

Theorem 1. The set of columns that can be generated by operation B is isomorphic to a set of n columns treated elementwise.

Lemma: The operation B acting on columns of 1 component must be the operation of symmetric difference (defined by the multiplication table

	0	1
0	0	1
1	1	0

Proof: By inspection the operation must have the form

	0	1
0	a	b
1	b	a

$(a \neq b)$.

From the duality of the elements 0,1, we may chose $a = 0$, $b = 1$ or $a = 1$, $b = 0$ without loss of generality. In order to make the choice conform with the use of 0 in extending columns (definition 3) we make the former choice.

Proof of Theorem. I. *Establishment of Group Property*
From definition 3 (symmetric discrimination) we have

$$\text{Buv} = \text{Bvu} \qquad (1)$$
$$\text{B(Buv)w} = \text{B B(vw)u} \qquad (2)$$

Here (2) is the extension of 1 to 3 terms. The brackets are actually unnecessary, and one can write BuBvw for Bu(Bvw). This formation has to be equivalent to first forming Buv and then BwBuv, etc. Hence all discriminators of 3 elements are identical. The process can now be continued to higher numbers of components.

Part I of the theorem now falls into three parts:

(i) There is only one designated column e.
Proof. Let Buu = c   Bvv = d
Suppose c $\neq$ d
Then Bcd = B(Buu)(Bvv) = B(Buv)(Buv).
Hence Bcd is designated, and therefore

$$\underline{c = d.}$$

(ii) e is the unit element of a group under the discrimination operation.
Proof. B(Bev)v = BeBvv = Bee = e
Hence B(Bev) is designated.
But Bev = v = Bve by symmetry, which specifies e as the unit element, and we may use S to denote the group.
(iii) Every element of S is of order 2.
Proof. Buu = e for all u ($u^2 = e$).
Hence $S = C_2 \times C_2 \times C_2 \ldots$, where C_2 is the cyclic group of order 2.

II. *Establishment of Isomorphism.*

We have to show that if we give one element of each factor then we get one element of the discrimination system.

Proof. The isomorphism of C_2 with the set (0,1) under symmetric difference

x	1	a
1	1	a
a	a	1

→

	0	1
0	1	0
1	0	1

is $\begin{bmatrix} 1 \rightarrow 0 \\ a \rightarrow 1 \end{bmatrix}$.

We show the ismorphism of $C_2 \times C_2$ with the set of columns $\left\{\begin{bmatrix} a \\ b \end{bmatrix}\right\}$ where $\begin{matrix} a \in \\ b \in \end{matrix} \begin{Bmatrix} 0,1 \\ 0,1 \end{Bmatrix}$ and where curly brackets mean « the set of ».
We have $C_2 \times C_2 = \{a \times a'\}$ where $a \in C_2$

$$a' \in C_2$$

Then, if we juxtapose symbols for the group operation in $C_2 \times C_2$ (the quadratic group) then we have

$$(a \times a') \times (b \times b') = (ab) \times (a'b'),$$

showing that the structure of the discrimination system is represented by $C_2 \times C_2$ if the elements are grouped in pairs under the isomorphism.

$$\begin{aligned} 1 \times 1 &\rightarrow \left[\begin{smallmatrix}0\\0\end{smallmatrix}\right] \\ 1 \times a' &\rightarrow \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \\ a \times 1 &\rightarrow \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] \\ a \times a' &\rightarrow \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] \end{aligned}$$

qed.

This proof can be directly extended to columns of any length.

Mappings

We first define a *discrimination system* to be a set of discriminable columns under the operation B of Theorem 1.

We shall now define a mapping of a discrimination system S consisting of n-columns onto itself

$$S \rightarrow S_1$$

as a correspondence which assigns to each element of S some other one element a' = Øa, such that

1) the discrimination operation B is preserved by the mapping, so that for any two columns a,b $B(a'b') = (Bab)'$.

2) Only the null column is mapped into the null column.

Theorem 2. The set of mappings of S into itself consists of non-singular square matrices aver the field of two elements.

Proof. Condition (1) on the mapping Ø implies that it acts linearly (if the operation B is interpreted as addition), and all linear operations on n-columns are provided by the $n \times n$ matrices. The condition that only the null vector is mapped into the null vector then requires the matrices to be non-singular. qed.

Theorem 2 shows that the set of mappings is a discrimination system taken from the set of n^2-columns. This is to say that there exists a discrimination operation between the mappings which can be defined by setting up an isomorphism between the $n \times n$ arrays defining the matrices and the set of n^2 columns. We now need to define

Discriminate Closure. Definition : Let S be a discrimination system, with binary operation B. Then a proper subset $S' \subset S$ is said to be discriminately closed under B if and only if for all x, y in S' we have

$$x \neq y \leftrightarrow B(xy) \in S'.$$

That is to say that the discriminately closed subset is the subset containing all the resultants of successful discrimination operations. The null column is not a member of the discriminately closed subset. Using this definition we can now define:

Linear independence. Definition: A discrimination system S will be said to be linearly independent if *no* element w exists such that

$$w \in \begin{bmatrix}\text{the discriminately closed subset} \\ \text{generated by } S - (w)\end{bmatrix}.$$

S will otherwise be called linearly dependent.

The object of the whole of the foregoing formalisation in this section is to be able to construct a *hierarchy* of discrimination systems in which the possibilities of discrimination available in one system are used to construct a more complex discrimination system, and so on. The different discrimination systems in the hierarchy will be called the *levels* of the hierarchy, and the more complex levels will be said to be lower, and the less complex, higher.

Accordingly, we assume the existence of a level S consisting of n-columns. Then at the next level we define a discrimination system whose elements correspond to certain subsets which are discriminately closed; these subsets being treated as single elements of the new discrimination system. It is then convenient to represent the single elements at the new level by $n \times n$ matrices, this being possible by Theorem 2, and in regarding the new level as a new discrimination system the matrices in turn can be written as n^2-columns or n^2-vectors, which can be discriminated in the usual way.

The detailed stages in the construction of a hierarchy are accordingly:

a) Let S be a linearly independent set of r elements
b) Let S_1 be any non-null subset of S
c) Let $\bar{S}_1$ be the discriminate closure of S_1
d) Let **B** be the non-singular mapping whose set of invariant elements is $\bar{S}_1$
e) Choose $\varnothing_1, \varnothing_2, \ldots \varnothing_N$ (where $N = 2^r - 1$)

to be linearly independent to constitute the next stage. Regard these as r^2-columns.

Notes on the construction

1. Linear independence is essential, since we have to write

$$S = \bar{S}_1 \cup (S - \bar{S}_1)$$

and since $\varnothing_1 a = a$ for *every* a in $\bar{S}_1$ and every $\varnothing_1$.

$\varnothing_1 a \neq a$ for *every* a not in $\bar{S}_1$ and every $\varnothing_1$

and so if w is not in S_1, it must not be in $\bar{S}_1$.

2. Non singular mapping ensures that the zero column is never mapped onto any column and thus preserves discriminate closure as distinct from simple closure.

3. Hierarchy construction can be carried on using the concepts of discriminate closure and mapping without the use of linear algebra. This is, in general, a difficult thing to do, but to illustrate the principle the two methods are carried out side by side for a simple case to show how their equivalence works.

Matrix Method for 2-Vectors

There are three non-null vectors, $\begin{smallmatrix}0\\1\end{smallmatrix}$, $\begin{smallmatrix}1\\0\end{smallmatrix}$, $\begin{smallmatrix}1\\1\end{smallmatrix}$. The d.c. subsets are

(a) $\begin{smallmatrix}0\\1\end{smallmatrix}$

(b) $\begin{smallmatrix}1\\0\end{smallmatrix}$

(c) $\begin{smallmatrix}1\\1\end{smallmatrix}$

(d) $\begin{smallmatrix}0\\1\end{smallmatrix}$, $\begin{smallmatrix}1\\0\end{smallmatrix}$, $\begin{smallmatrix}1\\1\end{smallmatrix}$

However, the prescription requires S to be linearly independent, so we must take two of the three vectors, say $\begin{smallmatrix}1\\0\end{smallmatrix}$, $\begin{smallmatrix}0\\1\end{smallmatrix}$ and then (c) is omitted.

We have to find non-linear mappings with a), a), d) as their invariant subsets.

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}b\\d\end{pmatrix}$$

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}a\\c\end{pmatrix}$$

so for (a) b = 0, d = 1 and if non-singular a = 1, and if $\begin{pmatrix}1\\0\end{pmatrix}$ is *not* invariant c = 1 giving $\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$.

For (b) a = 1, c = 0, so d = 1 (non-singular) and b = 1 : $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$

For (c) evidently the unit matrix is the only possibility. These 3 are linearly independent.

Mapping Isomorphism Method for 2-columns

There are 3 non-null vectors, a, b, a + b. A linearly independent set is a,b, The d.c. subsets are

a) a,

b) b

c) a, b, a + b

We have to find permutations leaving just these invariant.

For (a)

x	a	b	c = a+b
↓	↓	↓	↓
Øx	a	≠b	≠c

which involves

x	a	b	c
↓	↓	↓	↓
Øx	a	c	b

Similarly for (b):

x	a	b	c
↓	↓	↓	↓
Øx	c	b	a

The remaining Ø is the identity permutation, These are linearly independent.

Two further remarks on the equivalence of the mapping isomorphism and the matrix algebra may also be added.

a) The argument is not so easy for 3 4-vectors, partly because of greater numbers but mainly for the non-uniqueness, e.g., in finding the Ø for the d.c. subset {a, b, a + b} the only condition on the Ø of c, a + c, b + c, a + b + c and all these vectors involving the 4th basis vector d is that they are all changed by ∅. When there were only two left over (above) this was easy.

b) More significantly, we observe that *any* ∅ operating on a set of vectors defines in a *single* operation a whole lot of cycles. There is no need for repeated operations, though that is a convenient way of doing the computing. For example, take 3 vectors ; :

$$\mathrm{v} = \mathrm{a},\ \mathrm{b},\quad \mathrm{c},\ \mathrm{b}+\mathrm{c},\quad \mathrm{c}+\mathrm{a},\ \mathrm{a}+\mathrm{b},\ \mathrm{a}+\mathrm{b}+\mathrm{c}$$

$$\downarrow \varnothing(\mathrm{v}) = \mathrm{a},\ \mathrm{c}+\mathrm{a},\ \mathrm{b},\ \mathrm{a}+\mathrm{b}+\mathrm{c},\ \mathrm{a}+\mathrm{b},\ \mathrm{c}\quad ,\mathrm{b}+\mathrm{c} \qquad \text{say}$$

Here there is (i) a one-element invariant d.c. subset $\{a\}$,
(ii) a 4-cycle $c \to b \to c + a \to a + b \to c$,
(iii) a 2-cycle $b + c \to a + b + c \to b + c$.

We now have a discrimination theory in which the possibilities of discrimination may be increased from a set of n-vectors to a set of n^2-vectors given that the first set is ordered so as to be in effect labelled with the integers $1 \to n$. This assumption was explained on p. 81, where the successor relation of vectors at a given level was assumed if and only if a scheme of higher levels existed to reduce the successor relation to discrimination. So it was left without explanation (definition 2, p. 80) that a column of discriminators could be so written, which is equivalent to assuming that a column could be labelled with an integer. (A natural way to do this would be to label it with the binary integer defined by the components ; this could then be replaced if desired by the corresponding denary integer.) We shall escape from the difficulty of defining order on any level by working recursively. That is to say we know we can have a level K ordered if we can assume that the next higher level J is ordered. But we know (definition 1, p. 80) that the level of 2-columns is ordered. Hence we can define order in 4-columns, 16-columns, etc.

A few definitions will prove convenient :

df. We write (m) for 22m (m = 0, 1...)

df. We write $M_{(m)}$ for the set of all (m)×(m) matrices

df. and $K_{(m)}$ for the set of all (m)-vectors.

The effect of Theorem 2 is now that we can always represent any matrix in $M_{(m)}$ as an $(m + 1)$ -vector, since $(m)^2 = (2^{2m})^2 = 2^{2m.2} = 2^{\overline{2m-1}}$

df. We speak of « the level m » (m — 1, etc)

Moreover, if we replace vectors by columns in the discrimination system, then the effect of Theorem 1 becomes that $K_{(m)}$ is isomorphic to $M_{(m-1)}$.

We further describe the hierarchy for which all the above concepts hold because order is recursively defined, as the base hierarchy. Hence *df.* The *base hierarchy* is the hierarchy with vectors of length 2 at the first level.

We have now the important theorem due to A. F. Parker-Rhodes.

Theorem 3. In the base hierarchy the construction terminates with the level of $(256)^2$ vectors.

Proof. The base hiererchy has values of K_m

$$2, 4, 16, 256, (256)^2 \ldots$$

and into these are mapped discrimination systems of cardinal

$$3, 7, 127, 2^{127}.$$

This mapping cannot be carried out for the last case because of the inequality

$$2^{127} > (256)^2$$

and hence this mapping cannot constitute the construction of a new level.

The generality of Theorem 3 is now shown by

Theorem 4. The possibilities of discrimination afforded by the base hierarchy form an upper bound to those defined by any hierarchy.

Proof. Suppose a given hierarchy has a discrimination system of columns of length p at some level. Then for all p one can choose a value of K_m such that

$$K_m > p,$$

and such that there is not a K'_m such that

$$K_m > K'_m \geqslant p.$$

Then one can increase each of the p-columns by K_m-p zeroes, without affecting the possibilities of discrimination. The proof of Theorem 3 then applies, and hence the proof of Theorem 4 follows.

We must make a distinction between two cardinals, namely

1) The cardinal $K_{(m)}-1$ of the set on n-columns at a given level, i.e. the cardinal of the largest set of things that can be discriminated at a given level m, and

2) The cardinal of the set of discriminations necessary to order the discriminable things at level m. Call this $Q_{(m)}$.

From the recursion principle, it follows that $Q_{(m)}$ consists of the discriminations possible at all higher levels, and that without using the discriminations at the higher levels it will not be possible to order the elements of m. Hence

$$Q_{(m)} = \sum^{m} (K_{(m)}-1).$$

The importance of $Q_{(m)}$ is that it is the only cardinal that (a) is characteristic of the level m, and that (b) can be constructed using the discrimination *process*. *df.* $Q_{(m)}$ will be called the *multiplicity* of m. The quantities m, (m), $K_{(m)}$-1, $Q_{(m)}$ are given in the following table.

m	(m)	$K_{(m)}-1$	$Q_{(m)}$
1	2	3	3
2	4	7	10
3	16	127	137
4	256	$10^{38.2}$	$10^{38.2}$

4. What the Mathematics Shows

It is now possible to add some conditions regarding the shape of our theory of interactions to the physical picture given in § 2 by considering the mathematics of § 3. I shall not be concerned with numerical results at the moment. These will appear in § 5.

1. The numbers of discriminable elements in the interaction system increase very sharply in 4 stages, after wihch increase is impossible (see the table at the end of § 3).

2. The stages or levels are defined by the number of successive mappings that can exist.

3 In accordance with the principles of § 2, which are developed further in the course of establishing comparison with experiment in § 5, the number of discriminable elements in the interaction system up to a given level (i.e. the multiplicity of that level) determines the unit of mass measurement (and hence the coupling constant of the field associated with that level). This principle of identification, taken together with the mathematics, establishes the general mathematical form for the representation of particle mass. The multiplicities of levels are derived from a type of mapping in which each element of a discriminately closed set is mapped into some one other element in such a way that discriminate closure is preserved for each element of the set taken separately. This type of mapping leads naturally to a less exacting requirement — namely that subsets should exist within which discriminate closure is preserved under the mapping operation. In terms of the more convenient though less fundamental matrix representation of the hierarchy that is developed in § 3 such subsets are determined by the successive operation of a matrix A to generate a « cycle » of transforms on a vector v, say, of the form :

$$v, Av, A^2v........A^{n-1}v, v.$$

Future developments of the theory being propounded will include extensive analysis of these cycles as the origin of the discrete masses of particles. For the present it is sufficient to point out that the mathematics dictates a form for the particle mass concept.

4. Momentum is the first concept we introduce that has a direct spatial or temporal reference in observation. To define it we need to be able to de fine an order (successor relation) on a set of observable quantities. In § 7 we shall show that all that is really required to define the momentum concept as it is used in high energy physics (as distinct from classical physics and from low energy physics which proceeds by analogy with the classical concept) is that when a new observation is made we can set its result in a linear order with respect to other results. For the moment I shall assume this position. An ordered sequence of particle masses can be used to define momentum by considering that for any ordered pair m_1, m_2 a new quantity $m = m_1$-m_2 is defined which is available to use to define the momentum. It is true that other ordered pairs defined by the sequence may give different values, but such a situation would be interpreted merely as showing that the momentum had changed from one ordered pair to another. The idea that there must be an underlying entity whose momentum is constant independently of observation is a purely classical one that is not relevant to our investigation. Ordered sequences of particle masses canbe constructed in the present theory by considering nested mappings : these exist in the algebra without further principles being invoked. Two distinct mappings M_1, M_2 are said to be *nested*, with M, induded im M_2 ($M_1 \leqslant M_2$) if all subsets that remain discriminately closed under M_2 also remain discriminately closed under M_1.

With the definition of momentum given in the last paragraph, which carries with it a corresponding definition of velocity (and with length and time as derivative from velocity) the dimensionality of physical space can no longer be assumed to be simply that of the vector space that defines the manifold of possible observations. By contrast with this view of dimensionality which is excluded for us, we attribute the dimension number 3 of physical space to the fact that this number is the order of the simplest discriminately closed set in the hierarchy. We think that this identification is quite consistent with the operational evidence we have that physical space is 3-dimensional, in spite of the fact that our present theory is so strongly based on the concept of the particle, for we argue that if ordinary macroscopic physics itself were not deeply committed to an analysis of phenomena in terms of the particle concept, we should not have any really necessary reason for taking space as 3-dimensional. Different problems oculd use spaces of different dimensionality. This view of the dimensionality of spaec was expressed in a different form by the writer with C. W. Kilmister (1) who attributed it to a particular type of symmetry, which they analysed, in the *quadratic group*. The quadratic group is in fact isomorphic with the set of elements at level 1 in the algrebraic hierarchy of § 3 under the discrimination operation. The place of time in 4-dimensional space-time is a different problem which will be treated at length in later publications. For the present it may be noticed that our theory makes a basic difference between length and time regarded as coordinates.

5. Comparison with Experiment

In any of the levels defined in § 3, the multiplicity of the level determines a limiting value and hence a unit for measurements for the mass of a localizable particle. Any other mass will be determined by a choice of mapping which defines a subset in a level, and this subset will have a cardinal less than (in the case of a proper subset) the multiplicity. It is therefore possible to take the multiplicity to define the unit in terms of which all other masses at a given level are to be measured, and therefore as the scale constant for that level. If it is desired to introduce the concept of a field into this scheme then it will be natural to define the field characteristic of a given level as the unit of measurement common to all the measurable quantities in that level. In fact masses are the only quantities defined numerically at all, and it is conventional to ascribe any effect which is common to a class of masses as being due to a field acting on those masses, since measurements of a given mass (as distinct from comparisons of two masses) must always be deduced from accelerations of particles which are equivalent to the existence of fields. The field thus defined will be called the *characteristic field* of the level. These considerations also enable us to identify the multiplicity of a given level with the coupling constant of its characteristic field, provided that the coupling constant is taken as a measure of the *relative* strengths of the different characteristic fields. (The coupling constant *can*, of course, be so defined, in which case the measurable quantity is the ratio of two coupling constants. It would be a complex matter to relate the present theory to definitions of the coupling constant given, for example, in terms of the perturbation expansions of field theory, and no such relation will be attempted in this paper).

As the present theory does not assume the space-time continuum, there is nothing surprising in the fact that it can treat not merely strong and electromagnetic, but also gravitational fields on the same footing, and this fact can simply be taken as an advantage of the theory. Reference to the table at the end of § 3 shows that the theory provides for two types of strong interaction, associated with couplings 1/3 and 1/10. There is considerable doubt as to the right way to define strong coupling, and all one can say is that these constants are of the right order of magnitude, and that the fact that a division into two classes is indicated, has been held by some experts to be a desirable aspect of the theory. The famous electromagnetic coupling constant is given a value by the theory which is in good agreement with experiment (1/137 for 1/137.037); though the difference is outside the probable limits of experimental error.

I regard the discrepancy as probably due to the anisotropy of matter on the large scale in the universe, and it is of the right order numerically for this to be the case.

The gravitational coupling constant can be compared with experiment in the following way. We can take the ratio of the gravitational force F_g on the nucleon to the electric force F_e on it as

$$\frac{e^2}{\gamma m^2} = \frac{\alpha \hbar c}{\gamma m_n^2} \sim 1.238 \times 10^{36},$$

since

$$F_g = \frac{m_1 m_2}{r^2}, \qquad F_e = \frac{e_1 e_2}{r^2}$$

so that

$$\frac{F_e}{F_g} = \frac{e^2}{m_n^2} = \frac{\alpha \hbar c}{m_n^2}$$

This value confirms our theory to about 1% if we compare it with $\frac{2^{127} + 137-1}{137}$ which is the value obtained from our algebra on the basis of the rudimentary account I have given of a field. It would certainly be wrong for the reasons I have given already to attach too much significance to the closeness of the experimental agreement at this stage. Indeed I think we should be on safer ground merely to observe that we have a large dimensionless constant of the order 10^{38} associated with the upper limit to measurement, and that our theory gives a value of that order.

The weak interactions have an experimental coupling constant 1.01×10^{-5}. (This value is given by Feynman (2) by taking one of the possible particle masses, and is therefore subject to revision). The value we get from our theory is 1.53×10^{5}. This result is obtained by considering the forbidden mapping $2^{127} \rightarrow (256)^2$, and arguing that if a constraint of some unknown nature were to be imposed upon the elements of this level so as to reduce its effective multiplicity to $(256)^2$, this mapping would no longer be forbidden. Such a mapping would presumably correspond to unstable particles. The value $(256)^{-2}$ then would give the coupling.

6. Space, Time, Continuity, the Free Particle.

The next two sections of this paper use physical concepts in the way in which they are currently defined, and their purpose is to show that in the high energy sphere the vital concept of momentum is progressively coming to be used in a way consistent with the definition we have found to be necessary in the theory that has been presented in this paper. This demonstration renders less serious the gap which we have not been able to eliminate completely between the definitions of concepts in this paper and the corresponding definitions in current usage. The present section deals with continuity space and time as a preliminary to the treatment of momentum in § 7.in

Many physicists have questioned the validity of the space and time concepts in constructing theories of the microscopic structure of matter (1), and especially in attempting to explain the processes that become accessible to observation when high energies are available for the exploration of the microscopic structure, but these criticisms seem never to have been pushed home so as to make a real impact on theory. Instead, it has hitherto been thought possible (or perhaps inevitable, whether possible or not) to deal with the difficulties that arise from the invalid use of the space and time concepts in a piecemeal way. Use them and then limit their applicability, has been the policy.

It is easy to see that this policy — once adopted — would have the inhibiting effect that I have mentioned on thoroughgoing evaluation of the limits of applicability of the space and time concepts; you cannot meaningfully impose spatial or temporal limits on the applicability oi the concepts 'space' and 'time' (2). As, however, the policy *has* been adopted, we must look critically at the way it has been justified in current

(1) The most clinching demonstration known to me of this invalidity is due to Frisch (3), who pointed out that observations of pairs of particles emitted from certain disintegration processes show a coupling of the polarisation of the particles that is quite inexplicable unless the components of the process are postulated to have a kind of unity, in spite of their spatial separation, that is not due to any signalling phenomenon. Such observations cannot be reconciled with ordinary ideas of spatial separation.

(2) See also (4).

theory. The justification depends upon the way discrete and continuous variables are related.

The method of quantum theory for relating discrete to physically continuous variables depends paradigmatically upon the treatment of the free particle. In the form of theory originally proposed by Heisenberg we deduce the existence of a velocity

$$x = \frac{p}{m}$$

from a momentum p which occurs in a Hamiltonian $H = \frac{p^2}{2m}$ where H and p are matrices of numbers representing probability amplitudes. The original function of a matrix in quantum theory was to provide as concise as possible a summary of the structure of spectral lines. In the above equations, however, the influence of classical ways of thinking has already caused a transference of emphasis from the element as the primary dynamical quantity to the matrix as primary dynamical quantity — the thought process at work clearly having been that a dynamical quantity was essentially complex and specified a relation on a set of possible states.

From the history of the fundamental discussions of the quantum theory of the 1920's it is difficult to gather whether one is beginning with matrix elements as a new kind of physical entity and then establishing the existence of a concept identifiable with classical momentum, or on the contrary, assuming the meaningfulness of the classical concept and then demonstrating that there exist conditions under which that concept can be approximated to by a discrete set of observable quantities. Whatever the historical position has been, however, it is now certainly possible, without departing from recognised and current thinking, to assign priority to the elements. A quite unequivocal stand is taken on this question, for example, by Feyman (5), who asserts, « It has been found that all processes so far observed can be understood in terms of the following prescription : To every process there corresponds an *amplitude* (a complex number) ; with proper normalisation the probability of the process is equal to the absolute square of this amplitude ». Feynman's statement of the premises of quantum theory continues by demonstrating (in too great length to quote) that the « processes » mentioned in this quotation, upon which such a lot of weight rests, are such that there can be complex as well as simple ones, and that a *complete* specification of the complex ones is to be obtained by writing the amplitudes of the constituent simple ones in an array and then operating with arrays according to the well-known principles for manipulating arrays (matrix algebra).

Let us assume therefore that the logically primary representation of a dynamical quantity is a matrix element, and return to the paradigm problem of the the free particle. If we write

$$\psi\ (x,t) = a_0 e^{i(px-Ht)/h},$$

introducing a « wave-function » ψ, then we have done two things : firstly, we have got a form in which both position and time appear simultaneously in accordance with the requirements of any classical picture ; secondly, this equation defines an alternative mathematical device to that of matrix mechanics — namely, that of replacing the discrete matrix elements by the eigenvalues of a hypothetical continuously variable quantity, thus thinking of them as nodes in a wave. By extension, moreover, of these ideas it is possible to arrive at a formal specification of momentum as a differential operator, $ih\partial/\partial t$ which replaces momentum in the equations of motion and (via commutation relations) only allows interpretation of the « dynamical variable » in the case specified by the discrete matrix elements. One is thus encouraged to think of a continuum of values through which the momentum runs, with the observable values as kind of bus stops. This impression is further developed in the probability wave interpretation of the equation giving $\psi(x,t)$, according to which the formal equivalence of that equation with a wave equation (3) is used to give a picture which is classical in the sense that it appears to describe events localised in space and time (because x and t both appear explicitly). This picture is given an air of plausibility, as is well-known, by introducing the concept of a *probability* of a system being in a given state as the interpretation of $\psi(x,t)$. In fact, this interpretation involves a circular argument, since the probabilistic concepts required to establish an analogy with an equation of motion do not describe a situation in which operational meaning is given to the successive positions of a moving mass without the further assumption that the separate measurements of position that are required to specify a probability (on any of the theories of the nature of probability) may be brought together at a given time by an operationally well-defined pro-

(3) from the foregoing equations we have

$$\frac{h}{i} \cdot \frac{\partial\psi}{\partial\chi} = p\psi, \qquad \frac{h}{i} \cdot \frac{\partial\psi}{\partial t} = -H\psi,$$

Whence if $H = p^2/_{2m}$ we derive the wave-equation $\frac{h}{i} \cdot \frac{\delta\psi}{\delta t} = \frac{h^2\delta^2\psi}{\partial\chi^2}$ ·

cedure. This last assumption, however, would require appeal to the existence of classical space and time continua independently of the separate measurements, whereas it was precisely the validity of this appeal that the probabilistic argument was designed to justify (4).

We may conclude from the foregoing discussion, therefore, that a classical dynamical variable defined by the successive positions of a moving mass, is being used to provide a picture of quantum processes in a way that cannot be justified by appeal to the theory of the free particle without circularity.

We now ask the question "what was gained by appealing to the theory of the free particle?" The answer to this vital question is that the use of this classical concept implies the assumption that the actual states of the free particle from which its track has to be made up exist in a set to which access is available — the next state being ready to hand. If this assumption (which is very deep-seated from our training) is abandoned, then we have to envisage a type of theory in which the set of operations defined upon the set of states constituting a dynamical system, includes within itself the operation of construction or selection of each successive state from its predecessor. The expressions « construction » and « selection » sound very different: actually they amount operationally to the same thing within the present context, for the question whether a given state exists in the original set can only be settled by producing it. The question is analogous to the question whether each next operation in a computer program is constructed or selected.

I shall call the method of defining the physical continuum by ordering states which I have just described the *interpolation theory* of the physical continuum to refer to the fact that in it the physical continuum is defined as the field of progressive interpolation of new points. This theory is consistent with the theory of §§ 1-5 of this paper in which it is recognised that the order has to be explicitly introduced and in the present section it has been shown that appeal to our commonsense experience as in classical physics to provide this principle of ordering is not consistent with the

(4) A fundamental attack has recently been delivered by Dirac (6) on the validity of the familiar assumption that the Heisenberg (matrix) and the Schrödinger (wave) pictures are equivalent for the free particle. This attack is confirmation of the view here put forward to the extent that Dirac uses his disproof of equivalence to reject the Schrödinger picture. My position differs from that of Dirac, however, in that I do not think that the Heisenberg equations of motion are explanatory in their description of momentum for the reason I discuss, and I therefore think that some additional picture is necessary. Dirac says « The Heisenberg picture is a good picture; the Schrödinger picture is a bad picture ». I say: the Schrödinger picture is a bad picture; the Heisenberg picture isn't a picture at all.

principles of quantum theory. In § 7 a more detailed examination of the use of the momentum concept in high energy physics serves to re-emphasize the importance of defining order explicitly.

7. Momentum.

In § 6 reasons internal and central to quantum theory were advanced for regarding the current quantum-theoretical account of the continuous dynamical variable as inadequate. In the present section I shall describe attempts in high energy theory to counteract the effects of this inadequacy in ways less thoroughgoing than by a vigorous interpolation theory of the dynamical variable, namely, by retaining the momentum continuum while rejecting space and time. I shall try to show that nothing is gained by not going the whole hog.

It was shown by Bohr and Rosenfeld (7) that quantum electrodynamics and the principles of field quantization were consistent in detail with the requirement that all the quantities in which results were expressed should be referable to macroscopic procedures in space and time, in the sense defined in stating the complementarity principle.

The same critical enquiry made it clear that a size limit must at some stage be reached below which the classical picture of the electron as a moving charge would break down. The critical conditions determining this limit are reached when the radiative interactions become of the same order as the static interactions. This happens at the « classical electron radius » e^2/mc^2. (It is also found that this size coincides with nuclear dimensions, but in the nature of the case, the electromagnetic theory is not able to account for this coincidence, since the existence of nuclear fields constitutes a *deviation* from what is predicted by electromagnetic theory).

The importance of Bohr's type of analysis that it remains the only complete attempt to deduce properties of any field that is associated with an elementary particle from the basic structure of quantum theory itself. (It also has to be realised that there is no escape from Bohr's and Rosenfeld's conclusion by taking fields other than the electromagnetic ; for in the case of these other fields — such as the nuclear fields — no macroscopic procedures corresponding to the theoretical assertions about the field in question exist *at all*).

Chew, arguing the case for the analytically continued S-matrix approach and more particularly for the bootstrap hypothesis [5], has drawn attention (8)

(5) The « bootstrap » theory of the origin of an elementary particle was introduced by Chew and Frautschi (9) using the mathematical system of the analytically continued S-matrix approach to particle dynamics. A *bootstrap system* is defined as a particle

to the same « measureability limit » that occurs at lengths of the order 10^{-13} cm. in order to use it as an argument against the conventional application of the space-time continuum at distances below this value. Chew points out that no corresponding limitation exists on measurement of momentum, and in this way he justifies the use of the momentum continuum within the analytically continued S-matrix approach. His argument is that momentum still can be used, even in those circomstances from which length and time are excluded by the measureability limit.

It is important for me to analyse Chew's argument in some detail, since the bootstrap theory is the theory which comes closest to the model of the origin of particle mass that is presented in this paper.

In discussing Chew's argument, the first most general thing to notice is that from the point of view of the momentum concept that I have been elaborating, Chew has directed his attention to a symptom while failing to cure the disease that produced the symptom. The measureability limit of 10^{-13} cm appears within an overall theory in which the logical inadequacy of the quantum theoretical account of the continuous dynamical variable has been incorporated : the « measureability limit » is a device to stop it mattering. Hence the device of using the momentum continuum to which the measureability limit does not apply is likely to prove merely formal — a conclusion that can be reached independently from the argugent of § 6, since to use a momentum continuum which does not define order explicitly is to pre-suppose either the space or the time continuum. Let us now see, therefore, whether the extent of successes of the bootstrap theory bears out the truth of these remarks.

The crucial test seems to be whether or not any experimental results follow in the bootstrap theory from the use of the momentum continuum. It has been argued that none does and that one might as well work with a bootstrap theory devoid of continuum dynamics. This position has been taken recently by Ne'eman (10), who proposes one self-interacting system, with all the allowed spins, etc. He presumes this system to satisfy an equation of the type

$$(\square + j^2 + k^2 \ldots)\, \Phi = 0$$

where j^2, k^2, etc., are Casimir operators. Then, « if each particle is to be determined by any pair of particles capable of making (i.e., generating) it, within the symmetry restrictions, the term k^2 becomes k^2 ($\emptyset F \emptyset$) with F^a_{bc} a set of coefficients relating the multiplets to each other and projecting out, for example, the baryon octet (b) and meson octet (c) as one set of

assemblage in which the constituent particles (or elements) hold each other in existence by their own mutual interactions, and have no existence apart from those interactions.

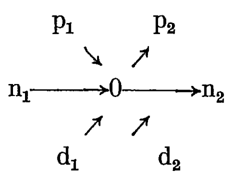
constituents contributing to the making of the baryon octet (a) in the eigenfunction ». In this way the non-linearity of the bootstrap system is introduced. Ne'eman thus explicitly introduces the symmetries and abandons the efforts of the bootstrap theorists to get them as a result of kinematics, as being unlikely to succeed.

In thus abandoning kinematics, however, Ne'eman has not abandoned appeal to an S-matrix-like philosophy. Yet logically we might have expected him to do just that, for he has his discrete symmetries to connect him with basic quantum theory. Accordingly, we must look back at Chew to discover just exactly what the bootstrap system really gets from the S-matrix philosophy (given that we are not prepared to swallow the whole story about momentum hook, line and sinker).

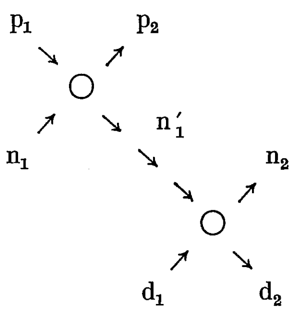
The essential step is the expansion of a complex particle reaction into two or more simpler reactions, using the Feynman graph idea, but embedding it in a new physical context appropriate to the momentum calculus. Thus the overall process

$$n_1 + p_1 + d_1 \rightarrow n_2 + p_2 + d_2 \qquad 7.1$$

described by the « connected part » graph



is re-expressed



corresponding to two successive separate reactions

$$\underline{n}_1 + \underline{p}_1 \rightarrow \underline{n}_1 + \underline{p}_2$$
$$n'_1 + \underline{d}_1 \rightarrow \underline{n}_2 + \underline{d}_2 \qquad 7.2$$

This succession of events is then, as it were, embedded in a dynamical structure by observing that the quantity

$$\frac{1}{(\underline{n}_2 + \underline{d}_2 - \underline{d}_1)^2 - m_n^2}$$

becomes large for $(\underline{n}_2 + \underline{d}_2 - \underline{d}_1)^2 - m_n^2$ **7.3**
Here m_n is the postulated mass of the newly introduced « particle » corresponding to the relation between the two constituent reactions.

A composite process of type 7.1 would be very improbable in general, but becomes probable for points for which the approximate equality 7.3 is satisfied. This observation is the basis of the application of the theory of Regge Poles to particle processes, and is hence the basis for the most successful existing attempts to map a dynamics onto a graph-type theory. At this point there are two ways one may argue. On the one hand one can say that there exists a Regge Pole dynamics of momentum and that it has been shown to apply successfully to the construction of a composite decay process from simpler ones ; on the other hand, one can say that the essential structure of the particle processes — namely that they can be decomposed into temporally ordered constituent processes — has dictated the choice of dynamics.

The second of these arguments is the one consistent with the point of view of the present paper since it gives central importance to the process of introducing a new experimental value of momentum. It is also the position taken by Chew (11) in the presentation of the theory that I have chiefly followed in my present summary. Moreover it is the argument that is consistent with the view shared by many workers that the nature of the high energy reesarch field gives the actual particle process — regarded as a primitive element in the theory — complete priority as the operational unit or ultimate contact with experiment.

If one adopts the second of the arguments just given then one must distinguish very sharply between two uses of the concept « momentum » in particle physics, namely :

1. Momentum as correlating concept for interpreting the results of experiments in terms of theory, but not itself part of the theory, and
2. Momentum as the central theoretical concept from which dynamics has to be built.

If we make this distinction rigorously then we conclude from the foregoing discussion that the S-matrix presentation of momentum as a theoretical concept, in the form presented by Chew in which that concept is freed as far as possible from classical ideas of space-time, still falls short of being the basis from which a dynamics can be built, in one essential respect. When two constituent processes are derived from a single complex process the order in which they take place is not specified. This basic requirement has been a major consideration in the application of the theory of § 4 and the form in which we shall derive this possibility of ordering will dictate the later applications though these applications will not be reached in the present paper.

8. Conclusion.

In this paper I have shown (a) that the scale constants of physics could be derived from a theory of discrete interactions, by constructing the theory, and (b) that although such a theory would require us to abandon a great many accepted results which would then have laboriously to be reconstructed on the new basis, nevertheless the theory would have the fundamental advantage that it would be free from a basic inconsistency that arises in current quantum physics and becomes disastrous in high energy particle physics in the use of the space-time continuum.

The next pieces of work to be done in following up the method initiated in this paper include the following:

(a) To relate the 2- and 4-vectors of the top level of the hierarchy to the spin vector representations of quantum theory, and to free the spin concept from its irrelevant connexion with Lorentz invariance.

(b) To investigate the invariant mappings that can be defined within the hierarchy in a way natural to its mathematical structure, and to compare the constants arising in this way with particle masses.

(c) To find which groups are representations of the invariances defined in the hierarchy, and compare these with the groups that have been proposed to represent particle symmetries. This investigation should be specially interesting in providing a different dynamics for assessing the physical significance of the different groups. (Thus it is known that SU_6 cannot be made Lorentz-invariant, and this is currently thought to be an insuperable difficulty).

TED BASTIN

References

1. Bastin, E. W. and Kilmister, C. W., Proc. Camb. Phil. Soc., 50, p. 278, 1954.
2. Feynman, R. P., Theory of Fundamental Processes, Benjamin, New York, 1961, p. 33.
3. Frisch, O. R., "Observation and the Quantum", Karl Popper Festschrift, Ed. K. Bunge, p. 309.
4. Bastin, E. W., Proc. Camb. Phil. Soc., 574, 848, 1961.

5. Feynman, R. P., « Theory of Fundamental Processes », Benjamin, New York, 1962, p. 1.
6. Dirac, P. A. M. Nature, 4941, 1964, 115.
7. Bohr, N., and Rosenfeld, L., Phys, Rev. 78, 794, 1950.
See also Ed. P. A. Schilpp. Norwestern University, U.S.A.
Bohr, N., Contribution to Einstein volume in Library of Living Philosophers.
8. Chew, G. F., « The Dubious Role of the Space-Time Continuum in Subatomic Physics ». Open Lecture, Cambridge University, 1963, and Contract Report W-7405 — eng — 48. University of California, June 1963.
9. Chew, G. F. and Frautschi, S. C., Phys. Rev. Letters 7, 394, and 8, 41 (1962).
10. Ne'eman, Y., Physics 1, 4, 1965.
11. Chew, G. F., Nuclear Democracy and Bootstrap Dynamics, 1963, Radiation Laboratory, Berkeley, workpaper.