MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES
3 4456 0060600 2

ORNL-5908

ornl

OAK
RIDGE
NATIONAL
LABORATORY

UNION
CARBIDE

Application of Group Theory to Data Reduction

F. G. Perey

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
CIRCULATION SECTION
4500N ROOM 175
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this report, send in name with report and the library will arrange a loan.
UCN-7969 (3 9-77)

OPERATED BY
UNION CARBIDE CORPORATION
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes—Printed Copy: A05; Microfiche A01

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL-5908

Engineering Physics Division

APPLICATION OF GROUP THEORY TO DATA REDUCTION

F. G. Perey

Manuscript Completed - July 15, 1982

Date Published - September 1982

Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37830
operated by
UNION CARBIDE CORPORATION
for the
Division of Basic Energy Sciences
U.S. DEPARTMENT OF ENERGY
Under Contract No. W-7405-eng-26

MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES

3 4456 0060600 2

CONTENTS

Preface v

Abstract 1

1. Introduction 2

2. Application of Group Theory to Data Reduction 8
   2.1 Physical Theories and Observations 8
   2.2 Data Reduction and the Scientific Method 9
   2.3 The Possibilities-Generating Group 12
   2.4 The Probability Measure 15
   2.5 Summary of the Proposed Method of Data Reduction 17

3. Concrete Realizations of Possibilities-Generating Groups 20
   3.1 The Coin Experiment 21
   3.2 The Playing Die Experiment 23
   3.3 The Balls in a Box Experiment 25
   3.4 When? 26
   3.5 How Fast? 27
   3.6 Von Mises' "Water and Wine" Problem 28

4. The Fundamental Theorems of Physical Probabilities 31
   4.1 Notation 31
   4.2 The Disjunction Theorems 35
   4.3 The Complement Theorem 36
   4.4 The ConjunctionTheorem 36
   4.5 Bayes Theorem 38
   4.6 The Sum Rule 39
   4.7 Remarks 40

5. Physical Probabilities of Outcomes and Limiting Frequencies 41
   5.1 Conceptual Sampling 41
      5.1.1 The number on the first ball drawn 42
      5.1.2 The color of the first ball drawn 43
      5.1.3 The first two balls drawn are black 44
      5.1.4 The first r balls drawn are black and then m-r white balls are drawn 45
      5.1.5 In m draws r black balls are obtained 45
   5.2 Conceptual Limiting Frequency 46

6. Physical Probabilities of Physical Probabilities 50
   6.1 Discrete Probability Space 51
   6.2 Continuous Probability Space 55
   6.3 Physical Probability Spaces and Their Group Structure 58

7. Bernoulli Trials 61
7.1 More Black and White Balls in a Tube 62
7.2 Criterion for Bernoulli Trials 66
7.3 General Remarks 68

8. Conclusions 70

9. Acknowledgements 73

References 74

PREFACE

A primer to group theory was drafted to be included as an appendix to this report. Then it was thought that it could be useful to experimentalists who might want to apply the proposed method of data reduction. This idea was abandoned when the draft grew to be the size of this report, and it was felt that there was a major shortcoming in it. We had not succeeded, to our satisfaction, in presenting in an elementary but meaningful fashion the essential notions of topology and measure theory required for the proposed use, in this report, of the group manifold and its invariant measure as well as deal with continuous groups. The essential difficulty with drafting this appendix was that we did not know what to assume as the mathematical background of its readers. If we set this background too low, we would have to write a book presenting the modern approach to mathematics instead of writing this appendix. We feel there are many people more qualified to write such a book if there is a demand for it. However, since an appendix was drafted, it is available in its current draft form from the author upon request.

APPLICATION OF GROUP THEORY TO DATA REDUCTION

F. G. Perey

ABSTRACT

The analysis within the framework of a theory of what was observed in experiments is essential to the testing of theories and is fundamental to physics. It is shown in this report how group theory can be used to provide a general method of data reduction whereby only the laws of a particular theory are used in the analysis of observations. This application of group theory involves introducing a group of transformations of the physical system upon which the observations were made. This group of transformations leaves invariant the entities of the theory corresponding to the observations made but transforms the entities that were not observed from what they are presumed to be in this theory into what they are not, called possibilities for what they are. This group of transformations is called the possibilities-generating group for the entities for which the observations are being reduced. Since possibilities for entities of theories so obtained are functionals of a known group of transformations, the subsequent use of these possibilities in the theories must be made consistent with the theory of representations of groups. There is a well-known invariant associated with functionals of group elements which is invariant with respect to the parametrization of this group. It is the normalized volume measure in the group manifold of the possibilities-generating group. In this report this invariant measure is called the physical probability of the possibilities since it is a probability measure which has a Borel algebra. The above proposed method of data reduction allows us to deal unambiguously with the uncertainties in entities of physical theories obtained from all of the observations made in experiments and the distinction between systematic and statistical uncertainties disappears. Concrete realizations of possibilities-generating groups are given and explanations in group theoretical terms are offered for several important intuitive notions related to probabilities.

1. INTRODUCTION

The broad issue addressed in this report is the concept of testability of physical theories by observations made in experiments. For centuries, physicists have drawn from the observations they make implications in the theories they propose. Therefore, this activity is not new and is a legitimate subject of inquiry for physicists. Possibly because physicists have had such a long experience at this activity, there is not much concern expressed in the physics literature about the manner in which this task is carried out today. Consequently, this report does not address an issue recognized as a problem in physics today. In fact, it may be called a dead issue and this forces us to approach it from fundamental considerations in this introduction.

The statement "physical theories are testable proposed explanations for observed physical phenomena" is a philosophical statement about physics and is not too controversial to be used in a physics report. It is the purpose of this report to show that if one accepts at face value the above statement one can justify, and even demand, that group theory be the tool used to analyze the implications in a theory of what was observed in an experiment. The drawing of implications in a theory from what was observed in an experiment is called here "performing the data reduction." The tool currently used to estimate uncertainties in data reduction is the theory of statistical estimation, but it cannot deal with some of the observations said to lead to systematic uncertainties. It is shown in this report how group theory allows us to deal quantitatively in a coherent fashion with the uncertainties in data reductions that arise from all observations.

In order to develop some operational procedures useful to do some physics from the statement that physical theories are testable proposed explanations of observed physical phenomena, it is expedient to consider briefly a possible philosophical justification for this statement. A little over 300 years ago, Leibniz developed a philosophical system that has had a profound influence in the development of science and, in particular, physics. This philosophical system is based largely upon two main principles of interest to us here: the Principle of Contradiction and the Principle of Sufficient Reason. The first of these principles states that every assertion is either true or false, and the second one states that nothing happens without a reason. The Principle of Sufficient Reason can be regarded as the basis for the

notion of causality used in physical theories that are deductive systems of assertions from some hypotheses. The Principle of Contradictions plays a role in the choice of axioms used to develop theories for analyzing observations by demanding that they do not give rise to internal contradictions. The Principle of Contradiction is also the basis for the notion of testability of physical theories by experiments, if one considers that what is observed in experiments establishes by some process of identification the logical truth of an assertion of these theories dealing with observables. While the above is not physics and likely not philosophy either, it is offered as a "justification" for two points which allow us to apply group theory to perform data reduction. The first point is that physical theories are causal since they are deductive systems of assertions based upon some axioms. In these theories every assertion is either true or false, and they are free of internal contradictions. The second point is that observations made in experiments are used to establish the logical truth of some assertions of physical theories by a process of identification with what is observed. Whether one considers that these two points have been given adequate justification is open to question, and more modern-sounding arguments could have been offered. The philosophical justification of these two points is not our main concern here. We want to analyze the implications for physics of these two points assuming that they appear reasonable enough to warrant such an analysis.

Group theory plays such an important role in contemporary physical theories that it is presumed familiar to the reader. The idea of using group theory to deal unambiguously with ambiguity is almost as old as group theory itself. Our concern in this report is with the ambiguities that arise when one uses a physical theory to analyze the observations made in experiments. As is well known, 150 years ago Evariste Galois made important contributions to the development of group theory, and the circumstances under which these were made are also well known.[1] In 1832, on the eve of the duel in which he was killed, Evariste Galois wrote a celebrated letter to his friend Auguste Chevalier - a letter from which his most important contributions to group theory are known to us. In the last paragraph of this letter dealing with technical matters, Galois states that his main preoccupation for some time had been the application of the theory of ambiguity to transcendental analysis. To our knowledge, no clues have been found as to what Galois was referring to here except for what he then states in the next two

sentences. These two sentences freely translated are: "The object was to determine *a priori* in a relation among quantities or transcendental functions what transformations could be made, what quantities could be substituted to the given quantities, without changing what the relation was. This enables us to recognize right away the impossibility of many expressions that one would otherwise seek." There is little doubt from these statements that Galois' theory of ambiguity was a mathematical theory and not a physical theory. However, one can conjecture, on the basis of those two statements alone, that Galois' theory of ambiguity may have been closely related to the method of data reduction proposed in this report.

The idea that Galois' theory of ambiguity was group theory and could be used to deal with uncertainties resulting from observations is not new. In a very interesting discussion of the Leibniz Principle of Sufficient Reason, George Birkhoff[2] draws a relationship between group theory and the Principle of Sufficient Reason. In this paper, after citing the above mentioned paragraph in Galois' letter, he states: "This passage shows how in Galois' mind the theory of groups and the theory of ambiguity were interrelated. Unfortunately, later students of the theory of groups have all too frequently forgotten that, philosophically speaking, the subject remains neither more nor less than the theory of ambiguity." Birkhoff then shows by an elementary example the idea of a group and of the associated ambiguity. In the terminology to be introduced formally in the next section, Birkhoff states that the possibilities-generating group for the positions of the vertices of a flat square object, a tile, observed to lie aligned on a square grid, is the group of symmetries of the square: the cyclic group of order 4. However, even though Birkhoff does mention probability theory in this lecture, it does not appear that he pursued further these considerations to the point of developing formally the quantification by rational numbers of possibilities generated by a group of transformations using strictly group theory as is done in this report.

Many different theories of probabilities have been proposed.[3,4] In a few of these, group theory is explicitly used,[3] but in many more, developed before and since 1832, invariance arguments of one form or another that could be easily construed to be group theoretical arguments are made use of.[3,5] The essential characteristic of all these attempts at using group theory, or invariance arguments of a group theoretical nature, is that group theory is used only as a device for assigning an *a priori* probability to

possibilities within the framework of theories of probabilities that owe little or nothing to group theory. To our knowledge, group theoretical arguments have only been used in connection with classical or Bayesian theories of probabilities based upon the notions of "degrees of beliefs" or "degrees of ignorance" measured by probabilities. Recently, Jaynes[6] proposed a formal principle to justify the application of group theory to carry out a type of invariance argument first introduced by Poincaré[7] in 1895. Jaynes calls this principle a Desideratum of Consistency, and it is to be used in the assignment of *a priori* probabilities. This Desideratum of Consistency is: "In two problems where we have the same *a priori* information we should assign the same *a priori* probability." Jaynes[8] advocates the use of these *a priori* probabilities with the algebra of probabilities of Cox.[9] Cox's algebra of probabilities is generated from three essential axioms by a type of invariance argument built upon the symmetry properties of the operators of Boolean algebra. This led him to some functional equations, earlier considered by Abel,[10] whose solutions became the "laws" of his algebra of probable inference. The Cox-Jaynes theory of probabilities represents, to our knowledge, the theory of probabilities that makes the greatest use of invariance type arguments to date and was the starting point of our investigations.

Poincaré's views on the nature and meaning of probabilities are well known from an essay he devoted to them and which he used as the introduction to the 1912 edition of his book.[7] It does not appear from this essay that he gave great weight to the role that group theory may play in probability theories. However, he does mention, as would Birkhoff later, its usefulness in a particular problem. This problem is the shuffling of a deck of playing cards. There is no other specific mention of group theory throughout this book except in the last chapter, also added to the 1912 edition. In this last chapter, Poincaré gives a detailed analysis of the problem of shuffling a deck of cards based entirely upon group theory, including the use of elements of group algebra. It is of some historical interest that at the end of this analysis Poincaré refers to one of his papers:[11] "On the algebraic integration of linear equations and the periods of Abelian integrals," the very subject of Galois' lost "third manuscript" which Galois mentions[1] in the letter quoted from earlier.

The above references on the use of group theory in probability theories do not constitute an exhaustive survey of the literature. They merely serve to indicate that the idea of using group theory to deal

with the notion of uncertainty or ambiguity, most often in connection with theories of probabilities, seems to have occurred to some notable mathematicians such as Galois, Poincaré, and Birkhoff. It seems fair, however, to say that these attempts have not yet provided a group theoretical foundation to the notion of probability as used in physics as is done in this report.

In Sect. 2 of this report a method of data reduction is presented in a completely general and abstract context. This method should, therefore, be applicable to the analysis of observations by any testable physical theory proposed as an explanation for the phenomena involved. Group theory is used to provide an explanation, within the framework of the physical theory used to analyze the observations, of the fact that these observations yield several values for an entity of this physical theory related to the observations, even though in this physical theory this entity should have only one value. This problem occurs because when one applies the proposed theory to analyze the abstraction corresponding to the physical system, using only what was observed of it, this abstraction exhibits elements of apparent symmetry with respect to the entity of interest that was not observed. Group theory is used to deal with this element of apparent symmetry. The group of transformations of the abstraction for the physical system in the theory used is called a possibilities-generating group for the entity of interest. Its elements generate, from the value which the entity of interest is postulated to have in the physical theory used, values which this entity does not have which are called possibilities for the value of this entity. The identity element introduces the value which the entity has in the set of possibilities. The possible values for the entity of interest then become in the physical theory used functions of the elements of a group and must be treated as such. Consequently, there are "weights" that must be attached to these possible values, weights which are uniquely determined by the group structure of the possibilities-generating group. These weights are now normally called in group theory the volume measure in the group manifold. The volume measure at each "point" in the group manifold can be normalized to unity when summed over the group manifold. The normalized volume measure at each "point" of the group manifold is called here the physical probability of the possibility for the entity of interest corresponding to this point.

Through the remainder of the report, properties of these physical probabilities are developed and interpreted.

In Sect. 3 concrete realizations of possibilities-generating groups are given, and it is shown that the proposed method of data reduction yields unique physical probabilities in specific situations where the theory of statistics cannot be used.

In Sect. 4 a notation is introduced for designating the normalized volume measure in the group manifold of a possibilities-generating group. This notation for the physical probabilities of possibilities for entities of theories obtained from observations is then used to show that their algebra is the algebra of probabilities: a Borel algebra. This is merely the consequence of the additive nature of the volume measure in a group manifold.

In Sect. 5 the physical probability of the possibilities for the outcome of an experiment which can be characterized by a frequency is obtained. It is then shown that, according to the theory used to analyze this experiment, the limiting frequency with which a particular outcome would be observed if the experiment were carried out repeatedly is equal to the physical probability of this outcome. This result shows that the normalized volume measure in a group manifold of the possibilities-generating group can be interpreted in the theory used as a conceptual limiting frequency.

In Sect. 6 it is shown that the normalized volume measures in a group manifold constitute an abstract topological space that has a well-defined group structure. This leads to the notion of an abstract physical probability space which is imbedded in a subspace of a real linear space. Consequently, one can always associate a physical probability distribution to the physical probability of a possibility for the value of an entity of a theory. It is suggested that this result is a proper explanation to the widely held intuitive notion that most probabilities are more or less uncertain; they are said to be estimated.

In Sect. 7 the concept of Bernoulli trials is briefly analyzed from the point of view of the proposed method of data reductions in order to show that this method can also deal with observations from repeated measurements.

2. APPLICATION OF GROUP THEORY TO DATA REDUCTION

A formal methodology is presented in this section for obtaining the implications in an existing physical theory of what was observed in an experiment. A physical theory is taken to be a proposed explanation of observable phenomena. Because what one obtains from observations in experiments is called data and the analysis of these observations is called data reduction, the aim of this section is to present a method of data reduction; it could also be called a method of data evaluation. Why one cannot develop a *theory* of data reduction but *only* a *method* of data reduction is first discussed. The brief philosophical discussion in the introduction is offered as a justification for the starting point of the discussion.

2.1. Physical Theories and Observations

In the introduction physical theories were considered as deductive systems of assertions from some hypotheses, or axioms. However, it is more conventional in physics to use the terminology of mathematics than the one of logic when discussing physical theories, and this convention is followed. From this point of view, physical theories are considered as consisting mostly of mathematical relations that are postulated to hold among some abstract entities of these theories. That physical theories can be so considered is the consequence of the adoption of a convenient and very powerful symbolic notation. In order for the mathematical formulae that stand for physical theories to be construed as an explanation of observed physical phenomena, some of the abstract entities appearing in them must be "observables" of these physical theories. That is to say, an equivalence relation must exist between what is observed in an experiment by physicists, ultimately via their senses, and some of the abstract entities called observables of physical theories. Only through the presence of such equivalence relations in physical theories can observations made by physicists be transcribed into data of physical theories and vice versa. These transcription rules are formally axioms of physical theories. They will be assumed to have been correctly applied to the observations.

Through the use of the above transcription rules, what is observed generates in the symbolic notation of a particular theory a mathematical relation that holds, or is logically true, in this physical theory

since some specific observations were made. These mathematical relations generated from observations involve only observables of physical theories. In the conventional language of experimental physics, these mathematical formulae are the data of physical theories obtained from experiments.

These two types of mathematical relations, the general relations of the physical theories and the relations derived from observations, must satisfy some conditions if they are to be a part of physical theories. Since this paper is not concerned with what physical theories are in any detail except for some specific examples used as illustrations, these details are not discussed here. The above description of physical theories and observations is not presumed to be controversial and serves to define a vocabulary whereby one can discuss in an abstract way physical theories and observations made in experiments.

2.2. Data Reduction and the Scientific Method

On the basis of the above considerations, the activity that has been called data reduction can be formulated in mathematical terms. Let us consider a particular physical theory, T, in which a mathematical relation written

$$T\ (D,C) = 0 \tag{2.1}$$

holds, where D is an observable of this theory and C may be, but need not be, an observable of this theory.

Let us now consider a specific experiment, E, for which the theory T is a proposed explanation. In this experiment a particular observation is made which, after using correctly the appropriate transcription rule of the theory T, yields the mathematical relation of the theory T:

$$E_T(D) = 0\ . \tag{2.2}$$

What is called data reduction in this paper is the analysis of the solutions to the system of relations (2.1) and (2.2). In the performance of this analysis one must remember that strictly speaking it is a physical problem that is being solved and not a mathematical one. What this implies is that whatever is done mathematically can and must, if necessary, be verbalizable in physical terms from the point of view of the postulated physical theory T being applied. The particular concern of this paper is that in

many, if not most, experiments there may be several mathematical solutions to the system of relations (2.1) and (2.2) for some entity of the theory T, symbolized by C above, that is postulated in this theory to have a unique if unknown value.

A fundamental task that faces physicists is to provide a physical explanation in terms of the theory T for the fact that when it is properly applied it yields several values for an entity which it postulates must have a single value, although the mathematical reasons for this are obvious. In this paper, it is chosen to consider that the above is an "apparent difficulty" until a suitable explanation for it has been found in terms of the theory T even though its mathematical explanation is no mystery. That the above result frequently occurs is a generic problem in science and we are used to it, that when it occurs we say we are uncertain of the value of the entity and in some situations this uncertainty is of no practical significance, that the mathematical reason for its occurrence is understood, and, finally, that when it occurs in an experiment this does not invalidate the theory used because it has not been falsified, are all statements of facts or value judgements and do not constitute an explanation for its occurrence in terms of the physical theory T.

Can an explanation of this apparent difficulty be provided entirely from the viewpoint of the proposed testable physical theory T, or must one invoke some general, or specific, physical considerations X that are external to the physical theory T being applied? A little reflection shows that if the physical theory T is claimed to be testable and the scientific method demands that it be, then an explanation for this apparent difficulty must be provided strictly on the basis of the physical theory T. The reason for this is that the physical considerations X, if they are not implied by the physical theory T but used with this physical theory T, transform the physical theory T into the physical theory TX. Therefore, the physical theory T ceases to be when one introduces physical considerations not implied by its axioms. Then it cannot properly be said that it is the physical theory T which is being used or tested.

The above simple argument suggests that there cannot be a theory of physical measurement and a theory of data reduction which are independent of physical theories. This statement can be rephrased to say that every physical theory must imply a method of measurement and a method of data reduction. Such conclusions have been reached before from a different rationale.[12] In this paper the method of

measurement is bypassed and its end product is considered: it must yield a statement of the physical theory of the type (2.2). This allows us to focus upon the issue of the method of data reduction.

Since the apparent difficulty which is the concern of this paper (the fact that in some situations there may be several solutions where the physical theory applied implies there should only be one) is related to the idea that observations lead to uncertainties in reduced data, the above conclusion is at first disturbing. It is disturbing because a theory of data reduction has been used in physics for a long time. This theory of data reduction is independent of the physical theories used that are presumed to be testable. It is used with all physical theories and allows physicists to deal with the notion of uncertainties in physical data obtained from observations. This theory of data reduction is the theory of statistical estimation based upon the theory of statistics. If one attaches any weight to the above arguments, then one must seek to eliminate from physics the theory of statistical estimation in spite of its apparent success, its pervasiveness, and its resulting wide acceptance in the physics community today. The theory of statistical estimation must be replaced by a method of data reduction that is applied with and implied by testable physical theories. Furthermore, this method of data reduction cannot have any physical content of its own. Whatever physical notions are involved in this method of data reduction, they must be provided by the physical theory being applied.

The fact that it appears one must give up, at least in physics, the theory of statistics and the theory of statistical estimation is not as disturbing as it may seem at first, even though these theories play an important role today in the testing of physical theories and in many physical theories including those perceived to be most successful such as statistical mechanics, quantum mechanics, etc. The arguments used to reach the above conclusions only pertain to the physical meaning one attaches to the formulae one identifies with the theories of statistics and statistical estimation. These arguments say nothing whatsoever about the nature of the formalism that may result from a method of data reduction. From the pervasive nature throughout physical theories and other sciences of the formulae one justifies today on the basis of statistical arguments, one should suspect that many of these formulae will reappear but be justified on a totally different basis. Because of this, it is important to find out whether the whole

exercise which this paper is about is not a purely semantic one. The arguments given in this paper will be a more useful semantic than the current one if it can be shown that it allows us to solve, in addition to the problems we can cope with today, a clearly recognized important class of problems for which there is currently no satisfactory solution if one adopts the viewpoint of the theory of statistics. In this paper it will be shown that the semantic point of view proposed allows the exact quantification of the systematic uncertainties, as well as the exact quantification of the "statistical" uncertainties from a unified point of view that involves no physical postulates, and this cannot be done with the theory of statistical estimation.

2.3. The Possibilities-Generating Group

A method of data reduction based upon group theory is now presented. The fundamental idea of the method is explained in an abstract context. That is to say, neither the experiment, nor the observations made in the experiment, nor the proposed testable physical theory being used to analyze the observations are specified. Concrete examples are given in the next section.

The basic idea of the proposed method of data reduction is to look at the physics experiment that has been performed strictly on the basis of what was observed and from the viewpoint of the theory used to analyze these observations. Group theory is used to deal with the fact that even though in the physical theory used some entity should have only a single value, on the basis of the observations there are several mathematical values for this entity. To deal with this problem means to provide a physical explanation for this situation within the framework of the physical theory being applied.

A mathematical equation can have several different solutions. In mathematics none of these different solutions have anything special over the other ones. All of the solutions satisfy the given equation, and they collectively form the set of solutions of the equation. A particular terminology has been introduced to refer to the several solutions that are found in the problem of concern to us. The set of solutions found on the basis of the observations is called an exhaustive set of mutually exclusive

solutions. This statement conveys the notion that only one of the several solutions to the mathematical relations can be the value of the entity of the physical theory since it postulates there is only one value to this entity. The statement that we find an exhaustive set of mutually exclusive values for the physical entity is neither a statement of mathematics nor a statement of the physical theory used, it is a commentary acknowledging what we have called an apparent difficulty. This commentary labels what the apparent difficulty is: the set of mathematical solutions is not the singleton that the physical theory postulates. This commentary states that we have run into a mathematical impossibility and because it is the physical theory that created it, the physical theory should resolve it. The mathematical relations that resulted when the observations were encoded into the physical theory, together with the body of mathematical relations that stand for the physical theory, contained the statement that the set of solutions had to be a singleton. This was mathematically impossible; therefore the requirement that the set of solutions had to be a singleton was ignored and made a commentary to the set of several solutions found. The difficulty arose because a mistake was made. Since the physical theory uses the language of mathematics, after the observations were encoded into mathematical terms the solutions were sought as if we were dealing with a purely abstract mathematical problem ignoring the requirement that the set of solutions had to be a singleton. Once the several solutions were found, the abstract mathematical problem was considered to be a physics problem and the condition that the set of solutions had to be a singleton inserted. This yielded a commentary to that effect: the set of mathematical solutions is an exhaustive set of mutually exclusive solutions to the physics problem. We must correct this mistake and consider at the outset that the set of solutions must be a singleton, this statement is an essential aspect of the physical theory for which the mathematical relations stand.

It is now shown how group theory provides the tool to accomplish the above task. Before the method is verbalized in physical terms, a well-known result of group theory is recalled. This result was established 150 years ago by Galois in an important contribution to group theory. Certain mathematical equations have several different solutions. It is always possible to associate with these equations a group of transformations whose order is equal to the number of solutions. This group of transformations has the property that by starting from any one of the solutions one can generate all of the other

solutions. Another way of saying this is that each of these equations is the expression of an underlying symmetry and if there are n different solutions it is because this symmetry is n-fold. It will not be proven in this paper that the kinds of equations we are dealing with in physics have this property. This will be taken for granted since in modern physics one thinks of the underlying symmetries referred to above as more fundamental than the equations, the "laws of physics," themselves. One even says that the equations, the "laws of physics," are derived on the basis of the symmetries rather than the reverse. From a mathematical point of view, there is an equivalence relation between the symmetries and the equations.

The above mathematical result is all that is needed to verbalize a method of data reduction entirely in terms of the physical theory being used. This physical theory must be assumed correct until it has been proven that it cannot be correct. Let us consider a particular theory and a particular physical situation in which something is observed. Let that which was observed be properly encoded into the language of this theory and assume that according to this theory there is an entity which was not observed but is related to the observations and that we are interested in its value. If, in this physical theory, it is possible to find a group of physical transformations whose elements change the value of this entity into values it does not have without changing what was actually observed in this physical situation, then several possible values will be found for this entity on the basis of what was observed. There will be as many possible values as there are ways of changing the physical situation in the physical theory used, plus the identity transformation. Each element of the group of physical transformations, except the identity transformation, changes the value of the entity into a value it does not have, generating a possibility for its value. The set of physical transformations forms a group in the usual sense. For obvious reasons, this group of physical transformations in the physical theory being used is called the possibilities-generating group for the value of the entity in question. An equivalent way of stating the above is that on the basis of what was observed from the point of view of the physical theory applied, there is an apparent symmetry in the physical situation with respect to this entity of interest. It is very important to emphasize that this symmetry in the physical situation is only apparent from the

point of view of the logic of the physical theory applied and what is identified in this physical theory with what was observed.

The above is offered as an explanation for why there can be several possible solutions on the basis of some observations for an entity of a physical theory being applied to analyze these observations when, in this physical theory, there can only be one solution for this entity. It is clear that two different physicists, having made the same observations in a physical situation and having applied the same physical theory, will find the same possibilities-generating group. It is also clear that what is observed in a particular physical situation, when analyzed by two different physical theories which have a corresponding entity, may very well yield a different possibilities-generating group for this entity in each physical theory. Since the logics being used are different the same physical situation may have different apparent symmetries.

The above constitutes a very broad generalization of the specific instances used by Poincaré and Birkhoff to indicate how group theory can be used to deal with situations that lead to uncertainty.

2.4. The Probability Measure

The preceding discussion focussed upon providing a rationale using group theory for the notion of an exhaustive set of mutually exclusive possibilities for an entity of a physical theory that result from observations. Because the set of transformations of the physical situation by the laws of the theory used forms a group, it is not possible to tell which of the possibilities is the value of the entity of the physical theory used. Every different possibility for this abstract entity is generated by an element of the group of transformations; each possibility can be associated with an element of the group. If the possibilities-generating group is of order n there are exactly n different ways of associating the elements of the group with the n possibilities. These different ways are provided by the group multiplication table of the group. Therefore, every possibility can be associated with, mapped into, every element of the possibilities-generating group. Consequently, every mathematical formula of this physical theory in which this entity enters must be regarded as a functional of the elements of the possibilities-generating group for this entity. These functionals of the group elements must have an invariant meaning in the physical theory whenever a particular possibility for the entity is used. It was, in fact, through the

analysis of particular functionals of solutions of algebraic equations that Galois was led to important discoveries in group theory.

At this time it is useful to recall a well-known mathematical result of group theory that is at the foundation of the theory of representation of abstract groups by linear substitutions. In order to provide meaning to functionals of group elements, it is necessary to adjoin to each abstract group an abstract space called the group space, or the group manifold. This group manifold is a topological space in which there is a one-to-one correspondence between some "points" in this topological space and the elements of the group. It is through the introduction of the group manifold that functionals of group elements acquire the meaning of functionals defined in a space that represents the abstract group. Without going through the mathematical details found in every textbook on the theory of representations of groups that deal with both finite and continuous groups,[13,14] the fundamental theorems upon which group algebra is based depend upon the introduction of a "volume measure" in the group manifold. This volume measure in the group manifold is defined as the density of points in the neighborhood of each point in the group manifold. Because the group manifold of a finite group is a disjoint topological space the volume measure in the group manifold of a finite group is uniform, i.e., the same at every point in the group manifold. In the case of a continuous group, the volume measure in the group manifold is not necessarily uniform over the group manifold. Of interest to us is the fact that the volume measure in the group manifold of a group, being defined as a density of points in the group manifold, is a positive quantity that is additive. This volume measure can be summed over all the points of the group manifold to yield a finite measure, in view of the closed nature of the group manifold that results from the closure property of the elements of the group. Therefore, it is always possible to normalize the volume measure in a group manifold to yield unity when algebraically summed, a Riemannian integral in the case of a continuous group, over all the group manifold. In mathematics a measure which is introduced in an abstract space and has the above properties of being positive definite, additive, and normalizable to unity when summed over the complete space is called a probability measure. The terminology of volume measure in the group manifold of a group that has been used above is often called

the "volume element" in the group manifold[14] or the "weight function" in group space.[13] By whatever name it is called in group theory, it is a probability measure[15] in mathematics. The normalized volume measure in the group manifold of a group is uniquely defined by the structure of a group. Although it plays an important role in defining the Frobenius algebra of a group, the volume measure is not an element of this group algebra in the conventional definition of the elements of this algebra. However, it is possible from the definition of the volume measure in a group manifold to define an algebra of this volume measure and, as it must be since this volume measure is a probability measure, its algebra is a Borel algebra.

The relevance of the above mathematical considerations to the method of data reduction based upon the application of group theory is that, since there is a unique group of transformations that generate the possibilities for an entity of a theory from the observations, there is also a unique probability measure that one can associate with these possibilities.

What has just been accomplished in a very general and abstract way is to provide a group theoretical foundation for the notion of possibilities, and their associated probabilities, for entities of theories determined from observations. Because the probabilities of the possibilities for entities of theories defined above occur as a result of the interpretation of what was observed in a physics experiment within the framework of a physical theory, it is appropriate to refer to these probabilities as physical probabilities to distinguish them from the usual statistical probabilities.

2.5. Summary of the Proposed Method of Data Reduction

The operational method that emerges from the above discussion is now briefly summarized. If there are any implications in a theory of what was observed in an experiment, it must be that in this theory what was observed is not everything that there is in this physical situation. For instance, according to the theory some other observations could have been made but were not made or could yet be made. It may also be that in this theory what was observed is related to some abstract entities that are not observables of this theory. Whatever the case may be, if the theory used postulates that there are entities which were not observed but are related to what was observed, data reduction is the analysis of

what can be said about these entities on the basis of what was observed and the postulated laws of this theory. The basic idea is to apply group theoretical methods to perform this analysis so as to use only the laws of physics of this theory. This is done by seeking a group of transformations that changes the physical situation which constitute the experiment from what it was, is, or will be, into what it was not, is not, or will not be. The transformations that constitute the elements of this group must be made according to the laws of this theory and leave invariant what was observed of this physical situation, the experiment. Usually one is not interested in analyzing all of the implications in the theory of what was observed. The analysis focuses upon some entity for which the data that are identified with what was observed are being reduced. The above group of transformations is called the possibilities-generating group for the entity of interest which was not observed. Each element of this group changes the value of the entity of interest into a possibility for the value, except the indentity element of this group which leaves unchanged the value but inserts it into the set of possibilities for the value. The normalized volume measure at each "point" of the group manifold of the possibilities-generating group is called the physical probability of the possibility generated by the corresponding group element.

The physical probabilities of possibilities for entities of physical theories determined in the above fashion are conditional probabilities in the usual terminology of probability theories. They are conditional not only upon what was observed but also upon the logic that was used to analyze the observations: the physical theory used. Because of how they were obtained, according to group theory, physical probabilities of possibilities for entities of theories are nothing but weights which must be assigned in the physical theories to the possibilities regarded as functionals of group elements. These weights must also be assigned in the theory to any functionals of the theory in which the possibilities enter. According to the manner in which they were obtained in the proposed method of data reduction since it is not a physical theory, physical probabilities are not observables of the method of data reduction. They are not observables of the theories being used to analyze experiments either as they are weights

attached to possibilities for entities that have not been observed. If there is any relationship between physical probabilities for entities of theories and observables of these theories, it can only be a purely conceptual one and must be derived, or derivable, using these theories. The conditions under which this relationship holds must also be established.

3. CONCRETE REALIZATIONS OF POSSIBILITIES-GENERATING GROUPS

In this section simple examples of data reductions are given to illustrate the concept of a possibilities-generating group. In the usual language of group theory, concrete realizations of possibilities-generating groups are given. The examples given in this section are selected for a number of reasons. They are so simple that there is no need to introduce a special formalism, as will be done in the next section to deal with more complex problems of data reduction. The different possibilities-generating groups are isomorphic to a variety of abstract finite and continuous groups. Most of the examples form the basis of more elaborate data reductions later in the paper and are used to derive some general results. There are some historical reasons for choosing these examples. Some are experiments based upon situations that have been used in the past to illustrate concepts of probability theories; some can be solved using the classical theory of probability;[16] and others have been used to show the internal contradictions of the classical theory of probability.

In each of the experiments now considered, the observations only result in placing an upper and a lower bound on the value of an entity of a physical theory, and there is no direct observation of the value of this entity inside these bounds. In these instances it is conventional to say that the uncertainty in the value of the entity, based on the observations only, is systematic in nature, or that there is a systematic uncertainty in the value of the entity. These situations are to be contrasted to those where there are direct observations of the value of an entity upon repetition of a measurement and a sample of values having a distribution of values is observed. In these latter instances, using the theory of statistical estimation, one can obtain an estimated statistical uncertainty in the value of the entity, based upon the sample of values, an uncertainty which is characterized in terms of statistical probabilities. It is likely that most statisticians would frown upon the use of statistical probabilities to characterize the systematic uncertainty when the data contain no sample of values, only a single 100% confidence interval where the value of the entity lies. It is fair to say, however, that some notable physicists have expressed opinions that they would not hesitate to quantify these systematic uncertainties by probabilities on the basis of some "common sense" arguments starting with Laplace,[16] Poincaré,[7] Jeffreys,[17] and Feynman.[18] As mentioned in the introduction, Jaynes[8] used group theoretical arguments to solve problems of data

reduction using a rationale different from the one proposed in this paper, but he had failed to solve one of them using this rationale.[19]

All examples of data reductions in this section are the analysis of some thought experiments. For simplicity, classical physics is the theory used and the details of measuring instruments are omitted. In all examples, some definite observations are identified with some entities of classical physics which become the data reduced for a specified entity of classical physics that was not observed in the experiment.

3.1. The Coin Experiment

In the first experiment considered, a coin is observed lying flat on a horizontal surface. It was not observed how this specific coin came to be in this position or which side of the coin was facing up. This physical situation is analyzed using classical mechanics. The observations made in the experiment must first be encoded into an abstract entity of classical mechanics. The coin is encoded into a flat rigid body having two identifiable sides and being at rest in a three-dimensional euclidean space. Because the flat body must have two identifiable sides, these can be brought into correspondence with the numbers 1 and 2 to become side 1 and side 2. A set of orthogonal axes can be introduced to describe the three-dimensional euclidean space with the z-axis as "up" and the x-y plane in the plane of the flat rigid body. The observations made, including the fact that the position of the sides has not been observed, are equivalent to placing an observer in the x-y plane.

For an observer not located in the x-y plane which side is facing up can be observed in classical mechanics. Either side 1 or side 2 is facing up; both sides cannot be simultaneously facing up. In classical mechanics there is no symmetry for this physical situation with respect to which side is facing up. However, this situation must be analyzed from the vantage point of what was observed, namely from an observer located in the x-y plane. In classical mechanics for an observer in the x-y plane the physical situation has an apparent symmetry since the coin collapses into a two-dimensional object with parallel sides, and the labels on the sides can no longer be seen. The above data can be reduced, or analyzed, for the side that is facing up.

The element of symmetry exhibited in classical mechanics by the physical situation, which is the experiment when all of the observations have been encoded into it, can be analyzed in the usual fashion by introducing a group of transformations. The elements of this group change what is the side facing up in the physical situation into what is not the side facing up, and leave unchanged what is observed which is the flat rigid body lying in the x-y plane. The physical transformations in classical mechanics which accomplish this are rotations of 180° and 360° about an axis joining the center of the flat object to the position of the observer in the x-y plane. Rotation by 360° about this axis is the identity element of this group which is isomorphic to the abstract group of order 2. This group generates two possibilities for the side facing up: the side 1 and the side 2. The volume measure in the group manifold of a finite group is the same at every point of the group manifold since it is a disjoint space. Therefore, the normalized volume measure is $1/2$ at both points in this case. Consequently, the physical probability of the two possibilities for the side of the coin facing up is $1/2$. This completes the data reduction for the side of this coin facing up since we can formally use in classical physics the two possibilities for this entity called "side 1 up" and "side 2 up" if we attach to them a weight of $1/2$, called their physical probability, because these can and must be regarded as functionals of the elements of a group of transformations of classical physics. In this paper we terminate the data reduction for an entity of a physical theory related to observations made in an experiment with the determination of the physical probabilities of its possibilities since from then on these possibilities for the entity can be treated as any other abstract entity of this physical theory that is a functional of a group element of a group of this theory.

The above analysis does not say anything about the phase of the moon, which is another physical situation, because the analysis of this experiment is irrelevant to the analysis of another experiment. In particular, the above analysis does not tell us what would happen if this coin were tossed in the air and one subsequently observed which side was up when it came to rest. There is no random or statistical element introduced in the above arguments; these are concepts of another physical theory. There is also no element of "physical symmetry" in the coin in the above discussion as one might attempt to convey by saying the coin was "fair". The physical transformations that constitute the elements of the

possibilities-generating group are mental transformations allowed by the laws of the theory used. In the above analysis, the coin was never actually rotated about any axis.

It should be noted that in the above experiment if the side of the coin which was facing up had been observed but the side of the coin facing down had not been observed, it would be possible to reduce the data for what had not been observed in exactly the same fashion. The possibilities-generating group would be a rotation of 360° of the coin about any fixed axis and its order is unity. There is only one possibility - the side that is not facing up - and its physical probability is unity.

3.2. The Playing Die Experiment

The next experiment analyzed involves a playing die. The physical situation is very similar to the one in the previous example: the orientation of a solid object is observed but the markings on its faces are not observed. Therefore, the possibilities-generating group for the location of the faces of the object will be the point group which is the group of symmetry of a geometrical figure that is an abstraction corresponding to the encoding of the observations made into the physical theory used. The reason for considering in some detail this experiment is to provide a physical interpretation in terms of the observations made of the fact that the possibilities-generating group is decomposable into the direct product of two possibilities-generating groups. This feature of possibilities-generating groups, as will be shown later, plays an important role in most data reductions.

An ordinary right-handed playing die is the physical object involved in this experiment. By this it is meant that it has six faces numbered 1 through 6; opposite faces are parallel; the sum of the numbers on opposite faces is seven; and the normals to the faces having the numbers 1, 2, and 3 define a right-handed orthogonal coordinate system. It is subsequently determined that the die is at rest with one side facing "up" and another side facing "north;" however, the numbers on the sides were not observed nor how it came to be in this position. We are interested in reducing the above data for the markings on the various faces of the playing die. The above data do not define the playing die of this experiment as a "perfect" cube or a "fair" die. However, one can introduce an imaginary cube whose sides are parallel to those of the playing die and whose fixed point coincide with the fixed point of the playing die. The

numbers on the sides of the imaginary cube being projections from those on the corresponding sides of the playing die. In the following data reduction we deal with this imaginary cube, and at any time we can mentally project the numbers on its faces onto the faces of the playing die.

The possibilities-generating group for the markings on the sides of the cube is the total rotational symmetry group of the cube. The group keeps invariant the fact that there is a side facing up and one facing north. This group is known as the octahedral group[14] (O), and has 24 elements. Each of the possibilities generated by this group can be uniquely labelled by naming the number of dots on the top side and on the north side of the cube. The physical probability of each of these possibilities is equal to 1/24 since this group is finite and has order 24.

One could also reduce the data in this experiment for the number of dots on the side facing up only, if that is what one wanted to do. In this case, the possibilities-generating group is the dihedral group, D_3, which has six elements built upon any one of the four three-fold symmetry axes which constitute the space diagonals of the cube. The rotations are by 120° about these axes and the reflections are in a plane perpendicular to these axes. These reflections are equivalent to a rotation of 180° of the cube about a space diagonal normal to the space diagonal used for the 120° rotations above. The six possibilities generated by any one of the above dihedral groups, D_3, are labelled 1, 2, 3, 4, 5, and 6 dots on the top side, and 1/6 must be assigned as the physical probability of each of these possibilities. Each of these possibilities for the number of dots on the top side can be called a "marginal" possibility since they ignore which side of the die is facing north. There are 4 different possible dihedral groups that can be used in this data reduction. This "ambiguity" will be discussed later and should not be of concern at this stage since any one of these groups gives the same answer. The fact that the same answer is obtained for all these groups will be taken as an indication of the internal consistency of the method of data reduction proposed in the sense that if there are several different ways of reducing the data, the final result should be invariant of the way chosen.

Now let us suppose that one is interested in the number of dots on the side facing north, conditional upon the fact that the side facing up has a particular number of dots on it, say 1 dot. The possibilities-generating group for the number of dots on the north side must keep invariant the side of the die having

one dot facing up in the data reduction. This possibilities-generating group is the cyclic group of order 4, C_4. The generator of this group is a rotation by 90° about the vertical four-fold symmetry axis through the cube. The four possibilities for the number of dots on the side facing north can easily be named since the side having 6 dots has the physical probability of unity for being the side facing down. Therefore, conditional upon the fact that the side having one dot is facing up, there is a physical probability of 1/4 for the possibility that the side having 2, 3, 4, and 5 dots is facing north.

It should be noted that the octahedral group, O, is the direct product of the two groups D_3 and C_4. Therefore, it was unnecessary to do the three data reductions above if we had made use of this fact.

3.3. The Balls in a Box Experiment

All of the remaining examples of data reduction in this section are the analysis of physical situations that are closely related to the measurement of entities of physical theories where the observations lead to systematic uncertainties in the "measured" values. In the example now considered, the entity of interest can only have a discrete value; the following examples will deal with continuous parameters of physical theories.

The physical situation now considered is defined by observations that determine there are a certain number of balls in a box and this number of balls is not less than N_a but not more than N_b, where N_a and N_b are two integers with $N_b > N_a$. If the number of balls in the box is denoted by N, then the observations yield the data: $N_a \leqslant N \leqslant N_b$. This situation is the typical one where it is said today that the measured value of N has a systematic uncertainty.

In the method of data reduction proposed in this paper, since N was not observed but the data are related to N, $N_a \leqslant N \leqslant N_b$, these data can be reduced for the entity N. To reduce the data $N_a \leqslant N \leqslant N_b$ for the value of N, it is necessary to introduce a group of physical transformations that changes the number of balls in the box, but leaves unchanged the fact that there are no less than N_a balls in the box and no more than N_b balls. The physical operations that correspond to the elements of the group consist of adding and removing balls from the box.

That the above operations form a group is very well known since the above problem is nothing but an "occupation number" problem. It is usual to treat those problems by introducing "creation-anihilation" operators that change the number of "particles" (balls) in the "state" (the box) thus forming a group. The possibilities-generating group for the number of balls in the box is therefore isomorphic to the cyclic group of order $N_b - N_a + 1$; it is isomorphic to the group obtained by the addition modulo $N_b - N_a + 1$ of the integers O, 1, 2, ..., N_b-N_a. The possibilities-generating group for the number of balls, N, in the box generates from the number of balls in the box the possibilities N_a, $N_a + 1$, ..., $N_b - 1$, N_b. Because the order of the group is $N_b - N_a + 1$, the physical probability of each of these possibilities for N is $1/(N_b - N_a + 1)$.

3.4. When?

The remaining examples of data reduction involve possibilities-generating groups that are Lie groups. The next two examples deal with the "measurements" of the two parameters of the abstract time-space of classical physics. The time space of classical physics is a one-dimensional affine space. The two parameters associated with this time-space are called interval of time and instant of time. The parameter instant of time is the position parameter in this space; that is to say, it is characterized by translation invariance. The parameter interval of time is associated with the metric of this space and is, therefore, characterized by scale invariance.

The first experiment involves the determination of when an event occurred. There are two observations made in this experiment: the event of interest has been observed (1) not to have occurred before a known time, t_1, and (2) to have occurred not later than another known time, t_2. In classical physics which is the theory used to reduce the data, $t_1 \leqslant t \leqslant t_2$, the event of interest must have occurred at some specific time t. Since the time t has not been observed and is the entity for which the data are reduced, the elements of the possibilities-generating group must transform it to a new time t′ but must leave invariant what was observed: $t_1 \leqslant t \leqslant t_2$. As pointed out above, since an instant of time is defined in classical physics to be a continuous translation invariant parameter, the possibilities-generating group

for the time t is the one parameter Lie group T_1 where t is transformed into t′ according to $t' = t + a$, and where a is the continuous parameter of the group. The group manifold or parameter space of this group is a closed compact topological space. In the data, $t_1 \leqslant t \leqslant t_2$, t_1 and t_2 are irreducible representations of the known times. They define the closed parameter space of the possibilities-generating group for the instant of time t. It is well known[14] that the volume measure in the group manifold of T_1 is constant; therefore the physical probability that the event of interest occurred in an interval of time, dt, at any time, t, between t_1 and t_2 is $dt/(t_2 - t_1)$, and zero elsewhere.

3.5. How Fast?

The measurement of the scale parameter of the time-space of classical physics is considered in the form of the measurement of the period of a cyclic phenomenon τ. Here again, there are two observations in this experiment: the period of the phenomenon of interest was observed to be (1) not less than the known period τ_1 of a particular phenomenon and (2) not more than the known longer period τ_2 of another phenomenon. The data to be reduced for the period τ can be written $\tau_1 \leqslant \tau \leqslant \tau_2$.

The data reduction for this experiment parallels the one for the previous experiment. The only difference is that the possibilities-generating group for the period τ is the scale transformation, $\tau' = b\tau$. The parameter space for this Lie group is also closed. The volume measure in the parameter space of this group is no longer constant but proportional to $1/b$. Therefore, the physical probability for the possibility that the period of the phenomenon of interest is in the time interval, $d\tau$, about τ, in the range $\tau_1 \leqslant \tau \leqslant \tau_2$, is $d\tau/(\tau \ln (\tau_2/\tau_1))$, and zero elsewhere.

The last two data reductions involving continuous parameters solve a fundamental logical problem that arises when one tries to apply the theory of statistics to continuous parameters. It is then required that one deals with a sample of values of the continuous parameter. It is a logical impossibility to have a sample of values for a continuous parameter since a finite number of digits must be used. To cope with this problem the notion of significant figures is introduced. Every datum in the sample must therefore be of the type just reduced, but it is said that such datum cannot be dealt with using the theory of statistical estimation because this datum leads to systematic uncertainties. The fact that for some

specific applications it does not matter where the value of the parameter is in an interval is irrelevant from a purely logical point of view: one cannot obtain a sample of values for a continuous parameter.

It is often said that classical physics is deterministic. One interpretation of this statement is that it is possible in principle to determine the initial condition of a system such that its behavior at any subsequent time can be predicted. Since time is a continuous parameter in classical physics, the last two data reductions indicate that the above statement cannot be correct if one considers that the initial conditions would have to be determined by observations.

3.6. Von Mises' "Water and Wine" Problem

The example of data reduction now considered is the analysis of a problem proposed by Von Mises.[20] This problem is of interest because it is one of the last remaining "Bertrand Paradoxes." Bertrand paradoxes are problems which yield contradictions when the principle of insufficient reason is applied to solve them in two different ways. Von Mises' "water and wine" problem is analyzed here because Jaynes,[19] who used transformation group arguments based upon his desideratum of consistency to solve the Bertrand chord problem, conjectured that Von Mises' problem may be "ill posed." From an experimental physics point of view, we deal with the measurement of the ratio of two scale parameters that produce a dimensionless entity.

As originally stated by Von Mises, the problem is the following: "We are given a glass containing a mixture of water and wine. All that is known about the concentration of the liquid is that the ratio of water to wine is not less than 1 and not more than 2; this means that the mixture contains at least as much water as wine and, at most, twice as much water as wine." Von Mises shows that if the principle of indifference is applied to the ratio of water to wine, one obtains a median value for the probability density function of 3/2. However, if one applies the same principle for the inverse ratio, the median value for the probability density function of the ratio of water to wine is 4/3. The classical theory of probabilities is therefore internally inconsistent when one uses the principle of insufficient reason in this fashion to assign probabilities. Inconsistencies such as the ones above have been used by Von Mises, among others, to argue that definitions of probabilities not based upon limiting frequencies of observations (the statistical definitions of probabilities) were doomed to failure.

In order to reduce the data corresponding to Von Mises' stated observations in the experiment, it is necessary to specify what attribute is used to express the concentration of the liquid. This is required in order to know what is the "law" of the physical theory that applies to this attribute since this law determines the allowed transformations of the physical system. Classical physics is used to analyze the situation of Von Mises, and masses are used as the relevant attributes. Since a definite mixture of water and wine is observed, it has a definite mass M which is equal to the sum of the mass of the water, M_1, and the mass of the wine, M_2. Therefore from classical physics one has the datum, $M_1 + M_2 = M$, where M must be kept constant. The second datum is $1 \leqslant M_1/M_2 \leqslant 2$ which is the same thing as $1/2 \leqslant M_2/M_1 \leqslant 1$. These data have to be reduced for the ratio M_1/M_2, or M_2/M_1, which was not observed. The elements of the possibilities-generating group for the ratio M_1/M_2 must change this ratio from what it is in the mixture while keeping all of the data unchanged. Since one datum is $M_1 + M_2 = M$ where M is fixed, the elements of the group must change simultaneously M_2 and M_1. If a certain mass of water is added to change the ratio, M_1/M_2, the same mass of wine must be withdrawn. Since mass is a continuous parameter in classical physics, the group of transformations is a one-parameter Lie group where M_1 and M_2 are transformed into M'_1 and M'_2 according to $M'_1 = M_1 + a$ and $M'_2 = M_2 - a$. From this group of transformations which is the possibilities-generating group for the masses M_1 and M_2, one can easily calculate the physical probability density function for the ratio, M_1/M_2, and find that its median value is neither 3/2 nor 4/3, but is 7/5. If one uses "Jeffreys rule"[17] one must conclude in his theory of probabilities that the median value of the ratio M_1/M_2 is $\sqrt{2}$ since the ratio of the two masses can only be in the domain 0 to ∞. This is wrong since in classical physics it is equivalent to treating the ratio of the masses, M_1/M_2, as a scale parameter which it is not.

The above analysis of Von Mises' problem requires that a particular attribute of wine and water in a physical theory be used for expressing the "concentration of the liquid." This analysis only holds for attributes that are additively conserved in the mixture in the theory used. The answer given would not hold if volumes were used and the wine contained alcohol. However, a unique answer prevails for any

specified law of volume additions, but it varies with the law. We conclude that after all Jaynes was right that the problem was ill-posed and we have given the reason why and it differs from Von Mises'. If we look at Von Mises' problem as one of data reduction in a physics experiment, it is well posed since a particular attribute of a physical theory must have been used and needs to be specified, and this determines the transformation law which allows us to reduce the data uniquely.

4. THE FUNDAMENTAL THEOREMS OF PHYSICAL PROBABILITIES

In this section a notation is introduced to designate the physical probability of the possibilities for entities of physical theories; that is to say, to designate the normalized volume measures in group manifolds of possibilities-generating groups. It is then shown that when this notation is used to express the fact that the normalized volume measures in a group manifold are additive (countably additive) one obtains formulae which have been used for centuries as the axioms of theories of probabilities and are considered today as the physical laws of random phenomena: the laws of statistics. Mathematicians describe the formulae given in this section by saying that the algebra of the normalized volume measures in a group manifold is a Borel algebra.

4.1. Notation

The notation used to denote the physical probabilities of possibilities for entities of physical theories must of necessity be complex since it must indicate not only to which possibilities they refer to but also which possibilities-generating group is involved. However, to adopt too complex a notation defeats the very purpose for which a symbolic notation is used: conciseness.

It would be desirable to use five symbols to indicate the possibilities-generating group involved. These symbols would be used to indicate: (1) the observations made in the experiment, O; (2) the physical theory being applied, T; (3) the data of the theory identified with the observations, D; (4) the abstract entity C of the theory T for which the data D are reduced; and (5) some standard designation for the abstract group isomorphic to the possibilities-generating group.

Although it should always be possible to provide all of the above information, it is not practical to do so in the notation. One must rely upon the context in which the notation is used to provide much of that information. The only element among those listed above that is retained is the data D of the physical theory which are left invariant under the transformations of the group. Occasionally it is useful to indicate the specific entity C of the physical theory for which the data D are being reduced. The fact that the theory being used does not appear in the notation of the possibilities-generating group is a departure from tradition since data reduction is used to test physical theories considered as hypotheses

and in the usual tradition of "hypothesis testing," the hypothesis being tested appears explicitly in the notation.

No shortcuts can be taken in the notation to designate which possibility, or sets of possibilities, the physical probability applies to since one cannot rely upon the context in which the formulae are used to provide the information. A physical probability can only apply to a specific possibility of a set of possibilities, a subset of this set, or to the whole set. Consequently, the notation of set theory could be used but this will not be done since, in general, one is interested in dealing with the power set of a set of possibilities. It is therefore more convenient to use the notation of lattice theory[15] which has the further advantage of being frequently used in connection with Borel algebras. However, since the terminology of lattice theory is not commonly used in physics, the notation adopted will be explained in terms of the presumed, more familiar operators of Boolean algebra.

The sign Λ is called a "conjunction" and is a binary relation; it is the usual "logical and" in the sense of "and also." The sign V is called a "disjunction" and is also a binary relation; it is the usual "logical inclusive or" in the sense of "and/or." The sign $\sim$ is called "complement" and is a unary relation in the sense that $\sim$A is the complement of A. The notation $\sim$A is only used in a context when A designates a subset of a set of possibilities and $\sim$A is a subset of the set of possibilities where the set A V $\sim$A is the complete set of possibilities and A Λ $\sim$A is the null set. What is called the complement of a set is therefore the relative complement of a set. Readers familiar with lattice theory will recognize that mathematically we are dealing with Boolean lattices in which the σ-measure introduced in the lattices to generate a Borel algebra is uniquely defined since these Boolean lattices are generated from the group manifold of a group. It must also be noted that in this method of data reduction the Lebesgue extension of a σ-additive measure in a σ-field of sets is not introduced as is normally done in the theory of statistics.[15]

If in a particular physical theory on the data D the possibilities-generating group for the entity C is of order n it generates a set of possibilities $\{C_i\}$, i = 1, 2, ..., n, then the notation

$$(C_i|D) \tag{4.1}$$

is used to denote the physical probability of the possibility C_i. The symbol to the left of the vertical bar indicates to which possibility, or subset of the set of possibilities, the physical probability applies to; the symbol to the right of the vertical bar indicates the possibilities-generating group, although only the data D left invariant by this group is indicated in the notation. Occasionally, the notation (4.1) for a physical probability is extended to show explicitly which entity C of the physical theory the data D are being reduced for as follows:

$$(C_i|D:C) \quad . \tag{4.2}$$

If one denotes by B a particular subset containing the elements C_1, C_2, ..., C_m, of the set of possibilities $\{C_i\}$, i = 1, 2, ..., n, the Boolean notation

$$B = C_1 V C_2 V \ldots V C_m \quad , \quad m \leqslant n \tag{4.3}$$

is used to denote B. The physical probability of B is written:

$$(B|D) = (C_1 V C_2 V \ldots V C_m|D) \quad , \tag{4.4}$$

or

$$(B|D:C) = (C_1 V C_2 V \ldots V C_m|D:C) \quad . \tag{4.5}$$

If it is necessary to indicate that the data D consist of two sets of data, D_1 and D_2, then the Boolean notation

$$D = D_1 \Lambda D_2 \tag{4.6}$$

is used for D in the symbol for the possibilities-generating group. For instance, if this were the case one would write the physical probability (4.1):

$$(C_i|D_1 \Lambda D_2) \quad . \tag{4.7}$$

It is important to note that the disjunction sign cannot be used on the right-hand side of the vertical bar in the symbol for a physical probability, and if the conjunction sign appears between two symbols they cannot be mutually contradictory in a physical theory because a group of order zero does not exist. By definition, the symbol $(C|C:C)$ is unity since the possibilities-generating group in this case is of order one: the identity transformation.

If the possibilities-generating group is the direct product of two groups, as was the case in the example of the playing die experiment in Sect. 3, each possibility of the set of possibilities $\{C_i\}$, i = 1, 2, ..., n, can uniquely be identified by two labels, or indices, one from each group. There is a conflict of notation between the standard notation of group theory where $C_{j;k}$ would be used and the notation of lattice theory where (A_j, B_k) would be used (the elements of one lattice being denoted A_j and the elements of the other B_k). This is also in conflict with the standard notation of probability theories where a "compound event" would be written A_jB_k. Even though it may create some confusion for those familiar with lattice theory, in the case of possibilities generated by a group which is the direct product of two groups, the notation adopted in this paper is:

$$C_i = A_j \Lambda B_k \quad , \tag{4.8}$$

where C_i is a possibility generated by the direct product group and A_j, B_k are the possibilities generated by the two groups. There is no conflict in this use of the conjunction sign and its use in (4.6) because of its meaning of "logical and". In the case of a group of transformations that is the direct product of two groups, one may write:

$$(C_i|D:C) = (A_j \Lambda B_k|D:C) \quad . \tag{4.9}$$

It is often useful in this instance to indicate for which entity of the theory the data D are being reduced.

It was implicit in the above notations that the possibilities-generating group was a finite group. This notation can be extended when it is a continuous group. When the possibilities-generating group is continuous, it is preferable to identify the possibilities for which the physical probability is given by the

corresponding elements in the group manifold. This is done by using a representation of the possibilities-generating group as was done in the two examples of continuous groups given in Sect. 3. Using a one-parameter continuous group as an example, if one denotes by x a specified faithful representation of the possibilities-generating group for the entity X and by dx the volume element in the group manifold at x, then the physical probability for the possibilities corresponding to the volume element dx is written:

$$(dx|D{:}X) \equiv (x|D{:}X)dx \quad , \tag{4.10}$$

where $(x|D{:}X)$ is the density function of group elements at x normalized to unity when integrated over the group manifold and is called the physical probability density function of the possibilities at x.

Because the notation for continuous groups [the left-hand side of (4.10)] is formally identical to the notation for finite groups and the results apply to both, there is no need to differentiate between the two types of groups in the rest of this section. In the case of continuous groups, the summations are to be replaced by Riemanian integrations over the group manifold.

The notation adopted makes it apparent that, in the method of data reduction proposed, a physical probability of a possibility for an entity of a physical theory is a conditional probability, in the usual terminology of probability theories.

4.2. The Disjunction Theorems

What are called here the disjunction theorems of physical probabilities are the expressions, in the notation just defined, of the fact that the volume measure in the group manifold is additive. That is to say, the physical probability corresponding to a subset of the set of possibilities is the sum of the physical probability of each element of the subset.

If $\{C_i\}$, i = 1, 2, ..., n are the n distinct possibilities for an entity C generated by a possibilities-generating group of order n on the data D, then by definition of physical probabilities one has:

$$(C_1 V C_2 V \ldots V C_n|D{:}C) = \sum_{i=1}^{n} (C_i|D{:}C) = 1 \quad . \tag{4.11}$$

If one denotes by A a subset of the set $\{C_i\}$, i = 1, 2, ..., n, say the subset $\{C_j\}$, j = 1, 2, ..., m with m < n, then one obtains directly:

$$(A|D:C) \equiv (C_1 V C_2 V \ldots V C_m|D:C) = \sum_{i=1}^{m} (C_i|D:C) \quad . \tag{4.12}$$

If B denotes another subset of the set of possibilities $\{C_i\}$, then it follows that:

$$(AVB|D:C) = (A|D:C) + (B|D:C) - (A \Lambda B|D:C) \quad , \tag{4.13}$$

where AVB denotes the union of the two subsets A and B, usually written $A \cup B$, and $A \Lambda B$ denotes the intersection of two subsets A and B, usually written $A \cap B$.

The use of the sign $A \Lambda B$ to designate the subset of possibilities which is the intersection of the two subsets A and B of the set of possibilities $\{C_i\}$ is consistent with the meaning of the symbol Λ in boolean algebra where it stands for the possibilities which are in the subset A *and also* in the subset B. It is used here in the same sense as "meet" in lattice theory where the elements of the lattice are subsets and the binary relation is set inclusion.[15]

4.3. The Complement Theorem

Given a set of possibilities $\{C_i\}$, i=1, 2, ..., n and A denotes a subset of this set. By definition, the subset of the set $\{C_i\}$, called the relative complement of the subset A is $\sim A$ and $\sim A$ has no element in common with A. Therefore, applying the three disjunction theorems above one obtains directly:

$$(C_1 V C_2 V \ldots V C_n|D:C) = (AV \sim A|D:C) = (A|D:C) + (\sim A|D:C) = 1 \quad . \tag{4.14}$$

The last equality in (4.14) is called the complement theorem.

4.4. The Conjunction Theorem

As previously stated, if the possibilities-generating group for an entity C on the data D is the direct product of two groups, each possibility C_i can be specified by two indices, one coming from each of the two groups. It follows that the number of possibilities C_i, n, is the product of the orders l and m of the

two groups: n = lm. Since for a continuous group the volume measure in the group manifold corresponds to the number of elements of the group in the neighborhood of the point,[13] the volume measure in the group manifold of a continuous group which is a direct product group is equal to the product of the volume measures of the two groups and the group manifold is two-dimensional.

In the notation adopted, one writes for the possibility C_i of the direct product group $A_j \Lambda B_k$, where A_j is one of the possibilities associated with one of the two groups and B_k is one of the possibilities associated with the other group. It therefore follows that one has:

$$(C_i|D{:}C) = (A_j \Lambda B_k|D{:}C) = (A_j|D{:}A)(B_k|A_j \Lambda D{:}B) \quad . \tag{4.15}$$

The formula (4.15) is called the conjunction theorem. In most theories of probabilities, the formula corresponding to (4.15) is called the law of conditional probabilities, an inappropriate terminology here since all physical probabilities are conditional.

It must be noted that in the conjunction theorem (4.15) there are three different possibilities-generating groups involved, as is seen from the symbols on the right-hand side of the vertical bars. The possibilities-generating group D:C is the direct product of the possibilities-generating group D:A, and the possibilities-generating group $A_j \Lambda$D:B. The elements of the group labelled D:A leave D invariant and generate the possibilities for A making complete abstraction of the entity B of the physical theory. The elements of the group $A_j \Lambda$D:B leave invariant D and A_j, as if A_j was the value that A has in the physical theory, and generates the possibilities for B.

If G_1 and G_2 are two groups, then the direct product group G_1 x G_2 is the same as the direct product group G_2 x G_1. The groups G_1 and G_2 are both invariant subgroups of the group G_1 x G_2 or G_2 x G_1. Since it is purely conventional how one labels the elements of a direct product group in terms of the elements of the two groups it is the direct product of, the conjunction theorem (4.15) can be rewritten:

$$(C_i|D{:}C) = (B_k \Lambda A_j|D{:}C) = (B_k|D{:}B)(A_j|B_k \Lambda D{:}A) \quad . \tag{4.16}$$

It must be emphasized, as is clear from the notation, that the possibilities-generating group D:A is very different from the possibilities-generating group $B_k \Lambda D{:}A$. Although the set of possibilities generated by the group $B_k \Lambda D{:}A$ is a subset of the possibilities for A generated by the group D:A, the group $B_k \Lambda D{:}A$ is not generally a subgroup of the group D:A.

A physical interpretation of the groups that enter in the conjunction theorem is provided by the playing-die experiment analyzed in Sect. 3 if one denotes by B the side of the die which is facing north and by A the side of the die facing up.

4.5. Bayes Theorem

The two formulae (4.15) and (4.16), collectively called the conjunction theorem, have two important corollaries. The first one has been given the name of Bayes theorem, even though it was not first proposed by Bayes, and holds in all theories of probabilities where the equivalent of (4.15) and (4.16) hold, however they are justified.

What is called Bayes theorem is the formula that results from equating the right-hand sides of (4.15) and (4.16), since they are both equal to $(C_i|D{:}C)$. We can rearrange the terms:

$$(A_j|B_k \Lambda D{:}A) = (A_j|D{:}A)\frac{(B_k|A_j \Lambda D{:}B)}{(B_k|D{:}B)} \quad , \tag{4.17}$$

or

$$(B_k|A_j \Lambda D{:}B) = (B_k|D{:}B)\frac{(A_j|B_k \Lambda D{:}A)}{(A_j|D{:}A)} \quad . \tag{4.18}$$

The above formulae can be obtained directly from group theory. They are important because one can obtain algebraically the volume measure in a group manifold without having to exhibit explicitly this group of transformations; for instance, the group $B_k \Lambda D{:}A$. In many problems of data reduction it is much easier to obtain the physical probabilities of the possibilities on the left-hand sides of (4.17) or (4.18) from the physical probabilities of the possibilities that appear on their right-hand sides. Although the price paid to do so appears to be high, the following corollary reduces the number of transformation groups one has to deal with from three to only two.

4.6. The Sum Rule

The second important corollary to the conjunction theorem is obtained by making sure that the physical probabilities are the normalized volume measures in the group manifold, a fact which resulted in the first disjunction theorem (4.11). This second corollary is often called the sum rule. It is obtained by summing both sides of (4.15) over all the possibilities A_j of the set of possibilities generated by the group of transformations, D:A, of which there are m. Thus one obtains:

$$\sum_{j=1}^{m}(B_k \wedge A_j | D:C) = \sum_{j=1}^{m}(A_j | D:A)(B_k | A_j \wedge D:B) \quad . \tag{4.19}$$

It should be recalled that the group D:C is the direct product of two groups, and what one obtains by summing the physical probabilities on the left-hand side of (4.19) over all of the m possibilities A_j is just the physical probability that was denoted $(B_k|D:B)$ in (4.16). Therefore, if $(B_k|D:B)$ is substituted for the left-hand side of (4.19), one obtains:

$$(B_k | D:B) = \sum_{j=1}^{m}(A_j | D:A)(B_k | A_j \wedge D:B) \quad , \tag{4.20}$$

which is called in this paper the sum rule of physical probabilities.

It should be noted that the left-hand side of (4.20) is just the physical probability that appears in the denominator of (4.17) and that the physical probabilities that appear on the right-hand side of (4.20) are those in the numerator of (4.17). The result (4.20) (the sum rule) shows that one does not need to exhibit explicitly the group D:B in order to obtain the physical probability $(B_k|D:B)$ if one already has the physical probabilities $(A_j|D:A)$ and $(B_k|A_j \wedge D:B)$. If one uses (4.20) in (4.17), the Bayes theorem becomes:

$$(A_j | B_k \wedge D:A) = \frac{(A_j | D:A)(B_k | A_j \wedge D:B)}{\sum_{J=1}^{m}(A_j | D:A)(B_k | A_j \wedge D:B)} \tag{4.21}$$

Often Bayes theorem is only quoted in the form shown in (4.21) rather than in the form shown in (4.17).

4.7. Remarks

The formulae given in this section are merely the encoding of the fact that the normalized volume measure in a group manifold is an additive measure into a notation which is adapted to the method of data reduction proposed. These formulae establish explicitly that the algebra of physical probabilities of possibilities for entities of physical theories is a Borel algebra. This fact was implicit in the additivity of the volume measure in a group manifold. The Frobenius algebra of groups has long been used in physical theories. The fact that there is a Borel algebra associated with a group does not seem to have been widely recognized or made explicit use of in developing physical theories, although many useful statistical physical theories have been proposed. It should be noted that usually one tends to associate Borel algebras with semigroups rather than groups.

The formulae given in this section have been called the fundmental theorems of physical probabilities because they are useful in performing most data reductions as will be shown explicitly in the next section when an important result is derived at the same time.

5. PHYSICAL PROBABILITIES OF OUTCOMES AND LIMITING FREQUENCES

The purpose of this section is to provide an illustration of the use of the algebra of physical probabilities to perform a data reduction. In the physical situation now analyzed, the data will be reduced for an entity of the physical theory that could be called the outcome of the experiment if the experiment were to be performed. It will be shown that, from the point of view of the theory used to analyze the physical situation, the limiting frequency with which a certain type of outcome should be observed to occur (if the experiment were repeated an infinite number of times) is numerically equal to the physical probability of this type of outcome based upon the data that describe that physical situation and does not contain any observed outcome. It is only in Sect. 7 that the reduction of data containing observations of outcomes in repeated experiments is considered.

The type of physical situation now analyzed played an important role in the development of theories of probabilities. From about 1650 to approximately 1750 the use of probability theories was mostly restricted to the analysis of thought experiments such as the ones analyzed in this section. It was only after approximately 1750 that theories of probabilities were explicitly used in the analysis of data obtained in experiments, including the observation of their outcomes.

5.1. Conceptual Sampling

The thought experiment now analyzed is based upon and analyzed with classical physics and can be characterized by the name of conceptual sampling without replacement. The experiment involves a long tube and N distinguishable spherical balls. As usual the distinguishability of these spherical balls will be explicitly accomplished by considering that these balls are numbered 1 through N. In this experiment, the balls numbered 1 through n, $n < N$, are black, and the remaining balls are white. The balls are all approximately of the same diameter. The inside diameter of the tube is larger than the diameter of any of the balls but smaller than the sum of the diameter of any two balls. The tube is held vertical and has a release mechanism at the bottom whereby balls placed inside the tube could be released or drawn out of the tube one at a time. According to classical physics, balls should be released from the bottom of the tube in the order in which they were inserted into the tube at the top. In the

language adopted in this paper, however, if the order of the loading of the balls into the tube has been observed, there is, based on these data, a physical probability of unity that they will come out of the tube in that order, if the balls have not yet been drawn from the tube.

The data reduction that concerns us in this section is when the observations mentioned above have been made but the order in which the N balls were loaded into the tube has not been observed. This can easily be accomplished by having someone else load the balls into the tube. The above geometry for this type of experiment is chosen over the conventional urn from which balls are drawn blindfolded to avoid any connotation that there is an element of chance involved.

5.1.1 The number on the first ball drawn

On the basis of the observations interpreted in classical physics, the first ball drawn would be the one at the bottom of the tube and there is a number on it. The number that is on the first ball drawn will be considered the outcome of the experiment and the data are reduced for this entity. The relevant data here are that the N balls are in the tube and the entity for which these data are reduced is denoted F for the first ball drawn but it is implied that the relevant attribute is the number on this ball. The possibilities-generating group for F is denoted N:F. This group of transformations corresponds to an element of apparent symmetry in the physical problem considered from the point of view of classical physics in which only that which has been observed appears. What this apparent element of symmetry is can easily be found. There are N balls in a column each with a different number on them. There is no element of symmetry if we visualize the column of balls with the numbers visible to the observer; however, for an observer placed on the opposite side of the column of balls so that the numbers on the balls could not be seen, this same column of balls exhibit a great amount of "symmetry." What this element of symmetry is corresponds to the fact that two different columns of N balls, which are different because the order of the balls in the columns are different seen from a point in space where the numbers on the balls cannot be observed, appear to be the same. There are N! such different columns of N balls that would appear to be the same. The complete apparent symmetry in this column is described by the symmetric group S_N. The symmetric group of degree N, S_N, is not the one that interests us in the data reduction considered here because we only care about the number that is on the

bottom ball of the column. The possibilities-generating group N:F that is needed is the cyclic group of order N, C_N which is an invariant subgroup of S_N. A physical interpretation of the elements of the group, N:F, must be provided. This is easily done. The generator of the cyclic group, C_N, is the transformation of the column of N balls where the bottom ball is placed at the top, or its inverse transformation where the top ball is placed at the bottom. There are N different elements of this group of transformations. Each element of this group of transformations applied to the column of balls that is in the tube changes the number that is on the bottom ball and, therefore, changes the outcome of the experiment. Let us denote by $\{F_i\}$, i = 1, 2, ..., N, the set of possibilities for F, where F_i stands for the first ball drawn has the number i on it. Because the group N:F is finite and of order N the normalized volume measure in the group manifold of N:F is 1/N. Therefore, the physical probability of each possibility F_i for the outcome of this experiment is:

$$(F_i|N{:}F) = 1/N \quad , \; i = 1, 2, ..., N \; . \tag{5.1}$$

5.1.2 The color of the first ball drawn

To consider that the outcome of the experiment analyzed is the number on the first ball drawn is permissible, but one can also consider the color of the first ball drawn as its outcome. The observations made in this experiment include the data that the balls with numbers less than or equal to n are black and the others white. These data have now to be considered as well, and for this data reduction the data will be written N,n. Since we are still considering the first ball drawn from the tube, the possibilities-generating group is still the one found in the previous data reduction which will now be denoted N,n:F. If one denotes by B_1 the subset $\{F_j\}$, j = 1, 2, ..., n, of the set of possibilities for F, $\{F_i\}$, i = 1, 2, ..., N, in Boolean notation we have:

$$B_1 = F_1 V F_2 V ... V F_n \quad , \tag{5.2}$$

and B_1 can be read: "the first ball drawn is black." The relative complement of the subset B_1, $\sim B_1$, can be written W_1 to be read: "the first ball drawn is white." Using the disjunction theorem of the algebra one obtains for the physical probability of the possible outcome B_1:

$$(B_1|N,n:F) = (F_1VF_2V...VF_n|N,n:F) = \sum_{i=1}^{n}(F_i|N,n:F) = \frac{n}{N} \quad . \tag{5.3}$$

5.1.3 The first two balls drawn are black

The emphasis in the rest of this section is on the application of the algebra of physical probabilities to calculate the physical probability of some specific possible outcome of the experiment, where the outcome of the experiment is considered the drawing of several balls from the tube. The idea is to focus upon some specific possible outcomes rather than generating all of the possible outcomes for a certain number of drawings of balls from the tube, by giving explicitly the relevant possibilities-generating group. It is for these types of data reductions that the algebra of physical probabilities is most useful.

The notation B_i is introduced to indicate that the outcome of the ith draw results in a black ball and the notation W_i denotes that it results in a white ball. Assuming that in this experiment $n > 2$, the data N,n are now reduced to calculate the physical probability of the outcome denoted $B_1 \Lambda B_2$. It is emphasized that no balls are drawn from the tube; what is being calculated is the physical probability of a specific outcome of a contemplated experiment. In the rest of this section only the data available to perform the data reduction, N,n, are used to symbolize the possibilities-generating group. What is now sought is $(B_1 \Lambda B_2|N,n)$. Using formally the conjunction theorem of the algebra one obtains:

$$(B_1 \Lambda B_2|N,n) = (B_1|N,n)(B_2|B_1 \Lambda [N,n]) \quad . \tag{5.4}$$

The first term on the right-hand side is given by (5.3). The second term on the right-hand side can readily be obtained from the considerations that yielded (5.3). The symbol $B_1 \Lambda [N,n]$ means that there were N balls in the tube of which n were black and then a black ball was drawn from the tube. The state of the system at this stage can be described by saying there are N-1 balls in the tube of which n-1 are black. Repeating the same argument that led to (5.3), but with N replaced by N-1 and n replaced by n-1, the second term of (5.4) is just (n-1)/(N-1). Therefore, the desired answer is:

$$(B_1 \Lambda B_2|N,n) = \frac{n(n-1)}{N(N-1)} = \frac{n!(N-2)!}{(n-2)!N!} \quad , \tag{5.5}$$

where n! is the gamma function $\Gamma(n+1)$.

By induction it is readily shown that:

$$(B_1 \Lambda B_2 \Lambda \cdots \Lambda B_r | N,n) = \frac{n!(N-r)!}{(n-r)!N!} \quad . \tag{5.6}$$

Similarly one has:

$$(W_1 \Lambda W_2 \Lambda \cdots \Lambda W_s | N,n) = \frac{(N-n)!(N-s)!}{(N-n-s)!N!} \quad . \tag{5.7}$$

The possibilities for the outcomes (5.5), (5.6), and (5.7) are some of the possibilities generated from the unknown order in which the balls are in the tube, by the direct product of cyclic groups.

5.1.4 The first r balls drawn are black and then m-r white balls are drawn

From the results (5.6) and (5.7), applying the conjunction theorem yields for the physical probability of the possibility that r black balls are drawn and then m-r white balls are drawn for the outcome of drawing m balls from the tube:

$$(B_1 \Lambda B_2 \Lambda \cdots \Lambda B_r \Lambda W_{r+1} \Lambda W_{r+2} \Lambda \cdots \Lambda W_m | N,n) = \frac{n!(N-n)!(N-m)!}{(n-r)!N!(N-n-m+r)!} \tag{5.8}$$

This result holds for the particular sequence indicated. However, it can readily be shown that it also holds for any particular specified order of obtaining r black balls in m draws.

5.1.5 In m draws r black balls are obtained

From the result (5.8) one can obtain the physical probability of the possible outcome in which exactly r black balls are obtained in m draws regardless of the order in which the black balls occur in the drawing. To do this requires only consideration of the group of permutation of r B symbols and m-r W symbols. The order of this group is well known to be given by the binomial coefficient:

$$\binom{m}{r} = \frac{m!}{r!(m-r)!} \quad . \tag{5.9}$$

If in analogy with the notation N,n used to indicate that there are N balls in the tube of which n are black and the remaining ones white, one denotes by m,r the possible outcomes of the drawing of m balls and r of them are black and the remaining ones white, then one obtains by multiplying (5.8) by (5.9):

$$(m,r|N,n) = \frac{\binom{m}{r}\binom{N-n}{m-r}}{\binom{N}{m}} \quad . \tag{5.10}$$

The right-hand side of (5.10) is known as the hypergeometric distribution.

The result (5.10) is the physical probability of the possible outcome of the experiment if one considers that the outcome of the experiment would be the drawing of m balls from the tube and this outcome is characterized by the number of black balls that would be obtained. Since the drawings of a ball would be the outcome of pressing the lever of the release mechanism at the bottom of the tube, one can also view the drawing of m balls as the collection of the outcomes of pressing the lever m times. If one considers the drawing of a black ball as a result of pressing the lever as one type of outcome and the drawing of a white ball as another type of outcome, making abstraction of the fact that every time the lever is pressed it is a different ball that is obtained, then the ratio r/m can be called the frequency with which a particular type of outcome (a black ball) would be obtained if the lever was pressed m times. Consequently, the result (5.10) can be viewed as the physical probability of the possible frequencies for a particular type of outcome.

It must be recalled that since there is no drawing of balls actually done in the experiment, the frequency $f_m = r/m$ is a ratio of observables of the physical theory used to analyze the physical situation and that it is not a ratio of observations in this situation.

5.2. Conceptual Limiting Frequency

In the theory of statistics, the concept of the limiting frequency of observed outcomes of a certain type plays an important role since it is the definition of a statistical probability. Physical probabilities in this paper are not so defined but, having obtained in (5.10) what can be interpreted as the physical probability for possible frequencies of outcomes in the physical situation analyzed, this situation is now modified to allow the definition of what will be called a conceptual limiting frequency.

The ratio r/m was called the frequency of possible outcomes where a black ball was drawn in m draws and denoted f_m. In order to define a limiting frequency one must generate an infinite sequence of values of f_m. Because in any actual experiment of the type analyzed N and n can only be finite, and

also consequently m and r, any sequence of frequencies f_m, however defined, must be finite. There is, however, no essential reason why the integers N, n, m, and r cannot be made very large, at least conceptually in a thought experiment. That is to say, one can treat mathematically the formulae just derived and continue to assign to the symbols that appear in them the meaning they have in classical physics for the situation they were obtained from.

One can take the formula (5.10) and let both N and n tend to infinity while keeping the ratio n/N constant. What one then obtains is:

$$\lim_{\substack{N\to\infty \\ n\to\infty \\ n/N=p}} (m,r|N,n) = (m,r|p) = \frac{m!}{r!(m-r)!}p^r(1-p)^{m-r} \quad . \tag{5.11}$$

If one does the above, all of the analysis that lead to (5.10) still holds. In particular, the physical probability of the possibility for drawing the first ball black is still n/N which is now p. However, what has been obtained in the process of letting N and n tend to infinity, while leaving their ratio stay constant and equal to p, is that the physical probability of the possibility for drawing a black ball at any time the lever is pressed now stays constant and is given by p. The physical meaning of this is clear. When n and N are very large the ratio of black balls to the total number of balls that are in the tube remains very closely the same after any number of drawings very much smaller than N and n.

Having obtained (5.11), one can now let both m and r be large. This allows us to use the Stirling approximation to the factorials in (5.11). Introducing the notation f_m for the ratio r/m, the formula (5.11) becomes:

$$(f_m|p) = \left[2\pi m f_m(1-f_m)\right]^{-1/2} \left[\left(\frac{p}{f_m}\right)^{f_m} \left(\frac{1-p}{1-f_m}\right)^{1-f_m}\right]^m \quad . \tag{5.12}$$

The mathematical analysis of the distribution (5.12) as a function of m has been performed in great detail long ago (Laplace,[16] Chapter III, Book II, and Poincaré,[7] Chapters IV and V). The purpose of this analysis is to show that the sequence of distributions (5.12) as a function of m tends to a limiting distribution which in modern terminology is the Dirac delta distribution δ(f-p). This result can be verbalized in physical terms as follows. According to the physical theory used, the conceptual limiting

frequency with which the outcome of a particular type would be observed, if the experiment were carried out indefinitely, is equal to the physical probability of this outcome on the data.

It must be emphasized that the above interpretation of the fact that the limiting distribution (5.12) is the Dirac delta distribution $\delta(f\text{-}p)$, is a statement made from the point of view of the physical theory used about entities that have not been observed and says nothing about what will be observed. Any assertion about what will be observed in an experiment is an assertion of a physical theory since it is a statement about what is and has not been observed and must therefore be either true or false.

Although the above derivation of the relation between conceptual limiting frequencies and physical probabilities of possibilities for the outcome of experiments is the traditional one given in the classical theory of probability,[16] it only applies to observables of physical theories. However, data of physical theories can be reduced for entities that are not observables, and it is not very meaningful then to speak of limiting frequencies of what would be observed if the experiment was carried out indefinitely. For every data reduction in which possibilities for entities of theories are generated by a group of transformations, there is another conceptual limiting frequency that one can introduce and which is numerically equal to the normalized volume measure in the group manifold (the physical probability of the possibilities). This conceptual limiting frequency arises from the group axioms. Each possibility for an entity of a theory is generated by applying conceptually a transformation of the physical system defined by the observations (the data). Each different transformation produces a physical system that has the same elements of apparent symmetry on the data; therefore one can conceptually apply indefinitely these transformations and count the number of times a particular possibility is generated as well as count the total number of times these transformations have been applied. In the limit where this is done an infinite number of times, each possibility will occur the same number of times and the ratio of the number of times a particular possibility occurs to the total number of times these transformations have been applied is a limiting frequency which is equal to the normalized volume measure in the group manifold. There seems very little doubt that the above conceptual physical interpretation which can be given to the normalized volume measure in a group manifold may have been at least intuitively perceived when it was introduced and is reflected in the manner it is called: it is a countably additive measure that can

be associated with each group element of the group. It may not be unreasonable to conjecture that the above physical interpretation given to the volume measure in a group manifold may play some role in the apparently long held and deep-seated conviction that there is a relationship between probabilities and limiting frequencies. What is argued above is that these frequencies are conceptual as opposed to being real and result from apparent symmetries that arise from the interpretation in a theory of what a physical system is on the basis of what is perceived to have been observed.

6. PHYSICAL PROBABILITIES OF PHYSICAL PROBABILITIES

In all of the examples of data reductions performed so far in this paper, one could easily find in the theory used a group of transformations of the physical system analyzed whose elements changed what was the entity of interest and left invariant the data corresponding to the observations. In the type of data reductions now considered such a direct relationship is lacking between the data corresponding to the observations and the entity of interest. What happens in these situations is that an element of symmetry of the physical system involves not the entity of interest but some other entity. One can therefore introduce a group of transformations whose elements change what this other entity is and leave invariant the data corresponding to the observations. This group of transformations is therefore a possibilities-generating group for this other entity and one can determine the physical probability of each possibility for this other entity. Now one finds that some, if not all, of the possibilities for this other entity give rise in the physical situation to an element of symmetry that involves the entity of interest when they are individually adjoined to the data corresponding to the observations. One can now introduce a possibilities-generating group for the entity of interest. The elements of this group of transformations leave invariant not only the data corresponding to the observations but also one of the possibilities for the other entity. Since this second possibilities-generating group is a function of the elements of the first possibilities-generating group one must assign to this second possibilities-generating group the physical probability of the possibility for the other entity that enters into it. It is appropriate in these situations to view the first group of transformations as a possibilities-generating group for possibilities-generating groups for the entity of interest. In these situations it is also appropriate to say that the physical probability of the possibilities for the entity of interest, determined by the second group of transformations, have a physical probability determined by the first group of transformations.

In the proposed method of data reduction, it is the above mechanism that explains and allows us to deal formally, therefore quantitatively, with the intuitive notion that all probabilities which are numerically equal are not the same. This also explains the intuitive notion that probabilities are uncertain. Two examples of data reductions involving possibilities-generating groups for possibilities-generating groups are first given. Then the situation is analyzed formally from the point of view of group theory.

6.1. Discrete Probability Space

The first example considered involves a finite group of transformations that generates, on the data in the experiment, possibilities for an entity which is not the one of interest. Then each of the possibilities for this entity so obtained is used to determine a possibilities-generating group for the entity of interest.

The physical situation that forms the basis of the data reduction is the same as the one considered in Sect. 5. That is to say, there are N numbered balls, numbered 1 through N, in a vertical tube which has a release mechanism at the bottom. However, instead of having observed in this experiment the number of balls in the tube, N, it has only been observed that this number of balls is not less than $N_1 > 0$, but not more than N_2, where N_{12} and N_2 are two integers with $N_1 < N_2$. As in the previous section, the order in which the balls have been inserted into the tube from the top has not been observed. The data corresponding to the observations in this experiment is written symbolically $N_1 \leqslant N \leqslant N_2$. The object of the data reduction is to find the physical probability of the possibility that the first ball drawn would have the number 1 on it. In the notation of Sect. 5, one seeks the physical probability $(F_1|N_1 \leqslant N \leqslant N_2)$.

According to classical physics, there is a number of balls in the tube and let us denote this number by N_j. If this number of balls, N_j, had been observed, then as derived in Sect. 5, one would have:

$$(F_1|N_j) = \frac{1}{N_j} \quad . \tag{6.1}$$

The result (6.1) holds for all values of $N_j > 0$ and therefore it holds for $N_1 \leqslant N_j \leqslant N_2$. Consequently, one can formally write:

$$(F_1|[N=N_j]\Lambda[N_1 \leqslant N \leqslant N_2]) = \frac{1}{N_j} \quad . \tag{6.2}$$

The physical probability (6.2) is obtained from the cyclic group of order N_j, as explained in Sect. 5. The possibilities-generating group in (6.2) is one of the possible possibilities-generating groups for the ball at the bottom of the tube; it is the one when the number of balls in the tube is N_j. One must formally generate by a transformation group all of the possible values for N. Therefore this

possibilities-generating group for N generates all of the possible possibilities-generating groups for the ball at the bottom of the tube. This group of transformations will also consequently establish the physical probability of these possible possibilities-generating groups.

According to classical physics, there must be a particular number of balls, N, which is the number of balls in the tube. The data corresponding to the observations are $N_1 \leqslant N \leqslant N_2$, and they must be reduced for the value of N though not the entity of interest. We must therefore find a group of transformations of the physical system whose elements change the value of N and leave invariant the data $N_1 \leqslant N \leqslant N_2$. This possibilities-generating group was given in Sect. 3 in terms of "creation-anihilation" operators but the element of apparent symmetry was not discussed in classical terms; this will now be done. As is always the case, this group of transformations corresponds to an apparent element of symmetry in the physical system when it is viewed from the point of view of the physical theory and only what was observed appears. In this particular situation what is this element of apparent symmetry can easily be visualized. There is a column of balls, N_1 of which are seen, but there is an obstruction that prevents balls in this column to be seen, if there are more than N_1, and this obstruction is finite in extent, such that if there were more than N_2 balls one would see the column extend above the obstruction. Therefore, every column of balls placed behind the obstruction which has at least N_1 balls but no more than N_2 balls, will appear the same. However, this element of symmetry is only apparent in the physical theory. If we did look from behind the obstruction at the column of balls the number of balls, N, that is in this column of balls would be seen. The possibilities-generating group for this data reduction, as discussed in Sect. 3, is isomorphic to the group of addition modulo $N_2 - N_1 + 1$ of the integers O, 1, 2, ..., N_2-N_1. The physical meaning of the elements of this group correspond to adding and subtracting balls to the column, such that the number of balls in the column before and after each transformation is at least N_1 balls but not more than N_2 balls, except for the identity element of the group that does not change the number of balls in the column.

The above possibilities-generating group is isomorphic to the cyclic group of order $N_2 - N_1 + 1$ and the possibilities for N that it generates are $N_1, N_1 + 1, ..., N_2 - 1, N_2$. Each of these possible values for N have the physical probability

$$(N_j|N_1 \leqslant N \leqslant N_2) = \frac{1}{N_2 - N_1 + 1} \quad . \tag{6.3}$$

To each of the possible values N_j for the value of N, there is a different possibilities-generating group for the number on the ball at the bottom of the tube, and the physical probability (6.3) is the physical probability of each of these possibilities-generating groups.

We are now in possession of all the elements needed to calculate the desired physical probability $(F_1|N_1 \leqslant N \leqslant N_2)$. Since we are not interested in which of the possible cyclic groups of order N_j generated the possibility F_1 and all of them can:

$$(F_1|N_1 \leqslant N \leqslant N_2) = \sum_{N_j = N_1}^{N_2} (F_1 \Lambda [N = N_j] | N_1 \leqslant N \leqslant N_2) \quad . \tag{6.4}$$

Using the conjunction theorem of the algebra of physical probabilities, one can expand each term in the sum on the right-hand side of (6.4) to obtain:

$$(F_1|N_1 \leqslant N \leqslant N_2) = \sum_{N_j = N_1}^{N_2} (N_j|N_1 \leqslant N \leqslant N_2)(F_1|[N = N_j] \Lambda [N_1 \leqslant N \leqslant N_2]) \quad . \tag{6.5}$$

Since the two factors in each term of the sum (6.5) are given by (6.3) and (6.2) respectively, the final result is:

$$(F_1|N_1 \leqslant N \leqslant N_2) = \frac{1}{N_2 - N_1 + 1} \sum_{N_j = N_1}^{N_2} \frac{1}{N_j} \quad . \tag{6.6}$$

Had we been interested in the physical probability of the possibility that the first ball drawn had a number on it that was less than or equal to i, then we would be interested in the physical probability of the subset of possibilities for the ball at the bottom of the tube: $\{F_1, F_2, ..., F_i\}$. Using the disjunction theorem of the algebra of physical probabilities yields; if $i \leqslant N_i$:

$$(F_1 V F_2 V ... V F_i | N_1 \leqslant N \leqslant N_2) = \sum_{k=1}^{i} \sum_{N_j = N_1}^{N_2} (N_j|N_1 \leqslant N \leqslant N_2)(F_k|[N = N_j] \Lambda [N_1 \leqslant N \leqslant N_2]) \quad . \tag{6.7}$$

The formulae (6.5) and (6.7) are essentially the "golden rule" of quantum mechanics and were derived from an analysis of a classical experiment without using the theory of statistics but applying group theoretical methods.

It seems clear that on the basis of the observations made in the experiment considered in this section, $N_1 \leqslant N \leqslant N_2$, it is meaningful to ask: what is the physical probability of the possibility for the value N to be a particular value between N_1 and N_2? In classical physics one says that N must have a value. As shown in Sect. 5, having answered this question it is meaningful to ask then: what is the physical probability of the possibility for the number on the bottom ball to be the number "one"? This physical probability is a function of the answer to the previous question. The set of $N_2 - N_1 + 1$ numbers $\frac{1}{N_1}, \frac{1}{N_1+1}, \ldots, \frac{1}{N_2-1}, \frac{1}{N_2}$, are the different answers that one can give to this question depending upon the particular answer given to the previous question. This set of $N_2 - N_1 + 1$ numbers are the different possibilities for this physical probability and each of these has a physical probability. Therefore, in this data reduction one can associate a physical probability distribution with the physical probability $(F_1|N_1 \leqslant N \leqslant N_2)$, this physical probability distribution is the set of physical probabilities of the possibilities for the value of N. As shown by (6.6), $(F_1|N_1 \leqslant N \leqslant N_2)$ is the first moment of this distribution. The mathematical meaning of this is quite clear, the possibilities-generating group $N_1 \leqslant N \leqslant N_2$:F, where F is the bottom ball in the tube, is a very complex group involving the direct product of the cyclic group $C_{N_2-N_1+1}$ and the cyclic groups $C_{N_1}, C_{N_1+1}, \ldots, C_{N_2-1}, C_{N_2}$. Each of the cyclic groups $C_{N_1}, C_{N_1+1}, \ldots, C_{N_2-1}, C_{N_2}$ generates a possibility to which the label F_1 can be attached as well as a different N label from the cyclic group $C_{N_2-N_1+1}$. In the case of the physical probability of the possibility for the number on the bottom ball to be 1, in the data reduction of Sect. 5, there was no such physical probability distribution that it could be associated with it, except the Dirac delta distribution. In the proposed method of data reduction the Dirac delta distribution corresponds to a possibilities-generating group of order one; that is to say, the absence of symmetry.

6.2. Continuous Probability Space

In the data reduction just performed, the group of transformations that generated all of the possible possibilities-generating groups for the entity of interest was finite. This group of transformation is continuous in the data reduction now performed.

The physical situation is again the same one: a vertical tube containing balls with a release mechanism for balls at the bottom of the tube. However, the observations now made on the system are that the balls are of two types: some are black and the remaining ones are white, as in Sect. 5; but the observations on the number of balls yield the data of classical physics: there are at least two balls in the tube and at least one of each color. The data corresponding to the observations are written $N \geqslant 2$ and $0 < n < N$, where N is the total number of balls and n is the number of black balls. N may therefore be arbitrarily large. These data are going to be reduced for the color of the ball at the bottom of the tube which would be, according to classical physics, the first ball drawn.

On the basis of the analysis of this physical situation in Sect. 5, one has:

$$(B_1|N,n) = \frac{n}{N} \quad . \tag{6.8}$$

Therefore a knowledge of both n and N is not required in order to obtain the physical probability for the possibility that the bottom ball is black. It is sufficient to know the ratio n/N which is again denoted p. If the observations in the experiment had only determined the ratio p of black balls to total number of balls in the tube we would have:

$$(B_1|p=n/N) = p \quad . \tag{6.9}$$

From the data in this experiment we have:

$$0 < p = \frac{n}{N} < 1 \tag{6.10}$$

where the limits 0 and 1 must be excluded since $0 < n < N$. What is therefore sought in this data reduction can be written $(B_1|0<p<1)$. The physical probability (6.9) can be rewritten:

$$(B_1|[p=n/N]\wedge[0<p<1]) = p \quad . \tag{6.11}$$

If W_1 denotes the possibility for the bottom ball to be white one always has:

$$(B_1|0<p<1) + (W_1|0<p<1) = 1 \quad . \tag{6.12}$$

From classical physics $n/N = p$ must be a rational number and so is $q = (N - n)/N$; therefore (6.11) and (6.12) can be rewritten:

$$p + q = 1 \quad . \tag{6.13}$$

Although not necessary, it is useful at this stage to consider the above two rational numbers p and q to be the cartesian coordinates of a point P in a real linear two-dimensional space in which one has introduced an orthogonal coordinate system. Because p and q correspond to physical probabilities, the relations (6.10) and (6.13) can be said to define a physical probability space for this data reduction. This physical probability space is embedded into and is a subspace of the above real linear two-dimensional space. In the real linear two-dimensional space, the relations (6.10) and (6.13) define the segment of line joining the end points of the two basis vectors (1, 0) and (0, 1), but the two end points of this segment of line are excluded. Not all of the points of this segment of line are in the physical probability space, only those having rational coordinates. The physical probability space thus defined is a compact topological space.

According to classical physics there must be a number N of balls in the tube and n of these are black; therefore there is a ratio n/N for the specific physical system considered. Consequently, one can say that there is a point P in the physical probability space defined above that corresponds to what is and was not observed. The possibilities-generating group for the ratio n/N in the data reduction must be isomorphic to the group of transformations in the physical probability space that transforms a point

P, (p,q) into the other points P′, (p′,q′), of this space. Because of the topology of the physical probability space, the possibilities-generating group sought can only be the continuous one parameter Lie group that translates any point P of the probability space into any other point P′ of this space. This Lie group is usually denoted T_1 and is the projective group of transformations defined by the transformation:

$$p \rightarrow p' = p + m \tag{6.14}$$

where m is a rational number. Since we have $0 < p < 1$ and $0 < p' < 1$ the parameter space, or group manifold, of this group is closed. The volume measure in the group manifold of this group is a constant. Normalizing the volume measure to unity over the closed group manifold one obtains:

$$(dp|0<p<1) = (p|0<p<1)dp = dp \ , \ 0<p<1 \quad , \tag{6.15}$$

and the physical probability density function $(p|0<p<1)$ is unity.

We have now all the elements needed to calculate the physical probability $(B_1|0<p<1)$ from the physical probability distribution (6.15) associated with it. From the sum rule of the algebra of physical probabilities one obtains:

$$(B_1|0<p<1) = \lim_{\substack{a\rightarrow 0\\ b\rightarrow 1}} \int_a^b (B_1 \wedge [p = \frac{n}{N}]|0<p<1)dp \quad . \tag{6.16}$$

Expanding the integrand by the conjunction theorem yields:

$$(B_1|0<p<1) = \lim_{\substack{a\rightarrow 0\\ b\rightarrow 1}} \int_a^b (B_1|[p = \frac{n}{N}] \wedge [0<p<1])\ (p|0<p<1)dp \quad , \tag{6.17}$$

and the two factors in the integrand are given by (6.11) and (6.15) respectively. Performing the integration and taking the limits indicated gives the numerical value:

$$(B_1|0<p<1) = \frac{1}{2} \quad , \tag{6.18}$$

and from (6.12) one also obtains $(W_1|0<p<1) = 1/2$.

The physical probability $(B_1|0<p<1)$ may be written $(B_1|[N \geqslant 2]\Lambda[0<n<N])$. From Sect. 5 one has $(B_1|[N=2]\Lambda[n=1]) = 1/2$ which is the data reduction for a physical situation related to the one considered in this section. This latter physical probability has the same magnitude as the one just derived but the physical probability distribution associated with it is the Dirac delta distribution.

6.3. Physical Probability Spaces and Their Group Structure

In the last two data reductions, it was necessary to introduce a possibilities-generating group for an intermediate entity of the theory before a group of transformations involving the entity of interest could be found. In the usual language of group theory, one had to adjoin to the data corresponding to the observations another datum of the theory before a group of transformations could be found whose elements generated the possibilities for the entity of interest. This other datum of the physical theory, since it was not among the data corresponding to the observations, could only be a possibility for another entity of the theory and a possibilities-generating group for this other entity had to be found and was found. It was then a simple matter to sum over all of the possibilities for this other entity to obtain the desired result. Because there was a physical probability distribution corresponding to the possibilities for this other entity, the end result was that there was a probability distribution "associated" with the physical probability of each possibility for the entity of interest. In fact, the physical probability of each possibility for the entity of interest was the first moment of its associated physical probability distribution. The above feature is found in most data reductions and is largely responsible for the fact that it is frequently not obvious at first sight that there is a group of transformations of the system that transforms the entity of interest but leaves invariant the data corresponding to the observations. There is, however, an aspect of the data reductions in the last two examples which seemed to have been intuitively perceived: it is the notion that there is more to the physical probability of a possibility for an entity than its magnitude, that physical probabilities of a given magnitude may be more or less uncertain. It seems clear that this notion is intimately related to what has been called the physical probability distribution associated with the physical probability of the possibility for the entity of interest.

Jaynes[8] discovered on purely algebraic grounds that the algebra of probabilities was such that one could formally always introduce a probability distribution associated with a probability if Bayes hypothesis held. He calls the probability distribution the A_p distribution and relates it to the "strength of the evidence" upon which the probability p is based.

In the last data reduction, a physical probability space for this data reduction was introduced. This physical probability space was embedded into a subspace of a real linear space in which an orthogonal system of coordinates was introduced. It was found that there was a group structure associated with this physical probability space regarded as a topological space. It will now be shown that such physical probability spaces can always be introduced in data reductions. The group structure of such probability spaces play an important role in many data reductions; for instance in the analysis of the shape of an angular distribution or the shape of an energy spectrum.

Given a data reduction in which the possibilities-generating group for the entity of interest is of order m, where m is infinite if the group is continuous, one can always form mutually exclusive subsets of the set of m possibilities. Let there be n such subsets whose union forms the complete set of possibilities for the entity. It is then said that we have an exhaustive set of n mutually exclusive possibilities and, if the order of the possibilities-generating group is greater than n, some of these n possibilities are decomposable into several mutually exclusive possibilities for the entity of interest. Let us denote by p_i the physical probability of the ith subset of the set of possibilities, then we necessarily have:

$$\sum_{i=1}^{n} p_i = 1 \quad , \tag{6.19}$$

where the p_i's are positive and rational numbers. If one now considers the n p_i's as the components of an n-dimensional vector P, then one can consider the set of all such vectors P as defining an n-dimensional physical probability space which is embedded into an n-dimensional Euclidean space. By an n-dimensional Euclidean space is meant a n-dimensional linear real space in which a system of orthogonal coordinate has been introduced with the usual Euclidean metric. Such a physical probability

space lies in an n-1 dimensional subspace of the n-dimensional Euclidean space and is in a closed domain in this subspace. This domain is bounded by the lines that join the "points" of the Euclidean space corresponding to the position of the unit vectors. These lines define a n-sided regular polygon in an n-1 dimensional subspace. All of the points inside this n-sided regular polygon in the Euclidean n-1 dimensional subspace are not in the n-dimensional physical probability space: only the points whose coordinates are rational numbers.

It is clear that the n-dimensional physical probability space defined as above is a compact topological space and that there is a group structure associated with it. This group is an invariant subgroup of GL(n,R), it has n-1 essential parameters and is usually called the translation group T_{n-1} and the group manifold of this group is closed. As is well known, the volume measure in the group manifold of this group is uniform over the group manifold.

From a purely mathematical point of view, what has just been shown is that one can consider the normalized volume measures in the group manifolds of groups as components of abstract probability vectors and these vectors define a topological probability space embedded into a subspace of a linear Euclidean space. There is a particular group structure associated with such a probability space. The relevance of the above for physical theories is quite clear. Physical probabilities of possibilities for entities of physical theories can be treated in physical theories as any other abstract entities of these physical theories in the sense that they are abstractions associated with a group structure in an abstract linear space. This seems to have been clearly realized by Laplace[16] who treated physical probabilities on the same footing as any other entities of classical physics assigning probabilities to their possibilities on the basis of the Principle of Insufficient Reason as any other position parameter of classical physics. This was earlier perceived by Bayes[21] but he seemed far more reluctant to do this and he found that he did not have to do it in the particular problems he was interested in solving. The fact that one can define probability spaces that have a definite group structure, as just shown, establishes the very controversial result of the classical theory of probability known today as "Bayes hypothesis." When one defines physical probabilities as done in this paper, Bayes hypothesis becomes a theorem of group theory.

7. BERNOULLI TRIALS

In all the thought experiments analyzed, classical physics was the physical theory used, and the data were reduced for an observable of classical physics that had not been observed, as it must be. One can conceive that these entities for which the data were reduced since they were observables of the theory used are the outcomes of the experiments. If one takes this point of view, the experiments were not complete since their outcomes were not observed; for instance, the side of the coin facing up, the number of balls in the tube, the number on the ball at the bottom of the tube, etc. This is a legitimate point of view, and, in most experiments, there is among the observations one of them that is called, with some legitimate reason, the outcome of the experiment. Most such experiments are repeated in the sense that the observations that are not called the outcome are said to define the conditions under which the experiment is conducted and produces the observation called the outcome. It is then found that when one repeats the experiment in the above sense, the observed outcomes vary: there is a distribution of outcomes observed. These so-called repetitions of the experiment can only be made a finite number of times. It is for the treatment of such series of outcomes that the theory of statistical estimation is used today. It cannot properly be used to deal with observations that are not considered a series of outcomes in the above sense. It is, however, perfectly legitimate to consider everything that was observed in an experiment as the observations made in one experiment and to analyze this experiment on the basis of a theory. This latter point of view involves properly encoding all of the observations made into data of a theory. Then these data can be used to analyze their implications in this theory for some related entity of this theory. It is by taking this second point of view that experiments involving a so-called set of observed outcomes are analyzed in the proposed method of data reduction.

In the usual point of view where one applies the theory of statistical estimation, the question that is answered first is: what would be the limiting frequencies with which the different outcomes would be observed to occur if one were to repeat the experiment an infinite number of times and if one assumes that the repetitions of the experiments were random samples from a "parent" distribution? Having obtained an estimated "parent" distribution, one can then answer any number of questions related to random sampling from this "parent" distribution using the theory of statistics. In particular, if one asks

what are the probabilities for all the different outcomes at the next repetition of the experiment, the answer is given by the estimated "parent" distribution. Therefore, to show that the proposed method of data reduction is capable of dealing with data that contain a sample of outcomes in the experiment, the question that will be asked is: what is the entity of the physical theory that corresponds to the next outcome on the basis of all the observations?

A concrete example is given before the general technique, known as Bayesian method of data reduction, is briefly discussed.

7.1. More Black and White Balls in a Tube

Once more a tube containing black and white balls is used to provide a concrete example of a repeated experiment, but this time some outcomes will be observed. The outcomes here are the color of the balls that are drawn from the bottom of the tube. First, one must define the observations that determine one is indeed repeating an experiment. It is the fact that the data of the physical theory corresponding to the observations at the beginning of each trial are the same, as far as the outcome of interest is concerned, that determines one is repeating an experiment. In the specific example considered now, the data of classical physics identified with the observations are: (a) the number, N, of black and white balls in the tube is considerably larger than any number of balls one may contemplate drawing from the tube; and (b) both the number, n, of black balls and N-n of white balls in the tube are also considerably larger than any number of balls one may contemplate drawing from the tube. What is precisely required of these data, called the initial data, are that the initial ratio of black balls, n, to total number of balls, N, $p = n/N$, remains practically unchanged as one draws balls from the tube, whatever their color happens to be. In other words, one requires that if it is contemplated drawing altogether m balls from the tube, the initial data are such that: $p = n/N \simeq n/(N-m) \simeq (n-m)/(N-m)$. By choosing such initial conditions, if neither n nor N have been observed then one can select the ratio of black balls to total number of balls in the tube as the entity of interest for which the data are to be reduced. Each drawing of the same number of balls constitutesa repetition of the experiment (a trial) and the different outcomes can be taken as the number of black balls which is observed among this set of balls.

In this experiment the initial conditions before each drawing of a set of balls yield the data of classical physics $0<p<1$ where p is the ratio of black balls to the total number of balls in the tube at each drawing. These data are the very same ones that were used in the second example in Sect. 6 and were reduced to obtain the physical probability of the possibility that the first ball drawn is black. In order to be more general, these data are now reduced for the physical probability of the possibility for obtaining r black balls when m of them are drawn from the tube. In this fashion one can consider the drawing of m balls from the tube as one outcome or a set of m outcomes. Therefore what one seeks is $(m,r|0<p<1)$ and on the basis of the analysis in Sect. 5, from formula (5.11), one obtains:

$$(m,r|[p=n/N]\Lambda[0<p<1]) = \binom{m}{r} p^r \ (1-p)^{m-r} \quad . \tag{7.1}$$

What is sought can be written using the sum rule of the algebra of physical probability:

$$(m,r|0<p<1) = \lim_{\substack{a\to 0\\ b\to 1}} \int_a^b ([m,r]\Lambda p|0<p<1) \ dp \quad . \tag{7.2}$$

Using the conjunction theorem one can expand the integrand to yield:

$$(m,r|0<p<1) = \lim_{\substack{a\to 0\\ b\to 1}} \int_a^b (p|0<p<1)(m,r|[p=n/N]\Lambda[0<p<1]) \ dp \quad . \tag{7.3}$$

The density function $(p|0<p<1)$ was calculated in Sect. 6 and the other term in the integrand is given by (7.1). Consequently, one obtains for (7.3):

$$(m,r|0<p<1) = \binom{m}{r} \lim_{\substack{a\to 0\\ b\to 1}} \int_a^b p^r(1-p)^{m-r} \ dp \quad . \tag{7.4}$$

Since $0 \leqslant r \leqslant m$, the integral in (7.4) is the well-known beta function of arguments $r+1$ and $m-r+1$:

$$\int_0^1 p^r(1-p)^{m-r} \ dp = \frac{\Gamma(r+1)\ \Gamma(m-r+1)}{\Gamma(m+2)} \quad . \tag{7.5}$$

Then the result sought is:

$$(m,r|0<p<1) = \frac{1}{m+1} \quad . \tag{7.6}$$

This result is a generalization of the data reduction performed in Sect. 6 where B_1 stands for m = 1 and r = 1 in the notation now adopted. This result can also be interpreted as giving the physical probability for a set of m outcomes of drawing one ball from the tube, where this set of m outcomes is only characterized by the number of times a particular type of outcome is observed in this set.

Now, for the first time in this paper, let the first m balls be drawn from the tube and let r be the number of black balls found among these m balls. For someone who has observed the number, r, of black balls in this set of m balls, the result (7.6) ceases to have any meaning whatsoever because it refers to possibilities and what has been observed is a fact, not a possibility. For someone who knows because he has observed that m balls have been drawn from this tube but has not observed how many of them are black, then (7.6) holds as the physical probability of the possibilities for the outcome, or set of outcomes. What is also remarkable is that the result (7.6) also holds for someone that knows nothing about the initial conditions of the experiment upon which (7.6) was derived; namely that N, n, and N-n are all large compared to m. The only necessary condition for (7.6) to hold is that this person be told that m outcomes have been observed for an experiment and that there were only two types of outcome in this experiment. Then to this individual there are m + 1 possible sets of outcomes where each possible set of outcome is defined by the number r that it contains of outcomes of one type. The data of physics that this information corresponds to is: $m\Lambda(0 \leqslant r \leqslant m)$. This problem was considered before and yields as the group of symmetry the cyclic group of order m + 1, which yields (7.6) for every possible set of outcome.

Let us now consider the data reduction for someone who has observed m balls being drawn from the tube and that r of them are black. For this individual, his data concerning the number of black balls, n - r, and total number of balls, N-m, in the tube still yield, in classical physics for the ratio of black balls to total number of balls in the tube, the same value p that prevailed before the m balls were drawn and $0<p<1$. Therefore, as far as the ratio of black balls to total number of balls is concerned, the physical system is in the same state as it was before the m balls were drawn. If he were to draw m more balls

from the tube he could consider it a repetition of his first experiment in both the sense of one outcome or in the sense of m outcomes. However, he cannot ignore the fact that he has made the observations that there were r black balls when m of them were drawn. Any further data reduction on his part must now be for the data: (m,r)Λ(0<p<1).

For generality's sake, let us now reduce the data (m,r)Λ(0<p<1) for the possibility of drawing exactly l black balls if k more balls were to be drawn from the tube now. In order to show that no new concepts, or notions, are involved in the proposed method of data reduction when, as it is now said, there is a sample of outcomes in the data to be reduced, the data reduction now performed follows very precisely the one just completed. What is now sought is (k,l|[m,r]Λ[0<p<1]) and what corresponds to (7.1) is:

$$(k,l|[p = \frac{m-r}{N-m}]\Lambda[m,r]\Lambda[0<p<1]) = \binom{k}{l}p^l(1-p)^{k-l} \quad . \tag{7.7}$$

The formula (7.2) becomes:

$$(k,l|[m,r]\Lambda[0<p<1]) = \lim_{\substack{a\to 0\\ b\to 1}} \int_a^b ([k,l]\Lambda p|[m,r]\Lambda[0<p<1])dp \quad . \tag{7.8}$$

Expanding the integrand as for (7.2) yields:

$$(k,l|[m,r]\Lambda[0<p<1]) = \lim_{\substack{a\to\infty\\ b\to 1}} \int_a^b (k,l|p\Lambda[m,r]\Lambda[0<p<1])(p|[m,r]\Lambda[0<p<1])dp \quad . \tag{7.9}$$

The first factor in the integrand is given by (7.7), but the density function (p|[m,r]Λ[0<p<1]) remains to be calculated and is no longer (p|0<p<1) as was the case in (7.3).

It is straightforward to calculate the density function (p|[m,r]Λ[0<p<1]) using the algebra of physical probabilities. The integrand in (7.2) was expanded in a particular fashion to become the integrand in (7.3) but it can also be expanded as:

$$(m,r|0<p<1)(p|[m,r]\Lambda[0<p<1])dp \quad . \tag{7.10}$$

Setting (7.10) equal to the integrand in (7.3) yields immediately:

$$(p|[m,r]\Lambda[0<p<1]) = (p|0<p<1)\frac{(m,r|p\Lambda[0<p<1])}{(m,r|0<p<1)} \quad , \tag{7.11}$$

which is Bayes theorem applied to physical probability density functions. Every term on the right-hand side of (7.11) has now been calculated: the density function $(p|0<p<1)$ is equal to unity, the numerator is given by (7.1) and the denominator by (7.6). Therefore one obtains the density function:

$$(p|[m,r]\Lambda[0<p<1]) = (m+1)\binom{m}{r} p^r(1-p)^{m-r} \quad , \tag{7.12}$$

which is considerably different from the uniform value it had before the m balls were drawn and it was observed that r of them are black.

The integration in (7.9) can now readily be done since it is also expressible as a Beta function. Rearranging the factorials into binomial coefficients one obtains:

$$(k,l|[m,r]\Lambda[0<p<1]) = \frac{\binom{r+l}{r}\binom{m+k-r-l}{m-r}}{\binom{m+k+1}{k}} \quad . \tag{7.13}$$

The expression (7.13) can be used to calculate the physical probability of the possibility for drawing a black ball on the first draw after having drawn m balls from the tube and observed that r of them were black. This result is obtained from (7.13) by letting $k = 1$ and $l = 1$, and this may again be written B_1 and we have:

$$(B_1|[m,r]\Lambda[0<p<1]) = \frac{r+1}{m+2} \quad . \tag{7.14}$$

The result (7.14) is known as Laplace's "Rule of Succession" and has been considered very controversial since derived by Laplace on the basis of Bayes hypothesis nearly 200 years ago. As it must, the result (7.14) reduces to the result obtained prior to the drawing of any balls by letting $m = r = 0$ in it.

7.2. Criterion for Bernoulli Trials

The notion of repetition of experiments under presumed identical conditions plays an important role in physics and is now briefly analyzed from the point of view of the proposed method of data reduction.

It may appear at first sight that the situation just analyzed is not a good example of the repetition of an experiment if one were to draw a second set of m balls from the tube after having drawn the first set of m balls. Since, after having drawn the first m balls from the bottom of the tube, the system is missing these m balls and there is now a different set of m balls at the bottom of the tube, the system is indeed different before and after the first m balls were drawn. The system is in two different states according to the theory used before and after the drawing of the first m balls. However, if one considers the original data and what they were reduced for: what is the number of black balls among the bottom m balls in the tube? and one takes into account the fact that m balls were drawn but exclude the fact that the outcome is known; then, whatever the outcome was, the same question being asked before the first drawing of m balls asked again after the m balls are drawn will yield the same set of possibilities for the answer with the very same set of physical probabilities. It is because one knows from the point of view of the theory applied and all of the observations made that the system is different, except for the same element of apparent symmetry with respect to the entity of interest, and will be so after the second outcome is observed that it is meaningful to ask the same question and reduce the data for the entity of interest. However, as was shown in the data reduction just performed, when one uses the exact observed outcomes of the trials (the number of black balls among the total number of balls drawn) the element of apparent symmetry of the physical system as viewed in the theory applied on the basis of all the observations changes dramatically and is reflected in the physical probabilities of the possibilities for the outcome of a subsequent trial. In particular, if one calculates the physical probability of the possibility for drawing one black ball at the next draw, since the ratio of black balls in the tube to the total number of balls would be this number if it had been observed, then one is calculating this ratio on the basis of the observations which do not yield this ratio directly but are related to it. As shown in Sect. 6, there is always a physical probability distribution that one can associate with the physical probability of a possibility for an entity related to the observations in a physical theory; there is therefore a physical probability distribution that one can associate with the calculation of the ratio of black balls to the total number of balls in the tube.

On the basis of the above discussion, the data reduction just completed, far from being a purely academic exercise, bears a very close relationship to what is called a physical measurement in a traditional physics experiment. There is a repetition of some observations. At each repetition of the "measurement" the system is not identical to what it was at previous trials, either by necessity or design, but it exhibits, on the basis of all observations at each trial if one excludes only the precise outcomes obtained at previous trials, the same element of symmetry with respect to the entity of interest in the theory applied.

7.3. General Remarks

The analysis technique used in the specific example just analyzed is well known and is referred to frequently[22] as a Bayesian method of inference. This technique can always be applied in the proposed method of data reduction because the data, corresponding to the observations which determine the experiment is being repeated, can be reduced to determine the physical probabilities of the possibilities for the entity of the physical theory that corresponds to what is called the outcome of the experiment. In the proposed method of data reduction, there is no uncertainty as to what to use as the "estimator," using the terminology of the theory of statistical estimation. This is because a physical theory is used to analyze the data. This physical theory, whatever it may be, determines the element of apparent symmetry which the physical system analyzed is presumed to have when viewed on the basis of the laws of this physical theory. This element of apparent symmetry determines uniquely the possibilities-generating group for the entity of interest.

When there is an observed sample of outcomes from repetitions of experiments, there are many other questions that one can try to answer than the one considered: what are the physical probabilities of the possibilities for the outcome of the next trial? For instance, one could ask if there are any meaningful regularities in the observed sequence of outcomes? or, is there any "memory effect" whereby the outcome at a particular trial influences to some degree the subsequent outcome or outcomes? If there

were no such effects, the trials could be considered independent and here again group theoretical techniques can be used to analyze these points since there should be no elements of symmetry in the sequence of outcomes. There are also other well-known notions of the theory of statistical estimation that can be investigated using group theoretical methods, such as the quantification of the extent to which the sample of outcomes supports one or the other of two proposed physical theories that compete as an explanation of the physical phenomenon. These questions will be investigated in a subsequent paper.

8. CONCLUSIONS

Independent of the philosophical justification offered for it and even though it may not be without some merit on cultural grounds, the method of analysis of observations presented in this paper solves a fundamental problem of scientific methodology. Given that scientific theories are proposed as explanations for observed phenomena and that there may be several different competing explanations for a particular phenomenon, this problem is: How can these theories be applied to the analysis of observations for which they are proposed explanations, strictly on the basis of their own concepts?

Group theory applied in the manner shown in this paper solves this problem for physical theories, since it makes use of only the physical concepts of the particular theory being applied to analyze the observations and can deal with the notion of ambiguity that results from most experiments. This application of group theory provides a rationale, within the framework of the physical theory used in the analysis of the observations, for formulae that have long been applied successfully to deal with only a certain type of observations. These formulae were previously justified on the basis of a specific theory, the theory of statistics, whose concepts are not necessarily coherent, in a two-valued logic system, with those of the physical theories being used with it. This meant that all physical theories had to be given a statistical interpretation when they were not based upon stochastic considerations, such as theory of measurements, classical mechanics, quantum mechanics, relativity theories, etc. The mathematical foundation of these formulae in group theory is as secure as their foundation in the Lebesgue extension of Borel measures in a σ-field of sets. Therefore, the issue of what interpretation to give these formulae, which strongly influences for which application it is perceived legitimate to use them and what one perceives to derive from their use, is not decidable on purely mathematical grounds.

The group theoretical interpretation of physical probabilities that results from the analysis given in this paper appears to have greater powers of explanation than the statistical interpretation since it allows us to deal in a unified fashion with the different types of observations made in actual experiments that lead to the well-accepted notions of "systematic" and "statistical" uncertainties. How to deal with observations that lead to "systematic" uncertainties is one of the great unsolved problems of the theory of statistical estimation since it is a "rare event problem." Although the "rare event problem,"

which is the problem of dealing in the theory of statistical estimation with events whose probabilities are such that none have been observed so far, has received a lot of attention lately, in view of its societal importance in risk analysis, limited progress seems to us to have been made so far to solve it on the basis of statistics. The "rare event problem" of risk analysis is formally identical to the situations analyzed in this paper where there was no direct observation of the entity of interest, and it was said there is a systematic uncertainty. This lack of progress is largely responsible for the revival of the notion of Bayesian probabilities during the last two decades.

The physical probabilities of this paper, the normalized volume measures in the group manifolds of possibilities generating groups, are neither observables of the method of data reduction nor of the physical theories that gave rise to them. Probabilities that are not physical observables are said today to be Bayesian probabilities and said to express degress of beliefs, or ignorance, of individuals, therefore argued to have no usefulness in science. Although the physical probabilities introduced in this paper are Bayesian probabilities, since they are not observables, far from being unscientific are a necessary consequence of the rigorous application of the scientific methods and are elements of deductive logic. Among other things the scientific method requires that scientific theories be deductive systems of assertions, possessing observables so that they be capable of making predictions about observables, on the basis of observations, in order to be falsifiable. It is therefore a requirement of the scientific method that deductive logic be used to generate from observations made in experiments assertions about observables that have not yet been directly observed. We have seen in this paper that deductive logic forces us to assign weights to assertions of theories dealing with entities of theories that have not been observed but are related to the observations. Although not observables, these weights (called physical probabilities in this paper) can be interpreted in the theory as conceptual limiting frequencies of future observations that should be made on the basis of what was observed. These conceptual limiting frequencies, dealing with future observations, cannot be observables for an essential reason: future observations cannot be ever observed. As soon as one makes an observation, this changes the physical probabilities that must now

be invariant with respect to all the observations that were made. Physical probabilities are therefore only elements of deductive logic that are relative to a set of observations being analyzed from within a deductive system of assertions: a physical theory.

This paper does not say anything about the usefulness of the concept of stochastic physical phenomena to construct a physical theory. However, the use of the mathematical theory of statistics to deal with observations transforms this abstract deductive system of assertions into a physical theory since it introduces into it observables. It is in essence argued in this paper that when this is done the theory of statistics should be treated as a physical theory and judged on the basis of its experimental verification and explanative powers. It is argued that it seems unscientific to promote a physical theory, the theory of statistical estimation, to a priviledged status acting as a filter that must be used in the testing of other physical theories. The method of data reduction presented in this paper makes it in any event unnecessary.

9. ACKNOWLEDGMENTS

I want to thank Professor E. T. Jaynes for sending me a copy of the fragmentary 1974 edition of his lecture notes on probability theory.[8] I also thank Professor Garrett Birkhoff for drawing my attention to G. D. Birkhoff's Rice lecture[2] on Leibniz Principle of Sufficient Reason. I thank my many colleagues at ORELA, in particular Gerard deSaussure who refused to accept as satisfactory my lengthy epistemological considerations during the last three years and thereby forced me to seek a more conventional representation of the ideas. Finally, I am grateful to F. C. Maienschein and R. W. Peelle for their continued support during this research.

REFERENCES

1. *Ecrits et Memoires Mathematiques d'Evariste Galois*, ed. by R. Bourgne and J. P. Azra, Gauthiers, Villars, Paris, 1962.

2. G. D. Birkhoff, *Rice Institute Pamphlet* 28:1 (1941), reprinted in *George D. Birkhoff Collected Mathematical Papers*, vol. III, p. 778, American Mathematical Society, 1950.

3. T. L. Fine, *Theories of Probability*, Academic Press, 1973.

4. I. J. Good, *Probability and the Weighing of Evidence*, Charles Griffin & Co., 1950.

5. I. Todhunter, *A History of the Mathematical Theory of Probability*, first edition, Cambridge, 1865. Reprinted Chelsea Publishing Co., 1965.

   J. M. Keynes, *A Treatise on Probability*, MacMillan and Co., first edition, 1921; reprinted in 1952.

6. E. T. Jaynes, *Trans. Sys. Sci. and Cybernetics* 2, 227 (1968).

7. H. Poincare', *Calcul des Probabilités*, Gauthiers-Villars, Paris, 1896. Revised edition, 1912.

8. E. T. Jaynes, *Probability Theory*, Washington University, fragmentary edition of 1974.

9. R. T. Cox, *Am. J. Phys.* 17:1 (1946) and *The Algebra of Probable Inference*, John Hopkins University Press, 1961.

10. *Oeuvres Complétes de Niels Henrik Abel*, vol. I, p. 61, ed. by L. Sylow and S. Lie, 1881. Johnson Reprint Corporation, 1973.

11. H. Poincare', *Journal de Mathematiques* 9:139 (1903). Reprinted in vol. III of *Oeuvres de Henri Poincare'*, Gauthias-Villars, 1934.

12. C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, Chapt. 3, W. H. Freeman and Co., Inc., San Francisco, 1973.

13. E. P. Wigner, *Group Theory*, translated from the German by J. J. Griffin, expanded and improved edition, New York Academic Press, 1959.

14. M. Hamermesh, *Group Theory*, Addison-Wesley Publishing Co., Inc., 1962.

15. G. Birkhoff, *Lattice Theory*, American Mathematical Society, Providence, Rhode Island, first edition, 1940; third edition, 1967.

16. Laplace, *Théorie Analytique des Probabilités*, first edition 1812, reprinted as *Oeuvres Complétes de Laplace*, vol. 7, Ganthiers-Villars, Paris, 1886. (Reprint of the 1820 edition).

17. H. Jeffreys, *Theory of Probability*, Oxford University Press, first edition, 1939; revised third edition, 1967.

18. R. P. Feynman, *The Feynman Lectures on Physics*, vol. I, chapt. 6, Addison-Wesley Publishing Company, Inc., 1963.

19. E. T. Jaynes, *Found. Phys.* 3, 477 (1973).

20. R. Von Mises, *Probability, Statistics and Truth*, first edition, 1928; English translation, The MacMillan Company, 1939.

21. T. Bayes, *Phil. Trans. of the Royal Society* 53:370 (1763). Reprinted in *Biometrika* 45:296 (1958).

22. M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, Hafner Publishing Company, New York, 1958.

D. V. Lindley, *Introduction to Probability and Statistics From a Bayesian Viewpoint*, Cambridge University Press, 1965.

INTERNAL DISTRIBUTION

1-2. L. S. Abbott
3. R. S. Alsmiller, Jr.
4. J. Barhen
5. D. E. Bartine
6. D. G. Cacuci
7. G. de Saussure
8-12. J. K. Dickens
13. G. L. Flanagan
14. M. C. G. Hall
15. R. L. Macklin
16-20. F. C. Maienschein
21. B. F. Maskewitz
22-26. E. M. Oblow
27-31. R. W. Peelle
32-36. C. M. Perey
37-41. F. G. Perey
42. R. B. Perez
43. C. R. Weisbin
44. A. Zucker
45. P. W. Dickson, Jr. (consultant)
46. H. J. C. Kouts (consultant)
47. W. B. Loewenstein (consultant)
48. Richard Wilson (consultant)
49-50. Laboratory Records
51. Laboratory Records-RC
52. ORNL Patent Office
53. Central Research Library
54. ORNL Y-12 Technical Library Document Reference Section

EXTERNAL DISTRIBUTION

55-81. Technical Information Center (TIC)
82. Office of Assistant Manager for Energy Research and Development, Department of Energy, Oak Ridge Operations, Oak Ridge, TN 37830
83. S. L. Whetstone, Division of Nuclear Sciences, Office of Basic Energy Sciences, U.S. Department of Energy, Rm G-355, Washington, D. C. 20545