International Journal of Theoretical Physics, Vol. 8, No. 3 (1973), pp. 155–170

Perception and Operation in the Definition of Observable

E. A. B. COLE

School of Mathematics, University of Leeds, Leeds, LS2 9JT

Received: 19 *July* 1972

Abstract

A discussion of the roles of assumption and abstraction in physics is presented, with emphasis on the fact that if abstractions are used long enough it becomes difficult to disentangle them from what is actually observed. The unclear meaning of the word *same* in the context of two observers measuring the same observable is also discussed. An observable is defined operationally as a list of instructions, and this definition also defines the word *same* in this context. Methods for the combination of these observables are presented. An observer decides whether or not a list of instructions is physically realisable by examining the states of the system, and it is shown how these states are based on the observer's powers of perception. Finally, an account of the operational nature of cellular space–time is given, the fuzzy nature of the cell boundaries is discussed, and it is shown how measuring rods consisting of 'straight' strings of cells may be constructed.

1. *Introduction*

The object of this paper is to analyse in part the role of the assumptions that are made in physics, in particular in the field of turning observed results into abstractions which then enter our theories. Certainly these abstractions are necessary if we are to classify observations and predict future observations, but the danger is that if these abstractions are used long enough it becomes more and more difficult to disentangle them from what is actually observed as with, for example, the notion of continuous space-time. As Atkin (1971) points out, it leads to the idea of '... an *ideal*, or *absolute*, world which scientists must always find elusive'.

In this connection, Meredith (1966) points out '... the abstractions change from generation to generation and it would be a pity if physics had to be described as a science of fashionable abstractions. Perhaps, then, abstractions are not the subject-matter of physics. But then what is physics about? And if the abstractions—which certainly loom large in any treatise on physics—are not its subject-matter, what is their status? Only a study of

Copyright © 1973 Plenum Publishing Company Limited. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission of Plenum Publishing Company Limited.

the history of the subject and of what physicists do can answer these two questions. And there need be no obscurity in the answer—for physics is a perfectly concrete science and the obscurity is only in men's minds'.

Thus the abstractions, the assumptions which lead to them and the very operations of measurement should be looked at very carefully, because it is only when these distinctions are recognised that we are able to jump from one theory to a more general one. In this paper we take up Meredith's suggestions of studying what physicists do by presenting an account of the nature of observation and measurement from an operational point of view.

One abstraction already mentioned is the continuum. After many years of traditional teaching it really is difficult to think of the natural world as being other than, in a very crude sense, a series of physical phenomena taking place in a continuous space-time. Yet whenever we observe the world we never use a working continuum—each observer is limited by the degree of refinement of his instruments, which in turn is limited by the amount of energy he supplies to the instruments for this purpose. Thus in any observation, a cellular space-time is always used, the exact structure of which is observer dependent (Cole, 1972a, b). And as it is never possible to 'observe' the continuum it must be regarded purely as a device which may make our theories easier to handle, and which may be discarded if it becomes too cumbersome.

Another abstraction arises when we come to handle and interpret a set of measurements. For example, suppose we try to measure a property of a system which we know to be stationary. Measurements are made over a time interval and at the end of the interval we have perhaps a spread of observed values. It is then convenient to abstract from this set of values one single representative value which can be inserted into our current theory—single values are much easier to use than sets of values. We do this so often that we begin to regard the abstracted value as representing the *actual* value of what we are measuring, and the individual values of the set are then regarded as being in *error* in some way. Having abstracted this value we then mentally, if not physically, destroy the set and use only the abstracted value. What is often forgotten is that the single representative value is merely a formal device which may make the theory easier to use, an abstraction from the only concrete quantities available to us—the individual elements of the original set which have been obtained by direct observation. (In this sense, experimentalists have been tricked into talking of *experimental error*, as though the observations which are the products of themselves and their apparatus are somehow faulty and the only correct value is some abstract quantity thought up by the theoretician.) Having recognised the fact of this abstraction, it is then a small step to realising that a theory may be constructed which handles sets of readings rather than one representative value from each set. Of course, such a theory will have its own abstractions, but to recognise the abstractions is to recognise the assumptions which lead to those abstractions, and hence may lead to a theory with weaker assumptions.

Observations are always made through real, non-ideal measuring instruments, each of which has a degree of refinement governed by the finite amount of energy available for refining the apparatus. Very often at the output of the apparatus, at the observer–apparatus interface, the readings are those of a distance in which the reading may be recorded as a line on a sheet of paper. On using a predetermined scaling factor, the length of the line is transformed into another value which gives a measure of what is being observed. Since the sheet of paper and any measuring rod will have cellular structures, this distance is actually a *distance set* (Cole, 1972a) in which the elements depend on both cellular structures. Now, the resulting set of observed values can be compared directly with the traditional way of presenting the results of an observation, usually of the form $x \pm \Delta x$. In this latter case it is usual to think of a 'real, actual' value lying somewhere in the continuous range $(x - \Delta x, x + \Delta x)$, which again confuses the abstraction with the reality. A way out is to deal with the complete set of observed values as it stands and not to abstract a representative value. Measurements by different observers could still be compared: suppose two observers measure the same observable (a detailed discussion of the word 'same' is presented in Section 2) and arrive at sets of observations d and d' which may or may not have elements in common. The sets may then be compared by testing whether or not they overlap on the real line; that is, the segments $(\min d, \max d)$ and $(\min d', \max d')$ overlap on the real line. If the two sets do overlap in this sense we may write $d \circ d'$, where $\circ$ is an *overlap relation*. This relation is reflexive and symmetric and can be compared with the *fuzzy* relation of Fuzzy Geometry (Poston, 1971). Note that in the definition of the overlap relation the notion of a continuous real line has been used. This is quite acceptable as long as it is remembered that it is merely a device for comparing the two sets in this way, and it leaves the door open for further detailed comparison of the individual elements of the sets because these sets have not been destroyed by the abstraction of a single value.

All this supposes that the observer knows just what it is he is measuring and that it is possible for him to measure the same quantity as another observer. But a detailed analysis of the word 'same' in this context shows up a great many difficulties which are not apparent in the fleeting analysis we usually give to everyday words. To illustrate this with an example, suppose that two observers A and B attempt to measure the momentum of a projectile. Observer A, knowing its mass, finds its speed by timing it over a fixed distance, and hence finds its momentum by multiplying its mass by the speed. Observer B has fixed a known weight to a vertical string and measures the amount by which the weight rises when struck by the projectile, and and thereby calculates its momentum using conservation principles. By saying that the projectile has a momenumt and that both observers are measuring it we have already assumed that they are measuring the same quantity. But if we take a step backwards and say that A and B are measuring quantities momentum (A) and momentum (B) respectively, that is, split the definition of the observable, it can then be recognised that it is a further assumption to say that momentum (A) is the same as momentum (B). Further, the assumption is very vague, because what do we mean by the word 'same' in this case? Nevertheless, this very vague notion of the sameness of observations is useful for communication processes between observers.

On the other hand, we can remove the problem of the meaning of the word 'same' by thinking not of a particular quantity, for example momentum, as being attached to the particle (Bohm *et al.*, 1970), but that the final values which emerge from the act of measurement refer to the whole particle–apparatus complex. In effect this is close to saying that each piece of apparatus defines its own observable and that it is meaningless to talk of two observers, each with their own apparatus, measuring the same thing. Thus not only are the observed values observer-dependent, but also the definition of the observables are observer-dependent. All this would be very well if the observer worked in isolation with his apparatus—he could make observations and construct very personalised theories which relate only to his apparatus, but in practice he must communicate with other observers.

Thus what is needed is a definition of observable between these two extremes, such that on the one hand it is not to be thought of as being attached to what is being observed independently of any observing process so that no assumption is made regarding the equivalence of two different processes, and on the other hand it is to be thought of as being more than a purely single-observer-dependent entity because we want several observers to be able to communicate their results in a meaningful way. This problem will be tackled in Section 2 with a purely operational approach to the definition of observable.

Several previous attempts have been made at a discussion of the purely operational nature of observation. Antoine & Gleit (1971) describe the structure of an experiment performed by one observer. But despite their discussion of apparatus being transformed in the sense of being translated, rotated, etc., this is not very useful operationally when we want to consider how two observers, each using a different set of apparatus, can compare their results. Again, Mielnik (1969) becomes trapped in his own vicious circle. As he points out, '... This description, although plausible, is not operational. In order to cheek whether a device does or does not change the properties of certain particles we must first be able to measure these properties. But in order to do this we must first have measuring devices (filters)! To avoid the vicious circle we need an operation definition of filters which would not assume any a priori ability of analyzing the beam'. This is correct, but when he then goes on to define two transmitters as being equivalent when they behave identically in all experiments on the optical bench, he falls into his own trap because in order to test that the transmitters are behaving identically the same beam must be passed through both transmitters, that is, the same beam must be used on at least two occasions. And we run into the same problems over the word 'same' in this context as we did in the previous context. Thus this description of equivalence is not operational. This problem arises again in his description of a linear detector d: it is linear if

$$d(x+y) = dx + dy$$

where x and y are any two beams. In order to operationally check whether a detector is linear in this sense it is again necessary to use beam x (as well as beam y) on at least two separate occasions which again supposes that the two beams are the same.

In the same way, although Davies & Lewis (1970) are correct in their desire of removing the repeatability hypothesis in measurement, they again have a distributive condition in their definition of an instrument and this again is not directly operationally verifiable.

With pitfalls of this nature in mind we go on in the next section to define an observable on a purely operational basis. In Sections 3 and 4 the observables are linked with the possible states of the system, and to be consistent an account of these states is given from an operational viewpoint; these states are not regarded as being intrinsic to the system but are related to the degree of perception of the individual observer. While an attempt is made to remove as many assumptions as possible from the discussion one assumption will still remain: the assumption of the constancy of language in its meaning to different observers. A detailed discussion of this assumption would be very troublesome, but it is an assumption explicitly made and as such must be discussed, if only briefly.

2. *The Operational Definition of Observable*

We define an *observable* as an ordered, physically realisable list of instructions. Two observers measure the *same* observable if they both carry out the same list.

Several points emerge from this definition. Firstly, as well as being a definition of observable it is also a definition, in this context, of the word *same*. In order to make it a sensible definition this supposes that for any two people who interpret a given list, neither will argue with the other's interpretation. This supposition is not strictly foolproof and will be discussed more fully in Section 4. Secondly, sets of instructions may have many different visual forms—as a set of verbal instructions as, for example, a tape recording of spoken instructions, as a set of written words, as a set of diagrams or as any other communication process. Thirdly, there will be no idea at this stage of two distinct sets of instructions being equivalent in any sense, but this idea may always be introduced at some later stage in theories which deal with specialised physical phenomena. Fourthly, an observer will be restricted to carrying out the instructions in the order in which they are presented. The instructions may also give orders telling how to construct apparatus as well as how to use it. In addition, all objections from a third observer that, for example, two other observers have not used the same temperature in their experiments must be overruled if the list of instructions makes no mention of temperature (or rather, if there are no instructions in the list which the third observer would regard as 'temperature' control), for they would still be carrying out the same list of instructions and hence would be measuring the same observable. The third observer may write in the extra instructions if he wishes, but then the new complete list will be regarded as a different observable. Finally, not all lists of instructions may be physically meaningful to an observer; for example, after carrying out one of the instructions it may be impossible to carry out at least one of the later instructions. Those that can be carried out are called observables, with no distinction being made between whether or not results are actually tabulated. The formal distinction between physically realisable and unrealisable lists of instructions will emerge in the following analysis.

A list of instructions will be specified by the product form $a_1 a_2 \ldots\ldots a_n$ for any $n \geqslant 1$, where the a_i are the separate instructions in the list, and the convention will be that a_i is performed before a_{i+1} $(i = 1,\ldots,n-1)$. The separate a_i are themselves lists consisting of one instruction. Define C to be the set of all possible lists, which contains all possible physically unrealisable as well as realisable lists of instructions. Then we may think of the identity element e of C, not necessarily unique, such that its insertion in any place in any list will not operationally change that list. In words, it could take the vague form 'ignore this instruction'. Thus for any list $a_1 \ldots\ldots a_n \in C$, we have

$$a_1 \ldots\ldots a_i e a_{i+1} \ldots\ldots a_n \equiv a_1 \ldots\ldots a_i a_{i+1} \ldots\ldots a_n \tag{2.1}$$

Further, we will use the convention that a, b, c, etc. refer to elements of C, while A, B, refer to subsets of C.

A useful concept is the relation $\leqslant$, defined such that if $a_1 \ldots\ldots a_n \in C$ then $x \leqslant a_1 \ldots\ldots a_n$ if and only if x is of the form $a_1 \ldots\ldots a_i$ $(i \leqslant n)$; that is, the elements of x are the same as the first few elements of $a_1 \ldots\ldots a_n$. Clearly, if a and b are elements of C then $a \leqslant a$ and

$$a \leqslant b,\, b \leqslant a \;\Rightarrow\; a = b$$

Further, since $ea_1 \ldots\ldots a_n \equiv a_1 \ldots\ldots a_n$ it would be quite natural to write $e \leqslant a$ for all $a \in C$.

Products of two elements of C can be defined such that if $a, b \in C$ with $a \equiv a_1 \ldots\ldots a_n$ and $b \equiv b_1 \ldots\ldots b_m$, then ab can be defined as the complete ordered list $a_1 \ldots\ldots a_n b_1 \ldots\ldots b_m$. In general, $ab \not\equiv ba$-equivalence would hold if the lists ab and ba were identical or if they were postulated to be equivalent in some superimposed theory.

For each $a \in C$ now introduce the subset $g(\{a\})$ of C defined by

$$g(\{a\}) \equiv \{x : x \leqslant a\} \tag{2.2}$$

Note that it is written as $g(\{a\})$ and not $g(a)$; this is because the results derived later are written much more simply in terms of sets rather than in terms of single elements, although in some sense it is more desirable to work with single observables rather than with sets. Thus, for example,

$$g(\{a_1 a_2 a_3\}) \equiv \{a_1 a_2 a_3,\, a_1 a_2,\, a_1,\, e\}$$

Definition (2.2) can then easily be extended to the form

$$g(A) \equiv \bigcup_{a \in A} g(\{a\}) = \{x : x \in C, x \leqslant a, a \in A\} \tag{2.3}$$

defined for all $A \subset C$. Note that $e \in g(A)$ for all $A \subset C$, except that we define $g(\varnothing) = \varnothing$, where $\varnothing$ denotes the empty set.

The following results are then easily proved: for all A, $B \subset C$ and all $a, b \in C$,

(i) $g(A \cup B) = g(A) \cup g(B)$;
(ii) $g(A \cap B) \subset g(A) \cap g(B)$;
(iii) $B \subset A \Rightarrow g(B) \subset g(A)$;
(iv) $b \leqslant a \Rightarrow g(\{b\}) \subset g(\{a\})$;
(v) $A \subset g(A)$;
(vi) $g(g(A)) = g(A)$;
(vii) $g(\{e\}) = \{e\}$;
(viii) $g(C) = C$.

Attention will now be focused on the subset C_0 of C such that each element of C_0 is an observable; that is, a physically realisable list of instructions. Clearly, we must allow $e \in C_0$, and although an element a of C which is not an element of C_0 will not be physically realisable, at least one element of $g(\{a\})$ will be physically realisable. That is, by removing some of the offending elements from the end of a the remaining list becomes physically realisable, if only because $e \in g(\{a\})$. Further, if $a_1 \ldots\ldots a_n$ is physically realisable then it is possible for $a_1 \ldots\ldots a_i$ $(i = 1, \ldots, n)$ to be carried out. Thus for each $a \in C_0$,

$$g(\{a\}) \subset C_0 \tag{2.4}$$

It is now possible to obtain a necessary condition on C_0. Result (2.4) gives

$$\bigcup_{a \in C_0} g(\{a\}) \subset C_0$$

and so by definition (2.3), $g(C_0) \subset C_0$. But by result (v), $C_0 \subset g(C_0)$. Hence

$$g(C_0) = C_0 \tag{2.5}$$

This result is a necessary condition on C_0 but it is not sufficient, because by results (vii) and (viii) $\{e\}$ and C also obey this condition, and by result (vi) any subset of the form $g(A)$ obeys this condition.

There are many different ways of combining two lists of instructions a and b of C. As already seen the straight products ab and ba can be formed, and further, because a and b both consist of a finite number of elements, they can be split up and interwoven with each other. A natural way of doing this is to define the cross product between $a \equiv a_1{}^1 a_2{}^1 \ldots\ldots a_n{}^1$ and $b \equiv a_1{}^2 a_2{}^2 \ldots\ldots a_m{}^2$ as

$$\{a\} \times \{b\} \equiv \{x_1 x_2 \ldots\ldots x_{m+n} : x_i = a_j{}^k \ (k = 1 \text{ or } 2);\ x_{i_1} = a^k_{j_1}, x_{i_2} = a^k_{j_2} \text{ and } i_2 > i_1 \Rightarrow j_2 > j_1 \ (k = 1, 2)\}. \tag{2.6}$$

That is, the individual instructions of a and b are interwoven in such a way that the orders of the separate instructions of a and b are separately preserved. Thus, for example,

$$\{a_1 a_2\}\times\{b_1 b_2\} = \{a_1 a_2 b_1 b_2, a_1 b_1 a_2 b_2, a_1 b_1 b_2 a_2, b_1 a_1 a_2 b_2, b_1 a_1 b_2 a_2, b_1 b_2 a_1 a_2\}$$

Again, we deal with the sets $\{a\}$ and $\{b\}$ consisting of individual elements a and b rather than with the elements themselves, because the following results may be more clearly written. It also enables us to extend the definition (2.6): for each A and $B \subset C$, define

$$\begin{aligned} A \times B &\equiv \bigcup_{a\in A} \bigcup_{b\in B} \{a\}\times\{b\} \text{ if } A, B \neq \varnothing \\ A \times \varnothing &\equiv \varnothing \end{aligned} \tag{2.7}$$

Thus, for example,

$$\{a_1 a_2, c_1\}\times\{b_1 b_2\} = [\{a_1 a_2\}\times\{b_1 b_2\}] \bigcup \{c_1 b_1 b_2, b_1 c_1 b_2, b_1 b_2 c_1\}$$

The following results may then be easily proved: for all subsets A, B and D of C,

(ix) $A \times B \subset C$;
(x) $A \times B = B \times A$;
(xi) $A \times \{e\} = A$;
(xii) $(A \bigcup B) \times D = (A \times D) \bigcup (B \times D)$;
(xiii) $(A \bigcap B) \times D \subset (A \times D) \bigcap (B \times D)$;
(xiv) $(A \times B) \times D = A \times (B \times D)$;
(xv) $A \subset A \times g(B)$ for $B \neq \varnothing$;
(xvi) $ab \in \{a\}\times\{b\}$;
(xvii) $g(A \times B) = g(A) \times g(B)$.

Several points emerge from these considerations. Firstly, if A and $B \subset C_0$, nothing can be generally said about whether or not the elements of $A \times B$ are physically reliasable, because the individual instructions making up the elements of A and B may be such that it is not possible to combine them into a physically realisable list. Secondly, if the element e had not been included in the definition of $g(A)$ then result (xvii) would not hold and would have to be replaced with

$$g(A \times B) = g(g(A) \times g(B))$$

from which result (vi) would follow on putting $B \equiv \{e\}$. Thirdly, to illustrate why sets of elements are used in the notation instead of single elements, if the product $\times$ had been defined in terms of individual elements then result (xiv) would have to be written: for all a,b and $d \in C$,

$$\bigcup_{x\in b\times d} a \times x = \bigcup_{y\in a\times b} y \times d$$

which certainly does not illustrate the associative property of the product as satisfactorily as result (xiv). Fourthly, if a, b and ab are physically realisable it does not necessarily follow that ba is physically realisable. This is in contrast to the quantum mechanical result that if a, b and ab are hermitian operators then ba is hermitian. Finally, note that the distributive results (i) and (xii) are of a different nature to that of Mielnik so that the criticism in Section 1 does not apply in these cases.

3. *States Associated with Observables*

In the last section the status of physical realisability was discussed and a necessary condition was obtained for the set C_0 of such processes. But the question now arises as to how an observer actually distinguishes in his own mind between those that are and are not realisable.

When presented with such a list of instructions the observer reads through it and, in effect, by direct observation or by delving into his past experiences he tries to find a situation to which the list may be meaningfully applied. It is not necessary for him to have actually experienced this situation, for the situation he finds could well be compounded from his past experiences in such a way that he is certain that such a situation could exist. If no such situation exists for him then the list is not physically realisable; that is, not an observable. But if he is aware of at least one situation to which the list may be applied then that list is an observable, and hence is an element of C_0. For a list to be an observable, it is then not necessary that it is capable of being applied to all physical situations, but only to at least one. In this way, the emphasis of recognising the set C_0 is shifted to the set of physical situations, or *states*.

Let Ψ be the set of all these physical states—the exact operational character of these states will be discussed in the next section after the structure of the states has been built. Then for each $a \in C_0$ there will be a non-empty subset Ψ^a of Ψ such that the operation of a may be meaningfully applied to any state of Ψ^a. In fact, this definition can be extended to any $a \in C$ such that $\Psi^a \equiv \varnothing$ if $a \notin C_0$. After the list a of C_0 has been carried out the system will be in a new state. Let $\overline{\Psi^a} \subset \Psi$ be the set of all states which are possible after a has been carried out, and again we may put $\overline{\Psi^a} \neq \varnothing$ if and only if $a \in C_0$. Clearly it may be possible for some $a \in C_0$ that $\Psi^a \cap \overline{\Psi^a} = \varnothing$; that is, once a list a has been applied to any situation it may not be possible to re-apply it as, for example, in destructive testing. Note also in the definition of $\overline{\Psi^a}$ that we do not restrict the definition to states immediately after the application of a because, since any act of perception always takes a non zero time, the idea of 'immediately after' loses its meaning.

To aid the analysis we now introduce the formal product φa defined for each subset $\varphi \subset \Psi$ and each $a \in C$. The product is a subset of Ψ and is interpreted as the set of all possible states allowable when list a has been applied to all the elements of φ common to Ψ^a. Consistent with the meaning of $\overline{\Psi^a}$ we then require that

$$\overline{\Psi^a} = \Psi^a a \tag{3.1}$$

The following formal properties of the product then ensure, as can be seen from the results which follow, that $\overline{\Psi^a}$ as defined by (3.1) does have the meaning required of it. For all subsets ψ and φ of Ψ and all $a, b \in C$, ψa is a product in Ψ such that

(a) $(\psi \cup \varphi) a = \psi a \cup \varphi a$;
(b) $(\psi \cap \varphi) a = \psi a \cap \varphi a$;
(c) $(\psi a) b = \psi(ab)$;
(d) $\psi e = \psi$;
(e) there exists $\Psi^a \subset \Psi$ such that
  (α) $\Psi^a \neq \varnothing$ if and only if $a \in C_0$,
  (β) $\psi a = \varnothing$ if and only if $\psi \cap \Psi^a = \varnothing$,
  (γ) $\Psi^{(ab)} \subset \Psi^a$.

Then using (3.1) as the definition of $\overline{\Psi^a}$ the following results can be easily derived. For all $\psi \subset \Psi$ and all $a, b \in C$,

(i) $\varnothing\, a = \varnothing$;
(ii) $\psi a \subset \overline{\Psi^a}$;
(iii) $\varphi \subset \psi \;\Rightarrow\; \varphi a \subset \psi a$;
(iv) $\overline{\Psi^{(ab)}} \subset \overline{\Psi^b}$;
(v) $\overline{\Psi^a} \cap \Psi^b \neq \varnothing \;\Rightarrow\; a \in C_0, b \in C_0, ab \in C_0$;
(vi) $\Psi^e = \overline{\Psi^e} = \Psi$.

Result (iv) shows that the total set of afterstates of the instructions of a followed by b is contained in the set of afterstates of b alone. In result (v) the condition on the left-hand side implies that $\overline{\Psi^a}$ and Ψ^b are both non-empty, and that at least one of the afterstates of a is a state to which b may be applied so that ab is physically realisable.

Further, as an extension of result (v) it follows that $a_1 a_2 \ldots\ldots a_n \in C$ is an observable if and only if

$$\overline{\Psi^{a_1 \ldots\ldots a_i}} \cap \Psi^{a_{i+1}} \neq \varnothing \qquad (i = 1, \ldots, n-1) \tag{3.2}$$

This result is an expression of what the observer does when he reads through a list of instructions to decide whether or not it is an observable. He first looks at a_1 and a_2 and decides if it is possible that at least one of the afterstates of a_1 is a state to which a_2 may be applied. If the answer is yes (that is, if $\overline{\Psi^{a_1}} \cap \Psi^{a_2} \neq \varnothing$) he then considers if it is possible to follow $a_1 a_2$ with a_3. If the answer is again yes (that is, if $\overline{\Psi^{a_1 a_2}} \cap \Psi^{a_3} \neq \varnothing$) he continues along the list. If he comes to an a_i which cannot follow $a_1 \ldots\ldots a_{i-1}$, or equivalently $\overline{\Psi^{a_1 \ldots a_i = 1}} \cap \Psi^{a_i} = \varnothing$, he concludes that the list is not an observable, but if the inequality holds right along the line he concludes that the list is an observable.

4. *Perception and States*

As pointed out in Section 2, the definition of observable given there is not as foolproof as we might wish, because if attention is restricted to a group of observers, each observer must be satisfied with every other observer's interpretation of the lists of instructions. That they will agree is not always certain, because their language of communication is based on only a finite number of shared experiences. Thus, although their interpretations of the words of the instructions will mostly always satisfy one another there must always be a certain looseness in their interpretations which could possibly lead to disagreement, and this possibility is increased when a list of instructions is translated from one language to another. Whether or not an observer's interpretations of the lists is acceptable must then be a group decision.

In Section 3 the emphasis was transferred from the observables themselves to the set of states Ψ to which the observables may be meaningfully applied, because it is through these states that the observer is able to recognise just which lists of instructions are physically permissible. Now, this set is the total set of *distinct* states, and the capability to distinguish between states rests on the observer's powers of perception through both his unaided eye and his instruments. Thus the actual operational construction of Ψ is observer dependent and depends on how much energy he has available for refining his instruments. Thus, for example, where one observer may distinguish two states, another observer with a keener sense of perception brought about by more refined apparatus may distinguish three or more states. In fact, if an observer is allowed extra energy for expending on his apparatus he may use it in two distinct ways. Firstly, it may be used for a finer distinction of states in, for example, enabling him to distinguish three states where previously he could distinguish only two (microscopic distinction). Secondly, keeping the distinction between states the same he may use the energy to extend his field of vision (telescopic distinction). In both cases the effect is to increase the number of elements of Ψ, but with the distinction that in the first case some of the elements of the set disappear and will be replaced by a greater number of completely new elements, while in the second case new elements are added to the existing set. In fact, when first setting up his apparatus the observer must balance to his satisfaction this microscopic and telescopic perception.

There are also elements of an observer's Ψ set which will not change—these are the situations perceived during his past experiences which are stored in his memory, and also situations communicated to him by other observers. Again, the problem of the constancy of language enters into this communication, and it may be that the imprecision of meaning between observers will somehow have to be built into physical theory. In this connection, an imprecision of a slightly different nature already exists in the formal transformation between the cellular space-times of two observers (Cole, 1972a); this imprecision is brought about by the observers exchanging only a limited amount of information concerning their placements of shared events within their own cellular structures.

5. *The Operational Natures of Space-time Structure and Distance Measurement*

As pointed out in Section 1, the space-time of every observer has a cellular structure brought about by the limited amount of energy he has available for refining his apparatus. In this section we will discuss more deeply the operational aspects of the cell boundaries and of extension measurement using measuring rods. We will also discuss just what is meant by a measuring rod which in some way is to be regarded as a 'straight' string of cells.

In fact, each observer has at his disposal two types of cellular structure. Firstly he has a *fixed* structure such that no matter how he moves relative to his surroundings this structure, once set up, remains fixed relative to the events which originally described it. For example, if he sets up a cellular structure to describe the events in a room he relates the cells to fixed objects in the room such as doors, windows, etc., and then describes other events by statements of the form 'the chair is near the window' which effectively puts the chair in the cell defined by the window. Then no matter how he moves around the room, this cellular structure will remain fixed relative to the door, window, etc., Another case of this fixed structure is that imposed on the space around the Earth by the graduated scale of a telescope sighting mechanism which is fixed to one point of the Earth: no matter how an observer moves away from the telescope this cellular decomposition always remains the same. In effect, a fixed cellular decomposition is one which the observer is unable, or unwilling, to move at will.

In contrast to this the observer also has at his disposal a *portable* cellular structure which he is able to move at will and which he is able to superimpose in any way on events. Such a portable structure is the basis of the rods he uses for extension measurement, and extensions between several pairs of events may be compared by using such a structure. An observer usually uses both types of structure at once; for example, a reading through his telescope may be plotted as a mark on a sheet of graph paper and the extension of this mark may then be recorded in terms of the cells of his measuring rod.

Just as we rule out the continuum in the sense that no observer uses a working continuum in his measurements, so we must rule out the idea of *point* as an infinitesimally small observable quantity (although Poston does retain the idea of point and defines distance between points in terms of a portable cellular structure). We must therefore rule out the *line*, regarded as an infinitesimally thin collection of points, as being a non-observable entity. Thus operationally the boundaries between an observer's cells cannot be regarded as lines which give hard edges to the cells. Rather, the boundaries will have a fuzzy nature and there will then be many situations in which the observer is unable to decide whether to place an event in one cell or another. In this way an uncertainty enters into the description of events and it may be that future theories will have to take this uncertainty into account. However, this uncertainty may be reduced by not forcing the observer to decide whether, say, an event occurs in cell n or cell m but by allowing him to say that the event is in the region defined by the set $\{n, m\}$ of the two cells. This regions is just a cell in some coarser cellular structure, and we should be able to construct our theories so that they can handle these regions and the basic cells in an equivalent manner. (Compare this with the very different natures of point, line and plane in a continuous geometry.)

In the following discussion the following notation will be used: n, m, etc. will denote the basic cells of the structure; N, M, etc. will denote sets of these basic cells, and $\bar{N}$, $\bar{M}$, etc. will denote regions defined by the sets N and M; these regions are themselves cells in some coarser cellular structure: $\mathscr{S}$ will denote the set of all such regions in the fixed cellular structure.

The idea of two regions overlapping may now be introduced, and the symbol o for the overlap relation may again be used because of the very close link with the sense of that symbol in Section 1. The overlap relation may be defined between any regions of $\mathscr{S}$ or the portable structure:

$$\left.\begin{array}{lll} \bar{N} \circ \bar{N} & & \\ \bar{N} \circ \bar{M} & \Rightarrow & \bar{M} \circ \bar{N} \\ \bar{N} \circ \bar{M} & \Rightarrow & \bar{N} \circ \bar{M}_1 \text{ for all } M_1 \text{ such that } M \subset M_1 \\ \bar{N} \circ \bar{M} & \Rightarrow & \text{there exists } n \in N \text{ such that } n \circ \bar{M} \end{array}\right\} \quad (5.1)$$

Operationally, the observer decides that two regions overlap if he is able to find at least one event which he is able to place simultaneously in both regions.

It is then possible to use a portable reference frame, denoted by r, to tabulate extensions between regions of $\mathscr{S}$. For each $\bar{M}$ and $\bar{N}$ of $\mathscr{S}$ and each portable reference frame r, define

$$D_r(\bar{M}, \bar{N}) \equiv \{(i, j)\colon i \text{ and } j \text{ are cells of } r \text{ such that } i \circ \bar{M} \text{ and } j \circ \bar{N}\} \quad (5.2)$$

Operationally this means that the observer places r over the regions $\bar{M}$ and $\bar{N}$ and notes that cells i and j of r overlap with $\bar{M}$ and $\bar{N}$ respectively. He then tabulates all such pairs (i, j) of incidences by moving r and hence forms the set $D_r(\bar{M}, \bar{N})$. This is perfectly possible operationally because since every r has a finite extension (Antoine & Gleit, 1971) and hence a finite number of cells the sets $D_r(\bar{M}, \bar{N})$ will all have a finite number of elements.

It then follows from (5.1) and (5.2) that

$$D_r(\bar{M}, \bar{N}) = \bigcup_{m \in M} \bigcup_{n \in N} D_r(m, n) \quad (5.3)$$

and that for each cell k of r,

$$(k, k) \in D_r(\bar{M}, \bar{M}) \quad (5.4)$$

Further, attention will be confined to those portable structures which have an unlimited freedom of movement in the sense that, operationally,

$$D_r(\bar{M}, \bar{N}) = D_r(\bar{N}, \bar{M}) \tag{5.5}$$

Next, two regions $\bar{M}$ and $\bar{N}$ are said to be *adjacent* if they do not intersect but it is possible to arrange any other portable cell q such that $q \circ \bar{M}$ and $q \circ \bar{N}$ simultaneously; again, this is perfectly acceptable operational definition. (But see the last paragraph of this section also.)

Having said that extensions between regions of $\mathscr{S}$ may be tabulated using a portable structure it is possible to move closer to the idea of such a structure as a 'straight' string of cells. To this end, define an ordered collection of cells $s_1, s_2, \ldots, s_i$ of a portable structure to be a *string* σ, denoted by $\sigma \equiv (s_1, s_2, \ldots, s_i)$, if s_j and s_{j+1} are adjacent. Thus a string is an ordered set of adjacent cells, and the definition allows the string to twist back on itself in the sense that s_j and s_k may overlap if $|j-k| > 1$.

Next, if $\sigma \equiv (s_1, s_2, \ldots, s_i)$ is any string, define the set

$$\Sigma(\sigma) \equiv \{\sigma' : \sigma' \text{ is a string } (s_1', s_2', \ldots, s_i') \text{ such that for each portable structure } r,\ D_r(s_k', s_k') = D_r(s_k, s_k) \quad (k = 1, \ldots, i)\} \tag{5.6}$$

That is, $\Sigma(\sigma)$ is the class of all strings which have the same number i of cells as σ, and corresponding cells of the strings all have the same extension with reference to all portable structures. Effectively, we form the class of strings which have the same number of cells and same 'total length' in this sense. It follows from (5.6) that

$$\sigma \in \Sigma(\sigma)$$

and

$$\sigma' \in \Sigma(\sigma) \Rightarrow \Sigma(\sigma') = \Sigma(\sigma)$$

It is then possible to define the 'straightest' string of the class Σ, called a *ruler*, as that element $\rho \equiv (\rho_1, \ldots, \rho_i)$ such that for each region $\bar{L}$ of any fixed cellular structure,

$$\bigcup_{k=1}^{i} \{m : m \in \mathscr{S},\ s_1 \circ \bar{L},\ s_k \circ m \text{ for all } (s_1, \ldots, s_i) \in \Sigma(\rho)\} \subset \bigcup_{j=1}^{i} \{n : n \in \mathscr{S},\ \rho_1 \circ \bar{L},\ \rho_j \circ n\} \tag{5.7}$$

This says that with one end ρ_1 of the ruler overlapping with $\bar{L}$, the region swept out by the ruler is not smaller than that swept out by any other string in $\Sigma(\rho)$. This definition allows for the fact that there may be more than one ruler in a class.

It is really an assumption to state that such a ruler exists in each class, but in a list of instructions for the construction of a ruler this assumption will not be stated explicitly. It is then up to the observer who tries to carry out the instructions to decide whether or not the assumption, according to his experience, is valid.

Having defined a ruler it is possible to define a *uniformly celled* ruler $\rho = (\rho_1, \ldots, \rho_i)$ such that

$$D_r(\rho_k, \rho_k) = D_r(\rho_j, \rho_j) \qquad (k, j = 1, \ldots, i) \tag{5.8}$$

for each portable structure r. For each uniformly celled ruler ρ it is then possible, by defining a real parameter ξ_ρ for each ρ, to use it to tabulate the extensions between regions of $\mathscr{S}$ in terms of sets of real numbers in the form

$$d_\rho(\bar{M}, \bar{N}) \equiv \{\xi_\rho |n - m| : (\rho_n, \rho_m) \in D_\rho(\bar{M}, \bar{N})\} \tag{5.9}$$

It follows from (5.5), (5.4) and (5.3) respectively that for each uniformly celled ruler ρ and all $\bar{M}, \bar{N} \in \mathscr{S}$,

$$d_\rho(\bar{M}, \bar{N}) = d_\rho(\bar{N}, \bar{M})$$
$$0 \in d_\rho(\bar{M}, \bar{M})$$
$$d_\rho(\bar{M}, \bar{N}) = \bigcup_{m \in M} \bigcup_{n \in N} d_\rho(m, n)$$

and we may add the further assumption that

$$\min d_\rho(\bar{M}, \bar{N}) \leqslant \max d_\rho(\bar{M}, \bar{L}) + \max d_\rho(\bar{L}, \bar{N})$$

These are just the results produced without derivation by Cole (1972a). In this way, the continuum notion of a single real-valued distance between two points is replaced by the notion of a set of real valued distances between regions. Allowing for the fact that two observers each with their own rulers will want to compare their distance sets (possibly by means of the overlap relation of Section 1), the values of ξ_ρ may be adjusted by comparing their rulers with some standard structure. The greater the number of cells in a ruler the smaller will be the value of ξ_ρ, and so this value is an observer-dependent parameter, the value of which is related to the amount of energy the observer has for refining his ruler.

Finally, it is necessary to discuss the operational status of the equations appearing in this section. All the quantities used—for example, m, $\bar{M}$, $D_r(\bar{M}, \bar{N})$ and $d_\rho(\bar{M}, \bar{N})$—are the actual quantities that the observer uses, sets up observationally, or is able to calculate from his observations. But phrases of the form 'for all portable structures r' must have more subtle interpretations if they are to be considered operationally. In the usual mathematical sense, 'for all r' is interpreted as 'for all r that has been used and all possible r that could ever be used'. But in practice an observer is not able to wait an indefinite period to test the validity of his equations with every possible r. Thus, for example, in the definition (5.8) of a uniformly celled ruler, he tests both sides of the equation with the limited number of portable structures he has available. The equality is inserted only when he is reasonably satisfied that it holds for all the r available to him. Exactly the same consideration applies to the words 'any other portable cell q' in the definition of adjacent regions. The words 'for all' and 'any other' which appear in all the equations must then be interpreted in this operational sense.

References

Antoine, J.-P. and Gleit, A. (1971). *International Journal of Theoretical Physics*, Vol. 4, No. 3, p. 197.

Atkin, R. H. (1971). *Quantum Theory and Beyond*, ed. Bastin, E. W., Cambridge University Press.

Bohm, D., Hiley, B. J. and Stewart, A. E. G. (1970). *International Journal of Theoretical Physics*, Vol. 3, No. 3, p. 171.

Cole, E. A. B. (1972a). *International Journal of Theoretical Physics*, Vol. 5, No. 1, p. 3.

Cole, E. A. B. (1972b). *International Journal of Theoretical Physics*, Vol. 5, No. 6, p. 437.

Davies, E. B. and Lewis, J. T. (1970). *Communications in Mathematical Physics*, **17**, 239.

Meredith, P. (1966). *Instruments of Communication*. Pergamon Press.

Mielnik, B. (1969). *Communications in Mathematical Physics*, **15**, 1.

Poston, T. (1971). *Manifold*, **10**, 25 (Universiy of Warwick).