Number 8

A COLLECTION OF MATHEMATICAL PROBLEMS

By S. M. Ulam

INTERSCIENCE PUBLISHERS, INC., NEW YORK
INTERSCIENCE PUBLISHERS LTD., LONDON

A COLLECTION OF MATHEMATICAL PROBLEMS

S. M. ULAM

Los Alamos Scientific Laboratories, Los Alamos, New Mexico

INTERSCIENCE PUBLISHERS, INC., NEW YORK

INTERSCIENCE PUBLISHERS, LTD., LONDON

First published 1960

Library of Congress Catalog Card Number 60-8486

Interscience Publishers, Inc.
250 Fifth Avenue, New York 1, N.Y.

For Great Britain and Northern Ireland:

Interscience Publishers Ltd.
88/90, Chancery Lane, London, W.C.2, England

PRINTED IN THE NETHERLANDS
BY DIJKSTRA'S DRUKKERIJ N.V., VOORHEEN BOEKDRUKKERIJ GEBR. HOITSEMA, GRONINGEN

In memory of
J. SCHREIER

PREFACE

In introducing the collection of problems forming the substance of this work, it is perhaps necessary to offer more explanations than is usually the case for a mathematical monograph. The problems listed are regarded as unsolved in the sense that the author does not know the answers. In this sense the structure of this small collection differs inherently from that of the well-known collection of problems by Pólya and Szegö [1]. The questions, drawn from several fields of mathematics, are by no means chosen to represent the central problems of these fields, but rather reflect the personal interests of the author. For the main part, the motif of the collection is a set-theoretical point of view and a combinatorial approach to problems in point set topology, some elementary parts of algebra, and the theory of functions of a real variable.

In spirit, the questions considered in the first part of this collection belong to a complex of problems represented in the *Scottish Book*. This was a list of problems compiled by mathematicians of Lwów in Poland before World War II, also containing problems written down by visiting mathematicians from other cities in Poland and from other countries. The author has recently translated this document into English and distributed it privately; the interest shown by some mathematicians in this collection encouraged him to prepare the present tract for publication. Many of the problems contained here were indeed first inscribed in the *Scottish Book*, but the greater part of the material is of later origin beginning with the years spent at Harvard (1936–1940) and a large proportion stems from recent years, appearing here for the first time. Many of the problems originated through conversations with others and were stimulated by the transitory interests of the moment in various mathematical centers. In addition, several problems were communicated by friends for inclusion in this collection. The last few chapters have a different character: the stress is on computations on calculating machines with examples of problems whose study through the use of this modern tool would have, in the author's opinion, great heuristic value.

Most of the problems were seriously considered and worked upon, but with different degrees of attention and time spent on attempts to solve them. Some have been studied by other mathematicians to whom they were communicated orally but others have not been thoroughly investigated and it would not surprise the author if a number admitted trivial solutions. Most of the problems are, so to say, of medium difficulty. A majority of them should definitely not fall into the category of mere exercises to be solved by routine applications of known lemmas and theorems. In fact one of the aims was a selection of "simple" questions in various domains of mathematics; simple, for example, in the sense that no elaborate definitions beyond those used in general courses on set theory, analysis, and algebra would be necessary for their understanding. The author believes that, on a purely heuristic level, a survey of this sort, if properly enlarged and deepened by others, could bring out the possible general and typical common "reasons" for the difficulties encountered in quite diverse branches of mathematics.

The present situation in mathematical research is perhaps different from that of previous epochs in its very great degree of specialization. The connections between different fields are growing more tenuous, or else so general and purely formal, that they become illusory. It has been said that unsolved problems form the very life of mathematics; certainly they can illuminate and, in the best cases, crystallize and summarize the essence of the difficulties inherent in various fields. The very existence of mathematics can be considered as fruitful only because it produces simple and concise statements whose proofs are much more complicated in comparison. Moreover, Gödel's discovery [1] of the existence of undecidable propositions in every consistent system of mathematics, including arithmetic, renders the "probably true" propositions all the more precious. The intriguing possibility which now exists *a priori* of undecidability lends an additional flavor, to some at least, of the unsolved mathematical problems (cf. Weyl [1]).

The separation between mathematical research stimulated by pure mathematics alone and the ideas stemming from theoretical physics has been increasing in the development of these fields during the last few decades. This may seem at first sight surprising, since the ideas and models of reality employed nowadays in physics tend toward increasing abstractness. However, it appears that on the whole, applied mathematics, so-called, deals at the present time in the majority of cases, with questions of classical physics—or else, when it concerns itself with the new theories, its role is restricted to a purely technical intervention. On the conceptual level one does not have enough, it seems, of cross fertilization of ideas! In the author's opinion it appears likely that in the near future the large class of concepts which have their origin in Cantor's set theory [1], which have influenced so many of the purely mathematical disciplines, will play a role in physical theory. The difficulties of the phenomena of divergence in present formulations of field theory may indicate the need for a type of mathematics capable of dealing with physical problems employing actual infinities *ab initio*. Several elementary problems are included here which are intended to indicate the nature of such possible formulations and the kind of mathematical schemes which may be of use in some future physical theories.

The set-theoretical motivation underlying the selection of questions in the various fields to which the problems refer influenced the choice of the more elementary problems and made the illustration of the more sophisticated ideas of recent years, in topology or algebra for example, impractical.

It is impossible to give detailed credit to all who have indirectly contributed to the set of ideas illustrated in the list of problems, but I would like to acknowledge in particular the pleasure of past collaboration with Banach, Borsuk, Kuratowski, Schreier, and

Mazur in Poland, and John von Neumann, Garrett Birkhoff, J. C. Oxtoby, P. Erdös, and C. Everett in this country. Thanks are due to Mrs. Lois Iles and Miss Marie Odell for their work in preparing the manuscript for publication.

S. M. ULAM

Fall, 1959

CONTENTS

Chapter I

Set Theory

1. Introductory remarks 1
2. The operation of direct product 3
3. Product-isomorphisms and some generalizations 5
4. Generalized projective sets 9
4a. Relations between products of different orders 11
5. Projective algebras 12
6. Generalized logic 13
7. Some problems on infinite sets 15
8. Measure in abstract sets 16
9. Nonmeasurable projective sets 17
10. Infinite games 23
11. Situations involving many quantifiers 24
12. Some problems of P. Erdös 25

Chapter II

Algebraic Problems

1. An inductive lemma in combinatorial analysis 29
2. A problem on matrices arising in the theory of automata . 30
3. A fundamental transformation in the "theory of equations" 30
4. A problem on Peano mappings 32
5. The determination of a mathematical structure from a given set of endomorphisms 32
6. A problem on continued fractions 32
7. Some questions about groups 33
8. Semi-groups 35
8a. Topological semi-groups 35
9. A problem in the game of bridge. 36
10. A problem on arithmetic functions 36

Chapter III

Metric Spaces

1. Invariant properties of trajectories observed from moving coordinate systems 37

2. Problems on convex bodies 38
3. Some problems on isometry 39
4. Systems of vectors 39
5. Other problems on metrics 40

Chapter IV

Topological Spaces

1. A problem on measure 43
2. Approximation of homeomorphisms of E^n 43
2a. On the approximability of transformations in three dimensions by compositions of cylindrical mappings 44
3. A problem on the invariance of dimension 45
4. Homeomorphisms of the sphere 47
5. Some topological invariants 48
6. Quasi-fixed points 48
7. Connectedness questions 50
8. Two problems about the disk 50
9. Approximation of continua by polyhedra 51
10. The symmetric product 51
11. A method of proof based on Baire category of sets 53
12. Quasi-homeomorphisms 54
13. Some problems of Borsuk 55

Chapter V

Topological Groups

1. Metrization questions 57
2. Universal groups 58
3. Basis problems 59
4. Conditionally convergent sequences 61

Chapter VI

Some Questions in Analysis

1. Stability 63
2. Conjugate functions 69
3. Ergodic phenomena 70
4. The Frobenius transform 73

5. Functions of two variables 75
6. Measure-preserving transformations 76
7. Relative measure 77
8. Vitali-Lebesque and Laplace-Liapounoff theorems 78
9. A problem in the calculus of variations 79
10. A problem on formal integration 80
11. Geometrical properties of the set of all solutions of certain equations 80

Chapter VII

Physical Systems

1. Generating functions and multiplicative systems 83
1a. Examples of mathematical problems suggested by biological schemata 85
2. Infinities in physics 89
3. Motion of infinite systems, randomly distributed 91
4. Infinite systems in equilibrium 96
5. Random Cantor sets 97
6. Dynamical flow in phase space 104
7. Some problems on electromagnetic fields 107
8. Nonlinear problems 109

Chapter VIII

Computing Machines as a Heuristic Aid

1. Introduction 115
2. Some combinatorial examples 116
3. Some experiments on finite games 118
4. Lucky numbers 120
5. Remarks on computations in mathematical physics 121
6. Examples from electromagnetism 122
7. The Schrödinger equation 123
8. Monte Carlo methods 125
9. Hydrodynamical problems 128
10. Synergesis 135
Bibliography 145

CHAPTER I

Set Theory

1. Introductory remarks

The outstanding unsolved problem of set theory is the well-known continuum hypothesis of Cantor which asserts that the power $c = 2^{\aleph_0}$ of the set of all real numbers is equal to the power $\aleph_1$, the common power of all well-ordered noncountable sets, all of whose segments are finite or countable. We shall not discuss it here; Sierpiński's book [1] in the Polish Monograph Collection deals extensively with various formulations of this hypothesis and with problems which, on the surface, appear more "concrete," but are equivalent or logically related to it. Gödel [2] has investigated the question from the point of view of certain special axiomatizations of set theory. The principal result is that in many such systems, Cantor's hypothesis is either true or else forms an independent statement. Also true or independent is the proposition that certain subsets of the real number system with "paradoxical" properties are projective sets in the sense of Lusin [1]. The problem of the continuum cannot yet be regarded as settled, since none of the axiomatic formulations of set theory can be considered as "definitive" or all-comprehensive, and it is at present impossible to assert that the "naive" set theory, or the intuitive conception of what set theory should be, has found a definitive axiomatic formulation. Apparently it is Gödel's present impression that in a suitable large and "free" axiomatic system for set theory the continuum hypothesis is false. This feeling, based on indications provided by results on projective sets and the abstract theory of measure, has been shared by the author for many years.

The weaker hypothesis: c is less than the first inaccessible aleph, suffices to establish certain results, e.g., in measure theory, valid under the assumption of the continuum hypothesis. For example, the *nonexistence* of a completely additive measure vanishing for sets consisting of a single point and defined *for all* subsets of the interval follows from this hypothesis (Ulam [1]). It is the author's feeling that, in "reasonable" systems of axioms for set theory, even this weaker hypothesis may be false.

Another, perhaps less well known, problem in the general theory of sets is the following question of Suslin: Let C be a class of sets such that every two sets of this class are either disjoint or else one is contained in the other. Every subclass of C consisting only of mutually disjoint sets is countable, as is every subclass such that for every pair of its sets one contains the other. Is C a countable class?

There are many equivalent formulations and some interesting partial results, but the problem must be regarded as unsolved (cf. G. Birkhoff [1], p. 47). The present state of Suslin's and related problems is thus one more of many indications that abstract set theory is far from forming a complete or "dead" field. On the contrary, *the combinatorics of the infinite*, abounding in problems, lead to a vast study which now seems only in its beginnings and is not even systematically formulated in a general form. Indeed, it is possible to generalize and reformulate many problems such as one finds in Netto's book [1] on combinatorics or MacMahon's treatise [1] so as to obtain nontrivial problems about infinite sets. Coming back to Suslin's problem, one can formulate it equivalently in an abstract Boolean algebra or more generally in terms of lattice operations (cf. G. Birkhoff [1]).

Difficult problems arise already in rather simple commutative structures, e.g., in the study of infinite (say countable) Boolean algebras, the equivalent of the propositional calculus, but especially in the more general algebras in which projection operators corresponding to the logical quantifiers are introduced in addition to Boolean operations (cf. Everett and Ulam [1]). Such algebras present ever deeper questions which can be still regarded as concerning pure set theory (Halmos [1]).

Going still further, one employs in set theory operations going beyond all the above, namely, a passage to variables of higher type, i.e., the formation of classes of sets, classes of classes, etc., so that it is possible to treat algebraically mathematical systems of still greater generality. One can formulate problems in naive set theory dealing with properties of *repeated* passages to *variables of higher type*, studied *per se*. We shall content ourselves in stating one:

A "super-class" K of objects is imagined which is closed with respect to the operation of forming the class of all subsets. Starting, say, with the set S_0 of integers, one forms the class of all subsets of this set. This is a set which we denote by S_1; the class of all subsets of S_1 will be denoted by S_2, etc. In addition, the class K is closed with respect to the following construction: if S_i, $i = 1, 2, \ldots$, is any countable collection of classes of sets belonging to K, we form the class of all possible sequences of sets $s_i \in S_i$; we postulate that this class Σ of all such sequences $\{S_i\}$ also belongs to K. Imagine K is the smallest class closed under the two operations. The question now arises of classifying the objects of K by means of transfinite ordinals and of determining the powers of sets forming the elements of K.

2. The operation of direct product

One deals with the operation of the direct product, in a more or less explicit form, in every mathematical theory involving more than one variable. It is used quite explicitly in topology, group theory, measure theory, in the theory of metric spaces, and it also occurs in one form or another in many algebraic theories. It seems, however, that a *general* investigation of the properties of this operation for its own sake, on a set theoretical basis, has not been undertaken in spite of the many common features, presented by problems of "many variables," apparent in the various theories referred to.

The notion of *phase space* in mechanics is essentially that of a product space. The state of a system of particles is represented by a point in a product space, namely a direct product of spaces valid under the assumption of the continuum hypothesis. For example, the *nonexistence* of a completely additive measure vanishing for sets consisting of a single point and defined for *all* subsets of the interval follows from this hypothesis (Ulam [1]). It is the author's feeling that, in "reasonable" systems of axioms for set theory, even this weaker hypothesis may be false.

Another, perhaps less well known, problem in the general theory of sets is the following question of Suslin: Let C be a class of sets such that every two sets of this class are either disjoint or else one is contained in the other. Every subclass of C consisting only of mutually disjoint sets is countable, as is every subclass such that for every pair of its sets one contains the other. Is C a countable class?

There are many equivalent formulations and some interesting partial results, but the problem must be regarded as unsolved (cf. G. Birkhoff [1], p. 47). The present state of Suslin's and related problems is thus one more of many indications that abstract set theory is far from forming a complete or "dead" field. On the contrary, *the combinatorics of the infinite*, abounding in problems, lead to a vast study which now seems only in its beginnings and is not even systematically formulated in a general form. Indeed, it is possible to generalize and reformulate many problems such as one finds in Netto's book [1] on combinatorics or MacMahon's treatise [1] so as to obtain nontrivial problems about infinite sets. Coming back to Suslin's problem, one can formulate it equivalently in an abstract Boolean algebra or more generally in terms of lattice operations (cf. G. Birkhoff [1]).

Difficult problems arise already in rather simple commutative structures, e.g., in the study of infinite (say countable) Boolean algebras, the equivalent of the propositional calculus, but especially in the more general algebras in which projection operators corresponding to the logical quantifiers are introduced in addition to Boolean operations (cf. Everett and Ulam [1]). Such algebras present ever deeper questions which can be still regarded as concerning pure set theory (Halmos [1]).

Going still further, one employs in set theory operations going beyond all the above, namely, a passage to variables of higher type, i.e., the formation of classes of sets, classes of classes, etc., so that it is possible to treat algebraically mathematical systems of still greater generality. One can formulate problems in naive set theory dealing with properties of *repeated* passages to *variables* of *higher type*, studied *per se*. We shall content ourselves in stating one:

A "super-class" K of objects is imagined which is closed with respect to the operation of forming the class of all subsets. Starting, say, with the set S_0 of integers, one forms the class of all subsets of this set. This is a set which we denote by S_1; the class of all subsets of S_1 will be denoted by S_2, etc. In addition, the class K is closed with respect to the following construction: if S_i, $i = 1, 2, \ldots$, is any countable collection of classes of sets belonging to K, we form the class of all possible sequences of sets $s_i \in S_i$; we postulate that this class Σ of all such sequences $\{S_i\}$ also belongs to K. Imagine K is the smallest class closed under the two operations. The question now arises of classifying the objects of K by means of transfinite ordinals and of determining the powers of sets forming the elements of K.

2. The operation of direct product

One deals with the operation of the direct product, in a more or less explicit form, in every mathematical theory involving more than one variable. It is used quite explicitly in topology, group theory, measure theory, in the theory of metric spaces, and it also occurs in one form or another in many algebraic theories. It seems, however, that a *general* investigation of the properties of this operation for its own sake, on a set theoretical basis, has not been undertaken in spite of the many common features, presented by problems of "many variables," apparent in the various theories referred to.

The notion of *phase space* in mechanics is essentially that of a product space. The state of a system of particles is represented by a point in a product space, namely a direct product of spaces each describing the state of one particle. The component spaces themselves may be infinitely dimensional, as in quantum theory, where the state of a single particle requires a function for its description. In dealing with infinitely many particles, as in the physics of continua, it is necessary to introduce a direct product of an infinity of component spaces. A somewhat different type of operation, the "symmetric product," also arises in physics in connection with the Fermi-Dirac statistics, and requires a notion of the direct product as a basic substructure.

But in the foundations of mathematics itself the direct product enters implicitly in every theory involving the logical quantifiers (the "there exists" and "for all" expressions; cf. the work of Kuratowski and Tarski [1]). A mathematical interpretation of the existential quantifier is the operation of *projection* of a set, located in a product space, on one of the component spaces. The theory of projective sets due to Suslin and Lusin [1], and Sierpiński [2] exhibits some of the difficulties of this operation in problems of point set theory, the sets being considered in a topological space. It seems, however, that the real causes of the difficulties in the theory of projective sets have their origin already in general set-theory including the *general* theory of the operation of direct product, rather than in the topology of the real line or the Euclidean space where the projective sets were originally defined.

The importance for mathematical logic of the study of the direct product and the contingent operations of projection in a purely algebraic spirit is manifest. Just as the study of the algebraic properties of Boolean algebras, their structure isomorphisms and representations give a mathematical counterpart of the elementary logic of the calculus of propositions, the theory of such algebras, widened to include the direct product and projection operators, may provide a mathematical representation of logical systems where quantifiers are admitted, and thus afford an adequate algebraic structure for "constructive" mathematical theories.

We now formulate a few definitions and problems concerning the direct product of sets.

3. *Product-isomorphisms and some generalizations*

The direct product $A \times B$ of two sets A and B is the set of all ordered pairs (a, b) with a in A and b in B. Analogously the product ΠA_i is the set of all sequences $\{a_1, a_2, \ldots\}$ with a_i in A_i. In case all $A_i = A$ and $i = 1, \ldots, n$, we shall write $\Pi A_i = A^n$.

Two subsets A and B of a product E^2 are said to be product-isomorphic in case there exists a one-one transformation $f(x)$ on E to all of E such that the resulting transformation

$$(x, y) \to (f(x), f(y))$$

of E^2 to itself takes A into all of B. The relation of product-isomorphism is reflexive, symmetric, and transitive, and thus constitutes an equivalence relation on subsets of E^2 which divides the class of all such subsets into mutually disjoint subclasses of sets, product-isomorphic among themselves.

The first questions that arise in connection with this relation concern enumeration properties. It is obvious that sets of different cardinal numbers cannot be product-isomorphic.

What is the power of the equivalence class of all subsets of E^2 product-isomorphic to a given subset A of E^2 (in case E has, for example, the power c of the continuum)? In general, it is 2^c; of course it cannot be greater, since the power of the class of all subsets of E^2 is 2^c. In special cases it will be smaller; for example, if A should consist of only one point (a, a) on the "diagonal" of E^2, then every product-isomorphic set also consists of a single point of this diagonal, and the number of such sets is c. Moreover, if $A = E^2$, then A is product-isomorphic only with itself.

A product-isomorphism of a subset A with itself is called a product-automorphism. The number of product-automorphisms of a subset A of E^2, different on A, is in general 2^c when E has power c; this is true, for example, when $A = E^2$. One easily constructs examples of sets A which have only a finite number of product-automorphisms, in particular, some which admit only the identity as such an automorphism. Does there exist, for every n, a set having exactly n product-automorphisms?

Consider now the class K of equivalence classes of product-isomorphic subsets of E^n, where $n \geqq 2$ and product-isomorphism is defined in the obvious way by generalization from the case $n = 2$. The power of K is 2^e when the power e of the set E is infinite, a result which follows from a theorem on the power of nonisomorphic relation sets. What is the power of K when E is finite?

Most of the enumeration questions are difficult in such cases. Specifically, what is the power of: (a) the class of subsets of E^n product-isomorphic to a given one, (b) the class K of equivalence classes of product-isomorphic sets, (c) the set of product-automorphisms of a set with itself? Even good inequalities on such powers should be of interest. Of course, the power of (a) above cannot exceed $e!$ nor can the power of K be less than that of e^n, but the "best possible" bounds may not be easy to find.

The concept of product-isomorphism bears an interesting relation to that of isomorphism for various mathematical structures. Suppose that under an operation O the elements of an abstract set E form a group G. In the set E^3 consider the set $\mathfrak{G}$ of all points (x, y, z) such that $xOy = z$. We may call $\mathfrak{G}$ the "representation" of G in E^3. If H is also a group defined on the elements of E with representation $\mathfrak{H}$, then G and H are group isomorphic if and only if their representations are product-isomorphic in E^3.

The wide applicability of the notion of product-isomorphism is obvious since the definition of isomorphism of mathematical structures depends only on the number and kind of operations and not on their special properties. Thus, if G is a partially ordered set defined on the elements of an abstract set E by an order relation $(<)$, we may take as its representation $\mathfrak{G}$ the set of all pairs (x, y) in E^2 for which $x < y$, and state that two partially ordered sets G, H over E are order-isomorphic if and only if their representations in E^2 are product-isomorphic.

A greater complexity of the system naturally demands higher exponents n of the basic set E for its "representation." Thus a "ring isomorphism = product isomorphism" statement similar to the one above for groups and relation sets obviously holds when a "representation" is constructed in E^6 (the idea of a ring involving *two* binary operations). The question of the minimal dimension n of the product space E^n necessary to afford a precise representation $\mathfrak{G}$ of a system G is a fundamental one. For example: Can one "represent" a group already in G^2, that is to say, attach *effectively* to every group of the same cardinal power as G a subset in G^2, so that the "representations" would be product-isomorphic if and only if the groups are isomorphic in the usual sense?

Topological systems may be characterized by representations in E^∞. Suppose, for example, that G is a Fréchet space defined on a set E, with $a_0 = \lim a_\nu$ defined for certain sequences $\{a_\nu; \nu = 1, 2, \ldots\}$. We may define the representation $\mathfrak{G}$ of G in E as the set of all points $\{a_0, a_1, a_2, \ldots\}$ where $a_0 = \lim a_\nu$. Two Fréchet spaces G and H defined on E are homeomorphic if and only if their representations are product-isomorphic in the obvious sense. The extension of such procedures to combinations of algebraic and analytic structures like topological groups is manifest.

Questions concerning the representations of mathematical systems arise immediately. For example, suppose G is a group defined on the unit interval $E = [0, 1]$. (All this means, of course, is that G is a set of the power of the continuum.) Its representation $\mathfrak{G}$ is then a subset Z in the unit cube E^3. Subsets of E^3 can be classified as follows: A sequence $\{A_n\}$, $n = 1, 2, 3, \ldots$, of sets in E is given; one considers sets B belonging to the Borel field over $\{A_n\}$. (On the line this sequence is usually taken as the sequence of rational intervals.) The simplest sets in E^3 are subproducts, i.e., sets Z of the form $Z = B_1 \times B_2 \times B_3$ where B_1, B_2, B_3 are B-subsets of E. One can then consider sets which are complements of these. All these will be called of "class 0." The next class 1 of subsets would be countable sums of sets of class 0 and their complements. Class 2: again sums of sets of class 1, also their complements, and so on. We get an analogue of the Borel classification for subsets of E^3. One problem is of existence of algebraic structures whose representation $\mathfrak{G}$ is of minimal Borel class α, α given—for *any* choice of the sequence A_n, $n = 1, 2, \ldots$. (For example: Does there exist a group G of power of continuum, whose representation $\mathfrak{G}$ would be of class $\alpha > 3$, for any choice of a countable sequence of the "elementary" sets?).

Starting with a given sequence $\{A_n\}$ of the "elementary" sets in E one can of course go beyond the Borel classification and define "projective sets" in E^3 in a manner completely analogous to that of defining the familiar projective sets of Lusin.

The definition of product-isomorphism of two sets A and B in E^n suggests a generalization which leads to interesting questions about the abstract systems G and H. Let us say that the two sets A and B are *weakly* product-isomorphic if there exist biunique transformations $f_i(x)$ on E to all of E such that the induced transformation

$$(x_1, \ldots, x_n) \to (f_1(x_1), \ldots, f_n(x_n))$$

maps A onto all of B. One may then define the *weak* isomorphism of two abstract mathematical structures G and H on the elements of E as the weak product-isomorphism of their representations $\mathfrak{G}$ and $\mathfrak{H}$ in E^n. Thus the weak isomorphism of two groups G and H over E is tantamount to the existence of three biunique correspondences $x = u(a)$, $y = v(b)$, $z = w(c)$ on G to H such that $c = ab$ in G implies $xy = z$ for the corresponding elements in H.

A different generalization may be obtained by defining two subsets A and B of E^n as (weakly) product-isomorphic under decomposition in case there exist decompositions

$$A = A_1 + \ldots + A_m,\ B = B_1 + \ldots + B_m,\ A_i A_j = 0 = B_i B_j \quad (i \neq j)$$

such that A_i and B_i, $i = 1, \ldots, m$, are (weakly) product-isomorphic. This leads naturally to the concept of "the weak isomorphism of two structures G and H under decomposition" defined in terms of the (weak) product-isomorphism of their representations $\mathfrak{G}$ and $\mathfrak{H}$ under decomposition. For example, are the groups S (all permutations of the set of integers) and H (all homeomorphisms of the unit interval of real members) isomorphic under decomposition?

Given a "representative" set Z of an algebraic structure (over a set of power c) one can ask whether it is Borelian or projective—in the sense of definitions given in the last paragraph. These refer to a given basic sequence of sets A_n in E. We call a sequence of abstract sets A_n in a set E *measurable* if it is possible to define for all sets S of the Borel field over $\{A_n\}$ a real-valued measure function $m(S)$ with the following properties:

1. $m(E) = 1$, $m(S) = 0$ if S consists of a single point.
2. $m\left(\sum_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} m(S_i)$ if $S_i \cdot S_j = 0$ for $i \neq j$.

If the set Z is Borelian or projective with respect to a measurable sequence A_n we call the given algebraic structure abstractly Borelian or abstractly projective.

Among the first problems that arise is that of *existence* of a group defined on a set of power c whose representation would *not* be Borelian. More generally, for other algebraic structures—e.g., lattices or rings—how far can sets, closed under these operations, still exhibit a set-theoretical "pathology" of their representative set $\mathfrak{S}$ of n-tuples?

The motivation behind the above definitions is to have provision for a connection between the purely algebraic properties on one hand and the topological or "analytic" properties on the other—of structures which are given combinatorially, not dependent on a given topology of the given group (or relation algebra, etc.).

4. Generalized projective sets

Let A_n be a class of subsets of a set E, the latter having the power of the continuum. The projections on E of the sets of the Borel class *over* the "rectangles" $A_m \times A_n$ in E^2 constitute the projective sets of class $k = 1$ over $\{A_n\}$. Continuing inductively, one defines the projective sets of class $k = 2, 3, \ldots$ over A_n. The problems of this section are based upon this definition (cf. Sierpiński [2]).

Is it true that, for every countable sequence of sets A_n in E, there exists a countable sequence of sets B_n such that the Borel class over B_n contains all projective sets over A_n?

Does there exist a sequence of sets A_n in E with the properties:

(a) the Boolean algebra generated by the sets A_n contains a noncountable set of atoms, and (b) all projective sets over A_n are $G_{\delta\sigma}$ sets relative to the sets $\{A_n\}$?

Does there exist a sequence of sets A_n in E with the properties: (a) the Borel class over A_n contains sets of arbitrarily high (Borel) class number, and (b) all projective sets over $\{A_n\}$ are Borel sets over the $\{A_n\}$?

More specifically, it is true that, for every positive integer k, there exists a sequence of sets A_n with the property (a) of the preceding problem, and the property (b) *all* projective sets over the A_n are sets of projective class k?

Given a sequence of sets A_n in E, and a transformation f on E to E, we shall say that f is a Borel transformation relative to $\{A_n\}$ in case the counter-image of every Borel set over the A_n is again such a Borel set. Does the product-isomorphism of two Borel sets over the class of rectangles $A_m \times A_n$ in E imply their product-isomorphism under a Borel transformation relative to A_n?

Given an arbitrary sequence of sets A_n in E, does there exist a one-to-one mapping of E into E^2 such that the Borel sets over A_n in E go into Borel sets over $A_m \times A_n$ in E^2 and conversely?

In the following problem, the term analytic has its classical connotation. Can every analytic subset of the unit square be obtained by Borel operations from "rectangles" $A \times B$ where A, B are analytic subsets of the unit interval?

The motivation for investigating the Borel operations and, beyond it, the projective operations when one starts with a *general* sequence of sets A_n—instead of the usual one which is the sequence of rational intervals or binary intervals—lies in the following possibility. There might exist a sequence of sets such that the number of its *atoms* is noncountable (i.e., still "nontrivial") and yet such that the projective class over this sequence is "*simpler*" than the "classical" projective class. For example, a sequence such that one could define a completely additive measure function for all sets of this projective class—this is impossible, according to a result of Gödel, for the familiar projective sets: i.e., it is free from contradiction in certain systems of axioms to assume that there exist projective sets which are nonmeasurable in the sense of Lebesgue. Even more generally, one can extend this result to show that *no* completely additive measure is possible for all projective sets (by a measure we understand a set function with the properties: (*1*) $m(E) = 1$; $m(p) = 0$ where E is the whole space, (p) is a set composed of any single point).

$$(2) \qquad m\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} m(A_i) \quad \text{if } A_i \cdot A_j = 0 \quad \text{for } i \neq j.$$

Paradoxically enough, it is conceivable that a measure function like the above could exist, if one starts with a sufficiently "wild" sequence of sets A_n in a class of projective sets "over" this sequence. Possibly all such sets could have the Baire property, that is each set of the class would be of first category—or complement of such?

4a. Relations between products of different orders

Consider an infinite set E and the sets E^n, E^m, $n \neq m$. In each of these we have a special class of subsets, the "R-set" subproducts, i.e., sets of the form $A_1 \times A_2 \times \ldots A_m$ and $A_1 \times A_2 \times \ldots A_n$, respectively where the A's are arbitrary subsets of E. What is a one-one mapping of E^m into E^n such that the R-sets in E^n become sets of lowest possible Borel class over the R-sets in E^m? If the power of the set E is greater than c (of continuum) does there exist such a mapping so that the R-sets in E^n go over into Borel sets over R-sets in E^m?

In the above question the A's are arbitrary subsets of E. An analogous problem exists for the case where the sets A are restricted to the Borel sets over a sequence of sets S_i, $i = 1, 2, \ldots$, given once for all in E.

The same problem for E^n, E^∞.

We mentioned before the "representations" of various algebraic structures through subsets of E^k (the index k being $= n + 1$ for an "algebra" involving a n-ary operation). The question arises whether such structures can be "represented" by sets in E^m with $m < n + 1$ (cf. section 5) and still "effectively" up to an isomorphism; in particular, e.g., whether the *groups* requiring offhand a representation in E^3 can be *effectively* put in a one-one correspondence with subsets of E^2 in such a way that isomorphic groups would have product-isomorphic sets in E^2 attached to them. Obviously the problem of existence of such a correspondence is trivial if we do not require *effectiveness* or *constructiveness* one can, using the axiom of choice, map all groups, isomorphic to each other into the same subset of E^2. However, if one requires that groups whose representation in E^3 is a Borel set (with respect to a given sequence of sets in E^1) correspond to sets in E^2 which are related to these sets in E^3 through one "Borelian" mapping, the relation with the problems above becomes obvious.

Questions of this sort are related in spirit to the recent results of Kolmogoroff [1], on the reduction of functions of n variables to superposition of functions of 2 variables, reducing transformations of E^m space into itself to superposition of transformations of a space E^n into itself, with $n < m$.

5. Projective algebras

Projective algebras form a generalization of Boolean algebras, and will permit, to a certain extent, an algebraic treatment of the logical quantifiers. For our present purposes it is sufficient to consider a representation of a projective algebra and, for simplicity, we shall restrict ourselves to the two-dimensional case, although the problems formulated below are meaningful in n-dimensions, $n > 2$.

Assume then that we have a class of sets situated in the Euclidean plane, closed under Boolean operations and under projection onto either axis, and containing the direct product $A \times B$ whenever A and B belong to the class and are situated on the X and Y axes respectively. Such a class constitutes the simplest example of a projective algebra (cf. Everett, Ulam [1]; MacKinsey [1]).

Given a countable class of sets in the plane, does there exist a finite number of sets which generate a projective algebra containing all sets of this countable class? Another statement might make this assertion for a countable class of sets given in E^m with the generating sets required to be in some E^n with $n < m$.

That this may be possible is shown by an example due to David Nelson in which two plane sets generate an infinite projective algebra. Note that the situation here is radically different from that of Boolean algebras, in which case one can obtain at most 2^n elements from n generating sets.

Does there exist a universal countable projective algebra, i.e., a countable projective algebra such that every countable projective algebra is isomorphic to some subalgebra of it?

Is it true that, for every positive integer k, there exists a projective algebra generated by k sets in the plane and which is free in the sense that no relations exist between the generated sets except those that are true in every projective algebra? Can every projective algebra be obtained by a homomorphism of a free projective algebra?

How many nonisomorphic projective algebras exist with k generators?

Many theorems in mathematics amount to stating that two sets, obtained by different sequences of Boolean operations and quantifications operating on a finite number of given sets are identical. It is, therefore, desirable to establish in some projective algebras a theorem to the effect that for two identical sets their identity can always be established by the rules of formal projective algebra. We are asking in particular whether there exists a countable sequence of sets in E^m such that the projective algebra generated by them is free; that is to say, whenever two sets, constructed from the generators by formal operations of projective algebra are identical, then they may be demonstrated to be so by formal projective-algebraic operations.

6. Generalized logic

The attempts to exhibit the nature of the essential difficulties in the foundations of mathematical logic in purely algebraic schemes have a long history. We will be concerned in this section with a type of problem which embodies some of the desired features of such a program, and at the same time seems to admit less familiar models, among them, what might be described as a system with infinitely many quantifiers.

One of the most striking weaknesses of projective algebra, as one naturally conceives it in generalizations from the plane to E^n, $n > 2$, is its limitation to a number of quantifiers bounded in advance. Ordinary logic, while it makes statements about only a finite number of variables at a time, suffers no such restriction. Moreover, it is apparent that the postulation of n projection operators along different axes is cumbersome and unnecessary. Instead, we might postulate *one* projection operator and one transformation of variables (say in E^3 $x' = y$, $y' = z$, $z' = x$). These two together will generate all projection operators.

Let us therefore define a more general type of "projective algebra" as a class R of subsets of a set E closed under Boolean operations, and under two operators P and T which we conceive of as a "projection" operator on the class of subsets to itself corresponding to projection of E^n along one "axis", and a one-to-one point-to-point transformation of E to E, corresponding to the permutation of the axes, respectively.

It seems possible to formulate a system of postulates in terms of such operators which, when applied to the special set E of all two-way infinite sequences $(\ldots, x_{-1}, x_0, x_1, x_2, \ldots)$, where x_i are real numbers, would properly contain ordinary logic. However, the possibilities of *extending* the formalism and still using this special set E, seem to include a logic of propositions with infinitely many variables. For an example of a statement involving infinitely many quantifiers see the infinite games of Chapter I, Section 11.

The formal structure would be the following: A class R of subsets of E is closed with respect to the Boolean operations and the operators T and P. In addition, we can require countable additivity in R.

Any such class could be considered as an infinitely dimensional projective algebra. It is then a class of sets contained in E such that:

1. If $Z \subset R$, then $(E - Z) \epsilon R$; if $Z_1 \epsilon R$ and $Z_2 \epsilon R$, then $(Z_1 + Z_2) \epsilon R$.

2. If $Z \epsilon R$, then $T(Z) \epsilon R$ and $P(Z) \epsilon R$; if $Z \subset R$, then $T^{-1}(Z) \epsilon R$.

3. If $Z_i \epsilon R$ for $i = 1, 2, \ldots$, then $\sum_{i=1}^{\infty} Z_i \epsilon R$.

It is in this way, by using *3*, that one will obtain sets defined by an infinite class of "quantifier" operations or sets of an infinite projective class.

This could be of course achieved through a more orthodox procedure, using a finite number of quantifier operators only, but then we would have to operate with spaces involving additional variables.

Perhaps the set-up above would have a greater "algebraic" homogeneity.

The first problems would involve a representation theorem — then the possibility of generating countable projective algebras of the above type from a finite number of sets, etc. — similarly to problems on the two-dimensional projective algebras.

7. Some problems on infinite sets

1. Let A and B be infinite sets which admit a transfinite sequence of point transformations $t_\xi(a) \epsilon B$, $a \epsilon A$, with the properties: (*1*) $t_\xi(X) \cdot t_\xi(Y) = 0$ for $X \subset A$, $Y \subset A$, and some ξ implies $t_\eta(X) \cdot t_\eta(Y) = 0$ for all $\eta > \xi$; (*2*) for every infinite subset $X \subset A$ there exists a ξ such that $t_\xi(X)$ contains at least two distinct points; (3) $X \cdot Y = 0$ for finite X, Y implies existence of η such that $t_\eta(X) \cdot t_\eta(Y) = 0$.

Is the power of A necessarily less than or equal to that of B?

2. Let C be a class of subsets of the interval (0, 1) with the following properties: (*1*) C contains the Borel sets (in the usual sense); (*2*) C is closed under complementation and countable unions; (*3*) For every decomposition of (0, 1) into disjoint sets, each containing at least two points, there exists a set in the class C which has exactly one point in common with every set of the decomposition.

Is the class C necessarily the class of all subsets of (0, 1)?

3. The power of the class of all additive subgroups of the real field R is 2^c, namely the power of the class C of *all* subsets of R. Is the Borel class of sets over the sets which are arbitrary subgroups R identical with the whole class C? In the event of a negative answer, one may ask a similar question about the analytic class (of all sets generated by analytic operations from sets which are subgroups) or the kth projective class. That is, is every set of real numbers obtainable by projective operations from sets which are subgroups? For partial results see Erdös, Kakutani [1].

8. Measure in abstract sets

It is known (cf. S. Ulam [1]) that no countably additive measure function $m(A)$ exists, defined for *all* subsets A of a set E of power $\aleph_1$, which vanishes for all subsets consisting of single points and for which $m(E) = 1$. Does there exist a class of measure functions, $m_\xi(A)$, ξ in a set of lower power than $\aleph_1$, such that every subset of E is measurable in at least one of these measures? This is a problem of Erdös and the author. Partial results have been obtained by Alaoglu and Erdös.

Let E denote the set of transfinite ordinals less than Ω, the first ordinal corresponding to a noncountable power. Thus E has the power $\aleph_1$. Is it possible to define a countably additive measure function such that all sets of the Borel field over the subsets of E which form arithmetic progressions shall be measurable? (The subclass of arithmetic progressions form an analogue in this set of the class of binary intervals for the set of real numbers). Similar questions may be asked concerning the measure for (a) analytic class over the set of arithmetical progressions in E, and (b) "projective" subsets of E (over the same class).

If a set E is of a power which is an inaccessible aleph $\aleph$, does there exist a countably additive measure for *all* subsets of E with $m(E) = 1$ and $m(p) = 0$ for all points p?

It is known that the results on the impossibility of defining a completely additive measure for *all* subsets of a set E hold true for all sets whose powers are *accessible* cardinals. It seems rather likely that at any rate the existence of a two-valued measure, countably additive and defined for all subsets of a set of an *inaccessible* power does not contradict the axioms of set theory. Indeed, probably a stronger additivity for all $\aleph < \overline{\aleph}$ can be obtained where $\overline{\aleph}$ denotes the power of the first inaccessible cardinal.

9. Nonmeasurable projective sets

Gödel has proved that the existence of projective sets which are nonmeasurable in the sense of Lebesgue does not contradict the axioms of set theory; that is to say, the statement that such sets exist is either true, within certain axioms of set theory, e.g., those of von Neumann, or is independent of these axioms. It would be interesting to prove that the existence of such sets follows from the continuum hypothesis.

In the sequel we shall sketch some possible lines of attack on this problem through the use of certain constructions in the product space.

Gödel's result shows really more: it is free of contradiction in such axiomatic treatments of set theory to state that no countably additive measure, vanishing on sets consisting of a single point, is possible for all projective sets.

A similar problem can be formulated about the existence (assuming the continuum hypothesis) of projective sets with other "paradoxical" properties, e.g., sets not satisfying the Baire property (a set satisfies this property if it or its complement is on every interval a sum of countably many nowhere dense sets).

We shall indicate how, given an effective or constructive decomposition of the interval into $\aleph_1$ sets each of Lebesgue measure 0, one can define constructively, i.e., projectively, a nonmeasurable set consisting of a sum of the sets used in this decomposition. A construction is given in the paper referred to (Ulam [1]) which amounts, in essence, to the following:

One can construct a doubly infinite matrix whose elements are subsets of a set of a power $\aleph_1$ with the following properties. Each row of the matrix represents a decomposition of E into $\aleph_1$ disjoint sets. There are countably many rows and noncountably many columns. The sum of the sets in each column gives the whole set E with the exception of, at most, countably many elements (which would be the original sets each of measure 0). The sum of each column, therefore, would be a set of measure 1. So in each row there would exist at least one set of positive measure. This leads to a contradiction because there are only countably many rows and noncountably many columns, there would exist at least one row with noncountably many sets, each of positive measure and all disjoint, which is impossible.

If we can, therefore, establish the existence of a constructive decomposition of the interval into $\aleph_1$ projective sets each of measure 0, the construction used in l.c. would turn out, as we shall prove, also projective and nonmeasurable.

Here is a suggestion of a possible way to obtain such a decomposition:

Consider a set X of measure 0 on the interval, and such that under a Peano mapping T of the interval into the square which preserves measure (i.e., linear measure in the interval = area of the image in the square and *vice versa*) its image is contained in the set X^2. Assume furthermore that $T(X)$ is a proper subset of the set X^2. We want to show the existence of a constructive set S_1 still of measure 0 and containing X as a proper subset.

Take $T^{-1}(X^2)$, and define X_1 as $X+T^{-1}(X^2)$. Since T is measure-preserving, $T^{-1}(X^2)$ is of measure 0 and X^{-1} is of measure 0 too.

Consider now X_1^2; this set contains $T(X_1)$. More generally, suppose that sets X were defined for all $\alpha < \eta$ and that X_α^2 contained $T(X_\alpha)$. Take X_η as $\sum X_\alpha$ for all α, then X_η^2 contains X_α^2 for all α and, therefore, also $T(X_\alpha)$ for all α. Take $T^{-1}(X_\eta^2)$. This will be a set containing X_η. Let us call it $X_{\eta+1}$. Proceeding in this fashion we shall obtain a well-ordered sequence of sets that are increasing and are all of measure 0.

Under the assumption that the power of continuum is that of $\aleph_1$ this process will stop at a transfinite ordinal of at most third class. We obtain thus the desired decomposition of the interval into a well-ordered sequence of projective sets all of measure 0.

We can use now the construction described in the paper in

Fund. Math. Vol. 16, by using instead of points, the sets of our sequence. The sets of our matrix would all be projective sets. At least one of these sets must be nonmeasurable. This proves our theorem, the existence of a projective set not measurable in the sense of Lebesgue.

REMARK 1. This method really could prove more. For every measure that is completely additive and such that there will exist a mapping of the interval into the square preserving the measure in the square, and such that Fubini's theorem holds for the measure, there will exist sets that are projective and not measurable in the sense of this measure.

REMARK 2. The same argument would yield the existence of projective sets which do not possess the Baire property, i.e., that there exist projective sets which are not of first category on any perfect set, and their complements are not of first category on any perfect set.

REMARK 3. One could possibly extend such a method to obtain the existence of a projective set which does not possess the λ-property of Lusin.

All these results would be valid under the assumption of the continuum hypothesis. Presumably they would still hold under the assumption of a weaker hypothesis, namely, that the power of continuum is less than that of the first inaccessible aleph (greater than $\aleph_1$).

PROBLEM. Does there exist a two-valued measure for all subsets of a set whose power is the first inaccessible aleph?

We shall summarize again our approach to the question of existence of projective sets, nonmeasurable in the sense of Lebesgue, and will sketch an alternative approach to establish the lemma on "projective" decompositions of the interval into $\aleph_1$, sets of measure 0. The basic role is played by the theorem, mentioned above:

There does not exist a completely additive measure function defined for all subsets of a set Z of power $\aleph_1$ which would assume the value 0 *for sets consisting of a single point and equal to* 1 *for the whole set Z* (cf. Ulam [1]).

The proof depends on the existence of a decomposition of the set Z into a doubly infinite matrix of sets (all subsets of Z) which would apply equally well if the set Z was of the power of continuum, but could be decomposed into $\aleph_1$ sets all of measure 0. The proof of the nonexistence of measure is contructive if the decomposition used in the paper quoted above could be assumed constructive.

It would be interesting to strengthen Gödel's result by proving that the existence of projective nonLebesgue measurable sets follows from the assumption of the continuum hypothesis or even, the weaker hypothesis, namely, that the power of continuum is less than that of the first inaccessible $\aleph$. Let us note that it is sufficient to show the existence of a decomposition of the interval 0,1 into $\aleph$ sets all of Lebesgue measure 0, where the $\aleph$ is a cardinal number smaller than the first inaccessible $\aleph$, and the decomposition would be constructive in the following sense: the sum of a subsequence of the sets of the decomposition corresponding to a constructive class of ordinals is understood here in the sense of, e.g., Kuratowski [2, 3] as "constructive". Given such a decomposition, the construction given in the paper by Ulam [1] yields a set which is obtained by projective operations from the given sets and, as shown, would be nonmeasurable.

The crucial point, therefore, is to show that one can decompose the interval into, for example $\aleph_1$ sets all of measure 0 and so that the decomposition itself would be "projective." An approach to such a construction, different from the one above, could be as follows:

Let us start with the well-known decomposition of Lebesgue of the interval into $\aleph_1$ disjoined sets of Borel sets and all having, in addition, the property that if a number z belongs to a set A_ξ, $\xi < \Omega$, then any number of the form $z + r$, where r is any rational number, belongs also to A_ξ. In virtue of this property, the sets A_ξ must all be of Lebesgue measure either 0 or 1. If all were of measure 0 we would have the desired decomposition so we can assume that one of these sets will have measure 1. The operations which we shall perform on sets from now on will lead always to sets that are invariant with respect to addition of a rational number so that these sets will also be of measure either 0 or 1 or else nonmeasurable. Since the sets will also be projective, we may assume that they are all measurable, otherwise our result would be already proved.

One needs the following general lemma:

Given a projective set that is uncountable, we shall attach to it "projectively" a proper subset of it. This will replace for our purposes the necessity of using the axiom of choice which, of course, in general destroys the constructive character of the set. Let $T(p)$ be a one-to-one transformation of the interval into the square. This $T(p)$ can be chosen as Borelian transformation, in fact, one of second Borel class. Consider a set Z contained in the interval. Consider the set Z^2 and the set $T(Z)$. We may consider the division of the unit square into the four sets Z^2, $C \times Z$, where CZ is the complement of Z to the whole interval, $(CZ)^2$ and $CZ \times Z$. If $T(Z)$ is not contained in any one of these sets, then the counter image of the two parts of $T(Z)$ lying in two different sets of this decomposition will give us a decomposition of Z into two nonvoid parts, which proves our assertion that there exists a projective proper subset of Z. Suppose then, that $T(Z)$ is wholly contained in one of the four sets mentioned above. We may assume that if $T(Z)$ is a subset of Z^2 it does not contain any points of the diagonal D of the square, because we could, by subtracting these points, obtain a proper subset of Z. Consider the part of the diagonal that corresponds to the set Z. This part, called Z_1, is contained in $T(CZ)$ in virtue of the remark that $T(Z) \cdot D = 0$. Consider, therefore, $T^{-1}(Z_1)$. If this set has a constructive proper subset we shall be able to define a constructive proper subset of the set Z itself, because of the existence of a constructive mapping of Z into Z_1

We may assume that Z_1 is contained wholly in just one of the nine sets into which the square is decomposed, by taking the decomposition of the interval into Z, $CZ - Z_1$ and Z_1 and multiplying it by itself which gives us the sets Z^2, $(CZ - Z_1)^2$, Z_1^2 and the six cross products of the three sets. Consider the set on the diagonal D corresponding to the set Z_1, call it Z_2. Any projective proper subset of Z_2 would give us immediately a projective subset of Z which would prove our theorem. It suffices, therefore, to find such a subset of Z_2. We can repeat the reasoning given before, and we shall arrive at a set Z_3. One can by transfinite induction prolong this construction for any ordinal of the second or third class. If we make the assumption that the power of continuum is only $\aleph_1$ this chain of sets Z must stop, which gives us a proper subset of the set Z given in the beginning. Our lemma is, therefore, proved.

Consider now for a given set Z its proper subset which we shall denote by $U(Z) = Z'$. Of the sets Z' and $Z - Z'$, one and one alone must have measure 1. We shall apply to it the lemma and obtain sets Z^2 and its complement. In this fashion we can define sets Z^n all of measure 1; that is to say, if they are measurable. The intersection of these sets would be a set still of measure 1, because of the additivity property of the measure function. Using our lemma for this set, which we shall call Z^ω, we shall obtain a set, still of measure 1, $Z^{\omega+1}$. We can continue our construction by transfinite induction remembering, however, that the axiom of choice is not used since we always select the proper subset effectively by taking the one which has measure 1. We can assume that the intersection of $\aleph_1$ of such sets is still a set of measure 1, otherwise we would have the desired decomposition into $\aleph_1$ sets of measure 0 by using the complements of these sets. Therefore, our construction can be prolonged to transfinite ordinals of class three. However, if we assume the continuum hypothesis, we must arrive at a vacuous set for some ordinal of the third class. Therefore, there must be a sequence of length $\aleph_1$ of sets that have measure 1 with a void intersection which proves our assertion.

We repeat that all the steps of our construction would be effective, i.e., the axiom of choice was not used, so the decomposition of the interval into $\aleph_1$ sets of measure 0 is *projective*, at least in the *wider sense* of the word.

10. Infinite games

The following combinatorial scheme was first proposed by S. Mazur around 1928. Imagine a set of points M contained in the unit interval (0, 1) and two indefatigable mathematicians A and B who will play the following game. Each will define an interval in turn, the first interval being given by A, and each successive interval being a subinterval of the preceding one but otherwise arbitrary. The game is won by A if the intersection of all intervals contains a point of M, otherwise by B.

In case the set M is a residual set in some interval, the player A can always win with the following strategy. The set M having a complement with respect to a certain interval which is of first category, A chooses this interval initially, and, regardless of the choices of B, always selects, at his nth turn, a subinterval disjoint with the nth nowhere dense set which figures in the decomposition of his opponent's set into a sum of countably many nowhere dense sets. Obviously A wins!

It is interesting that, using the axiom of choice, one can construct sets M such that for every subinterval neither M nor M' is of 1st category with respect to this subinterval. These are sets which do *not* have the so-called property of Baire. Banach has proved (unpublished) that for such a set M there exists no method of winning for either player.

One can generalize and vary Mazur's game in many directions. The problem occurring in all such games involving a subset M is: what is the class of sets M for which no method of winning for either player exists? It should be pointed out that, in all known games of this sort, a method of winning exists for either A or B when the subset M is *effectively defined*. For example, for the original game, up to the present no *effectively* constructed set is known which does not possess Baire's property. All known proofs of the existence of such sets utilize the axiom of choice.

A result of Gödel from which it follows that it is safe to assume, in certain systems of axioms for set theory, that such sets exist and are projective yields an interesting interpretation of Banach's result!

It is also interesting to note that games of this kind cannot easily be defined between three players without a trivial reduction to a game between two of the three players. The dichotomy inherent in such infinite constructions is to be observed in all constructive parts of set theory. Thus, for example, a Lebesgue measurable set possesses at almost all of its points a density which is either zero or one. A Baire set is of 1st category or a complement of such in all the points, etc.

The game first mentioned can also be specialized in various ways. Thus the choice of intervals open to A and B can be restricted by some rule. The simplest example might be one such that, given a set M, the players produce successive digits of a real number x in binary expansion, player A trying to make x belong to M, while B tries to make x belong to the complement of M. An analogue of Banach's result holds for these more special games. Compare Gale, Stewart [1]; Mycielski, Swierczkowski, Zieba [1]; Mycielski, Zieba [1]. See also the *Scottish Book* for original version of Mazur and modifications of Banach and Ulam.

An interesting variation of the latter situation would be provided by allowing a random element to enter: For a given subsequence n_i the play n_i is to be 0 or 1 determined by chance. A method of winning would be understood here as a strategy by means of which A succeeds in defining x in M for almost every sequence of 0's and 1's on the plays n_i.

11. Situations involving many quantifiers

It is in statements on existence of a winning strategy that a use of a large number of quantifiers appears most natural. In games between two persons they have, e.g., this form. For every move α_1 (of the player A), there exists a move α_2 for player B so that for every move α_3 there exists a move α_4 and so on so that after α_{2n} the situation is a win for B. In a notation of Kuratowski-Tarski:

$$\prod_{\alpha_1} \sum_{\alpha_2} \prod_{\alpha_3} \cdots \sum_{\alpha_{2n}} W(\alpha_1 \ldots \alpha_{2n})$$

where $W(\alpha_1 \ldots \alpha_{2n})$ is a "Boolean expression" describing a winning position.

Thus a mate announced in 5 moves involves, in this notation, 10 quantifiers. This has to be compared with the usual mathematical definitions, e.g., of uniform convergence, or the definition of an almost periodic function, etc. These, granting already the customary elementary mathematical notions and abbreviations, get by with 3–5 quantifiers.

In a study of the formal properties of repeated applications of quantifiers treated as mathematical operations (in the theory of projective sets of Luzin — or in projective algebras — cf. Section 5) the number of times one employs these operators is arbitrary. In the chapter on computing machines we shall mention the possibility of a heuristic investigation of combinatorial problems arising in this connection.

12. Some problems of P. Erdös

Several problems pertain to the theory of graphs. They have equivalent formulations in our terminology of product operation and can be stated as problems on subsets of a set E^2 through the obvious correspondence of considering pairs of points which are joined by a segment. When one considers vertices of triangles in a graph, one can consider the corresponding set of triplets in E^3 etc.

Some other problems belong to number theory but are of combinatorial character and we include them in this section.

1. Let S be a set of power $\aleph$; to each finite subset A of S there corresponds an element $f(A)$ of S and $f(A) \notin A$. A subset $S_1 \subset S$ is called independent if for every $A \subset S_1$, $f(A) \notin S_1$. Does there always exist an infinite independent subset of S? If S has power $\aleph_n$ this is false (a result of Erdös-Hajnal) but it may be true for $\aleph_\omega$.

2. A problem on infinite graphs (Erdös-Rado): Suppose an infinite graph is given whose vertices form an ordered set of type ω^2. If this graph does not contain a triangle, then its vertices have a subset of type ω^2, no two vertices of which are connected by an edge. This is denoted symbolically as follows: $\omega^2 \to (\omega^2, 3)^2$. Specker (*Comm. Helv.* 1957) showed that $\omega^n \to (\omega^n, 3)^2$ is false for $n > 2$. Does $\omega^\omega \to (\omega^\omega, 3)^2$ hold?

2a. Erdös-Rado: Let S be a set of power greater than c. All the finite subsets of S are divided into two classes (in an arbitrary way). Does there exist an infinite subset $S_1 \subset S$ so that for every integer k all subsets of S_1 having k elements belong to the same class (but the class can depend on k)? If the power of S is $\leqq c$ this is not true (Erdös, Hajnal).

3. Let g be a complete graph of power $\aleph_1$, $g = g_1 + g_2 + g_3$ and we assume that $g - g_i$, $i = 1, 2, 3$, does not contain a complete graph of power $\aleph_1$. Does there then exist a triangle, each edge of which is in a different g_i^2? Another problem: Let g be a complete graph of power $\aleph_1$, $g = \sum_\alpha g_\alpha$, $1 \leqq \alpha < \Omega_1$, so that $g - g_\alpha$ does not contain a complete graph of power $\aleph_1$ $1 \leqq \alpha < \Omega_1$ (such a decomposition is possible (Erdös-Rado)). Does there exist a complete subgraph g' of g of power $\aleph_0$, each edge of which is in a different g^2?

4. Erdös-Turan: Let $a_1 < a_2 < \ldots$ be an infinite sequence of integers. Denote by $f(n)$ the number of solutions of $n = a_i + a_j$. Assume that $f(n) > 0$ for $n > n_0$. Then $\limsup_{n=\infty} f(n) = \infty$. A still sharper conjecture is: let $a_k < ck^2$, $1 \leqq k \leqq \infty$, then $\limsup_{n=\infty} f(n) = \infty$.

5. Erdös-Turan: Denote by $r_k(n)$ the maximum number of integers not exceeding n in a set which does not contain an arithmetic progression of k terms. How large can $r_k(n)$ be? $r_k(n) < \frac{1}{2}n$ would prove van der Waerden's theorem according to which if one splits the set of integers into two groups, at least one of them contains an arbitrarily long arithmetic progression. (The best results so far are:

$$r_3(n) < cn/\log \log n$$

$$r_3 > n^{1 - c/\frac{1}{2} \log n}$$

$$r_3(n) > n^{1 - c/\log \log n}$$

Compare Roth [1] and references given there.)

6. Let $f(n) = \pm 1$, chosen arbitrarily. Prove that to every c there exists an m and a d so that

$$\left| \sum_{k=1}^{m} f(dk) \right| > c$$

This conjecture has connections with van der Waerden's theorem (cf. Khinchin [2]). Also it would imply that if $g(n) = \pm 1$, $g(n)$ multiplicative, then

$$\limsup_{n=\infty} \left| \sum_{k=1}^{n} g(k) \right| = \infty$$

7. Let $a_1, a_2, \ldots, a_n$ be n elements. $A_1, A_2, \ldots, A_k$ are sets formed of the a's. Assume that no A_i contains any A_j. Then $k \leqq {}^nC_{n/2}$. This is a theorem of Sperner [1]. Let now $B_1, B_2, \ldots, B_l$ be sets formed from the a's so that no B is the union of any two other B's distinct from it. What is the maximum l? Quite possibly $l < c\, {}^nC_{n/2}$. Erdös states that he cannot even prove that $l = 0(2^n)$.

8. Let $a_k \geqq 1$, $1 \leqq k \leqq n$. Consider all sums of the form

$$\sum_{k=1}^{n} \varepsilon_k a_k \text{ where } \varepsilon_k = \pm 1.$$

Erdös proves (sharpening a previous result of Littlewood and Offord), using the above result of Sperner, that the number of these sums falling in the interior of an interval of length 2 is $\leqq {}^nC_{n/2}$; equality for $a_k = 1$, $1 \leqq k \leqq n$. The conjecture is that the same holds if the a's are complex numbers of absolute value $\leqq 1$ and the interval is replaced by a circle of radius 1. The same result may even hold if the a's are vectors in a Hilbert (Banach) space.

9. Let $a_1 < a_2 < \ldots < a_k \leqq n$; $b_1 < b_2 < \ldots < b_l \leqq n$ be two sequences of integers such that all the products $a_i \cdot b_j$, $1 \leqq i \leqq k$, $1 \leqq j \leqq l$, are different. Prove that

$$k \cdot l < c\,(n^2/\log n).$$

If true this is the best possible result.

10. How many distinct residues $a_1, a_2, \ldots, a_k$ can one give (mod p) so that none of the $2^k - 1$ sums $a_{i_1} + a_{i_2} + \ldots + a_{i_j}$ should be $\not\equiv 0$ (mod p)? Clearly one can give $[\sqrt{(2p)}]$ such residues $(a_i = i,\ 1 \leqq i \leqq [\sqrt{(2p)}])$. No decent upper bound is known.

11. (Erdös and the writer). Let I be any finitely additive ideal in the set of all integers. Consider the Boolean algebra of subsets of integers mod I. Does one obtain $2c$ nonisomorphic Boolean algebras in this way?

CHAPTER II

Algebraic Problems

1. An inductive lemma in combinatorial analysis

We shall illustrate this lemma first on structures with a given binary relation. Suppose that in two sets A and B, each of n elements, there is defined a distance function ρ for every pair of distinct points, with values either 1 or 2, and $\rho(p, p) = 0$. Assume that for every subset of $n - 1$ points of A, there exists an isometric system of $n - 1$ points of B, and that the number of distinct subsets isometric to any given subset of $n - 1$ points is the same in A as in B. Are A and B isometric? This assertion is true for $n \leq 6$, as has been shown by P. Kelly [1] by examination of all possible cases.

Clearly the metric formulation is equivalent to a similar question about sets with a binary relation pRq, holding if and only if $\rho(p, q) = 1$. Can one infer the relational isomorphism of A and B from $n - 1$ level relational isomorphisms of subsets in the manner indicated?

Similar problems may be formulated in other algebraic systems. Specifically, suppose that G and H are groups of order n. We shall say two subsets $G_k \subset G$, $H_k \subset H$ of k elements each are conditionally isomorphic if there exists a one-to-one mapping f on G_k to H_k, such that whenever a, b and $c = ab$ are in G_k, then $f(c) = f(a) \cdot f(b)$. What is the minimum number $k(n)$ such that the conditional isomorphism of every G_k to some H_k implies the isomorphism of G and H? One might include the stronger hypothesis that if $\{G_k\}$ is any class of l distinct subsets G_k conditionally isomorphic to each other, then there are also l distinct subsets H_k of H, each conditionally isomorphic to the G_k sets.

2. A problem on matrices arising in the theory of automata

The theory of automata leads to some interesting questions which in the simplest case reduce to matrix theory formulations. Suppose one has an infinite regular system of lattice points in E^n, each capable of existing in various states $S_1, \ldots, S_k$. Each lattice point has a well defined system of m neighbors, and it is assumed that the state of each point at time $t+1$ is uniquely determined by the states of all its neighbors at time t. Assuming that at time t only a finite set of points are active, one wants to know how the activation will spread. In particular, do there exist "universal" systems which are capable of generating arbitrary systems of states. Do there exist subsystems which are able to "reproduce," i.e., to produce other subsystems like the initial ones? In a simple case, one would ask: Does there exist an infinite matrix $A = [a_{ij}]$ of zeros and ones with $\sum_j a_{ij} < B$ for all rows i, such that *every* possible finite matrix of zeros and ones will appear as a main-diagonal submatrix of some power A^p of A? A positive result would provide a simple example of a "universal" and "reproducing" system (in a very limited sense only).

More generally, an analogous question may be asked about matrices whose elements are integers modulo p.

A similar inquiry is pertinent in case of the "recursive functions." Can one obtain all recursive functions by a *prescribed* algorithm operating on a finite set of such functions? More generally, are all expressions in Gödel's system obtainable from a finite system of such expressions and a finite number of rules of composition performed in a prescribed order? That is to say, for example, application of two operations, applied in turn in an order given by one sequence of two symbols.

Perhaps there exists a logical analogue of our universal matrix model.

3. A fundamental transformation in the "theory of equations"

The transformation

$$T_n: \quad x_1' = -\sigma_1(x_1, \ldots, x_n),$$
$$x_n' = (-1)^n \sigma_n(x_1, \ldots, x_n),$$

where σ_j is the jth elementary symmetric function, gives the coefficients x_i' of the equation

$$z^n + x_1' z^{n-1} + \ldots + x_n' = 0$$

in terms of the roots x_i. The inverse T_n^{-1} of the transformation T_n "solves" the equation of n-th degree. This transformation can be considered operating on the n-dimensional real or on the n-dimensional complex space.

Many of the statements about algebraic equations are translatable into the elementary properties of this mapping. Thus Gauss' theorem on the existence of roots is simply the statement that T_n is a mapping (many-one) on E^n to *all* of E^n, where E is the complex plane. The points "constructable by ruler and compass" are related to those resulting from iteration of the inverse transformation T_2^{-1} where $n = 2$.

However, the topological nature of this transformation does not seem to have been very thoroughly investigated. For example, what are the nontrivial fixed points $p = T_n(p)$? The origin is always a fixed point, but there are others, e.g., $p = (1, -2)$ when $n = 2$. What are the invariant analytic manifolds $M = T_n(M)$? What points are periodic under T_n?

The impossibility of solving the general equation of degree $n \geqq 5$ "by radicals" means that the corresponding T_n^{-1} is not a transformation involving only field operations and extraction of roots.

Does there, however, exist a homeomorphism H of E^n such that the inverse of $S = H^{-1} T_n H$ would involve only such operations?

The solution of the equation of fifth degree may be made to depend on elliptic functions (Hermite) and such methods were generalized by Poincaré. Can one show that for any $n \geqq 6$ the transformation T_n^{-1} can be obtained by composition of such transformations T_m^{-1} of lower degree m operating on suitable subspaces of E^n?

Can one show that T_n itself is a composite of a finite number of mappings, each of which is a conjugate $HT_m H^{-1}$ of some T_m, $m < n$, operating on a suitable subspace of E^n?

4. A problem on Peano mappings

Let R be the set of positive rational integers with the usual operations $a + b = s(a, b)$ and $a \cdot b = m(a, b)$. Every one-to-one (Peano) mapping $c = p(a, b)$ on $R \times R$ to all of R may serve so associate with $s(a, b)$ and $m(a, b)$ two functions σ and μ on R to R by the definitions $\sigma(c) = \sigma(p(a, b)) = s(a, b)$, and $\mu(c) = \mu(p(a, b)) = m(a, b)$. Does there exist a Peano mapping $p(a, b)$ such that "addition commutes with multiplication" in the sense that $\sigma(\mu(c)) = \mu(\sigma(c))$ for all c of R? To illustrate, we note that the well-known Peano mapping $c = p(a, b) = 2^{a-1}(2b-1)$ fails. For, $\sigma(\mu(14)) = \sigma(\mu(2^{2-1} \cdot [2.4-1])) = \sigma(8) = \sigma(2^{4-1} \cdot [2.1-1]) = 5$, while $\mu(\sigma(14)) = \mu(\sigma(2^{2-1} \cdot [2.4-1])) = \mu(6) = \mu(2^{2-1} \cdot [2.2-1]) = 4$.

5. The determination of a mathematical structure from a given set of endomorphisms

One of the fundamental tasks of abstract algebra is the determination of the automorphisms or homomorphisms into itself (endomorphisms) of a given algebraic structure. The inverse problem, though not as familiar, presents many features of interest (cf. Everett, Ulam [4]).

Suppose that we are given the operation of ordinary *multiplication* on the rational integers $R = 0, \pm 1, \pm 2, \ldots$. What are all the possible operations of "addition" definable on the set R which, with the given multiplication, will yield a ring? It is easy to show that the characteristic of such a ring must be 0 or 3.

Less specifically, what are all possible rings with identity, countably many primes, and unique factorization up to units?

Given the class of homeomorphisms of a topological space, what other topologies exist on the same set which have these mappings as the class of all their homeomorphisms?

6. A problem on continued fractions

Apparently the explicit form of the simple continued fraction corresponding to a real algebraic number of degree exceeding two is not known in any individual case. The following special questions may, however, be more tractable. Does there exist an algebraic number of degree > 2 in whose continued fraction $n_1 + 1/n_2 + + 1/n_3 + \ldots$ the sequence n_i is not bounded? (Consider in particular the number ξ defined by $\xi = 1/(\xi + y)$ where $y = 1/(1 + y)$.) Or is it perhaps true that *every* real algebraic number of degree > 2 has an unbounded sequence $\{n_i\}$ for its continued fraction?

(The set of real numbers for which the sequence $\{n_i\}$ is bounded has, as is well known, measure zero (and is of first category), so that one might say that the *a priori* chance of a number x having the sequence $\{n_i\}$ unbounded is 1.)

7. Some questions about groups

Is every separable continuous group, considered solely as an abstract group, isomorphic to a subgroup of the group S_∞ of all permutations of the integers? It is obvious that S_∞ is "universal" for all *countable* groups in this sense (that is: every countable group is isomorphic to a subgroup of it), but one can also show that *some* groups of power c, like the additive group R of real numbers is isomorphic to a subgroup of S_∞. The proof is based on the fact that R is a rational vector space with a (Hamel) basis having the power of the continuum, and that S_∞ contains a free product of continuum many groups isomorphic to the rational numbers under addition.

Let G be a subgroup of S_∞ with the property that for every two sets of integers of the same power whose complements are also of the same power, there exists a permutation g of G which transforms one set into the other. Is $G = S_\infty$ (Chevalley, von Neumann, *et al.*)?

If in the symmetric group S_n on n integers two pairs of elements a, b and α, β are simultaneously conjugate, i.e., there exists an element x of S_n such that $\alpha = x^{-1}ax$ and $\beta = x^{-1}bx$, then obviously every element generated by a and b is conjugate to the corresponding element generated by α and β. Is the converse true? That is, if every combination of a and b is conjugate to the corresponding combination of α and β (through perhaps a variable

x depending on this combination) are then a, b simultaneously conjugate to α, β?

The following question is due to H. Auerbach. Let G be a group of $n \times n$ matrices g such that every cyclic subgroup of G: $\ldots g^{-2}, g^{-1}, e, g, g^2, \ldots$ is bounded. Is G bounded? The answer is affirmative for $n = 2$.

We shall raise here a question on some purely group-theoretic properties of certain important infinitely-dimensional continuous groups. Later in the discussion of topological groups we shall refer to the simplicity of the group of all homeomorphisms of the circumference of the circle — a result of J. Schreier and the author [2] — to the corresponding result of von Neumann and the author on the group of all homeomorphisms of the surface of the sphere and to the recent results of Anderson. This question, whether the group under consideration possesses no invariant subgroups (except the identity element) is of interest for the groups of isometric transformations of Banach spaces onto themselves.

Is the group of all measure preserving transformations of the interval (0, 1) a simple group?

Does there exist a universal constant c (independent of dimension) such that, for *every* irreducible group G of orthogonal $n \times n$ matrices g, there is a vector u of unit length, *some* n of whose images $g_1 u, \ldots, g_n u$ under G have separation c from each other, i.e.,

$$| g_i u - g_j u | \geqq c, \quad i \neq j,\ i,\ j = 1, \ldots, n$$

This constant could be greater than $\frac{1}{2}$. In fact, the group of rotations of a pentagon in the plane very likely has the minimum value of c.

The affirmative statement, if true, could serve as an important lemma in a geometric approach to Hilbert's problem [1] on the introduction of analytic parameters in a continuous group, recently solved by Gleason and Montgomery, cf. Montgomery [1].

8. Semi-groups

Let G be a semi-group (associative multiplicative system with identity). A semi-group H of G is *normal* if it has the property that, whenever $h = h_1g_1h_2g_2 \ldots h_kg_kh_{k+1}$, with $h, h_1, \ldots, h_{k+1}$ in H and $g_1, \ldots, g_k$ in G, then $g \equiv g_1g_2 \ldots g_k$ is in H. If G and H are groups, normality of H in this sense coincides with the usual definition. Two elements a and b of G are said to be congruent mod H (for H normal) if there exist elements $a_1, \ldots, a_l, b_1, \ldots, b_l$ in G and $h_1, \ldots, h_l,\ h_1', \ldots, h_l'$ in H such that $a = a_1 \ldots a_l$, $b = b_1 \ldots b_l$ and $h_1a_1, \ldots, h_la_l = h_1'b_1, \ldots, h_l'b_l$. This again coincides with the usual congruence in the group case. It would be of great interest to establish the analogues of the classical chain theorems culminating in the Jordan-Holder theorem. We mention the following statements, some of which may be proved easily.

The group S_∞ of all one-to-one transformations of the integers (permutations) is not a normal semi-group of the semi-group of T_∞: the set of all transformations of the integers into themselves, while S_m *is* a normal subsemi-group of T_n, the latter referring to the corresponding set of operators on the set of integers $1, \ldots, n$. In T_∞, the semi-group F consisting of all mappings $f(n)$ such that $f(n) \neq n$ for only a finite number of integers n is a normal subsemi-group. If N is a normal subsemi-group of T_n, which contains an element not in F, then $N = T$.

The homeomorphisms of the line form a normal subsemi-group of the semi-group of all continuous functions.

8a. Topological semi-groups

Problems of A. D. Wallace:

Let S be a compact, connected semi-group.

1. If S is finite dimensional, homogeneous and has a unit, is S a group? (Yes, if $\dim S = 1$.)

2. If S has a zero and a unit, does S have the fixed point property?

3. If S has a zero and if $S^2 = S$, can S be homeomorphic with an n-sphere? (No, if $\dim S = 1$.)

9. *A problem in the game of bridge*

Many of the problems of combinatorial analysis, especially those of the theory of probability, derive from situations arising in various games of chance or even games of "skill". The majority of such problems refer to given or fixed situations. We give here an example of a problem in the game of bridge involving, so to say, one more existensional quantifier than the usual problems of the game.

Does there exist an initial distribution of hands with the following properties? (a) East and West can make, against best defense, a grand slam (all 13 tricks) in every suit if this suit were trumps. (b) In a no trump contract, however, against a good defense, East and West are unable to make even a small slam.

More specifically, what is the greatest number of tricks that East and West can always make, even against the best defense, assuming property (a)? It seems almost certain, that a hand distribution with the property (a) guarantees at least 5 tricks in "no trumps". (J. Schreier and the author have found an example of a distribution with the property (a) such that a *grand slam* in "no trumps" cannot be made.)

10. *A problem on arithmetic functions*

The set of integer valued arithmetic functions $\alpha(n)$, $n = 1, 2, 3, \ldots$ forms a domain of integrity under ordinary addition, and multiplication:

$$\alpha\beta(n) = \sum_{\alpha/n} \alpha(d)\beta(n/d).$$

Is this ring a unique-factorization domain? (E. D. Cashwell, C. J. Everett, who have proved unique-factorization in case of functions $\alpha(n)$ on integers to a *field*.)

CHAPTER III

Metric Spaces

1. Invariant properties of trajectories observed from moving coordinate systems

Suppose that we have, given in a fixed cartesian coordinate system (x, y, z), n moving points describing given curves, $x_i(t)$, $y_i(t)$, $z_i(t)$, $i = 1, 2, \ldots, n$. Suppose now that we have another cartesian coordinate system x', y', z', which is in motion relative to the given system. The given curves will appear differently in the moving system. The motion of the second system with respect to the first one is a general rigid motion, that is to say, the origin of the coordinates moves on an arbitrary curve, and the *rotation* of the system x', y', z' with respect to x, y, z is quite arbitrary as a function of time.

The question arises: what are the invariants of the given system of trajectories in respect to the arbitrarily moving observer? It is clear that for just one trajectory nothing can be said. In a suitably moving system, this trajectory will appear as a stationary point. It suffices to put the origin of the coordinate system onto the moving point. For two given trajectories it is clear that we can move the system x', y', z' in such a way that, for example, one of the points will appear to be stationary, say, again the origin, while the second point is moving on a straight line, say, the x-axis. Likewise, for three points, the invariants with respect to arbitrarily moving coordinate systems are trivial. It is clear that the invariants are functions of the mutual distances between the moving points at any given time.

If, however, $m \geqq 4$, some more interesting questions begin to arise. For example, given arbitrary continuous motions of four points, can one move the system of coordinates in such a way that to an observer of this moving system, the given trajectories will all appear as convex plane curves or perhaps as conics? If we have a sufficient number k of moving points whose trajectories are enlaced, is it true that in any moving coordinate system at least some two of them will appear enlaced?

Analogous questions of invariants of systems of trajectories (or, for that matter, more general parametrically represented surfaces, etc.) could be studied for a given class of topological transformations of space, more general than the rigid motions of the coordinate system.

2. Problems on convex bodies

(Mazur): In the three-dimensional, Euclidean space there is given a convex surface W and a point in its interior. Consider the set V of all points P defined by: the length of the interval OP is equal to the area of the plane section of W through O and perpendicular to OP. Is the set V convex?

A solid S of uniform density ρ has the property that it will float in equilibrium (without turning) in water in every given orientation. Must S be a sphere? (In a two-dimensional version of this problem, H. Auerbach [1] found shapes other than the circle with the desired property.) In the limit $(\rho \to 0)$ one obtains the following problem: If a body rests in equilibrium in every position on a flat horizontal surface, is it a sphere?

Let C be a star-shaped closed plane curve, i.e., a polar curve given by $\rho = \rho(\theta)$, and suppose that $\rho(\theta)$ has a continuous derivative except possibly at a finite number of points. It can be shown that there exists a constant $k > 0$ such that the curve given by $\rho = \rho(\theta) + k$ is convex. An analogous remark applies to surfaces in n-dimensions. Suppose C is a curve $(\theta(t), \phi(t))$ in three-space contained in the surface S of a star-shaped region including the origin. Under what conditions is it true that the surface S can be expanded by adding a constant to each radius so that the curve which results from the given curve can be obtained as an intersection of convex surfaces?

3. *Some problems on isometry*

If A and B are metric spaces, then A^2 and B^2 may also be regarded as metric spaces, the metric of a product space A^2 being defined, for example, by

$$\rho\big((a_1, a_2), (a_3, a_4)\big) = [\rho^2(a_1, a_3) + \rho^2(a_2, a_4)]^{\frac{1}{2}}.$$

Does isometry of A^2 and B^2 imply that of A and B? By an isometry between two metric spaces is meant a bi-unique transformation of one space onto all of the other which preserves all distances. A similar question may be asked for other metrizations of the product space — instead of the "Euclidean" formula above, one may use the formula:

$$\rho\big((a_1, a_2), (a_3, a_4)\big) = \max\,[\rho(a_1, a_3), \rho(a_2, a_4)]$$

or another "Minkowski" gauge function. This is a metric version of problems concerning the "extraction of the square root" in algebraic structures, e.g., if the groups A^2 and B^2 are assumed to be isomorphic, does it follow that A is isomorphic to B? (Cf. Fox [1].)

Is Hilbert space characterized metrically among Banach spaces by the fact that its group of isometries is transitive on the unit sphere (Mazur)?

4. *Systems of vectors*

Let $V_1, \ldots V_n$ be a system of n vectors in k-dimensional space. We are interested here in "bound" vectors, that is to say, a vector V is defined by an ordered pair of points (A, B) in E^k. We allow three types of operations on vectors, namely, (a) replacement of a vector $V = (A, B)$ by a vector $V' = (A', B')$ obtained from it by a translation T along the line through A, B, i.e., $A' = T(A)$, $B' = T(B)$; (b) replacement of a pair of vectors $V = (A, B)$ and $V' = (A, B')$ with common origin by their sum $V'' = (A, B - A + B' - A)$; (c) the inverse operation to (b), i.e., the splitting of a vector V into any two vectors with the same origin as V whose sum is V. Any two systems of vectors obtainable one from the other by a finite number of such operations are said to be equivalent.

It has been shown that, if σ_k is an arbitrary k-simplex in E^k, every finite system of vectors is equivalent to a set of at most $k + 1$ vectors lying on the edges of σ_k and the latter system is uniquely determined by the original one. (A result of L. W. Cohen [1] and the author).

It would be interesting to prove an analogous representation for arbitrary countable systems of vectors in Hilbert space, allowing countably many operations and infinite summations in (b) and (c).

5. Other problems on metrics

Characterize subsets of the plane such that the distance between any two of their points has a rational value. (Can such a set be dense?)

A problem is mentioned elsewhere in this collection on introducing a metric in an abstract algebraic structure (e.g., group) in such a way that the group operations would be continuous in the metric and the topology resulting from the introduction of such a metric would be of a specified type. We shall raise here the vague question: given a metric space, can one introduce a *metric* in it which would lead to the given topology, the *metric* being the "most natural one" among all metrics giving this topology. One can try to formulate precise questions which would attempt to define concretely some aspects of the phrase "most natural." For example, given a topological space, can one find a metric in it so that the group of all isometric transformations under this metric would be maximal in the following sense: for no other metric (leading to the same topology) would the group of isometries contain this group as a proper subgroup? In particular, is the Euclidean metric defined on the surface of the n-dimensional sphere maximal in this sense? The same question for the Hilbert space sphere in the usual metric. Obviously, in general, a topological space will possess many different maximal metrics in the above sense. One could perhaps consider a metric introduced in a topological space as stable if transformations which are "almost isometric" must, of necessity, be near to strictly isometric transformations (cf. Chapter VI, Section 1). The question is now for which topological spaces can one introduce stable metrics in the above sense? One obviously would want such metrics to be also maximal. Without that requirement, the problem would not have much sense since, in general, one can find metrics for which only the identity would be the isometric transformation.

It is not without interest to consider, in certain algebraic structures, an introduction of metrics such that the algebraic automorphisms would be isometric transformations, but we shall not go into this subject.

CHAPTER IV

Topological Spaces

1. A problem on measure

Let E be a compact metric space. Does there exist a finitely additive measure $m(A)$ defined for at least all the Borel subsets A of E, such that $m(E) = 1$, $m(p) = 0$ for all points p of E, and such that congruent sets have equal measure (Banach-Ulam)?

Two sets A and B are called *congruent* if there exists an isometry between A and B alone, not necessarily a congruence under an isometry of the whole space E taking A into B. (If one postulates this latter more restricted notion of congruence, then such a measure is known to exist.) It is also clear that the term "finitely" cannot be replaced by "countably" for an affirmative solution. (Consider the set of points (x, y) in the plane where $0 \leqq y \leqq 1$, $x = 0,\ 1/n$; $n = 1,\ 2,\ 3, \ldots$.)

2. Approximation of homeomorphisms of E^n

Let E^2 be the Euclidean plane and G the group of biunique, bicontinuous transformations of E^2 to E^2 generated by the set of all such correspondences of the form:

$$X\colon \begin{cases} x' = f(x,\ y) \\ y' = y \end{cases} \qquad Y\colon \begin{cases} x' = x \\ y' = g(x,\ y). \end{cases}$$

An arbitrary homeomorphism can be approximated, arbitrarily closely (uniformly in every bounded part of the plane) by transformations belonging to G (cf. Eggleston [1]).

Similar questions may be posed for Euclidean n-space in various ways depending on the type of generators allowed for the group G. For example, in three dimensions we may permit generators of the form X: $x' = f(x, y, z)$, $y' = y$, $z' = z$ and its two analogues, or again, let G be generated by all homeomorphisms of the type

$x' = f(y, z)$, $y' = g(x, z)$, $z' = h(x, y)$. The question still open is: Are arbitrary homeomorphisms of E^n approximable by transformations of the above type?

The possibility of such approximations for the general case of E^n would be of considerable importance for topology, providing an inductive procedure for proving various topological theorems; for example, the famous conjecture of Alexander [1] on the approximability of arbitrary homeomorphisms by differentiable mappings, which in turn implies the possibility of triangulation of any topological manifold. The basic lemma here would be the Alexander conjecture for the plane (proved by N. Wiener and P. Franklin). Quite recently Moïse [1]) proved that three-dimensional homeomorphisms are approximable by simplicial homeomorphisms.

The problem of this section, even for $n = 2$, i.e., E^2, is open for the case where E is a more general topological space. Moreover, in the formulation as it stands, X, Y, etc., need not be restricted to homeomorphisms, and the approximability is of interest when the class of transformations is widened to include general, continuous, or even, say, Borelian mappings.

It would be interesting to attempt to utilize the recent results of Kolmogoroff [1] and Arnold [1] on representation of *functions* of any number of variables by composition of functions of two variables to obtain such results for topological, that is to say, one-to-one transformations. In other words, even theorems allowing only approximation, if not exact representation of homeomorphisms of n-dimensional spaces by compositions of homeomorphisms involving only two dimensions at a time, would be extremely valuable.

2a. On the approximability of transformations in three dimensions by compositions of cylindrical mappings

Suppose we consider the smallest group G of transformations containing the transformations of the form:

$$\begin{cases} (x + iy)' = W(x + iy) \\ z' = z \end{cases}$$

where W is an analytic function, and all rotations of the three-dimensional space. Can one approximate arbitrarily closely, by transformations from the group G, a mapping of a sphere onto a cube? Somewhat more generally, can one by means of such transformations obtain approximate mappings of a polyhedron to any other, topologically equivalent polyhedron, by compositions of two fixed transformations?

3. A problem on the invariance of dimension

There may be a possibility of strengthening Brouwer's theorem (cf. Hurewicz, Wallman [1]) on the invariance of dimension in the following way:

Does there exist for every integer $m > 1$ a one-one continuous mapping T on E^m to E^m, (E is the real line) such that for every biunique Borel transformation U of E^m to all of E^n, $n < m$, the transformation UTU^{-1} of E^n to E^n is *discontinuous*?

An affirmative answer would imply Brouwer's theorem on the nonhomeomorphism of E^n and E^m, $m \neq n$. For, if H should be a homeomorphism of E^m to E^n, $n < m$, then H is trivially Borelian and HTH^{-1} would be continuous. One may also consider the above questions with T continuous and (possibly) many-one. The special case $m = 2$, $n = 1$ is trivially verified.

As often in our problems, the statement conjectured above could have a wider applicability: for a general space E such that E^n and E^m are *not homeomorphic* for any pair $n \neq m$.

We may add here parenthetically that it suffices to show that E is not homeomorphic to E^n, $n > 1$, and that if E^{2n} is not homeomorphic to E^{2m}, then E^n is not homeomorphic to E^m. With this we may conclude that for all $n \neq m$, E^n and E^m are not homeomorphic. This is actually a purely arithmetical fact: If K is a collection of pairs of integers such that

1. from $(a, b) \in K$ and $(b, c) \in K$ it follows that $(a, c) \in K$,
2. from $(a, b) \in K$ it follows that $(a + 1, b + 1) \in K$,
3. from $(2a, 2b) \in K$ it follows that $(a, b) \in K$,

then, if K contains any pair (n, m), $n \neq m$, K must contain a pair $(1, n)$, $n > 1$.

4. *Homeomorphisms of the sphere*

The group of homeomorphisms of the surface of the sphere S in three dimensions has two components. The component of the identity forms a simple group G (a result of von Neumann and Ulam [1]). In fact, a stronger theorem holds: for every two homeomorphisms A and B of S different from the identity there exists a fixed number n of conjugates of A, i.e., $H_1AH_1^{-1}, \ldots, H_nAH_n^{-1}$, H_i in S, whose product is B. This number n does not exceed 23; to determine the minimum number seems very difficult.

Analogous theorems for the k-sphere, $k > 2$, are as yet unproved, as are theorems on the simplicity of groups of homeomorphisms (forming the component of the identity) of manifolds, other than the sphere.

Very recent work by R. D. Anderson [1] generalizes these results to groups of homeomorphisms of sufficiently homogeneous (setwise) spaces: In particular, the group of all homeomorphisms of the Cantor ternary set, the universal curve, the set of all rational and the set of all irrational numbers are simple, also some interesting partial results on the group of all (orientation preserving) homeomorphisms of S^n — the n-sphere in n-space.

A question considered by Borsuk and the writer [2] is: Given an arbitrary closed subset C of the surface S of the sphere in n-space, does there exist a sequence of homeomorphisms H_n of S to all of itself such that $\lim H_n(E) = C$, that is to say, such that for every p of S, $\lim H_n(p)$ exists and is in C, and every point of C is such a limit.

5. *Some topological invariants*

No algorithm has yet been found which would permit one to decide whether two given curves in three-dimensional space are mutually enlaced. A sufficient condition for enlacement is that the Gaussian integral over the two curves is different from zero (cf. Alexandroff, Hopf [1]). Elementary examples show that the condition is not necessary. (We consider here two curves as not enlaced if there exists a homeomorphism of the whole space under which the images of the two curves are contained in disjoint geometrical spheres.)

Let us draw from each point of one curve a vector to every point of the other curve. If these vectors are all referred to a common origin and are normed to one, we shall have a mapping of the torus, which is the direct product of the two curves, onto the unit sphere. Does enlacement of the curves imply that the vectors cover the surface of the sphere essentially, that is to say, the mapping of the torus to the sphere is not retractable to a mapping into a single point?

Suppose that a system of n-vectors in 3-space forms a closed polygon. The vector sum is zero. If we consider the total moment vector relative to some point, it is curious that this may be zero as well. In a plane, for a polygon forming a simple Jordan curve, the total outer moment of these vectors can never be zero; indeed its magnitude is twice the area enclosed by the polygon. However, it is easy to find a hexagon in 3-space whose total moment with respect to any point is zero. This points to a possibility that some topological invariants may be derived from the moments of the various polygons formed by edges of a complex, or from even more general tensorial expressions. So, for example: We have discussed (Chapter III, Section 4) an equivalence of a system of vectors in n-dimensional space to a unique system of vectors located on the edges of a fixed simplex in n-space. Consider now a simplicial subdivision of a complex C located in the n-dimensional space. The system of its edges, properly oriented, will form a system of vectors which we will "represent" on the edges of a fixed (but arbitrary) simplex σ. The question arises as to what properties of this representation remain invariant under a subdivision C' of C for the corresponding representation of the resulting systems of vectors of the edges of C' on σ.

Another construction which may lead to significant topological invariants is the following: Let E be a topological space, and f a real-valued continuous function $f(p)$, defined for all points p of E. Let $G(f, E)$ be the group of all homeomorphisms A of E into itself such that $f(A(p)) = f(p)$ for all p. There is a possibility of distinguishing between two nonhomeomorphic spaces E and F by producing a function f on E, whose group $G(f, E)$ is not isomorphic to $G(g, F)$ for *any* function g on F. For example, if E is the circumference of a circle, and F is the unit interval, one can easily define a function f on E (say of period $2\pi/3$) whose group $G(f, E)$, the cyclic group of order 3, is not isomorphic to any group of homeomorphisms of F, since the interval F admits no homeomorphism of order 3.

One can, of course, employ mappings f on E to other spaces X than the real line. A general conjecture would be: If E and F are two nonhomeomorphic manifolds, there exists a space X and a mapping f of E on X so that $G(f, E)$ is not isomorphic to any $G(g, E)$ for any mapping g of F on X.

Among the simplest questions are the following: What abstract groups can be realized as groups $G(f, E)$ for a given space E? For instance, can every finite group be realized as a group $G(f, E)$ where E is the plane? What are all *countable* groups $G(f, E)$ which can result for functions f on Euclidean n-space?

Let f and g be two commuting transformations of I into itself (I is the closed interval): $f(g) = g(f)$. Does there exist a common fixed point $p_0 = f(p_0) = g(p_0)$? Communicated by A. L. Shields — originally raised by E. Dyer. The answer is not known even for $n = 1$. There are interesting partial results of Shields, J. R. Isbell, R. E. Chamberlain, and others.

6. *Quasi-fixed points*

Let E be a topological space, $T(p)$ a continuous mapping of E into itself, and $\phi_1(p), \ldots, \phi_k(p)$ a set of k real-valued continuous functions defined on E. We shall say p_0 is a quasi-fixed point of T relative to the ϕ_i in case

$$\phi_i(T(p_0)) = \phi_i(p_0), \qquad i = 1, \ldots, k$$

As an example, let E be the Euclidean plane, $T(p)$ a continuous transformation of E onto (into the whole of) E, $\phi_1(x, y) = |x|$, $\phi_2(x,y) = |y|$. One can show then the existence of a quasi-fixed point (x_0, y_0) of T relative to ϕ_1, ϕ_2.

Surely one will have to restrict either T or the ϕ_i for a significant theorem. Obviously for some T there may exist points quasi-fixed relative to every set of ϕ_i, in particular if T has a fixed point $p_0 = T(p_0)$.

Does an orthogonal transformation T of the unit sphere E in n-space always possess a point p_0 quasi-fixed relative to a set of $n - 1$ (arbitrary) continuous functions? This is of interset, of course, only in case the determinant of T is -1, and if true would constitute a generalization of the "Antipodensatz" of Borsuk [1] and the author.

It would be worthwhile to obtain results on the existence of quasi-fixed points in case T is a transformation of a function space E into itself. For, suppose that U is a functional operator and we are interested in the existence of a solution f_0 of the equation $U(f) = 0$. This is equivalent to finding a fixed point f_0 of the transformation $T(f) = U(f) + f$. Now in cases where the existence of a fixed point is difficult or impossible to establish, we may be satisfied with the knowledge that $\phi_i(T(f_0)) = \phi_i(f_0)$ for some or all sets of k continuous functions ϕ_i on E. These functionals ϕ_i might be, for example, the first k coefficients of f in its Fourier or power series development, or the first k moments of f, etc. Of course, the f_0 will, in general, depend on the ϕ_i in case there is no true fixed point of T. Nevertheless it may be useful in some applications to know that a function f_0 exists which has the same ϕ_i-values as its transform $T(f_0)$.

If the $f_i(p)$ are *linear* real valued functions defined on the Hilbert space S, and T is a continuous transformation of S into itself, does there exist a point p_0 such that $f_i(T(p_0)) = f_i(p_0)$?

Let $T(p)$ be a homeomorphism of Euclidean n-space. Suppose that for every point p, the set of iterates p, $T(p)$, $T^2(p)$, . . . is a set of finite diameter d_p and the d_p are bounded: $d_p < B$ for all p. Does $T(p)$ possess a fixed point $p_0 = T(p_0)$?

Does there exist for every manifold M a constant B such that every continuous transformation T of M into part of itself, having the property

$$| T^\nu(x) - x | < B$$

for all iterates ν and all x of M, must have a fixed point $x_0 = T(x_0)$? More generally, does such a constant exist for every locally-connected continuum?

7. Connectedness questions

Suppose that $T(p)$ is a differentiable transformation of the plane. If for some point p, the closure C of the set of all iterates p, $T(p)$, $T^2(p)$, . . . is connected, is the set C necessarily locally connected (Borsuk)?

Let S_1 and S_2 be two topological spherical surfaces (i.e., sets homeomorphic to the surface of the geometric sphere) in Euclidean 3-space, with S_2 contained in the interior of S_1. According to the Jordan-Brouwer theorem, S_2 decomposes space into two regions, and we assume S_1 is contained in one of them. Does there exist a surface S_3, topologically spherical, containing S_2 in its interior, and contained in the interior of S_1. The problem is not trivial because of the well-known examples of Alexander which show that the interior of a topological sphere need not be homeomorphic to that of a geometrical sphere. Analogous problems exist for higher dimensions (Schreier-Ulam [1]).

One may generalize the situation: Given are three topological spherical surfaces S_1, S_2, S_3 so that S_3 is contained in the interior of S_2, which in turn is contained in the interior of S_1. Can one find a topological sphere S_4 *either* contained in the interior of S_1 and containing in its interior S_2, *or* contained in the interior of S_2 and having S_3 in its interior?

8. Two problems about the disk

Suppose that $T(p)$ is a homeomorphism of the disk D (all (x, y) with $x^2 + y^2 \leqq 1$) onto all of itself. Do there exist arbitrarily small "triangles," i.e., triplets of points p_1, p_2, p_3 congruent to the triangles formed by the images $T(p_1)$, $T(p_2)$, $T(p_3)$? Do such triangles exist with prescribed angles?

Given a metric space A, a topological space B, and a (many-one) continuous mapping T on A to all of B, we consider the diameter d_b of the set $T^{-1}(b)$ for every b in B, and denote by

η_T the least upper bound of all d_b. Does there exist, for every $\varepsilon > 0$, a continuous mapping T of the disk D onto the surface B of the torus with $\eta_T < \varepsilon$? *

9. Approximation of continua by polyhedra

Let C be a simple closed curve in Euclidean 3-space which is non-knotted, that is to say, there exists a homeomorphism of the whole space which transforms C into the circumference of a circle. Can C be approximated arbitrarily closely by nonknotted polygons? This problem, considered by Borsuk and the author in 1930, has connection with the problem (of Alexander) of approximability of arbitrary homeomorphisms of the n-dimensional space by simplicial ones. An affirmative answer can be obtained from a theorem recently established by Moïse [1]: *vice versa* from a positive answer to the above, i.e., from an approximability of nonknotted curves by nonknotted polygons, the approximability of homeomorphisms follows. In the spaces of higher dimension than three, an analogous situation may obtain, namely, approximability of nonknotted spheres by nonknotted polyhedra may be sufficient to prove the approximability of general one-one continuous transformations by one-one differential transformations.

A problem of Borsuk connected with the above concerns the possibility of approximation of a unicoherent continuum in 3-space by unicoherent polyhedra. (A unicoherent continuum is a continuum E such that for every decomposition $E = A + B$ into two continua, $A \cdot B$ is a continuum. The disk is unicoherent, the circumference of the circle is not.)

There are a number of unsolved problems dealing with the approximability of continua with various given properties by polyhedra having the same properties.

10. The symmetric product

By the symmetric product E^n_s of a set E with itself is meant the class of all subsets of at most n distinct elements of E. Thus

* *Note added in proof*: A negative answer has been demonstrated by M. K. Fort, Jr.

E_s^n may be obtained from E^n by identifying all n-tuples $(p_1, \ldots p_n)$ whose component points form the same *set*. The importance of this "phase space" lies in the fact that some quantum statistics, for example, the Fermi-Dirac statistics, operates in similar spaces, just as the Maxwell-Boltzmann statistics operates in the direct product E^n. Thus the phase space of a system in which certain particles are indistinguishable is a symmetric product of the spaces corresponding to the particles.

If E is a topological space, a metric may be introduced into E_s^n using the Hausdorff distance between two sets of points. The properties of symmetric products are less well known than those of the direct product of spaces. It hS_1. Accordown that, whereas the direct product of n circles result ce into t dimensional torus, their symmetric product, for $n = 2$ n. Doe Möbius band. Moreover, if E is a real interval, E_s^n, for ng S, 3, and **4**, is the corresponding E^n, whereas for $n \geq 5$ this is roh the case (Borsuk, Ulam [1]).

The metrized symmetric product E_s^n forms an nth order approximation to the space of all closed subsets of a compact space E (with Hausdorff distance as the metric for the latter space).

The exact topological structure and, in some cases, even some very general topological properties of E_s^n remain unknown (even when E is the interval and $n > 5$). For example, the existence of fixed points (for arbitrary continuous mappings) has not been established.

Can the quasi-fixed point (see Section 6) theorem be proved for E_s^n, for $k < n$ arbitrary real valued continuous functions, where E is the surface of the n-dimensional sphere? (Compare Bott [1].)

The symmetric product as defined above is only one of many different constructions possible on E^n. Products based on other rules for identifications of certain sets of n-tuples in E^n may lead to interesting spaces. Such possibilities have not been studied systematically.

While the formation of a topological space E_s^n from a topological space E is readily accomplished, it is not easy to see a possible symmetrization in the case of algebraic structures. There seems to be no simple way to obtain a symmetric product of a group with itself. As a matter of fact, the symmetric product of a group space forms often (as in the case of the Möbius band) a topological space which will not support any continuous group operation. In some cases, however, e.g., the 2nd and 3rd symmetric product of the line with itself, E_s^2 and E_s^3 are homeomorphic to E^2 and E^3, respectively. This means that the analogue of vector addition can be continuously defined here. This leads, however, to an algebraically rather artificial rule for "adding" elements in E^2 or E^3.

11. A method of proof based on Baire category of sets

The theorem stating that a complete metric space is not a set of first category (sum of countably many nowhere dense sets) has been used to advantage for obtaining *existence* proofs in modern mathematics, notably in the theory of functions of a real variable, and in topology. Thus, for example, in order to show the existence of continuous functions without derivatives at any point, one may prove that the set of such functions forms a residual set (complement of a set of first category) in the space of all continuous functions. Again, in order to show the existence of metrically transitive measure-preserving transformations, one proves that the set of all such mappings is residual in the space of all measure-preserving transformations. To show the existence of a homeomorphic image of an arbitrary n-dimensional set Z_n contained in the $2n+1$-dimensional Euclidean space E^{2n+1}, it is quite simple to show that, in the space of all continuous (possibly many-one) mappings of the set Z_n into E^{2n+1}, the one-one mappings form a residual set. There are many other examples.

A possibility exists of using another theorem on residual sets (or, more generally, of second category) to prove, instead of existence theorems, propositions involving the complementary quantifier, i.e., the "for all" theorems. Thus, a well-known theorem asserts that if G is a connected topological group and H is a subgroup of it which is not a set of first category with respect to G, then H contains all the elements of G.

Similarly, if G is such a group with a finite Haar measure defined for its subsets, and if H is a subgroup of measure > 0, then $H = G$.

To illustrate this possibility let us consider as an example the well-known theorem of van der Waerden which asserts that every partition of the set of all positive integers into two subsets N_0 and N_1 has the property (P): at least one of the sets N_0 and N_1 contains finite arithmetic progressions of unbounded length. Suppose that we make correspond to every partition a real number x, (regarded as a point on the circle of the circumference 1)

$$x = a_1 2^{-1} + a_2 2^{-2} + \dots$$

where $a_n = 0$ if n is in N_0 and $a_n = 1$ if n is in N_1. It is immediately apparent that the set H of all x for which the corresponding partition has property (P) is a set of measure 1 in the continuous group G of all x (under addition modulo 1). Since van der Waerden's theorem is true, it is clear *a fortiori* that $H = G$. It would be interesting if it could be shown by a relatively simple argument that our set H is a group, and thus that $H = G$.

There are quite a few combinatorial theorems and problems still unsolved, where such an approach could be tried.

12. *Quasi-homeomorphisms*

Let A and B be two manifolds (i.e., topological spaces such that neighborhoods of points are homeomorphic to n-dimensional Euclidean spheres); we suppose them to be metrized. We define A and B to be quasi-homeomorphic if for every $\varepsilon > 0$ there exists a continuous mapping T_ε of A onto B (all of B) such that for every a, a' in A whose distance exceeds ε, $T_\varepsilon(a) \neq T_\varepsilon(a')$ in B, and there exist similar transformations S_ε of B onto A. Are A and B necessarily homeomorphic?

This problem was proposed in a paper by C. Kuratowski and the author [1]. It would even be useful to show that some general topological invariants remain unchanged, in case of manifolds, for which there exist ε-mappings of both A into B and B into A.

The property of possessing a fixed point under every continuous mapping into itself could be such an invariant. Similarly the "quasi-fixed point" property (for any finite number k of arbitrary real valued continuous functions) may perhaps be shown to be invariant.

13. Some problems of Borsuk

In these problems AR-sets denote compact absolute retracts, that is to say, sets such that in every containing space they can be obtained as continuous images of the space by transformations which are equal to identity on these sets. ANR-sets denote compact absolute retracts of their neighborhood.

1. Are all homology dimensions (in the sense of Alexandroff) for ANR-sets identical with the usual dimension?

2. Is it true that every n-dimensional compactum is homeomorphic with a subset of an $(n + 1)$-dimensional AR-set?

3. Is it true that every n-dimensional ANR-set contains an $(n - 1)$-dimensional ANR-set?

4. Is it true that every $(2n + 1)$-dimensional ANR-set contains topologically every compactum of dimension $\leqq n$?

5. In the three-dimensional Euclidean space, are the AR-sets the same as acyclic locally contractible compacta?

6. Does there exist for every ANR-set a polytope having the same homotopy type (in the sense of Hurewicz)?

7. Is the factorization of a continuum into one-dimensional factors unique? (Factorization = decomposition into Cartesian product.)

Note added in proof: Compare the beautiful recent example of Bing [1].

CHAPTER V

Topological Groups

1. Metrization questions

One of the general problems of topological algebra is to determine all possible topological groups which are definable for a given abstract group G. This means the characterization of all topologies on the set G in which the group operation will be continuous. There are many special known results related to this problem. For example, it has been proved (Ulam, von Neumann) that there exists an abstract group of the power of the continuum such that no metrization, for which the group operation is continuous, will make this group separable. Such a group can even be chosen to be abelian (commutative). We may mention here a few open questions:

Can the group S_∞ of all permutations of the set of all integers be metrized in such a way that the group operation is continuous in the metric, and the group becomes a locally compact space? J. Schreier and the author have proved that no metrization is possible which would make it compact. This is due to the following purely group-theoretical property: Starting with any truly infinite permutation p, that is to say, one which changes the positions of infinitely many integers, one can obtain, by multiplying a fixed number N of suitable conjugates of this permutation, an arbitrary infinite permutation. In a compact topological group, given any integer N, one can find a neighborhood of the identity so small that the products of at most N conjugates of the elements of this neighborhood will still form a set which is not the whole group; in fact, by choosing ε sufficiently small, one can insure that elements of this form will be confined to another neighborhood of the identity. Now any neighborhood of the identity in this group would have to contain, were it compact, a truly infinite permutation.

There are the four obvious metrizations of the group S_∞ based on its invariant subgroups: (a) the identity permutation, (b) the finite even permutations, (c) the finite permutations, and (d) the whole group S_∞. The "natural" topology in the whole group S_∞ is the following one. A sequence of permutations converges to a fixed permutation if it converges weakly, that is to say, the images of every integer in these permutations become, after a while, equal to the image of the limiting permutation and stay constant. The same should be required about the inverses of these permutations. This topology is easily defined by means of a metric. The topology thus obtained leads obviously to a nonlocally compact space. An analogous topology introduced in the invariant subgroups will also lead to nonlocally compact spaces since it amounts to having the space of co-sets discrete.

Are there any continuous metrizations of S_∞ different from these four? Since all four lead to nonlocally compact topologies, a theorem stating that no other continuous topologies are possible would, combined with certain results of A. Weil [1], provide definite examples of groups in which no invariant measure (Haar-Weil measure) is possible.

2. *Universal groups*

The theorems of von Neumann, Pontrjagin *et al.* [1], on the representation of continuous groups imply that every compact topological group G is continuously isomorphic to some subgroup of the direct product of all finite dimensional rotation groups.

Does there exist a universal compact semi-group, i.e., a semi-group U such that every compact topological semi-group is continuously isomorphic to a subsemi-group of it?

Can one show that the group R of all rotations in the three-dimensional space is isomorphic (as an abstract group, not continuously, of course) to a subgroup of the group S_∞ of all permutations of integers? Or, perhaps quite generally: is every Lie group isomorphic (as an abstract group) to a subgroup of the group S_∞?

Does there exist a separable, locally-compact group U such that every locally compact group is isomorphic to a subgroup of U? The results of Gleason, Montgomery [1], and Zippin generalize those of von Neumann and Pontrjagin to *locally* compact groups.

A separable continuous group U, "universal" for all separable continuous groups G (in the sense that every such G be continuously isomorphic to a subgroup of U) may not exist.

3. Basis problems

A finite set of elements of a topological group which generate a subgroup dense in the whole space is called a finite basis. There exist isolated results on the existence of such bases, but there seems to be no systematic investigation of their existence and properties for general continuous groups. Nor has the minimal number of elements in a basis been determined in most cases where a finite basis had been shown to exist. We mention here a few of the known results.

Let S_∞ be the group of all permutations of the integers, with the "natural" topology, i.e., a sequence of permutations converges to a limit if the convergence is termwise, and the sequence of inverse permutations converges in the same sense to the inverse of the limit. There exist two elements in S_∞ such that the group generated by them is dense in the whole group. In fact, a specific set of products of powers of these two elements can be shown to form a dense set in S_∞ (J. Schreier, Ulam [4]). There is no scarcity of such pairs of permutations; almost every pair (in the sense of category) will serve as a basis.

An analogous situation obtains in the group of homeomorphisms of n-dimensional Euclidean space (J. Schreier, Ulam [1]). Does this group possess a basis consisting of differentiable transformations?

In the case of finite dimensional topological groups, one knows, for example, that the group of rotations contains pairs of elements generating dense subgroups.

In every semi-simple connected Lie group there exists a basis of four elements. It is not known whether this is the minimal number (Schreier, Ulam [3]).

One can find a pair of rotations of Hilbert space which generates a dense subgroup (in a weak sense) in the group of all such rotations (J. Schreier [1]).

Less is known about topological semi-groups, but the general situation seems to be quite similar to the one for groups, to judge from the evidence we have so far. Thus the class T_∞ of all (many-one) transformations of the set of integers into itself has a natural topology. There exists a finite set of elements generating a subset, dense in this topology, in the semigroup T_∞. A similar result is known for the semi-group of continuous transformations of the sphere into itself (Schreier, Ulam [1]).

Let H_n denote the group of all homeomorphisms of the n-cube I^n onto all of itself, metrized by means of the distance function:

$$\rho(h_1, h_2) = \max_{p \in I^n} \rho(h_1(p), h_2(p)) + \max_{p \in I^n} \rho(h_1^{-1}(p), h_2^{-1}(p))$$

the latter distances referring to the Euclidean metric in E^n. Is it true that almost every pair of homeomorphisms generate a group dense in H_n? ("Almost every" is used here in the sense of category, i.e., do these pairs form a residual set in the space of all pairs?)

Again, let H_n^k denote the kth direct product of the space H_n with itself, with the usual extension of the above H_n metric to a product space metric. Does almost every element $(h_1, \ldots, h_k)$ of H_n^k provide a set $h_1, \ldots, h_k$ which generate a free group of k generators in H_n? The same question may be asked about the semi-group G of all continuous, not necessarily one-one, transformations of n-space.

Given a pair of measure preserving transformations S, T of a space E into itself one may consider transformations such as S, $T(S)$, $T(T(S))$, $S(T(T(S)))$ ... etc. Specifically, to every real number $0 \leqq x \leqq 1$ in binary expansion $x = \alpha_1 \alpha_2 \ldots \alpha_n \ldots$, $\alpha_i = 0$ or 1, we may make correspond a sequence of transformations interpreting the symbol 0 as applying S and 1 as applying T in the order indicated. Is it true that, for almost every point p of the space E and for almost all x, the resulting sequence of images: p, $S(p)$, $T(S(p))$, $T(T(S(p))$, $S(T(T(S(p)))$ etc. will be *uniformly* dense?

The problem is: is this sequence uniformly dense for almost every (in the sense of category) pair of measure preserving transformations of E_n?

Is this true for a Haar group when S, T are transformations obtained by left multiplying the group by two elements s, t (if the Haar group is simple)?

It appears likely that there exist finite sets $T_1, T_2, \ldots, T_k$ ($k = 2$?) of measure-preserving transformations of the Euclidean cube generating a group dense in the space Σ of *all* continuous measure preserving transformations. If this were so, one could define a measure function in the space of all incompressible flows of say the three-dimensional space into itself. This measure would be obtained by considering the Lebesgue measure in the space of the x's corresponding to the transformations, as above.

4. Conditionally convergent sequences

Let V_n, $n = 1, 2, \ldots$, be a conditionally convergent series of vectors. It is well known that if we rearrange the terms in all possible ways and form their sums, these sums will form a linear manifold in the vector space. Garrett Birkhoff and the author have noticed that this well known theorem can be generalized if one considers the V_n to be elements of a compact group G; then the "sums" of all possible rearrangements of the sequence form a *coset* modulo a certain subgroup of G. Does a similar result hold for more general, noncompact topological groups?

CHAPTER VI

Some Questions in Analysis

1. Stability

We intend now to discuss, by means of a few examples, the notion of the stability of mathematical theorems considered from a rather general point of view: When is it true that by changing "a little" the hypotheses of a theorem one can still assert that the thesis of the theorem remains true or "approximately" true? The notion of stability arose naturally in problems of mechanics. It involves there, mathematically speaking, the continuity of the solution of a problem in its dependence on initial parameters. This continuity may be defined in various ways. Often it is sufficient to prove the boundedness of the solutions for arbitrarily long times, e.g., the boundedness of the distance between the point representing the system at any time from the initial point, etc. Needless to say, problems of stability occur in other branches of physics and, in a way, also, even in pure mathematics.

We shall not try to formulate a generally applicable definition of stability. One could attempt to do this by introducing suitable function spaces for physical theories and various metrics in them, but we shall be content instead to indicate some of the salient features of this concept as it appears in purely mathematical formulations. In particular we shall formulate some problems concerning the stability of solutions of functional equations.

For very general functional equations one can ask the following question. When is it true that the solution of an equation differing slightly from a given one, must of necessity be close to the solution of the given equation? Similarly, if we replace a given functional equation by a functional inequality, when can one assert that the solutions of the inequality lie near to the solutions of the strict equation? Instead of trying to define a general class of functional equations (one recalls a dictum of Poincaré's: "On ne peut guère définir les équations fonctionnelles en général") for this purpose, let us restrict ourselves to examples. A good illustration, in an elementary case, is provided by a result of Hyers' [1], resolving a problem of the author:

If $f(x)$ is a measurable real-valued function defined for all real x satisfying the inequality

$$|f(x+y) - (f(x) + f(y))| < \varepsilon$$

everywhere, one can show that there exists a function $l(x) = ax$ such that

$$l(x+y) = l(x) + l(y) \text{ and } |l(x) - f(x)| \leqq \varepsilon$$

everywhere. We say then that the functional equation of linearity

$$f(x+y) = f(x) + f(y)$$

is stable with respect to a change into an inequality. (By the way, even if $f(x)$ is not measurable, one can assert that the solution of the inequality is close to some — nonmeasurable perhaps — solution of the strict equation.)

One can ask, much more generally, for what metric groups G is it true that an ε-automorphism of G is necessarily near to a strict automorphism? (An ε-automorphism means a transformation f of G into itself such that $\rho(f(x \cdot y), f(x) \cdot f(y)) < \varepsilon$ for all x, y in G.) There should exist then a constant k, depending only on G and not on f, and an $a(x)$ depending of course on f, with $a(x \cdot y) = a(x) \cdot a(y)$ such that $\rho(a(x), f(x)) < k\varepsilon$ for all x. We require this to hold e.g., for all continuous or measurable, f. The above result of Hyers' answers the question when G is the group of real numbers relative to addition. Hyers also obtained results in the case when G is a more general vector space. Another paper, by Hyers and the author, answers the question, in an affirmative sense, for some infinite-dimensional vector spaces. In this and other examples, it should be pointed out that a formulation of stability "in the large" requires a metric in the space of functions. This is, of course, true in all the classical studies of the problem of stability in differential equations or systems of differential equations. One could ask for weak convergence of solutions — pointwise — or more strongly for uniform convergence in the norm (cf. Hyers, Ulam [1, 2]).

It is interesting that the notion of stability in the above sense can be introduced even for discrete structures; thus, for example, Shapiro [1] has answered a question of the author by the following result. If $t(a)$ is an automorphism "up to an integer k" of the set of all the integers mod p where p is a "large" prime and k is an integer "much smaller" than p, then under suitable conditions $t(a)$ is a strict automorphism. One could also ask for stability of configurations or constructions even in elementary geometry. Thus, to give an elementary and *ad hoc* example, and to depart from linear formulations and illustrate what we mean by a "quadratic" problem — if the constructions of the Pascal and Brianchon theorems for arbitrary sextuplets of points located on a continuous curve always result in points which are almost collinear — or in lines which are almost concurrent — is the given curve approximately a conic?

Still another illustration in metric geometry: D. Hyers and the author [1, 2] have shown that a transformation of Euclidean space which changes all distances by at most $\varepsilon > 0$ is of necessity near to an isometry, that is, a transformation *strictly* preserving all distances. It is not known for what general metric spaces the above statement remains true.

Many questions about transformations suggest themselves in the same vein. It can be shown that a transformation which very nearly preserves measure in a Euclidean space is close to a transformation which strictly preserves the measure of all subsets. Can one prove it in more general measure spaces? Is a transformation which is nearly laminar of necessity close to one which is strictly so? What can be said about transformations which are almost irrotational? Questions like the above on stability of properties involving differential flows lead to general problems about the stability of differential expressions. The following theorem was also noticed by D. Hyers and the author [4].

Let $f(x)$ be a real function on the line with its nth derivative $f^{(n)}(x_0) = 0$ for a certain n, and suppose that $f^{(n)}(x)$ changes sign in the neighborhood of x_0. For every $\varepsilon > 0$ there exists a $\delta > 0$ such that for every function $g(x)$ of class $C^{(n)}$ satisfying $|f(x) - g(x)| < \delta$, there is a point x_1 such that $g^{(n)}(x_1) = 0$ and $|x_1 - x_0| < \varepsilon$. It is perhaps remarkable that the hypothesis involves only the nearness of the functions f and g themselves, and the number δ is dependent on ε alone.

One can obtain a partial generalization of this result. Let $F(x, f(x), f'(x))$ be a continuous function of the three arguments, with $F(x_0, f(x_0), f'(x_0)) = 0$ for some x_0 and suppose that F changes sign in a neighborhood of x_0. Then again for every $\varepsilon > 0$ a $\delta > 0$ exists such that for every $g(x)$ of class $C^{(1)}$ closer than δ to $f(x)$ there is a point x_1 close to x_0 such that $F(x_1, g(x_1), g'(x_1)) = 0$.

Under what conditions can the previous theorem be generalized to $F(x, f(x), \ldots, f^{(n)}(x))$ or to $F(f(x, y); \partial f/\partial x, \partial f/\partial y, \ldots)$? The most interesting question of this kind concerning functions of two variables is the simultaneous vanishing of several partial derivatives at a point (x_0, y_0). Here one has to be careful about the meaning of "change of signs" in the neighborhood of a given point. It seems necessary to assume at least that all combinations of sign occur. Very little is known about such questions in the n-dimensional case.

A theory of the Calculus of Variations in the Large developed by Marston Morse [1] operates with qualitative or topological definitions of critical points of functions of several variables. This theory provides a general qualitative basis for the phenomena implied by the vanishing of first derivatives. Our remarks would seem to indicate that there may exist topological definitions relative to expressions involving higher derivatives.

We have dealt above with stability of solutions of some functional equations — when we change these equations into inequalities. One might study this question when an inequality is given *ab initio*. For example, there is the result of Hyers and the author [3] that a function which is almost convex lies near to a strictly convex function. In the most elementary case, it states that a solution of the functional inequality

$$f(x+y)-f(x)-f(y)<\varepsilon$$

is a function differing everywhere from a convex function by at most a fixed multiple of ε.

The subject of small deformations is treated explicitly in topology which, as once defined by Poincaré, is the study of those properties of figures which remain true even when the figure is drawn by unskilled draftsmen. This definition demands more than the invariance of properties of sets under 1-1 continuous transformations or homeomorphisms. It would seem to require such invariants under more general ε-deformations. Many topological properties are known to be invariant under such transformations. For example, two manifolds that are ε-deformable into each other for every $\varepsilon > 0$ have the same Betti numbers and other homology invariants. However, there are many topological properties for which this more general invariance is not yet demonstrated.

It is of interest not only to prove that for sufficiently small ε certain topological properties remain invariant under ε-deformations but actually to find the maximal value of this ε. More precisely, we have a given metric set and we consider all continuous mappings of this set into another set such that no two points distant by more that ε have the same image point. In other words, the transformation does not coalesce any pair of points whose distance is ε or more. Suppose, for example, the given set is an S_n (surface of a sphere in n dimensions with radius one). A theorem of Borsuk and the author [2] asserts that if this set S_n is mapped into a set in the Euclidean space E_n in such a way that no two points whose distance is greater than a certain l_n coalesce, then the image still cuts the Euclidean space into at least two regions. Determination of the number l_n which is the best (largest) constant has been made in this case. An analogous determination of such constants would be of some interest in the case of other topological properties. For example, suppose that T_n is a torus given metrically in the Euclidean space E_n as a product of two circumferences of the circle of radius 1. One should determine the maximum k_n such that if this set T is mapped into a Euclidean space in such a way that no two points more distant than k_n coincide, then the Betti numbers of the image must be greater than zero. (Compare our problems on quasi-fixed points in the previous chapter.)

The theorems and problems concerning the stability (of the vanishing) of differential expressions for functions of n variables, discussed above, can be considered for functionals. If we consider a typical elementary problem, that of finding an extremum y_0 of

$$I(y) = \int_a^b F(x, y(x), y'(x))\, dx$$

an analogous question to the one about derivatives of functions of a finite number of variables arises, namely, the conditions which guarantee that for every $\varepsilon > 0$ there exists a $d > 0$ such that for all sufficiently "regular" $G(x, y, z)$ with $|F - G| < d$, there exists a minimum $y_1(x)$ for

$$J(y) = \int_a^b G(x, y(x), y'(x))\, dx$$

where $|y_1 - y_0| < \varepsilon$. We assume here merely the proximity of F and G, and nothing is assumed about the proximity of their partial derivatives, occurring in the Lagrangian equations. Speaking descriptively, the question is: when is it true that solutions of two problems in the calculus of variations which correspond to "close" physical data must be close to each other? Affirmative theorems of this sort would ensure the stability of the solutions of physical problems even with respect to the introduction of small additional "hidden" parameters. Since this situation regarding the stability is obscure at present even for total differential equations, it is perhaps futile at this stage to speculate about more general formulations involving partial differential equations from this point of view. However, the following remark should perhaps be made. It seems desirable in many mathematical formulations of physical problems to add still another requirement to the well-known desiderata of Hadamard of existence, uniqueness, and continuity of the dependence of solutions on the initial parameters. Specifically one should have a stability in the stronger sense illustrated by us above: the solutions should vary continuously even when the operator itself is subject to "small" variations.

But even in problems of discrete mathematical structures, the notion of stability can be introduced quite generally. One could even speculate about defining a *distance* between statements in some formal system of mathematics in such a way that definitions of sets which are "close" to each other, say in the sense of a Hausdorff distance, would correspond to points with a small separation.

2. Conjugate functions

We call two transformations f and g of a space E into itself conjugate if there exists a biunique h such that $f = hgh^{-1}$. It is well known that two one-one functions f and g on an abstract set E to itself are conjugate if and only if the two decompositions of E into *cycles* under f and under g are similar. This means that the number of f-cycles of length l is the same as the number of g-cycles of length l for every cardinal l. In the case of functions f and g which are many-one a similar theorem may be proved, generalizing the concept of cycle to that of a "tree". By a tree is meant a minimal set T containing the image $f(x)$ and the complete counter-image $f^{-1}(x)$ for every point x in T. A tree may be represented as a graph containing at most one closed cycle. Different trees are disjoint. It is obvious what is meant by two trees being of the same type. A necessary and sufficient condition for two many-one functions on an abstract set E to itself to be conjugate is that the number of trees of each type be the same under f as under g.

A general investigation of conditions for conjugacy in case E is a given space and f, g, and h are of restricted character, e.g., where continuity for h is required, seems lacking. In particular we may ask the following questions:

If two transformations f, g of n-dimensional Euclidean space E^n are each given by polynomial forms and are conjugate under a continuous h, are they then conjugate under a linear transformation h_1 of E^n?

If two continuous transformations f, g are conjugate under a Borel transformation h, are they then necessarily conjugate under a continuous transformation h_1?

Under what conditions is a homeomorphism of E^n conjugate to a uniformly continuous transformation?

Restricting ourselves to one dimension: is every "smooth" function $f(x)$ on $(0, 1)$ to $(0, 1)$ (e.g., every such polynomial) conjugate to a suitable piecewise linear function? For example the parabola $f(x) = 4x(1 - x)$ is conjugate to the function defined on $(0, 1)$ by the "broken line"

$$g(x) = 2x, \qquad 0 \leqq x \leqq \tfrac{1}{2}$$
$$g(x) = 2(1 - x), \quad \tfrac{1}{2} \leqq x \leqq 1$$

under the biunique transformation

$$h(x) = 2/\pi \sin^{-1} \sqrt{x}$$

An affirmative answer to the above question would reduce the study of the iteration of such functions $f(x)$ to a purely combinatorial investigation of the properties of "broken line" functions.

It may be advantageous to consider a type of conjugacy (at least formally) weaker than the one defined. Let us say that two functions $f(x)$ and $g(x)$ are asymptotically conjugate if the behavior of iterates of points under f and under g is similar in the following sense: there exists a biunique function $h(x)$ on $(0, 1)$ to itself, such that for almost every a on $(0, 1)$, $h(R_a) = S_a$, where R_a is the set of all points which have identical sojourn time in $(0, a)$ under iteration of f, and S_a is the corresponding set for g.

Is it then true that every polynomial $f(x)$ is asymptotically conjugate to a broken-line function $g(x)$? These problems are, of course, not limited to the one-dimensional case and are indeed of greatest interest in higher-dimensional spaces.

3. *Ergodic phenomena*

In this section we shall be concerned with iteration of functions and transformations, more particularly with the asymptotic properties of the sequence of iterated images of points. The great advances in ergodic theory of the last few decades have clarified the mathematical basis of statistical mechanics to a considerable if not to a complete extent. Roughly speaking, the analogues of the laws of large numbers in the theory of probabilities do now exist in the form of ergodic theorems. The more detailed analysis of analogues of the Gauss-Liapounoff-type theorems is by far less complete. We should mention here parenthetically, that often it is important to deal with transformations of noncompact spaces, e.g., the entire Euclidean space, into themselves. Certain theorems formulated originally for the compact case can be generalized under suitable formulation for such cases. So, for example, the Kronecker-Weyl theorem on the existence of ergodic means for rotations in n-dimensional space can be generalized, to some extent, as follows.

Let L be an *arbitrary* linear transformation of the Euclidean n-space into itself. Let C be any cone of directions in space. For almost every point p the sequence of iterated images $L^n(p)$ has a sojourn time in C. In other words, the ergodic limit of angles exists for almost all initial points p.

If the transformation

$$T \begin{cases} x'_1 = f_1(x_1, \ldots, x_n) \\ \vdots \\ x'_n = f_n(x_1, \ldots, x_n) \end{cases}$$

of Euclidean n-space E^n to itself is linear:

$$x'_1 = \sum_j x_i a_{ij} \qquad A = (a_{ij})$$

with all coefficients $a_{ij} > 0$, it is well known (Frobenius-Perron) that there exists a unique positive characteristic root r and a unique unit invariant vector $\bar{v} = (\bar{x}_1, \ldots, \bar{x}_n)$ with positive components such that $\bar{v}A = r\bar{v}$. Moreover, for every vector $v = (x_1, \ldots, x_n)$ with positive components, the sequence of points on the unit sphere:

$$vA^n/|vA^n|, \qquad n = 1, 2, 3, \ldots,$$

converges to $\bar{v}$. These facts establish the existence of "steady-state" distributions in many problems involving the multiplication and diffusion of particles.

A simple result of this sort cannot be hoped for in case the transformation T is nonlinear. Such transformations occur naturally in various physical problems involving interaction between the multiplying and diffusing particles. For example, if one takes account of the depletion of the medium in a multiplicative process, the equation for the moments of the probability flow is nonlinear. Also, if there are particles of, say, two different types, and the multiplication of each type depends upon the number of both types present, the corresponding transformation is also nonlinear.

To illustrate the simplest type of question that arises in such cases, suppose that the transformation T has the form

$$f_i(x_1, \ldots, x_n) = l_i(x_1, \ldots, x_n) + q_i(x_1, \ldots, x_n), \quad i = 1, \ldots, n,$$

where the l_i are linear and the q_i pure quadratic forms in the variables $x_1, \ldots, x_n$.

The problem of interest concerns the asymptotic behavior of the sequence of directions of $T^\nu(x)$ generated by iteration of T on an initial vector x. The Frobenius theorem may perhaps generalize to the following, which we state as a conjecture.

1. Given a cone C of directions issuing from the origin and "almost any" vector $x = (x_1, \ldots, x_n)$, the "time of sojourn" of the iterates $T^\nu(x)$ in C exists.

2. For a given cone C the "sojourn time" of $T^\nu(x)$ in C may depend upon x but there are in this case only a finite number of values of such times.

The latter conjecture, if true, would mean that the space E^n splits into a finite number of disjoint subsets $S_1, \ldots, S_m$, all vectors x in S_i having the same sojourn time t for $T^\nu(x)$ in C.

The considerable variety of physical problems which can be cast in the form of a study of such transformations provides an interest in the investigation of their iterative properties. Thus the equation $\dot{x} = f(x)$ where x is a vector becomes

$$x^{\nu+1} = x^\nu + f(x^\nu) = g(x^\nu)$$

in difference form, which in turn leads to the transformation

$$x' = g(x)$$

and its iterations.

In similar fashion, a partial differential equation of the form

$$\partial u/\partial t = F(u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}, \ldots)$$

when written in difference form on a finite mesh can be regarded from exactly the same point of view, the function $u(x, y)$ being considered as a vector

$$[u(x_1, y_1), \ldots, u(x_n, y_n)] = [u_1, \ldots, u_n] = u$$

the function u' at time $t + 1$ being given in terms of u at time t by an equation of form

$$u' = G(u)$$

The problems proposed on quadratic transformations in spaces of two or more dimensions seem rather difficult. It is of interest to consider in more detail the one-dimensional case.

4. *The Frobenius transform*

Let $y = f(x)$ be a measurable nonsingular transformation of the unit interval into itself. O. W. Rechard [1] has studied the transformation T_f of $L^1(0, 1)$ into itself defined by

$$\int_A T_f \xi \, dx = \int_{f^{-1}(A)} \xi \, dx$$

For reasons that shall appear later, this might be called the Frobenius-Perron transform corresponding to f. It can be shown that the transformation T_f has a nonnegative invariant function $\mu(x)$ such that

$$\lim_{n \to \infty} m[f^{-n}(N)] = 0$$

(where N is the subset of (0, 1) on which μ vanishes) if and only if the set functions $m[f^{-n}(A)]$ are uniformly absolutely continuous with respect to the Lebesgue measure m. Under these circumstances, for every interval (a, b) on (0, 1) and for almost every x, χ denoting the characteristic function,

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{j=0}^{n} \chi_{(a,b)} [f^j(x)]$$

exists, and (if f is metrically transitive) is equal to $\int_a^b \mu\, dx$.

For example, if $g(x)$ is the broken-line function of Section 2, the transformation T_g is defined by

$$T_g \xi(x) = \tfrac{1}{2}\xi\left(\frac{x}{2}\right) + \tfrac{1}{2}\xi(1 - \tfrac{1}{2}x)$$

and the function $\mu(x)$ is identically 1. From this it follows that if $f(x)$ is the parabolic function $f(x) = 4x(1 - x)$, for which the transformation T_f is

$$T_f \xi(x) = 1/4\sqrt{(1-x)}\{\xi[\tfrac{1}{2}(1-\sqrt{(1-x)})] + \xi[\tfrac{1}{2}(1+\sqrt{(1-x)})]\}$$

then the corresponding function $\mu(x)$ is given by

$$\mu(x) = d/dx\,[2/\pi \sin^{-1} \sqrt{x}] = 1/\pi \sqrt{x}\sqrt{1-x}$$

These remarks suggest the following question. If a transformation $f(x)$ of the unit interval into itself is defined by a sufficiently "simple" function (e.g., a broken line function or a polynomial) whose graph does not cross the line $y = x$ with a slope in absolute value less than 1, does the corresponding F.-P. transform have a nontrivial invariant function? It is not even known if this is true for every transformation of the form

$$f(x) = 2x, \qquad 0 \leqq x \leqq \tfrac{1}{2}$$
$$f(x) = (2 - a) + 2(a - 1)x, \qquad \tfrac{1}{2} \leqq x \leqq 1$$

where $0 < a < \frac{1}{2}$. (For $a = \frac{1}{2}$ this is easily shown to be the case.)

The F.-P. transform corresponding to $f(x)$ can be thought of as the continuous analogue of the following transformation defined on the space of step functions on (0. 1). Let the unit interval be divided into n equal nonoverlapping subintervals $I_1, I_2, \ldots, I_n$ and define a_{ij} as that fraction of interval j which is mapped into interval i by $f(x)$. That is,

$$a_{ij} = m[I_i \cdot f^{-1}(I_j)]/m(I_i)$$

If now

$$\sigma_n(x) = \sum_{i=1}^{n} C_i \chi_I(x)$$

is a step function defined on the intervals I_j, then we define the transformed function

$$T_n \sigma_n(x) = \sum_{i=1}^{n} C'_i \chi_{I_i}(x)$$

where

$$C'_i = \sum_{j=1}^{n} a_{ij} C_j$$

That is the vector of coefficients (C'_i) is just the result of operating on the vector (C_i) with the matrix a_{ij}.

If A and B are two subsets of $(0, 1)$ each consisting of a finite number of intervals I_j and if $f^{-1}(A) = B$, then

$$\int_A T_n \sigma \, dx = \int_B \sigma \, dx$$

The matrix a_{ij} has its largest eigenvalue equal to 1 and a corresponding nonnegative eigenvector $(\bar{C}_i)$ which can be regarded as defining a step function $\bar{\sigma}_n(x)$ that is invariant under the transformation T_n. We conjecture that if the F.-P. transformation T_f has a nonnegative invariant function $\mu(x)$, then the invariant step functions $\bar{\sigma}_n(x)$ converge to $\mu(x)$ in $L'(0, 1)$ as n, the number of subdivisions of $(0, 1)$, becomes infinite.

Under fairly weak restrictions on the matrix (a_{ij}) (e.g., if some power $(a_{ij})^k$ contains only positive elements), the theorem of Frobenius-Perron asserts that the invariant step function $\bar{\sigma}_n(x)$ is the limit, as j becomes infinite, of the sequence of iterates $T_n^j \sigma_n(x)$, where $\sigma_n(x)$ is any nonnegative step function not identically zero. Does a similar result hold for the general continuous F.-P. transform? That is, if $\xi(x) \geqq 0$ is not identically zero, does the sequence of iterates $T_f^j \xi(x)$ converge in $L'(0, 1)$ to $\mu(x)$? Computational evidence for the case $f(x) = 4x(1 - x)$ and $\xi(x) \equiv 1$ suggests that this is the case.

5. Functions of two variables

The following conjecture of the author has been proved by Zahorski [1]: for every function $f(x)$ continuous on the unit interval (0, 1), there exists a function $g(x)$ analytic on (0, 1) and a perfect set C on this interval such that $f(x) \equiv g(x)$ for all x of C. Is the analogue true in the plane if the functions $f(x, y)$ and $g(x, y)$ are continuous and analytic, respectively, on the unit square, and if the set C is required to be a direct product of two perfect sets?

Let $f(x_1, \ldots, x_n)$ be a real-valued continuous function defined on the "unit cube" $0 \leqq x_i \leqq 1$. Does there exist an arc in the cube on which the function is constant? *

Consider a continuous function $f(x, y)$ of two real variables which is associative

$$f(x, f(y, z)) = f(f(x, y), z)$$

for example, $x + y$, xy, $(x^2 + y^2)^{\frac{1}{2}}$, etc. What further condition on f guarantees that there exist a finite number of such functions g_i such that every associative continuous function f is conjugate to one of these, in the sense that for some $L(z)$

$$f(x, y) = L^{-1}(g_i(L(x), L(y)))$$

Thus $f(x, y) = xy$ and $g(x, y) = x + y$ are conjugate under $L(z) = \log z$ for positive x, y.

Obviously, the class of associative functions of two variables, if one merely requires continuity is very large — is it ever true, perhaps, that every continuous transformation of a plane $T : x' = f(x, y)$, $y' = g(x, y)$ can be obtained by composing a finite number of transformations T_i where f_i and g_i would be associative? Compare the wonderful recent results of Kolmogoroff [1].

6. *Measure-preserving transformations*

Does there exist a square integrable function $f(x)$ and a measure preserving transformation $T(x)$, $-\infty < x < \infty$, such that the sequence of functions $\{f(T^n(x))\,;\, n = 1, 2, 3, \ldots\}$ forms a complete orthogonal set in Hilbert space? (Banach)

* *Note added in proof*: A negative answer follows from recent constructions of R. H. Bing.

Let the real numbers $x = \sum_1^\infty a_i 2^{-i}$ ($a_i = 0$ or 1) on $(0, 1)$ be represented by the sequences

$$\sigma(x) = \{\ldots, a_6, a_4, a_2; a_1, a_3, a_5, \ldots\} \equiv \{\ldots, b_{-3}, b_{-2}, b_{-1}; b_0, b_1, \ldots\}$$

The transformation

$$T(x) = \sigma^{-1}\{\ldots, b_{-3}, b_{-2}; b_{-1}, b_0, b_1, \ldots\}$$

(right shift of one place) is known to be measure preserving. Is it true that T and its iterate T^2 are not conjugate under any measurable transformation $H(x)$ (i.e., $T \neq HT^2H^{-1}$)? More strongly, is it true that T has no measurable square root $S(x)$: $T(x) = S(S(x))$?

7. Relative measure

Does there exist, for every set A of measure zero (say on the interval), a countably additive measure function m_A under which at least all Borel subsets of A are measurable, and which has the properties

1. $m_A(A) = 1$; $m_A(p) = 0$, p a point of A,
2. for $A \supset B \supset C$, $m_A(C) = m_A(B) \cdot m_B(C)$.

In other words we desire a *class* of measure functions with a possibility of relativising it "uniformly."

Suppose that the basic space is the circumference E of the unit circle, and $m(X)$ the Lebesgue measure, which can be equivalently defined as

$$m(X) = \lim_{N \to \infty} (N+1)^{-1} \sum_{\nu=0}^{N} \chi_X(T^\nu(x))$$

for almost all x, where χ_X is the characteristic function of the subset X of E and T is any irrational rotation. For B taken as a Borel subset of a set A of measure zero, when does the limit

$$\lim_{N \to \infty} \sum_{\nu=0}^{N} \chi_B(T^\nu(x)) \Big/ \sum_{\nu=0}^{N} \chi_A(T^\nu(x))$$

exist, and for which subclass does this provide a suitable $m_A(B)$ function?

Another suggestion, due to A. L. Shields, is to consider

$$m_A(B) = \lim_{\varepsilon=0} m(B_\varepsilon)/m(A_\varepsilon)$$

where m is the ordinary Lebesgue measure and B_ε is the set of points whose distance from B is less than ε (similarly for A_ε).

(It is clear *a priori* that we shall not establish in this way an m_A for *all* Borel sets A of measure zero with the desired properties — in fact, for every A dense in the interval the measure would coincide with the measure of its closure. It would be of interest, however, to construct a class of measure functions with the above properties *1* and *2* for a sufficiently large class of the sets A.)

8. Vitali-Lebesgue and Laplace-Liapounoff theorems

A classical result (the "Vitali-Lebesgue theorem") in the theory of functions of a real variable is that, in a set Z of positive measure $m(Z)$, almost every point p of Z has density 1, that is to say, if I_n is any sequence of intervals all with midpoint p, whose lengths $m(I_n) \to 0$ as n becomes infinite, then

$$\lim_{n\to\infty} m(I_n \cdot Z)/m(I_n) = 1$$

(Compare Saks [1].)

One could try to strengthen this theorem by proving a definite *rate* of this convergence to 1. For example, knowing that

$$d_n \equiv m(I_n \cdot Z) - m(I_n) = 0(m(I_n))$$

can one assert more; e.g., is it true that for every $\varepsilon > 0$ and almost every p,

$$d_n = 0(m(I_n))^{\frac{3}{2}-\varepsilon}?$$

One can obviously restrict the set Z to the class of G_δ-sets. If Z is an open set, then $d_n = 0$ for all n sufficiently large. For sets which are both G_δ- and F_σ-sets, such a strengthening of the density theorem might well be possible, even to the extent of estimating, for any k, the measure of the set of those points of Z for which

$$\lim_{n\to\infty} d_n/[m(I_n)]^{\frac{3}{2}} < k$$

A similar investigation is possible in the ergodic theory. Let E be a measure space (say, the Euclidean cube), $T(p)$ a measure-preserving, metrically-transitive transformation of E onto itself.

The ergodic theorem states that for almost all p, and for every set A of positive measure $m(A)$,

$$\lim_{N\to\infty} (N+1)^{-1} \sum_{\nu=0}^{N} \chi_A(T^\nu(p)) = m(A)$$

$\chi_A(p)$ being the characteristic function of A (Chapter VI, Sections 2, 4, 7) that is,

$$d_N = \sum_{\nu=0}^{N} \chi_A(T^\nu(p)) - m(A)\cdot(N+1) = o(N+1)$$

What can one say about $d_N/(N+1)^{\frac{1}{2}+\varepsilon}$? This limit is certainly zero for *some* measure-preserving transformations of the interval. For example, if $T(p)$ is the "shift" transformation of Section 6 of the present chapter, the "central limit theorem" of Laplace-Liapounoff applied to the "Bernoulli" case of the additive probability theory states that

$$d_N = 0(N+1)^{\frac{1}{2}+\varepsilon}$$

for every $\varepsilon > 0$. Can one assert the above relation for almost all $T(p)$ of the type postulated? (Compare Feller [1].)

9. A problem in the calculus of variations

Suppose two segments are given in the plane, each of length one. One is asked to move the first segment continuously, without changing its length to make it coincide at the end of the motion with the second given interval in such a way that the sum of the lengths of the two paths described by the end points should be a minimum. What is the general rule for this minimum motion? It is clear from the Euler-Lagrange equations of the variational problem that, locally, the motion will be a composition of rotations and translations. (The problem could be stated for two such intervals given in the 3-dimensional space.) One could require alternately that instead of the sum, the square root of the sum of the squares of the lengths described by the end-points should be minimum.

More generally, one could pose an analogous problem of the "most economical" motion given a geometrical object A and another B congruent to it and requiring the motion from A to B to be such that a sum or integral of the lengths of paths described by individual points be minimum. This bears a certain relation to the problem of Monge of "déblais et remblais," but differs from it in that we require here rigidity of A through the course of the motion. One motivation behind the consideration of such questions is that in certain problems of mechanics of continua, e.g., in hydrodynamics the motions that are most prevalent, are singled out by extremal principles not unlike the above; but of course operating in a space of infinitely many dimensions.

10. A problem on formal integration

Let $f_1(x), \ldots, f_n(x)$ be arbitrary continuous functions. Does there exist a rational function $R(x)$ constructed from the f_i by rational operations such that the "indefinite integral"

$$\int R(x)dx$$

is not again such a rational combination of the f_i and of functions obtained by superposing the f_i? (Mazur and Ulam).

We mention this rather special question as a small example of more general and interesting problems involving the algebraic properties of finite "formal analytic" algorithms. Compare the paper by L. Bieberbach [1] also of Ritt [1] and, more specifically for the problems of the above type, Kaczmarz-Turowicz [1].

11. Geometical properties of the set of all solutions of certain equations

The class of all solutions of a *linear* differential equation forms a linear manifold in function space. We think here of functions satisfying the equation and given boundary conditions as points in the space of all such continuous and differentiable functions. What can one say about the geometric properties of the set of solutions of a differential equation which is *quadratic* in the unknown function and its derivatives? If the equation is of the type $Q(y, y') = 0$, where Q is a positive-definite quadratic form, the set of solutions has the property: given any k solutions, no other solution lies inside the simplex formed in function space with these k solutions as vertices. In other words, the manifold of solutions lies on an intersection of convex ("ellipsoidal") surfaces. Can one assert that the manifold M of solutions of an algebraic differential equation is formed by the intersection of (possibly infinitely many) cylinders, each of which is erected over a finite dimensional algebraic manifold A_i? That is to say

$$M = (A_1 \times E_1) \cdot (A_2 \times E_2) \ldots (A_i \times E_i) \ldots$$

where the A_i are as above, E_i are linear, infinitely dimensional hyperspaces in the function space.

CHAPTER VII

Physical Systems

1. Generating functions and multiplicative systems

Let us consider a system of particles of t distinct kinds such that a particle of type i, upon transformation, has a given probability $p_1(i; j_1, \ldots, j_t)$ of producing $j_1 + \ldots + j_t$ new particles, j_i of type i. The probability of a particular population $(j_1, \ldots, j_t)$ in the kth generation of progeny from a single particle of type i is given by the coefficient $p_k(i; j_1, \ldots, j_t)$ of the product $x_1^{j_1} \ldots x_t^{j_t}$ in the kth iterate of the generating transformation $x' = G(x)$:

$$x_i' = g_i(x_1, \ldots, x_n) = \sum_j p_1(i;\ j_1, \ldots, j_t)\, x_1^{j_1}, \ldots, x_t^{j_t}$$

This theorem on iteration of generating functions allows one to calculate the first moments of the distributions by multiplication of the matrices whose terms are the first partial derivatives of the g_i evaluated at $x_1 = x_2 = \ldots = x_t = 1$. Higher moments can be computed also, but the expressions become increasingly complicated.

Unfortunately, it is very difficult to obtain precise information about the behavior of these coefficients except in the simplest cases. If only one type of particle is involved, one can explicitly study the iterates of a generating function of the type

$$g(x) = (ax + b)/(cx + d)$$

where the a, b, c, d are chosen so that the coefficients of the power series for $g(x)$ are nonnegative and have sum unity. The iterates are easy to compute since iteration of $g(x)$ leads to functions of the same form. Analogously the transformation

$$x_i' = \left(\sum a_{ij}x_j + b_i\right)/\left(\sum c_{ij}x_j + d_i\right)$$

may serve for systems of t types of particles. The question now arises whether there exist groups (or semi-groups) of specific transformations in the t-dimensional space with *more* parameters than are available for the mappings of the above type which, when developed in power series, can be considered as generating transformations (i.e., coefficients of g_i nonnegative and with sum $= 1$)? This would allow a greater variety of transformations whose iterates could be obtained in closed form.

The expected number of particles of type j in generation k from one particle of type i is given by the number in the ith row and jth column of the kth power J^k of the Jacobian of $G(x)$ at $x = (1, 1, \ldots, 1)$. The moment matrix J, when positive, has a unique positive eigenvector v of norm 1 (a theorem of Frobenius-Perron) and for "supercritical" systems, it may be shown that "almost all" genealogies terminate in death or approach, ratio-wise, the vector v, in the sense of a natural measure defined in the space of genealogies. This statement constitutes an analogue of the strong law of large numbers for the "case of Bernoulli" for multiplicative processes (cf. Everett, Ulam [2, 3]). These results, however, represent only a first step in the theory of such processes. It would be important to establish the analogue of the central limit theorem. What are the asymptotic properties of such a process if the basic probabilities $p_1(i; j_1, \ldots, j_t)$ are not constant in time, but change in a specified way either explicitly in time or dependent upon the existing population? If the limit of the product of the Jacobians of the generating transformations $G_1, G_2, \ldots$ exists, does the population approach the corresponding vector v or die out with probability 1?

The reader will find several problems on multiplicative systems in the papers referred to above. These are concerned with the iteration of generating transformations given by polynomials or power series with nonnegative coefficients in n variables and the number of particles of each type present in the kth generation. The problem of total progeny from the first to the kth generation and of systems with source may be so studied. The behavior of the coefficients of the kth iterate of such transformations has not been determined. (See also Bellman and Harris [1].)

1a. *Examples of mathematical problems suggested by biological schemata*

The combinatorial complexities and analytical questions suggested by problems of genetics and by problems of structure of organic materials present features of purely mathematical interest. The well-known work of Volterra [1] on the struggle for survival and the subsequent work of W. Feller [2] dealing with certain systems of quadratic total differential equations contained important results on special nonlinear systems.

We shall mention briefly some related problems, leading also to a system of infinitely many nonlinear differential equations also suggested by biological situations — of course treated in an extremely simplified and schematized way.

Imagine a system of N particles which reproduce in discrete units of time (generations). In the simplest version assume that the reproduction is asexual. Each of these particles possesses an index k denoting the number of its "characteristics." This number may increase in time due to mutations occurring at random at a fixed rate in the population. We assume an advantage in acquiring additional characteristics considered to be improvements leading to higher probability for survival of an individual. Specifically, there is a probability $\alpha \ll 1$ for each individual to acquire an additional improvement in the course of one generation — α^2 being the probability of acquiring in one generation two improvements, etc. Another constant β defines for an individual the differential advantage in its survival. In the simplest scheme one may assume that the differential advantage for survival is proportional to the number of these improvements—that is to say, if one particle has an index k and the other one $k+j$, then the relative chance of survival of the richer individual over the poorer one is proportional to j. In a numerical treatment of the problem one may assume the population to be always normalized to a constant number N. The first problem concerns the number x_i of the particles with i advantages as a function of time in its dependence on the two constants α and β. A simple system of equation would be:

$$\Delta x_k = -\alpha x_k - \alpha^2 x_k - \frac{x_k}{N} \sum_{j=k}^{l} (j-k)\beta x_j ,$$

$$\Delta x_{k+1} = -\alpha x_{k+1} - \alpha^2 x_{k+1} + \alpha x_k + \beta x_{k+1} - \frac{x_{k+1}}{N} \sum_{j=k}^{l} (j-k)\beta x_j ,$$

$$\Delta x_i = -\alpha x_i - \alpha^2 x_i + \alpha x_{i-1} + \alpha^2 x_{i-2} + (i-k)\beta x_i - \frac{x_i}{N} \sum_{j=k}^{l} (j-k)\beta x_j ,$$

$$i = k+2, \ldots, l-2,$$

$$\Delta x_{l-1} = -\alpha x_{l-1} + \alpha x_{l-2} + \alpha^2 x_{l-3} + (l-1-k)\beta x_{l-1} - \frac{x_{l-1}}{N} \sum_{j=k}^{l} (j-k)\beta x_j ,$$

$$\Delta x_l = \alpha x_{l-1} + \alpha^2 x_{l-2} + (l-k)\beta x_l - \frac{x_l}{N} \sum_{j=k}^{l} (j-k)\beta x_j$$

Here the species k is the first one of any importance that is present in the generation and l is the last. One may assume that in each generation the k is determined by making it equal to the index of the first number in the sequence, $\{x_i\}$, which is equal to or greater than 1 and the l is made equal to 2 plus the index of the last number of the sequence, which is $\geqq 1$. The numerical investigation made by C. Luehr and the author led to the following results. The solution of the system seems to approach a steady state in the following sense: the average index $\bar{\imath}$ of the population existing in one generation increases linearly with time, and the distribution of the i around the $\bar{\imath}$ appears to approximate the Gaussian. The parameters of this normal distribution depend in a simple fashion on the constants α and β.

A problem next in complexity, but still extremely simplified compared to the real biological situations involves bisexual reproduction — that is to say, production of particles by pairs of particles. The equations will now be *essentially* nonlinear. One should assume that the advantageous characteristics acquired by mutations can be transmitted from either parent to the offspring in the next generation. We may assume again that the total population is constant by normalizing it and, in the simplest case, assume only two kinds of "genes". The advantage for survival (or in the increase of the number of offspring) depends, say, on the sum of the number of the improvement genes of each kind. If the offspring can acquire independently the genes from each parent, obtaining certainly those that are present in both and each gene which is present in one parent only with the probability equal to $\frac{1}{2}$, the equations could be

$$x'_{ij} = y_{ij}/\sum y_{ij}$$

where

$$y_{ij} = -\alpha x_{ij} + \alpha/2(x_{i,j-1} + x_{i-1,j}) + \beta[i + j - \min(i + j)]x_{ij} + 1/N \sum\sum \gamma_i^{kl} \gamma_j^{mn} x_{km} x_{ln}$$

where the choice of

$$\gamma_i^{kl} = 1/2^{(k-l)} \binom{(k-l)}{i - \min(k, l)}$$

corresponds to our rule on the inheritance of the extra genes. A numerical study of this system was undertaken by P. Stein and the author. Again a steady state distribution seems to establish itself with the rather curious property that only a few species with adjoining indices co-exist in any one generation. The speed with which the average number of "improvements" increases is constant.

This mathematical formulation is still very naive and too simple as compared to the biological reality. The number of kinds of genes (or phenotypes) is much greater than 2. Also one should study such systems with more realistic rules, i.e., the distinction between dominant and recessive genes and the Mendelian properties. The mathematical properties of solutions of such systems seem to be akin to those of the phenomena encountered in the study of nonlinear systems describing a vibrating string, etc., mentioned in Chapter VII, Section 8.

P. Stein succeeded in proving the following:

Consider the quadratic transformation in n-space given by the quadratic terms alone:

$$x'_{ij} = \sum\sum \gamma_i^{kp} \gamma_j^{mn} x_{km} x_{pn}$$

assume

$$x_{ij} \geqq 0, \qquad \sum x_{ij} = 1.$$

About the γ's assume only:

$\gamma_i^{kl} > 0$ only for $k \leqq i \leqq l$; $\sum_{i=1}^{N} \gamma_i^{kl} = 1$ and $\sum_{i=k}^{l} i\, \gamma_i^{kl} = (k + l)/2.$

Under these assumptions the only fixed points of the transformation are of the form

$$x_{jj} = 1/(1+\tau)^2, \ x_{j+1,j} = \tau/(1+\tau)^2, \ x_{j+1,j+1} = \tau^2/(1+\tau)^2$$

all other x_{ij} being 0, τ is a parameter, $0 \leqq \tau \leqq \infty$. The iteration of the transformation, starting from any point, converges to one of these fixed points; the value of τ being computable from the initial conditions.

A still different class of problems, leading to a study of nonlinear (quadratic) transformations and their iterations originates from the following schema: imagine a large number of individuals (or particles) present in a given generation. Suppose these combine in pairs and produce, in the next generation, new particles, parents dying after procreating the new ones. Suppose the original particles are each one of N different types. A rule is now given for the type i ($i = 1, 2, \ldots, N$) produced by the mating of individuals of type j and k. In a random mating of particles the expected value of the fraction x'_i of particles of a given type in the next generation will be a quadratic function of the two fractions x_j and x_k. The equations would be

$$x'_i = \sum_{k,l=1}^{N} \gamma_i^{kl} x_k x_l \qquad i = 1, \ldots, N$$

where, if we assume that each pair produces exactly two new particles and specifically the rule of defining the type is that the γ's are either 0's or 1's, we may insist that for any index i not all γ_i^{kl} vanish and the system of equations really specializes to a form where each term in the product $(x_1 + x_2 + \ldots + x_N)^2$ will appear in exactly one row of the set of equations. For further simplification we may assume that the cross-products appear with the factor 2 (commutativity). So as an example we could have with $N = 4$,

$$\begin{aligned} x'_1 &= x_1^2 + x_2^2 + x_4^2 + 2x_1x_4 + 2x_2x_4 + 2x_3x_4 \\ x'_2 &= 2x_1x_3 + 2x_2x_3 \\ x'_3 &= 2x_1x_2 \\ x'_4 &= x_3^2 \end{aligned}$$

A study of all such transformations was made by P. Stein and the author in case $N = 3$. There are 97 nonequivalent possible "genetic" rules of this kind and all 97 corresponding transformations of this type were studied with regard to the properties of iterated sequences of each transformation. In some cases, starting with an arbitrary initial distribution (nondegenerate) one converges to a fixed point; that is to say, the ratios of the numbers of individuals of each type stabilize. In other cases the points seem to approach an oscillation between a finite number of fixed ratios. In every case, for $N = 3$ the first means, in time, of the ratios x_i exist for almost every point, it appears. The ergodic and asymptotic properties of the iterates of such transformations for $N > 3$ are unknown in general (see P. Stein and the author, [2]).

2. *Infinities in physics*

The simplest problems involving an actual infinity of particles in distributions of matter appear already in classical mechanics. A discussion of these will permit us to introduce more general schemes which may possibly be useful in future physical theories.

Strictly speaking, one has to consider a true infinity in the distribution of matter in all problems of the physics of continua. In the classical treatment, as usually given in textbooks of hydrodynamics and field theory, this is, however, not really essential, and in most theories serves merely as a convenient limiting model of *finite* systems enabling one to use the algorithms of the calculus. The usual introduction of the continuum leaves much to be discussed and examined critically. The derivation of the equations of motion for fluids, for example, runs somewhat as follows. One imagines a very large number N of particles, say with equal masses, constituting a net approximating the continuum which is to be studied. The forces between these particles are assumed to be given, and one writes the Lagrange equations for the motion of the N particles. The finite system of ordinary differential equations "becomes" in the limit $N = \infty$ one or several *partial* differential equations. The Newtonian laws of conservation of energy and momentum are seemingly correctly formulated for the limiting case of the continuum. There appears at once, however, at least one possible objection to the unrestricted validity of this formulation. For the very fact that the limiting equations imply tacitly the continuity and differentiability of the functions describing the motion of the continuum seems to impose various *constraints* on the possible motions of the approximating finite systems. Indeed, at any stage of the limiting process, it is quite conceivable for two neighboring particles to be moving in opposite directions with a relative velocity which need not tend to zero as N becomes infinite, whereas the continuity imposed on the solution of the limiting continuum excludes such a situation. There are, therefore, constraints on the class of possible motions which are not explicitly recognized. This means that a viscosity or other type of constraint must be introduced initially, singling out the "smooth" motions from the totality of all possible ones. In some cases, therefore, the usual differential equations of hydrodynamics may constitute a misleading description of the physical process.

On the other hand, the numerical solution of such a system of partial differential equations involves the use of a model of finitely many points approximating the continuum. The corresponding finite difference scheme must be carefully designed to insure not only that the distances between neighboring points are sufficiently small, but that various numerical stability conditions, e.g., so-called "Courant conditions," hold. This necessity shows again a number of implicit assumptions about the finite model approximating the mechanical system. The question whether the limit of the solutions of the approximating equations is in fact the solution of the limiting equation is, in the general case, open. The statement is probably false in the most general case.

Indeed, it may be that, in some future physical theories, the Euclidean continuum presently used as the exclusive model for distributions of matter will cease to be the *sole* convenient model for reality. It seems possible that in some cases spaces with the topology of the *Cantor* (perfect, nowhere dense) sets might serve to represent distributions of matter or energy.

3. *Motion of infinite systems, randomly distributed* *

We shall propose a few very simple mathematical questions illustrating the problems which will arise in the study of systems of this sort. They involve infinite assemblies of mass points with interactions assumed to exist between them. Such assemblies, while not finite, will *not* correspond to the continua which are presently employed in physical theories.

Suppose that we distribute a set of equal masses m on a line, on the integer points $0, \pm 1, \pm 2, \ldots$ by a probability scheme, placing a mass $m = 1$ on each point *n with probability* $\frac{1}{2}$ and leaving the point vacant with this probability. Since there is an obvious one-one correspondence between all possible initial distributions and real numbers on $(0, 1)$ in dyadic expansion, we can have a measure, e.g., the ordinary Lebesgue measure, in the set of distributions, and in what follows, the phrase "almost all" is understood in the sense of this measure. Let us further assume that between every two mass points there exists an attractive force inversely proportional to the square of their distance. Obviously the total force on each particle is well defined, the series of inverse squares of integers converging absolutely. We shall postulate also that colliding masses remain forever together, forming a single particle of mass equal to the sum of their masses, and that momentum is conserved under collision.

For times $t > 0$, the behavior of the system is described by an infinite system of Newtonian equations. We might remark parenthetically, that the various formulations of the principles of mechanics, all equivalent for finite systems, become in our case quite distinct. The total mass of our system being infinite, one has to use reformulations of the usual variational principles, or even the Lagrangian equations, to arrive at unequivocal statements.

The questions that arise concern the asymptotic behavior of such systems after long times. They can be put in this form. What is the measure of the set of initial distributions which will behave asymptotically in a specified fashion? Some such questions have

* The next three sections follow the author's article [2].

been answered. Many others suggest themselves (cf. Metropolis, Ulam [1]).

The situation becomes perhaps more "physically" interesting in two and three dimensions. The problem is still mathematically well defined. One can state that for almost every initial distribution it is true that, for every constituent mass point, the net force vector *exists* if the component forces are summed over successive spherical shells about the point; that is to say, the limit of these forces exists for almost all initial distributions. However, collision must now be understood in the sense of gravitational capture, i.e., the "colliding" points remain within some specified distance of each other for all time. Is it true that the series of forces on all masses of a distribution remain weakly convergent for all $t > 0$ if this condition obtains initially? The initial average density, in the obvious sense, of our system of particles, randomly distributed as they are at $t = 0$ is $= \frac{1}{2}$. Is it true that almost all distributions retain this density for all time?

In the case of one dimension there will be a tendency toward the formation of successively larger condensations. Is the same true in higher dimensions if by a condensation we understand a subsystem of points whose mutual distances all remain forever bounded? Will almost every distribution show a tendency toward the formation of "galaxies", "super-galaxies," etc.?

What force laws $F(r)$, or equivalently, what potential functions $V(r)$ have the property that, for almost every initial configuration, all forces remain well defined for all time under the ensuing motions, calculated from Newton's equations and assuming our conventions for collision?

Analogous but more difficult problems arise if we deal with countable systems, again randomly distributed initially but not restricted to the set of lattice points. The following general properties of our one-dimensional initially randomly distributed infinite systems are established very easily:

1. The masses appearing in the course of time will be unbounded. In other words, for almost every initial condition of our system there will exist for every M a time t such that a mass greater than M will appear after this time.

2. There will always exist single particles. In other words, for almost every system and for every t there will exist in the system points with unit mass.

3. The asymptotic density of our system remains constant and equal to the original density. We define the asymptotic density as follows. Consider the totality of particles contained in an interval from $-N$ to $+N$ and denote by M_N the total mass of all particles in this interval. The $\lim_{N=\infty} M_N/2N$ shall be called the asymptotic density if it exists. With our initial masses equal to 1, and the random placing of these masses on integer points, this limit (from Bernoulli's theorem) is equal to $\frac{1}{2}$. It is easy to see that this limit will exist and be equal to $\frac{1}{2}$ for all t. This is simply due to the fact that, given any t, the displacement of each particle will be bounded. If we take a sufficiently large interval, the flux across its ends will constitute an arbitrarily small fraction of the total number of particles and our assertion follows.

4. Arbitrarily large "holes" will appear in our system; that is, for almost every system and for all d there will exist a time t so that there will be infinitely many mass points separated by intervals larger than d. Moreover, for all greater times these long empty intervals will continue to exist.

These assertions are easy to prove in one dimension. In two or more dimensions, we shall *not* have, in general, collision between point masses and we would again have to define captures, that is to say, formation of double or multiple systems. The corresponding theorems on the existence of stable or semi-stable captures seem much harder to prove. An easier way to deal with an analogue of our system would be to give each point a finite size and then consider certain collisions as completely inelastic and leading to formations of larger masses. Property *2* and property *3* should then be easy to prove.

More interesting are the quantitative properties of such systems. For example, it would be interesting to calculate, even for systems in one dimension, the average mass of particles in our infinite collection at a given time t, and to determine the distribution of masses as a function of time. Another interesting question concerns the distribution of velocities of our particles as a function of time. (To define a meaningful average velocity one would have to introduce a cut-off in the distance between two particles, which approach each other, just before a collision; in the system consisting of mathematical points, these velocities become arbitrarily large during the collision process.) If we define a cut-off distance then the average velocity of our particles will have a meaning for all t and the question arises: what is the "temperature" of the system as a function of time?

In order to study these and other similar questions, a series of experiments were performed on a computing machine by John Pasta and the author. An attempt was made to imitate the infinite system by a finite one composed of a great number of masses placed on points of a regular subdivision of an interval with a decision, for placing or not placing mass points in successive positions, made by "throwing a die." In order to "approximate" an infinite system somewhat realistically, one has to imagine the two end points of the interval on which the points are located as coinciding; that is to say, we have a finite system of points on the circumference of a circle or a periodic structure. This attenuates the end effects. It is clear that such a finite system will imitate an infinite one only for a limited time. Given a finite system, it is certain that it will ultimately collapse to a single point, whereas in the infinite case, we saw that the asymptotic density will remain constant for all times. Therefore, in interpreting the results of a calculation made for a finite system, one has to carry them only up to a value T of time when the system still contains "many" points. In order to make a rigorous analysis, for the study of any given functional of the distribution of distances, masses, velocities, etc., of our system, one would have to give *a priori* inequalities for this T as a function of N and a $\delta > 0$; that is to say, with only a finite system of N points, we restrict T so that the functional of the system in which one is interested, computed up to time T, will differ in value by less than δ from the value of this functional for an infinite system (i.e., for almost all infinite systems). Many finite systems were computed (that is to say we started with many distributions of mass points given by our initial random procedure and in each case the subsequent motions were calculated with the hope of obtaining heuristic results on some functionals of such systems). The number of points initially taken was of the order of 1,000. Among the quantities calculated were the distribution of masses at any time t, the distribution of distances, etc. We shall indicate here merely very briefly some of the qualitative facts about them.

(a) The average mass of a particle appears to increase linearly with time.

(b) There is a suspicion, at least, that if one considers the distribution of masses of particles existing at time t in the units of the average mass at that time, this distribution tends to a fixed function. In other words a steady state may establish itself.

(c) A quantity which was called hierarchy was studied. This is defined as follows by induction. The original particles have by definition a hierarchy rank zero. When two particles of rank m and n collide, they form a particle whose hierarchy rank is equal to the greater of the two numbers m, n if $m \neq n$. In case $m = n$, the rank of the new particle is $m + 1$. This index gives an idea of the degree or hierarchy of conglomeration as distinguished from a mere increase of mass by accretion. The average hierarchy was increasing more slowly than the average mass, but presumably tends to infinity for an infinite system.

(d) The average kinetic energy was studied as a function of time. We have used in the computation a cut-off in the distance of approach in order to eliminate the arbitrarily high velocity just before collapse. The shape of this function is not known, but it is obvious that this average energy increases initially and then starts decreasing again, which, of course, is due to the fact that our system ultimately will end up as one big particle at rest. Presumably, our finite systems only imitate the infinite one up to the time when this average energy stops increasing. Nothing conclusive as yet, therefore, can be said about the change of "temperature" in time.

It should be pointed out here that the type of distribution (à la Bernoulli) which we have introduced in our numerical work could equally well be chosen differently. For example, we could assume that there is a fixed probability $\alpha \cdot dx$ of finding a unit mass in the interval dx. That is to say, the initial randomness could be chosen à la Poisson. It is easy to see that the convergence of the force on each particle of the system would be equally valid for almost all initial conditions with this set-up.

One could postulate a finite interval with infinitely many points of various masses distributed in this interval at time $t = 0$ by a given random process and then discuss the ensuing motions.

One could assume, of course, that not only the initial positions of our points on the line or in space are given at random, but also that there exist, at time $t = 0$, initial velocities of each of these points, given in a random fashion, say with a Maxwellian distribution.

4. *Infinite systems in equilibrium*

One of the qualitative differences in the behavior of finite and of infinite systems is that a system which has infinitely many mass points may exist in a state of static equilibrium, i.e., at rest, even when the forces between any two points are attractive. For example, assuming that any two points attract each other, the resultant force on each point may be still equal to 0: One can find a countably infinite set of point masses $m_1, m_2, \ldots$ and a set of initial positions $x_1, x_2, \ldots$ on the unit interval such that (a) $\sum_{i=1}^{\infty} m_i = 1$, (b) the masses attract each other according to the inverse square law, (c) the whole system is in static equilibrium, that is to say, the net force on each mass point exists and is zero. This equilibrium will not be stable, that is arbitrarily small displacements of such initial positions may lead to motions of the system which will make it collapse or, in any case, lead to configurations which with time increasing to infinity will differ more and more from the initial position.

It is easy to find distributions of such mass points with attracting forces like the above so that the initial motions will be expanding!

It would be of interest to find two "truly" three-dimensional infinite systems of points with total mass finite and so located that the net force on each point would be zero. One would like such systems to be of more than one dimension in the following sense. The set of all possible directions, i.e., angles between pairs of points should be dense on the circumference unit circle or — in a three-dimensional distribution of points — the set of directions between pairs of points should be dense on a unit sphere.

Can systems of the above sort be found which would be even dense on an interval or in a region of the plane or space?

5. *Random Cantor sets*

We have discussed an infinite but countable system of material points subject to forces acting on each of them. We shall now outline some models of mass distributions which combine a discrete character with certain properties of a continuum. One way to establish a rather simple distribution of this sort would be through the following process: Imagine a point with a mass equal to unity, located in the middle of the interval (0, 1). This mass point can now, either, with a probability p_1, remain forever in its original position, or, with a probability $p_2 = 1 - p_1$, split into two parts, each of mass $\frac{1}{2}$, which will be located in the positions $\frac{1}{4}$ and $\frac{3}{4}$, respectively. If the latter eventuality has occurred, we shall assume again that each of the two masses can, independently, just as before, either stay what it is, or each can independently split into two masses (equal $\frac{1}{4}$ each) which will be located at $\frac{1}{8}$ and $\frac{3}{8}$ for the first point, or $\frac{5}{8}$ and $\frac{7}{8}$ for the second point. This process is to continue indefinitely. We imagine that each of the points can split into two equal ones which will then be located to the left and right of it, with a probability p_2, or stay "dead" forever. We have thus a branching or multiplicative process which will define a possible distribution of masses on dyadically rational points of the interval. The process is defined by the two constants p_1, p_2. If p_2 should equal 1 and $p_1 = 0$, we would have a certainty of splitting every time and the process would lead to all rational binaries, each having, in the limit, mass zero. The closure of our set would be the full continuum of real numbers between zero and 1. If p_2 is less than 1, we shall get as a result of our continuing branching process a countable set of points whose closure will be, *with probability* 1, a nowhere dense set consisting of some isolated points and a perfect nowhere dense set of the kind defined by Cantor. There are three cases to distinguish, a subcritical, just critical, and supercritical system. In our simple set-up, these correspond to the cases $p_2 < \frac{1}{2}$, $=\frac{1}{2}$, $>\frac{1}{2}$, respectively. In the last case there is a finite probability for the process never ending and as a result, in addition to having a finite collection of points which are isolated, we shall obtain a perfect set on the interval as a closure of the "unending" part of the process. We have to make the sense in which we speak of a result "of one such process" more precise. What is meant, of course, is that one considers all possible outcomes of such a branching process. There exists a measure in the space of all possible branching processes defined in a rather natural way (cf. Everett, Ulam [3]). When we speak of the process leading with probability 1 to a set with given properties, we mean that the subset of the set of all processes with these properties has measure 1 in the space of all possible processes.

One can look upon the sets of points obtained by the above construction as describing "virtual positions of a physical object," or consider the space itself as being a collection of virtual symbols generated in such a way. These then will not, in general, form Euclidean continua. Neither will they consist of a discrete set of points. It is obvious that an analogous procedure can be effected in spaces of dimensions higher than one. One could, for example, perform our branching independently in all three dimensions, or one could imagine the following single process: A particle with mass 1, located at the center of a unit sphere splits with probability p_2 into two particles, each with mass $\frac{1}{2}$ and located on the opposite ends of an interval with length α_1. The direction of this interval can be obtained by a random process, say with isotropic distribution in space. Each of the ends with mass equal $\frac{1}{2}$ can again be subject to the same splitting possibility, say, with the same probability p_2 independently, and then split into two particles, each with mass equal $\frac{1}{4}$, located at the ends of an isotropically chosen interval of length α_2. If the process continues indefinitely, we shall obtain a three-dimensional analogue of the sets above. (It might be of interest to add here that it is by no means easy to determine the topological character of the resulting set. It is known since Antoine [1] that some perfect nowhere dense sets in 3-space are equivalent to sets located in one dimension under a homeomorphism of the whole three-dimensional space and others are not. The question which of the two is more likely under our process is not immediately answered.)

The above special way of constructing a space of symbols to correspond to a model of a physical situation is perhaps the simplest one of its kind and, so to say, applies to the "configuration space" of a particle. An analogous construction could be thought of as proceeding in the phase space. Not merely the positions, but the momenta or velocities of a particle could be generated by a process like the above, leading to a Cantor set structure of all "possible" values of the physical quantities in which we are interested. In the sequel we restrict ourselves to the configuration space alone. In addition, we shall confine ourselves here to classical, that is to say, nonrelativistic and nonquantum theoretical features of such models. Any pretended attempt to take the implications of such constructions for physical models more seriously would, of course, have to involve such constructions in Minkowski-Lorentz spaces.

We mentioned above some of the obvious properties of sets obtained by our branching process, e.g., that the sets will be, with probability 1, nowhere dense. To obtain more precise information about the nature of these sets seems difficult. Again exploratory numerical work seemed of some value and a series of computations on an electronic machine was undertaken as follows. Starting with the original point in the middle of the interval, the process was continued by the use of random numbers. That is to say, a great number of finite sets was determined, the process being stopped each time after a certain number of "generations." Given the probabilities p_1, p_2, one may obtain a number N of sets, produced by the splitting process, sufficient for a statistical study. If the number of generations k for the splitting process is kept constant, we shall obtain a variable number of points in the set. One can then compute the average value of any given functional of such a set. (The integration in the space of all possible branching processes is replaced by averaging on the N sets actually produced. One can easily justify such averaging as an approximation to integration for various functionals). The functionals studied were as follows.

1. Given a set of mass points at generation time k, one can compute its moment on inertia I. (The center of gravity of each system, as is obvious from its definition, is located in the original position at $\frac{1}{2}$, since our splitting preserves the center of mass.) If we average the value of the moment of inertia over all N sets which we have manufactured, we obtain an approximate value of the average moment of inertia of the infinite set.

2. Imagine that any two points of our set attract each other with a force proportional to the masses and inversely to the square of the distance. There arises the question: What is the value of the "gravitational self-energy"? Here again, if we compute this quantity for each of our N sets and take the average, we shall get an approximation to the average or expected value of this quantity in our infinite process. This value exists in the three-dimensional case.

In addition to obtaining the average, one can get a fair idea of the distribution of the value of this self-energy. Of course, in order to ascertain it with any accuracy, a very great number N of sets would have to be manufactured.

3. One could also ask about the mutual attraction of two systems of the above sort with n and m points located on an interval of length 1, but separated from each other, again assuming that any element of one set attracts any element of the other set with a force proportional to the product of the masses and inversely with the square of the distance.

The motivation for a calculation of quantities like self-energy is that it is *a priori* clear that the values of these will be less than the values for corresponding Euclidean continuum models and may have finite values in certain systems of our sort when these latter diverge. One may try to calculate a mass splitting process more general than the one outlined above. Namely, the probabilities for splitting which we have assumed to be constant from generation to generation and from point to point could be assumed to depend on the existence of another object of the above sort in proximity to our given system (cf. Chapter VIII Section 7). Indeed, if we assumed that a function $V(x)$ is given and the probability of multiplication is proportional to this function, we obtain a distribution u as a result of our splitting process which will obey an equation of the type of the Schrödinger equation

$$\Delta u + (E - V)u = 0$$

The process, as defined above, has the unsatisfactory feature that it still leaves mathematical points with finite masses. If one wants to insist on all mathematical points having masses zero, the following iteration of our procedure should be considered: Let L_ω denote the first passage to the limit of our process as defined above. We shall now iterate it in this fashion: each of the points which remain with finite mass will again be subject to a splitting procedure, say with the same probabilities, into two equal masses but this time located on the opposite ends of an interval shorter than the one in the first process by a fixed ratio R. For example, let $R = \frac{1}{100}$, then if it happened that the first mass point located at $\frac{1}{2}$ has not split during the first process, it will have a probability p_2 of splitting into two masses equal to $\frac{1}{2}$, but located at positions $(\frac{1}{2}) - (\frac{1}{100}) \cdot (\frac{1}{4})$ and $(\frac{1}{2}) + (\frac{1}{100}) \cdot (\frac{1}{4})$. Again, for example, if the point located at $\frac{3}{4}$ has not split in the first process, let it have probability p_2 of splitting into two masses each equal to $\frac{1}{4}$, and located at $(\frac{3}{4}) - (\frac{1}{100}) \cdot (\frac{1}{8})$ and $(\frac{3}{4}) + (\frac{1}{100}) \cdot (\frac{1}{8})$. If this continued, we should have a second limiting set $L_{2\omega}$. If, at the end of this procedure, some points still have finite masses, we continue, assuming now another still smaller ratio R for the splitting distance and repeat this splitting, obtaining $L_{3\omega}$ and so on. Now with probability 1, a sequence of these processes will lead to a Cantor-set of points so that at ordinal ω^2 all points will have mass zero and we shall obtain a distribution of density without finite values at any point.

Another generalization seems strongly suggested. In the processes described, the objects which are obtained by our branching are still algebraically of the nature of real numbers. It would seem that this is too special. The formalism of the new quantum theory would suggest considering each of the points "manufactured" as having spinlike properties. That is to say, there should be several *kinds*, $x_1, x_2, \ldots, x_n$ of these points. This is, in fact, already the case for the branching process as we described it in three dimensions.

In each of the processes of the above type one may consider a measure in the space of all their possible outcomes. This measure is defined in a natural way by defining it first for the set of all possible outcomes which have a specified appearance up to the kth generation for $k = 1, 2, 3, \ldots$. These special sets correspond to elementary intervals in the Lebesgue measure on the interval. Their measure is the probability of the specified special set of occurrences up to the kth generation. One may then extend it, in the usual way, to all sets of the Borel field over these elementary sets. This construction has been discussed for our simplest process L_ω and for an n-dimensional process, that is to say, particles of n types. Can it be generalized to our process of the type L_{ω^2}? Our set of all possible outcomes is, after closure, a continuum in which its own measure can be defined in a natural way. This allows one to integrate functionals of such a set: for example, of the points attracting each other according to a given law, one may define gravitational self-energy, etc.; then one can integrate the value of such a function of the set of all possible genealogies or outcomes in the sense of measure in this space. This way one obtains an average or expected value of such a quantity.

We repeat that mathematical studies of models of the kind suggested above may perhaps be of interest since the development of physical theories during the last few decades suggests a possibility of a continuing process of "atomization." The alternating vogues for a "field theory" and for the "elementary particles" points of view present, at a given time, either a topologically Euclidean continuum, the primitive mathematical entities being functions of continuous variables, or else a notion of fundamental particles the "insides" of which are not further analyzed. The interpretation of these ultimately small units of space evolve through stages: the atom becomes a nucleus surrounded by electrons, the nucleus in turn exhibits its inner components of nucleons, while at the present moment the protons and neutrons may be losing their right to a status of "particles" by exhibiting a definite substructure. All this, so to say, in the small -- whereas, in the direction of the distribution of the physical universe in the large, such iterative processes also seems to exist: stars appear in clusters, the clusters surround galaxies. There exist clusters of galaxies, i.e., super-galaxies and perhaps one might see an indication of an infinite hierarchy at the other end of the scale.

One might say that neither a quantization of fields nor a relativistic quantization of space-time arrests this tendency of the successive models of physical reality to replace elementary particles by systems of more elementary ones.

It may therefore be interesting to imagine such processes continuing infinitely and in particular to consider the cases where the subdivision of mass (or energy) goes on forever without leading necessarily in the limit to a real number system or a Euclidean continuum in which the field is defined. On the contrary, in general these limits will lead to schemas of mass distribution which will be Cantor-set like and have the topology of the p-adic rather than the real numbers. One has to repeat that models which could even remotely claim a physical interest would have to be constructed in space-time rather than in the ordinary space and in the phase space and not in configuration space alone. It seems, also, that a quantum theoretical viewpoint would have to be adopted implying in particular that the physical interaction between two elements at a given stage of the process is of a "shorter range" than that between elements of different stages.

One might note that such a geometrization of physics, if it ever takes place, would not consist of merely generalizing the Euclidean character of the distance in the large which is the content of Riemannian geometry. Since physical phenomena seem to radically change their properties in the small, it is clear that a differentiable metric will never bring out this feature. A more radical change, even that of local topology, as indicated above, would be necessary. The increasing recent evidence for the frequency of transmutation in properties of what at each stage was considered an elementary particle, and the multiplicative character of the phenomena in the small may make it amusing, at least for a mathematician, to consider models of this sort.

6. *Dynamical flow in phase space*

A dynamical system of n mass points may be represented, in the well-known fashion, by a single point in the $6n$-dimensional space. The totality of all possible initial positions and momenta will define a set of points in the $6n$-dimensional space — i.e., in the phase space — of the generalized coordinates and momenta. The change in time of the positions and velocities of the system defines then a measure-preserving flow of the phase space into itself. Mathematical work during the last few decades has brought a rigorous description of some properties of such transformations; in particular the ergodic theorems of von Neumann and G. D. Birkhoff (cf. Hopf [1], Khinchin [1]) provide a rigorous mathematical foundation for the ideas of statistical mechanics. One also knows that, among all possible measure-preserving continuous flows of a manifold into itself, the ergodic ones, i.e., those which are metrically transitive, form in a certain sense the general case (J. C. Oxtoby and S. Ulam [1]). These results show that almost all continuous measure-preserving flows possess the property that the time averages are equal to space averages, but this has not yet been proved in the general case of actual dynamical flows, that is to say, flows defined by differential equations with given Hamiltonians. In addition, there is a need for further information still of a qualitative kind about general properties of dynamical flows in phase space. One such property, first defined and studied by Poincaré, is that of "mixing". This involves the following property, broadly speaking: given any region A in phase space and any other region B, after a sufficiently long time the measure of the image of the region A which is contained in the region B will have approximately the value $m(A)$ multiplied by $m(B)$.

A quantity of great physical interest is the *rate* at which mixing proceeds. This problem is, of course, of interest in an actual hydrodynamical flow, of a fluid in three-dimensional space. Imagine that the initial appearance of the flow is quite regular, say almost laminar, but with a small irregularity superimposed initially as a perturbation. As time goes on, the motion may become more and more irregular and after sufficiently long time very complicated and turbulent.

If one should consider the actual fluid approximated by a large number k of points and their phase space, then the region B of the phase space whose points correspond to highly irregular velocity distributions of the given three-dimensional fluid undoubtedly occupies a very large proportion of the total volume of the entire phase space. It is important for the theory of turbulence to know something about the rate at which the set A of points, each of which corresponds to smooth or regular velocity distributions of the actual fluid tends to penetrate into the much larger region B of the phase space corresponding to turbulent motions. This question is of course a special case of the general problem: how to estimate for a dynamical system the rate of transition in the flow of phase space from one region into another.

The mathematical treatment of ergodic properties of mechanical systems is largely measure-theoretic. The definition of mixing is also of that nature. It would be interesting, in order to provide an additional description of the behavior of such systems, to introduce *metric* notions. The notion of measure is natural in phase space — Liouville's theorem refers to such measures established for any Lagrangian system of coordinates. The topology of the phase space is also given quite naturally by these coordinates. When it comes to defining a metric, that is to say distance between any two points of phase space, no such unique definition seems apparent. In order to treat both the coordinates and the momenta on a comparable basis one must reduce, say, the latter dimensionally to quantities in units of length since distance will involve additively both coordinates and momenta. If this were done, then the notion of mixing could be investigated, in addition to its measure-theoretic behavior, from a more geometric point of view. One could, for example, demand that, for a general flow, two fixed points at a given initial distance traverse space in such a manner that their distance as a function of time, averaged in time, tends to the average distance between *any* two points of phase space. For the latter to be defined, the phase space has to be bounded (compact). More strongly yet, one could define a "metric mixing" by requiring the analogue of the above property for turbulence on k-tuplets of points.*

Another illustration of the type of problem involving the estimation of the rates is the following: Consider a problem of three attracting bodies, say with equal masses, and with given total energy — the initial conditions being such that the kinetic energy is roughly equal to the potential energy of the system. For certain special initial conditions, the three points may, for all time, remain in a bounded portion of the configuration space. With other initial conditions one of the points may, after a certain time, start escaping from the remaining two (in the sense, for example, that its distance from the center of mass of the other two points will increase without limit). The first question is, given

* An analogue of the above property of mixing has been investigated in the finite case: Instead of a continuous space, we have a finite set of points. One may consider a transformation, i.e., a permutation of such a set of points. One may then consider, e.g., the average distance in the positions of two particular points. For a random permutation this turns out to be asymptotically equal to $n/3$ when n is the number of points in the space, for the obvious distance $|i - j|$ between any two points i, j of the set. The notion of a metric in a phase space would be useful for "geometrization" of various other physical properties of the flow. A useful metric will have to depend on the particular Hamiltonian describing the given system.

the total energy of the system, what is the volume of phase space corresponding to conditions which will guarantee its boundedness in space for all time. In case of unstable systems, what is their average lifetime, that is to say, time after which the escape in the above sense will take place? In the language of the phase space one is interested in the rate at which the volume, occupied by points corresponding to bounded configurations, goes into the volume corresponding to systems in which one of the three bodies is escaping from the other two.

Since hardly any quantitative work exists on problems of rates in the above sense, numerical work on computing machines may be of heuristic value in suggesting the kind of theorems which one might try to prove. The usual arguments on "relaxation" times are based on the size of volumes in the phase space alone and are in general lacking in rigor. In the last chapter we shall discuss some heuristic possibilities now open through computations on electronic machines

7. Some problems on electromagnetic fields

The mathematical investigations of Poincaré, Birkhoff, and their followers, of general qualitative and asymptotic properties of motions of dynamical systems and the corresponding ergodic theories are now in need of generalization to systems of infinitely many degrees of freedom. Recent developments in the new field of magneto-hydrodynamics — a study of the motion and behavior of ionized matter in electromagnetic fields — stimulate corresponding investigations on a more complicated topological level. In particular, a theory is needed to describe the topological properties of families of magnetic lines of force in space; in the most elementary case, these may be imagined as due to steady electric currents on a fixed system of curves (wires).

There are interesting problems arising in this connection. To start with the simplest: consider a single closed curve and a steady electric current stowing through it. Does there exist such a curve with at least one single line of magnetic force g dense in some region of space? (It is easy to find systems of wires such that magnetic lines will be, in general, dense on surfaces; statements sometimes made to the effect that a magnetic line is either closed or else goes from infinity to infinity are obviously incorrect. For example, if the currents are on the two curves: the z-axis and the circumference of the unit circle in the x, y plane, the lines of force will be dense on tori, except for a countable set of surfaces.)

It would be interesting to describe the system of magnetic lines of force due to a current flowing on a knotted (infinitely thin) wire. In particular, suppose the current flows through a "clover-leaf" knot. Does the system of lines of magnetic force in space surrounding the knot reflect topologically the "knottedness" of the curve? Such systems of curves may exhibit considerable topological complexity even when generated by currents flowing on straight lines, as shown by calculations on the properties of lines of force due to currents flowing on the three straight lines $x = 1$, $y = 0$; $y = 1$, $z = 0$; $z = 1$, $y = 1$.

It would be of interest to know the ergodic properties of a line of force (ergodic with respect to a particular parametrization); e.g., the behavior of the sequence of points one unit of length apart on a line of force. For example, are such sequences uniformly dense in some regions of space?

Answers to questions of this sort would, of course, give merely preliminary material towards a theory which should generalize the existing work on dynamical systems of n particles. Clearly, systems involving infinitely many degrees of freedom lead to consideration of phase spaces of infinitely many dimensions. In the absence of a definition of invariant measure (invariant under the flow defined by the Hamiltonians of the problem) the first attempts would be to approximate the infinitely many variables by systems with a finite number and then attempt a passage to the limit for the functionals in question. The additional difficulties due to the fact that Hamiltonians involve forces dependent on velocities (e.g., Lorentz forces) require, in the finite approximation, a study of nonholonomic systems.

8. Nonlinear problems

In some previous sections we have mentioned a few questions referring to nonlinear functional equations. A short discussion will be given below of a specific problem of some physical interest. This problem was investigated through computations on an electronic machine and was meant to be the first of a systematic sequence of such problems, increasing in complexity, involving non-linear equations. This work was planned by E. Fermi, J. Pasta, and the writer [1] — a preliminary account of the motivation for this program and the part of it already performed will be given below.

The first problem involved the case where the nonlinear terms were small compared to the main, linear ones and could be treated as perturbations for the initial range of the parameters. The first such system studied was that of a vibrating string with fixed ends, with forces between elements which contained in addition to the usual elastic linear terms, nonlinear ones, for example, terms quadratic in the displacements. The linear problem has the well-known periodic solutions; the presence of the additional terms provides, in time, a "mixing" of the states which the linear system would possess. The plan was to follow the motion for a long time and to obtain the *speed* with which such a system attained a statistical equilibrium, a state in which modes of vibration of all kinds are present. The calculations involved an approximation to the continuous string, by replacing it with a finite number of particles, and were performed during the summer of 1953 on an electronic computer in Los Alamos (the "Maniac," one of the first such machines built.)

This problem was to serve as the very simplest one. The ultimate aim was to discuss problems with more independent variables in the hope of obtaining material which would suggest some general features of behavior of systems with an infinite number of degrees of freedom, with nonlinear interaction terms as they occur between the oscillators in quantum theory or between the degrees of freedom in an electromagnetic field or a meson field. The mathematical possibilities concerning what might be called *quasi-states* (in the sense of Chapter VI, Section 3) were discussed with Fermi in this connection.

For the numerical work the continuum is replaced by a finite number of points (at most 64 in our actual computation) so that the partial differential equations defining the motion are replaced by a finite number of total differential equations. We have, therefore, a dynamical system of 64 particles. If x_i denotes the displacement of the ith point from its original position and α denotes the coefficient of a quadratic term in the force between the neighboring mass points (β that of the cubic term in other problems), the equations were

$$(1)\quad \ddot{x}_i = (x_{i+1} + x_{i-1} - 2x_i) + \alpha[(x_{i+1} - x_i)^2 - (x_i - x_{i-1})^2], \quad i = 1, 2, \ldots, 64$$

or

$$(2)\quad \ddot{x}_i = (x_{i+1} + x_{i-1} - 2x_i) + \beta[(x_{i+1} - x_i)^3 - (x_i - x_{i-1})^3], \quad i = 1, 2, \ldots, 64$$

The coefficients α and β were chosen so that, even at the maximum displacement time, the nonlinear part of the force was small compared to the linear term (e.g., of the order of one tenth of it). If we let the number of particles become infinite in the limit, we would obtain a partial differential equation containing, in addition to the terms in the usual wave equation, nonlinear terms of a complicated nature.

Another case studied later was:

$$\ddot{x}_i = A(x_{i+1} - x_i) + B(x_i - x_{i-1}) + C$$

where the parameters A, B, C are not constant but assumed different values depending on whether or not the quantities in parentheses were less than or greater than a certain value fixed in advance. This prescription amounts to assuming the force as a broken linear function of the displacement. This broken linear function may imitate to some extent a cubic dependence.

The solution of the corresponding linear problem is a periodic vibration of the string. If the initial position of the string is, say, a single sine wave, the string will oscillate in this mode indefinitely.

Starting with the string in a simple configuration, for example, in the first mode (or in other problems, starting with a combination of a few low modes), the purpose of our computations was to see how, due to nonlinear forces perturbing the periodic linear solution, the string would assume more and more complicated shapes, and, for t tending to infinity, how the total energy of the string would come to be distributed in all the Fourier modes. In order to observe this, the shape of the string, that is to say, x as a function of i and the total potential plus kinetic energy were analyzed periodically in Fourier series. This amounts to a Lagrangian change of variables: instead of the original $\dot{x}_i$ and x_i, $i = 1, 2, \ldots, 64$, we may introduce $\dot{a}_k$ and a_k, $k = 1, \ldots, 64$, where

$$a_k = \sum x_i \sin \frac{ik\pi}{64} \tag{4}$$

The sum of kinetic and potential energies in the problem with a quadratic force is

$$E_{\text{kin}} + E_{\text{pot}} = \tfrac{1}{2} \sum_i \left(\dot{x}_i^2 + (x_{i+1} - x_i)^2 + (x_i - x_{i-1})^2 \right)$$

or

$$E_{\text{kin}} + E_{\text{pot}} = \sum_k \tfrac{1}{2} \dot{a}_k^2 + 2a_k^2 \sin^2 \frac{\pi k}{128} \tag{5}$$

if we neglect the contributions to potential energy from the quadratic terms in the force which amounted in our case to at most a few per cent of the total.

In each problem reported here, the system was started with zero velocities at time $t = 0$. The length of a computational time cycle used varied somewhat from problem to problem. What corresponded in the linear problem to a full simple period of the motion was divided into a large number of time cycles (up to 500) in the computation. This is necessary for accuracy. Each problem ran through many "would-be periods" of the linear problem, so the number of time cycles in each computation ran to many thousands. That is to say, the number of swings of the string was of the order of several hundred, if by a swing we understand the period of the initial configuration in the corresponding linear problem. The distribution of energy in the Fourier modes (5) was noted after every few hundred of the computation cycles. The accuracy of the numerical work was checked by the constancy of the quantity representing the total energy. In some cases, for checking purposes, the corresponding linear problems were run, and these behaved correctly within one per cent or so, even after 10,000 or more cycles.

The calculation of the motion was performed in the x_i variables; after every few hundred cycles the quantities referring to the a_k variables were computed by the above formulae. It should be noted here that the calculation of the motion could be performed directly in the a_k and $\dot{a}_k$. The formulae, however, become unwieldy and the computation would take longer time. The computation in the a_k variables could have been more instructive for the purpose of observing directly the interaction between the a_k's.

One should now say here that the results of our computations showed features which were, from the beginning, surprising. Instead of a gradual, continuous flow of energy from the first mode to the higher modes, all of the problems show an entirely different behavior. Starting in one problem with a quadratic force and a pure sine wave as the initial position of the string, we indeed observed initially a gradual increase of energy in the higher modes as predicted e.g., by Rayleigh in a perturbation analysis. In the first problem, mode 2 starts increasing first, followed by mode 3, and so on. Later on, however, this gradual sharing of energy among successive modes ceases. Instead, it is one or the other mode that predominates. For example, mode 2 decides, as it were, to increase rather rapidly at the cost of all other modes and becomes predominant. At one time, it has more energy than all the others put together! Then mode 3 undertakes this role. It is only the first few modes which exchange energy among themselves, and they do this in a rather regular fashion. Finally, at a later time, mode 1 comes back to within one per cent of its initial value so that the system seems to be almost periodic. All our problems have at least this one feature in common. Instead of a gradual increase of all the higher modes, the energy is exchanged, essentially, among only a certain few. It is, therefore, very hard to observe the rate of "thermalization" or mixing in our problem which was the initial purpose of the calculation.

Looking at the problem from the point of view of statistical mechanics, the situation could be described as follows: the phase space required by a point representing our entire system has a great number of dimensions. Only a very small part of its volume is represented by the regions where only one or a few out of all possible Fourier modes have divided among themselves the bulk of the total energy. If our system, with nonlinear forces acting between the neighboring points, should behave as a good example of a transformation of the phase space which is ergodic or metrically transitive, then the trajectory of almost every point should be everywhere dense in the whole phase space. With overwhelming probability this should also be true of the point which at time $t = 0$ represents our initial configuration, and this point should spend most of its time in regions corresponding to the equipartition of energy among various degrees of freedom. From the results, this seems hardly the case. The ergodic sojourn times in certain subsets of the phase space may show a tendency to approach limits as guaranteed by the ergodic theorem. These limits, however, do not seem to correspond to equipartition even in the time average. Certainly, there seems to be very little, if any, tendency towards equipartition of energy among all degrees of freedom at any given time. In other words, these systems certainly do not show much mixing. Rather, the behavior of such physical systems suggests the existence of "quasi-states" in quadratic problems, discussed mathematically in Chapter VI, section 3.

CHAPTER VIII

Computing Machines as a Heuristic Aid

1. Introduction

In recent years a considerable amount of interesting heuristic mathematical work has been made possible by the development of fast electronic computing machines. Much more of this, primarily in mathematical physics, but also in combinatorial analysis and number theory can be expected in the near future. A knowledge of the behavior of special cases has always been useful and a background of experimental results has played an important role, notably in number theory; this was emphasized by Gauss himself. Means of greatly enlarging this background are now within reach and at a cost in time spent in such experimentations which has been enormously reduced in comparison with hand calculations. In the following sections we shall give a few examples of work already performed together with a rather arbitrary illustrative selection of possible future studies.

It might be thought that the interesting questions of combinatorial analysis lead immediately to numbers so large that memory limitations prohibit a useful study of such problems even on the present fast computing machines. Such is not always the case. Some problems have this general character: the number of elementary operations (e.g., additions and multiplications) may indeed be very large, i.e., of the order of many tens of millions, but the "memory" involved is only moderate, of the orders of hundreds to tens of thousands of positions. Indeed the rather widespread "fear" of enormously large numbers appearing in the study of combinatorial problems is often a misapprehension. While it is true that the fundamental combinatorial functions like 2^n or $n!$ increase rapidly with n, it seems that in many cases the formulae expressing the asymptotic behavior of certain interesting combinatorial functions are not only suggested but already rather well illustrated for small n. For example, the density of primes less than ten thousand gives a fair picture of the asymptotic density. Speaking loosely, many asymptotic formulae become valid within a small error for moderate values of n.

In such cases this may be due to a combinatorial fact which we shall vaguely describe as follows: Let Z be a subset of the set of all integers such that in its definition at most l operations of Boolean algebra and elementary arithmetical operations are used; the number of quantifiers and the operations of formation of direct product used in the definition is p. Let $f(N)$ be the number of integers in Z not exceeding N and suppose that $r = \lim_{n\to\infty} f(N)/N$ exists. In certain cases, given $\varepsilon > 0$, an estimate on $N(\varepsilon)$ could perhaps be given such that $|(f(N)/N)-r| < \varepsilon$ for $N > N(\varepsilon)$, *a priori* in terms of ε, l, and p. This will depend most sensitively upon p. Such an estimate cannot be given *a priori*, valid for *all* general recursive functions. This follows from the results of Gödel. In many practical cases, however, when p is small, one might expect some result like $N(\varepsilon) \sim (l^p/\varepsilon)$.

2. Some combinatorial examples

We shall mention at first a few questions of enumeration.

Given the set E of integers from 1 to n and two permutations S_1 and S_2 of this set selected at random, what is the expected number σ of elements in the group generated by these two permutations?

Analogously, let T_1 and T_2 be two random transformations of the set E into itself; what is the expected number τ of elements in the generated semi-group? Is $\sigma > \tau$?

These numbers of course do not exceed $n!$ and n^n, respectively. What is the relation between the sizes of τ and σ? If we postulate three or more permutations and the same number of transformations, does the relation between τ and σ reverse?

In any case, one can investigate such questions empirically for small n, on the machine. This can be done by using only a moderate number, say 100, of randomly selected pairs of permutations and even though 100 is very small compared to $(n!)^2$ even in cases where n equals, say, 7; the average values of σ and τ could be approximated well with a good probability from such sampling.

To give another example of such a sampling or "Monte Carlo" approach to combinatorial problems, one could mention the problem of the walk of the knight on the chess board. This consists of finding a succession of moves of the knight which would cover each position on the board exactly once. Euler found many solutions, but no general procedure for generating such walks is known even for the case of the ordinary 8×8 board. Certain practical recipes, e.g., the one (given by Euler) of moving into places which are already almost surrounded, seem to be successful. It is quite obvious that a computing machine can perform attempts to find solutions by trial and error even embodying rules like the above and also more complicated ones which involve retracting the last 1, 2, . . ., k moves in cases when blind alleys develop. The memory requirements in such a problem are small and it would appear that the advantage of the speed of the machine is considerable in situations of this sort where the advantage of a visual survey available to the human brain is very small. The number of attempted walks of the knight which can be performed on the machine in the course of a few hours is of the order of 10^6. A result of such statistics would be the proportion of successes in trials involving given rules. M. Wells has obtained results on the Los Alamos computing machines for this problem, comparing the 8×8 case with other $n \times n$ cases.

Heuristic results can be obtained on combinatorial aspects of an algebra of quantifiers in models of projective algebras (cf. Chapter I, Section 5). If E is an infinite set, it is known that starting with two sets A and B in $E \times E$ one can obtain by the operations of projective algebra an infinity of new sets. If E is finite, consisting of n points, the question arises of finding a basis for the class of all subsets of $E \times E$. Of course the n^2 individual points form such a basis. But what is the minimal number of sets of a basis? The projective algebras generated by two fixed randomly selected subsets A and B of $E \times E$ are of some interest. What is the expected number of elements in the projective algebra thus generated? This could be investigated on a computing machine for all small n, perhaps up to 7 or so. The obvious symmetry and conditions of the problem reduce the number of cases to be investigated from $(2^{(n^2)})^2$ to something like 2^{2n} for small values of n.

Still another example: We have mentioned before the problem of approximating a $1-1$ transformation $T: x' = f(x, y), y' = g(x, y)$ of a product space into itself by a composition of transformations of the two types U: *(1)* $x' = x$, $y' = k(x, y)$ and *(2)* $x' = y$, $y' = x$. The finite analogue would be to have the space E consist of a small number n of points, say $n = 8$ and try to obtain all transformations of the space $E \times E$ by composition of the transformations of the above type. A similar question for the three-dimensional space, i.e., E^3, involving transformations of the type indicated in Chapter 4, Section 2a could be settled by work on computing machines for small values of n.

The combinatorial illustrative lemma of Chapter 2, Secton 1 has been verified by P. Kelly for $n \leq 5$ by an examination of all cases. Probably a fast computer could examine all cases for $n = 8$ in a relatively short time.

The question whether the isometry of two spaces A^2 and B^2 implies the isometry of A and B where the latter are finite sets, with a metric such that the distance assumes only the values 1 or 2, has also been studied by Kelly [1] who verified it for small n. This problem also lends itself to investigation for somewhat larger n by the electronic computer.

3. *Some experiments on finite games*

As is well known, a game between two players can be considered as taking place according to the following schema: A combinatorial graph is defined starting with one vertex and branching successively into a set of segments. That is to say, we have n_1 segments issuing from the initial point corresponding to the choice of any one of n_1 possible moves by the first player. Then from the end of *each* of these segments we may assume there exist n_2 new segments (without loss of generality we may assume that this number is the same at the end of every initial segment). The choice among these is left to the second player. After that we have n_3 vertices issuing from the ends of these, and so on. In a finite game the game terminates after k moves. In the set of $N = n_1 \cdot n_2 \ldots n_k$ final vertices a subset W of these is defined as the winning positions. There might be given another subset D of points corresponding to the drawn games. Most of the games like checkers, chess, etc. can be thus formulated, although one should remember that what we consider as a position of the game may be a rather complicated structure involving the sequence of all previous positions, etc.

A winning strategy (for the first player) means the existence of an integer i_1 so that for every choice of i_2 there will exist a choice i_3 so that for every choice of i_4 and so on, the element constructed belongs to the set W. On the computing machines in Los Alamos some extremely simple cases of games of this sort were studied (P. Stein and S. Ulam [1], and W. Walden [1]). These experiments concerned comparisons between players of different strength, e.g., the computation of the winning chances for a player employing a perfect strategy versus one who "sees" only a limited number of moves ahead, when the set W was given at random.

Recently the writer has tried to find similar game formulations in some other mathematical situations. So, for example, the two players A and B, instead of selecting intervals as in the example above, could construct together, in turn, two permutations of the set of N integers. A selects n_1, B selects n_2, etc., up to n_N, obtaining a permutation E_1. After that they similarly construct a permutation E_2. If E_1 and E_2 generate the group S_N of all permutations on N integers it is the player A who wins, otherwise B wins.

To give a "topological" example, suppose a cube is subdivided into a great many smaller cubes and one of their vertices is given as a starting point. Two players, A and B, play as follows: The first player selects an edge of a small cube issuing from the starting point. The second player has to continue by selecting *another* edge whose beginning coincides with the end of the segment just given, otherwise arbitrary. Together the two players establish a path. This path must intersect itself at some time. If the closed curve thus obtained for the first time is knotted, let us say that the first player wins. In other words, A tries to close the curve so that it will have a knot in it; the second player tries to unwind the arc before it closes. Obviously some knowledge of topology is desired to plan a good strategy in this game! To study a game of this sort, preliminary explorations would be helpful on the computing machine which could try out tentative strategies.

The examples given are intended merely as arbitrary illustrations of the general possibility of converting special combinatorial studies to game situations. The desirable corresponding theorems would assert the existence of winning strategies. The writer found it rather amusing to consider how one can "gamize" various mathematical situations (or perhaps the verb should be "paizise" from the Greek word *παιζειν*, to play).

4. *Lucky numbers*

The sieve of Eratosthenes yields the primes among the natural integers. One can think of variations in the definitions of sieves. The following procedure was investigated numerically on the electronic computing machine in Los Alamos (Gardiner *et al.* [1]). In the sequence of all integers we strike out every second one, that is to say, all the even numbers. The first number remaining (apart from 1, which will not be counted) is 3. We shall now strike out every *third* integer, counting only the remaining ones, that is to say, this time we will strike out the integers 5, 11, 17, etc. In the remaining sequence the first number not used before is 7. Therefore we shall strike out every seventh number, counting among the remaining ones again. This will eliminate 19, etc. We proceed in this manner *ad infinitum*. The numbers that remain after the sequence of these procedures we might call, for example, lucky numbers. They are: 1, 3, 7, 9, 13, etc.

It turns out that many asymptotic properties of the prime number sequence are shared by the lucky numbers. Thus, for example, their asymptotic density is $1/ln(n)$. The numbers of twin primes and of twin luckies exhibit a remarkable similarity up to the integer $n = 100,000$, the range which we have investigated on the machine. The number of adjacent primes differing by 4 or by 6, 8, etc. is, in this range, very similar to the corresponding number of adjacent luckies. It also happens that within the range investigated every even number is a sum of two lucky numbers. The lucky numbers of the form $4n+1$ and $4n+3$ seem to be equally distributed, etc.

Another sieve would be a random one, that is to say, we retain the integer n with the probability $1/ln(n)$. In the class of sequences thus obtained one could try to prove theorems for almost every such sequence — statements about their distribution in detail such as representation of integers by sums of these, distribution of gaps between adjacent ones, etc. (See Erdös, Jabotinsky [1]; Chowla, Hawkins, forthcoming papers.)

It would perhaps not be without interest for number theory to randomize similarly procedures leading to primes in the quadratic fields. Analogously one could randomize the sequences of squares of integers or cubes of integers, etc., and consider the analogue of Waring's theorem. An experimental study on computing machines would quickly suggest statements likely to be true.

5. *Remarks on computations in mathematical physics*

The next few pages will deal with a few simple examples of physical problems which can be usefully studied through computations on modern machines. It is obvious that computational work on a large scale is immensely helpful in testing out theories on physical models; problems in general relativity, for example, lead to equations so complex that in most cases it will be numerical work which will test the validity of the models. Problems of constitution of the stars with assumed models of energy generation can be solved in special cases by such computations. One could name problems in atomic and molecular physics, etc., etc., where numerical checks of existing theories will be increasingly available. The mathematical complexity of theoretical physics has been steadily increasing during the last decades. This is so, not only because of its increased scope in embracing more complicated systems, but also because of changes in the very foundations of physics: the entities once treated as "points" (e.g., the molecules, the atoms, the nucleons themselves exhibit now a combinatorial substructure). In a number of problems one considers interactions between a considerable number of points (or fields with many degrees of freedom), that is, one deals with the n-body problem for sizeable values of n, but not large enough so that they can be handled by statistical or thermodynamical methods. There is little hope of obtaining solutions with the method of classical analysis alone, so that heuristic investigations involving an examination of many special cases seems indicated for the present.

The small selection of calculations described in the sequel is intended to illustrate some of the instances where understanding of the problems might be furthered by computation of special cases; the electronic computing machines enable one to make, as it were, "experiments in theory" by computations, either of the classical or of the "Monte Carlo" type computations (Section 8 contains a short account of the latter).

6. *Examples from electromagnetism*

Suppose a system of steady currents is given flowing through one or several curves fixed in space. One is interested in the field of the magnetic lines of force due to these curves. The calculation of these lines involves a numerical integration of a simple system of ordinary differential equations. The force field, at every point, is determined by curvilinear integrals (the law of Biot-Savart). We mentioned before some questions concerning the behavior in the large of such systems of curves — their topological and ergodic properties, for instance, in Chapter VII, Section 7, the problem of the field due to a simple cloverleaf knot. This can be realized for computational purposes by a space polygon with six sides. To begin with, simpler configurations were calculated on the electronic computer in Los Alamos. The first system studied was one given by currents flowing through two infinite straight lines, one the y-axis, the other wire parallel to the x-axis and intersecting the z-axis at the point $(0, 0, d)$. Starting with an arbitrary point in space, the magnetic line of force is traced out by computing the position of a sequence of points in the direction of the force at the preceding point. Taking the points sufficiently close and correcting for the changes in the field, one obtains an idea of the properties of the lines of force (tangent everywhere to the force vector). In order to obtain information about some qualitative properties of such lines in the large, the machine computed the number of times a line of force winds around each of the given straight lines. This is done by computing the Gauss integral. Some lines of force continue looping one of the wires; some others loop them both at once. Also the machine will compute the number of times a line of force crosses the surface of a given sphere. Other simple topological invariants, e.g., the Kronecker index of a mapping of directions onto a given sphere can be calculated and printed as a result of the computation. The next system involved three straight wires all skew to each other. These computations were considered as a preparation for the computation of a field due to the six-sided knotted polygon. The computation of such integral-valued invariants will not be affected by the accumulation of errors, unavoidably present in a numerical computation (see, e.g., a paper by Borsuk and the author [2] on ε-invariance of mappings). Roughly speaking, this is due to the fact that the computation may be so arranged that the field of error vectors is rotation free in the large.

7. *The Schrödinger equation*

The time independent Schrödinger equation

$$\Delta\psi + (E - V)\psi = 0$$

can be put into the form

$$\partial u/\partial t = \Delta u - Vu$$

through a substitution $u = \psi e^{-Et}$ by introducing a "time" parameter t. Both the lowest eigenvalue E and its characteristic function $\psi(x, y, z)$ can be obtained approximately by the following calculation: We imagine a population of fictitious point particles which multiply at the point x, y, z according to the given function $V(x, y, z)$ and which diffuse randomly through space, that is to say, perform a random walk as indicated in the equation by the Laplacian. The "game" played on the machine is then to process a great number of particles, each of which diffuses in space and at the same time multiplies its weight according to the given potential function. Processing a great number of the particles one obtains a population which asymptotically will tend to one possessing a density ψe^{Et}. The boundary conditions can be properly handled; the usual conditions at infinity are automatically fulfilled in such a model, and if they are prescribed on finite surfaces they may be satisfied by introducing suitable rules, e.g., absorption of particles on such surfaces. The procedure will give the lowest eigenvalue and its characteristic function. To obtain the successive eigenvalues and their characteristic functions becomes increasingly difficult by such methods. For example, to obtain the second characteristic function one would have to start with a population consisting of two kinds of particles ("black" and "red") corresponding to the positive and negative values of the function. One would start with a distribution orthogonal to the first eigenfunction (assuming it already determined) and then allow this population to undergo a diffusion-multiplication process with a suitable convention of mutual annihilation of the black and red particles whenever they appear at the same point (really a zone in space). The asymptotic population would then represent our second eigenfunction.

Such a statistical approach possesses an advantage in problems where a greater number of space dimensions is involved and the given potential function is complicated so that the usual methods of approximation become prohibitively lengthy. The main point, however, is as follows. One may only be interested in the value of certain given functionals $F_i(\psi)$, $i = 1, \ldots, k$, of the unknown solution ψ. In a classical approach one has to "know" ψ itself or at least know it at a sufficient number of points. Through a statistical approach like the above such functionals might be directly computed through sampling on those values of ψ which are obtained in the calculation. This is a general feature of the so-called Monte Carlo method, which we shall describe very briefly in the next section.

8. Monte Carlo methods

The sampling procedures in question might be said to consist in the "physical" production of models of combinatorial situations in mathematical problems, or of the playing of games with distributions of "particles" produced on computing machines, which distributions will constitute approximate solutions of physical problems.

Perhaps the simplest example is given by the problem of evaluation of a definite multiple integral. Suppose we want to find the volume of a region R contained in the unit cube in n dimensions defined by a given number of inequalities, that is to say, evaluating

$$\iint_{(R)} \cdots \int dx_1 \ldots dx_n ,$$

where the region R is defined by inequalities

$$\varphi_1(x_1 \ldots x_n) < 0, \ \varphi_2(x_1 \ldots x_n) < 0, \ldots, \ \varphi_k(x_1 \ldots x_n) < 0$$

The elementary numerical procedures would consist essentially of counting the lattice points of a subdivision of a cube in the space of n dimensions containing R and ascertaining the proportion of those satisfying the given inequalities. This is impossible in practice when n is moderately large, sometimes even when n is greater than 3, and the subdivision on each axis involves something like 100 points. The obvious statistical procedure would be instead to take a number of points *at random* in the n-dimensional cube, with uniform probability in this cube and count among those selected the points satisfying the given inequalities. Obviously for a precision of a per cent or so one needs to take only a number of points (of the order of 10^4) much smaller than the total number of all lattice points in a 100^n set. This is clear in case we know *a priori* that the volume of the region in question is not very small compared to that of the containing cube. In such a case special procedures are indicated. To illustrate them we shall take two examples of problems suggesting less simple-minded sampling.

Suppose one is interested in calculating the probability of a successful outcome of a solitaire (game of cards for one) where we assume that skill plays no role: a pure game of chance. In cases where the game has a very small probability of success most actual plays will end in failure and only an upper limit would be obtained for this probability. To get some idea of a lower limit > 0, suppose one obtains (still obeying the rules of the game) in a significant proportion α of tries, a situation where only rather few, let us say 10, cards remain to be played. It might be useful to try, in such cases, various permutations of the remaining 10 cards and continue the play from there. By examining a large number of such 10 factorial permutations one could obtain the number β expressing the chance that starting with one of the previous positions we obtain a win. A reasonable guess for the chance of success from the beginning without "cheating" would then be of the order of $\alpha \cdot \beta$. The idea is to decompose the problem into two or more stages in order to economize in the number of experiments performed for our sampling.

A similar problem is encountered in the transmission of particles (neutrons or gamma rays) traversing a series of materials serving as a shield for these particles. One is interested in cases where only an extremely small fraction of such particles which scatter at random and possibly produce other particles of this type, manage to traverse the whole shield. Again it is clear that in trying models of such random walks on the machine one could waste most of the experiments. The obvious thing to do is again to decompose the geometry of the shield or, more generally, the history of the process, into two or more stages, ascertain the proportion of traversals of the first stage; then starting with a typical situation after the traversal, try again the random walks from there on, etc.

An analogue of the Laplace-Liapounoff theorem is needed for estimates of errors in multiplicative and diffusive processes.

The computing machines are well suited for experiments of the above sort on the qualitative behavior of solutions of linear partial differential equations. Broadly speaking, one tries to change them to a form of integro-differential equation describing a diffusive and branching process, and uses the digital computer as an analogy machine as it were. Equations which are quadratic or of higher order in the unknown functions and their derivatives lead to a more complicated Monte Carlo procedure. As an example let us consider as a simple generalization of the Schrödinger equation discussed above, a bilinear system of two partial differential equations:

$$\partial u_1/\partial t = \alpha_1 \Delta u_1 + V_1(u_2)u_1,$$

$$\partial u_2/\partial t = \alpha_2 \Delta u_2 + V_2(u_1)u_2$$

u_1 and u_2 being unknown functions of the coordinates x, y, z and the time t; α_1, α_2, say, two given constants; V_1, V_2 given functions for simplicity linear in u_1 and u_2 and also involving the independent variables x, y, z. One would like to know the asymptotic form of u_1 and u_2 (for large values of t). This problem may be looked upon as a generalization of the diffusion model of the Schrödinger equation. It would correspond to a model of a system of two particles with the *potential* function depending only on x, y, z replaced by functions of the distributions. Such a system of linked particles presents, somewhat in the spirit of a field theory, a nonlinear problem. In general there will not exist eigenfunctions. The separation into a time independent system will not be possible. Instead one hopes for an almost periodical behavior, or at least that the unknown functions are summable by first means in time. One could again try a numerical approach in the study of such systems through a Monte Carlo procedure. In this case the fictitious particles whose density is to correspond to u_1 and u_2 are to diffuse and multiply, but this time not according to a given function V of coordinates but rather proportionally to a given function of the density of the particles of the other type. It is therefore necessary to operate with a frequent census of the particles of each type to ascertain the value of the "potential."

In practice an iterative procedure is necessary with the hope of converging to a self-consistent solution.

Prima facie sampling methods like the above work for problems involving distributions of nonnegative real numbers, specifically for Markoff chains involving iteration of matrices with nonnegative coefficients. Let us indicate, however, a way to study "experimentally" the behavior of matrices with arbitrary real terms, or even of more general ones. This possibility rests on the fact that real numbers may be considered as matrices with all terms nonnegative, by starting with the two matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

representing the numbers $+1$ and -1, respectively. This correspondence obviously preserves both additions and multiplications. Any system described by an $n \times n$ matrix of real numbers can be interpreted probabilistically using $2n \times 2n$ matrices with nonnegative terms. The sampling would then involve two kinds of particles ("black" and "red") with the transformation rules given by the matrices above. Again in this case, however, a frequent census of the result is necessary since the final results depend on the *difference* between the numbers of black and red particles at each point. One can realize stochastic models for matrices with complex terms by having 4 kinds of particles, or more general algebras over real numbers through a suitable number of *kinds* of particles. The Dirac equation could be studied in such a way (cf. Everett, Ulam [5]).

9. *Hydrodynamical problems*

Numerical work is essential in many problems of hydrodynamics of compressible flow since at the present time one cannot obtain solutions in closed form. Analytic work alone is unable, in general, to throw light on the asymptotic behavior of solutions or even on their simpler general properties.

In problems involving more than one space dimension, one encounters serious difficulties in planning numerical work. The differential equations of Euler and Lagrange can be replaced, of course, by approximate systems of difference equations, the time variable, too, can be dealt with by a discrete succession of steps in time. The practical difficulties of executing such computations are considerable: if the problem involves large displacements of the fluid or gas (large compared to the initial dimensions of the system) and if the distortions of shape are sizeable, the initial network of Lagrangian points will change its structure in time and very often cease to provide an adequate basis for an approximate representation of the differential expressions which one has to compute (e.g., to evaluate density and pressure gradients). It would seem that the Eulerian system of fixed divisions of space would be more suitable and indeed in many problems this is so. However, other difficulties arise in this treatment; for example, if the fluid moves to vacuous regions — or in problems involving two or more fluids — it is very hard to obtain an idea about the position of the boundary between the two fluids. This is so since in the Eulerian treatment the dependent variables are density and the velocity and from these alone it is impossible to distinguish between a region which is completely occupied by a rarified fluid and one in which half of the region contains a fluid at normal density. In the Eulerian treatment there will be a fictitious computational diffusion from such regions and one loses sight of the position of the true boundaries.

In practice, in problems of this sort, one has to proceed in time through discrete periods called cycles such that the relative displacements in these times are small or moderate, then refine or change the spatial subdivision of the continuum and shorten the time step, a process which may not converge beyond a finite value of time.

Calculations can be considered that are of less "orthodox" variety. Broadly speaking, at least two different ways to approach numerically the problems dealing with the dynamical behavior of continua suggest themselves. In what follows a brief account will be given of some work in this direction, performed by J. Pasta and the author, on computing machines. This work was followed by additional computational experiments of this sort. F. Harlow [1], of Los Alamos Scientific Laboratory, is mainly responsible for the evolution of such methods.

One could try to base the hydrodynamical computations on the kinetic model of the gas or fluid. Assuming as a model of physical reality the Boltzmann integral differential equation, one may calculate the properties in the large of a motion of N "points" — atoms or molecules which have given velocity distributions. These are subject to fluctuations and yield the macroscopic motions only as statistical averages. It is well known, e.g., that the Navier-Stokes equations can be derived from this model. The question is whether it is computationally feasible to obtain the macroscopic quantities like density, pressure, and velocity as functions of time on this basis. The answer, it seems, is in the negative for the time being. If one takes the Boltzmann equations literally and considers the individual points of calculation as representing atoms, then to obtain the average velocity of a point of the gas one would require, say, k particles per cell in a space resolution in which we are interested. Let the number of cells in the computation be l. If on a linear scale the number of intervals is d, $l = d^r$, where r is the number of dimensions of the problem (for problems in 1, 2, or 3 space dimensions without special symmetries, $r = 1, 2, 3$), the total number of points in the calculation would be equal to kd^r. In a statistical study if one wants a mean error of the order of 1 per cent, then $k \sim 10^4$. If the linear resolution is to be of the order of, say, 5 per cent, then d is 20. Even in problems in one space dimension one would have to consider 200,000 points. This is much too high for present computers and makes it clear that the "points" of the calculation should be thought of as representing not individual atoms, but rather large aggregates of them. The dynamical behavior of such a point has to be schematized so as to represent a statistical average of a great number of atoms. The numerical work will not use the Boltzmann equations, but simpler equations which are its consequences. The implicit assumption in such a setup is that our conglomeration of atoms remains coherent during the entire course of the problem. In other words, the small globules of the fluid do not become too much distended and distorted. One should now define the density ρ of the fluid, and the pressure through the equation of state. In this outline we may specialize, say, to either an isothermal or adiabatic process, that is to say, pressure $p = f(\rho)$. The pressure gradients which are available through gradients of density will be calculated by estimating these gradients through the instantaneous appearance of our system of points, representing, we state again, the positions of centers of mass of fluid globules. The problem is then one of finding a rule or recipe to estimate the density at a point of space, given only a finite system of points (rather widely separated in practice). In the limiting case of a very large number of points one could simply count the number of points in a square or cube of a fixed mesh and this number will be proportional to the density. In practice we are limited to a moderate number of points for the whole fluid, the question is how to estimate this density in a most reliable way. Consider the case of two dimensions. One can think of the points located initially in a regular array, for example, on the vertices of a rectangular division, or better, vertices of a triangular subdivision of space. Of course, after some cycles of the computation the geometry of the system will change and one could, for example, estimate the density as follows: enclose each point by triangles with the closest points as vertices. The smallest (in area) triangle in whose interior a given point is located would give, through the ratio of its area to that of the original triangle at least an idea of the change in density at that point.

This procedure suffers from several drawbacks. They are due to the question of computational stability of our calculation. The selection of the nearest points leads to discontinuity in time of the area. In the equations of motion for each of our points we need the gradients of the density. And in this crude way of calculating the densities themselves the computation of *differences* in fixed directions in space is not straightforward. The nearby points may not be sufficiently close and these gradients may be very inaccurately estimated. In the case of points on the boundary of the fluid one would need special prescriptions. These errors would accumulate very seriously with time.

Another way that suggests itself is to introduce numerically the forces due to pressure gradients. One could imagine repulsive forces acting between any pair of our points. These would, in simple cases, depend upon the distance alone (the forces should derive from potentials if we assume scalar pressure, i.e., no viscosity or tensor forces, etc.). The form of the potential will, of course, depend upon the equation of state. In a one-dimensional problem the nature of this correspondence is clear. It suffices to have forces between neighboring points only. The continuity of motion guarantees the permanence in time of the relation of neighborhood. The situation is, however, completely different in one or more space dimensions. The neighbors of a given point will change in the course of time. If one tried to calculate the resultant force on a point due to *all* points in the problem, not merely the neighboring points, the computational work increases — it will grow with the square of the number of points considered. Therefore a "cut-off" for the force is necessary at some distance so that we need not calculate it, if the distance between two points exceeds a certain constant. The pressure gradient would then be given directly as a resultant of all the forces acting on a point due to its neighbors and thus depends on the actual positions without reference to the initial configuration, *"forgotten" by the system*. Let us summarize briefly such a computational scheme.

The particles represent small parts of the fluid. The forces due to pressure gradients are introduced directly by imagining that neighboring or "close" points repel each other. The dependence of this force on the distance between points is so chosen that in the limiting case of very many points, it would represent correctly the equation of state. That this is possible, in principle, is clear *a priori*: the density is inversely proportional, in the limit of a very large number of points, to the square (in two dimensions) or cube (in three dimensions) of the average distance between them. The pressure is a function of density, and this being a function of the distances, we obtain an analogue of the equation of state by choosing a suitable distance dependence of the force. The Lagrangian particles are at

$$P_1\colon\ (x_1,\ y_1,\ z_1),\ P_2\colon\ (x_2,\ y_2,\ z_2),\ \ldots,\ P_N = (x_N,\ y_N,\ z_N)$$

The forces (repulsive) between any two are given by

$$F_{i,j} = F(d(P_i,\ P_j))$$

where $d(P_i,\ P_j)$ is the distance between the two points.

The average value of $\bar{d}$ at a point of the fluid is, in the limit of $N \to \infty$, a function of the local density: $\bar{d} \cong \rho^{-\frac{1}{3}}$ for three dimensions.

The pressure p is a function of ρ alone in, say, isothermal or adiabatic problems. The pressure gradients are thus replaced by forces acting directly on point masses.

There is so far no general theory and the convergence of such finite approximations to the hydrodynamical equations remains to be proved, but even more important than that would be an estimate of the speed of convergence.

In order to test some of these proposals on actual problems, some numerical computations were performed on an electronic machine, the "Maniac" in Los Alamos and on its prototype at the Institute for Advanced Study in Princeton. (Cf. Pasta and Ulam [1]). The main purpose of these calculations was to test the feasibility of such numerical schemes rather than to obtain quantitative precision of the results. One of the problems concerned the study of the motion of a heavy fluid on top of a lighter one, both contained in a cube — with an initial irregularity present at their boundary. The problem involved following the development of this unstable situation — known as Taylor's instability — for motions in the large — beyond the infinitesimal stage — leading to gradual mixing of the two fluids.

With the comparatively crude network of points, representing the fluid, one could hardly expect that the details of the motion would be correctly represented; it was hoped that the behavior of a few functionals of the motion would be more accurately depicted.

One of the functionals which was computed as a function of time was the total kinetic energy of the particles divided into two parts: the kinetic energy of the horizontal and vertical motions separately. Even though the motion itself was very irregular, these quantities seemed to present rather smooth functions of time. In the case of a stable configuration with the lighter fluid on top and the heavy fluid on the bottom, there would be only a periodic interchange of kinetic and potential energies. In the unstable case, there should be an increase of kinetic energy which persists for a considerable time. The hope is that with many more calculations one could try to guess from such numerical results the form of an empirical law for the increase of this quantity as a function of initial parameters.

Another quantity of interest is the spectrum of angular momentum which may be defined, for example, in the following way. We draw around each particle a circle of fixed radius and calculate the angular momentum in each such region. We then take the sum of the absolute values or the sum of the squares of these quantities. This gives us an over-all measure of angular momentum on a scale given by the radius. Varying the radius, these numbers may then be studied as a function of both time and the radius. It is interesting to consider the rate at which the large scale angular momentum is transferred to small scale eddies — also, the spectrum in the asymptotic state, if it exists — all this, of course, as a function of the parameters of the problem: the ratio of the densities of the two fluids and the external force. It is clear that very many numerical experiments would have to be performed before one would trust such pragmatic "laws" for the time behavior of mixing or increase in vorticity.

Another reason for selection of this problem was an interest in the degree of spatial mixing of two fluids, starting from an unstable equilibrium. To measure quantitatively such mixing in configuration space, a functional was adopted similar to the one just mentioned for the spectrum of angular momentum. At a given time t, a circle is drawn around each point P. In each circle, we look at the ratio r of the number of heavy particles to the total number of particles in this region. We take the quantity $\mu(P, t) = 4r(1 - r)$. This gives an index of mixing of the fluid in the circle, being equal to zero if only one fluid is present and equal to 1, if particles of both fluids are equally present in it. We average this quantity over all the circles. Initially this average is very close to zero, the only region where it differs from zero being around the interface. As the mixing proceeds, our average measure of mixing increases with time. Again one might expect that after a sufficient number of numerical experiments, one would obtain an idea of the time behavior of this "mixing functional," or at least of the time T taken for the over-all mixing (which is initially close to zero) to become of the order of $\frac{1}{2}$ or $1/e$. This time, for dimensional reasons, must depend *inter alia* on $\sqrt{(L/a)}$ where L is the depth of the vessel and a the acceleration of the lighter fluid into the heavier one.

A comparison with experiments actually performed in Los Alamos in the problem of Taylor's instability for the case of a heavy gas on top of a light gas with an irregularity in the interface between them showed configurations not unlike those developed in the calculation. Of course, if one wanted to put credence in calculations like the above, the form of the proper force law imitating the real equation of state would have to be carefully chosen. The actual behavior, in time, of our functionals of the motion which we have calculated was very smooth despite the complicated nature and increasing "turbulence" of the motion itself.

A series of calculations was performed by F. Harlow and M. Evans [1], using a somewhat different method but still employing instead of the "classical" difference equation a finite approximation to the motion of the continuum by calculating motions of points representing globules of the fluid. The results obtained agreed in elementary cases with the analytic results and in other cases agreed quite well with experiments.

10. *Synergesis*

In addition to, and in some aspects quite beyond the above type of use of electronic computing machines one can conceive more general possibilities for large scale *experimentation* on problems of pure mathematics or in exploration of tentative ideas in physical theories. The program which will be outlined here requires certain equipment additional to the existing electronic machines and certain changes in their operation. These are not yet available but will soon exist since engineering work in this direction is in progress.

Speaking very broadly, the present machines operate on a set of given instructions (a flow diagram and a code) which once given make the machine proceed autonomously in solving problems of mathematical analysis or physics. The course of the operations is completely prescribed and the limited flexibility of the machine consists, roughly speaking, in taking one or the other course of computation depending on the value of numbers just computed. These so called *decisions* made by the machine in practice, so far, involve only a limited and prescribed set of changes in the logical course which is given.

One could conceive of a more general plan. Instead of using the machine as a robot or, as it were, as a player piano whose tunes are written in advance, the machine could be kept in constant communication with an intelligent operator who changes even the logical nature of the problem during the course of a computation, at will, after evaluating the results which the machine provides. Of course such possibilities exist already, but to a *very limited* extent. We shall try to indicate by some examples how a much more intimate relation could provide useful results. It is obvious for one thing that a computing machine can very quickly provide examples of geometrical or combinatorial situations which can be studied by a mathematician. Thus it obviously can play a role analogous to but vastly more efficient and helpful than scratch paper and pencil work which a person studying a problem uses to provide himself with a visual or a memory aid. Certainly the machine could perform quickly some of the drudgery of elementary algebraic of analytic calculations. In a search for examples or counter-examples the machine could calculate and then display on a screen elements of geometrical objects envisaged by the working mathematician and provide a responsive "scratch paper" if not an audience on which one could try one's ideas.

Obviously for activities of this sort, a rapid access to the machine is necessary and to work a computer in such a fashion would require a way to change its program more substantially than by just changing some constants of the computation in a short time. Also, the machine has to provide a quick illustration and display of the computed quantities and figures. In other words, the problem is of constructing ways of informing the machine of the desired course of computation and conversely of having the machine communicate the results obtained.

One rather wide field of application of a display technique will be in the study of properties of functions of several variables. In many problems it is important to find the critical values of a function of several real variables $f(x_1, x_2, \ldots, x_n)$, the function f being given analytically, e.g., as a polynomial in elementary functions or quotients of such polynomials. It is well known that the usual procedure employed to find, numerically say, local minima or maxima of a single function requires a search which is extremely time consuming. If the number of independent variables is large, for example 5 or more, no really efficient method of finding all the critical points has been proposed. Imagine now that the values of the function can be quickly computed on a grid of points which forms, say, a two-dimensional section in the given space and the resulting function of two variables, i.e., a surface displayed as in a projection, in perspective, on a tube. (A code for calculating axonometric projections is readily made.) The eye quickly notices the region or the regions where minima are likely to be assumed. By a quick change of scale and a magnification, that is to say, a subdivision of the region in question into a greater number of points, a computing machine can act as a microscope of arbitrary power. All we are saying is that in the search for critical points one can utilize, instead of the blind recipes embodied in a search code, the visual perception of the brain which is still much quicker than any now-known automatic code for "recognition."

By critical points we mean, for example, all of the points of space where all the first partial derivatives of the given function vanish. Beyond that, the recognition of other properties of a function (say, of 2 variables, for example), the finding of valleys, ridges, etc., immediately striking the human eye, would be, at the present time, almost impossible to code in a reasonable time for an automatic search by the machine.

Moreover, for functions of 3 or 4 variables one should also have a quick way to instruct a machine to select a desired two-dimensional section, to establish on it a set of independent grid points, to compute the value of functions on these points and display it in perspective on a screen. More important still will be the ability to change the scale of the independent and dependent variables by a general linear transformation. If T is a transformation, one wants to "see" also the transformations of the form HTH^{-1}, where H is a given one-one change of coordinates — i.e., the conjugate transformation. In this still elementary type of problem, it would be convenient to have a quick way of displaying the Fourier transform of the function. Also a display of the functions defined by the derivatives and gradients of the given functions for final verification of the character of interesting points would be convenient. The next step should involve a study of implicit functions, i.e., surfaces in n-dimensional space. Given a function $f(x_1, \ldots, x_n) = 0$, one would like to display the appearance of all desired two-dimensional sections through the n-dimensional surface. An idea of the three-dimensional sections could be obtained by moving two-dimensional sections continuously on the screen. Obviously there are technical problems, e.g., of persistence of image on the screen.

The next aim, still more ambitious, would be to provide for a series of "experiences" with problems computed on the machine so that its operator would acquire, after some practice, a *feeling* for the four-dimensional space as a result of such experimentation. Imagine, for example, that we consider the problem, in three-dimensions, of threading a given solid through a given closed space curve, an exercise which involves "*trying*" since no simple criteria about projections seem to be sufficient to decide whether or not one can push a given solid through a given curve. The physical process of effecting such motion could be imitated by the machine by making it compute the successive positions of the solid following given manual instructions (to be quickly transmitted numerically) about the rotations and translations in the three-dimensional space. The contact between the two sets would be tested after each trial displacement.

An analogous problem is: to thread a four-dimensional solid through a closed two-dimensional surface in E, something for which no tactile or visual experience exists. It is possible that a certain facility for such tasks could be developed, to some extent at least, by continued experimentation on machines which would compute the course of motion of the given solid and display to the operator the appearance of three-dimensional sections and at least inform the operator of collisions between the given surface and the object which has to be disjoint with it. The above is perhaps an arbitrary and already very complicated example of what should be a systematic approach to acquiring, by practice, a feeling for four-dimensional geometry. The problem above is only partially topological, and principally metrical. A systematic attempt to familiarize oneself with the purely *topological* properties of complexes in four dimensions would, however, be more difficult to formulate and to code.

But even in two dimensions, if one studies, instead of single real valued functions, *transformations* defined on the plane into itself interesting heuristic methods seem to be possible. For a given transformation $T(p)$ of the plane we may want to study the properties of the iterates of this transformation, that is to say, the asymptotic properties of the sequence of points $T^k(p)$, to find its fixed points, i.e., $p = T(p)$, involution points, i.e., points p such that $T^k(p) = p$, $k > 1$, and invariant subsets S of the plane, i.e., sets such that $T(S) \subset S$, etc. If T is given analytically, it is a simple matter to compute $T(p)$ for a great number of points p. We can then imagine a way to connect the given points to their images by arrows on a display tube. Here, as in all previous problems, the question of which is the most instructive method of display is not easily answered. Perhaps one could select the given points on some curve, e.g., on a circle, and then display the curve consisting of the image of this circle. Suppose we want to find the fixed point of a transformation. The proposal would be to start with a circle, look at its image, then decide upon a suitable motion of the given circle so that its image will fall essentially inside or essentially outside itself. Then from Brouwer's fixed point theorem we would know that the search for the fixed point is narrowed to the interior of the curve. There should then be a rapid manual method of "steering" the initial curve whose image is displayed, in any desired direction, of changing its size, and possibly the eccentricity if we start with an ellipse, etc.

All this applies, mutatis mutandis, in the n-dimensional space, where we would of course be able to display only various plane sections, or, perhaps better, a 2-dimensional section of an initially given sphere and its image. One transformation for which such a study was undertaken is the one discussed in Chapter II, Section 3. This transformation contains in itself all the information about the general algebraic equation of the nth order. There are many still unsolved questions concerning its fixed points, invariant manifolds, etc. John Ackley has coded the "Whirlwind" machine at the Lincoln Laboratories of M.I.T., in Cambridge, Massachusetts, for a preliminary study of this sort.

Computations on a machine are particularly well suited for the display of asymptotic properties of iterates. It is obvious that in order to tabulate only the interesting points, it is preferable to have a visual evaluation of the iterated properties of the many points rather than print out an enormous number of numerical values. The latter have to be evaluated by a person in order, one by one, and it is in general very hard to guess in advance what properties of the sequence one wants to look at — it seems difficult to foresee and put in a code ahead of time the desired criteria for selecting significant special cases.

Another example of such interplay between the operations performed on the machine and displayed visually, and the decisions made by the operator upon evaluation of the results: such cooperation might be useful in finding solutions to differential equations with boundary values. In the simplest cases, those of the ordinary differential equations, say of the 2nd order, with prescribed boundary conditions at 2 points, one could satisfy the conditions at the first given point, then by integrating on the machine presumably follow the trend of the solution before it is computed at the second boundary and by sensing the trend, so to say, modify the arbitrary parameter assumed at the 1st point, e.g., the value of the 1st derivative. It is clear that a saving in time can be obtained by an intelligent intervention quite beyond any automatic coding of obvious changes.

But even in the study of systems of equations of the 1st order, by exhibiting the vector field on a screen (in 3 or a higher number of dimensions showing the vector field in projections) one could guess the location of singular points "at a glance", and also perhaps form an idea of the behavior of solutions in the large. To give an example: suppose the problem concerns the behavior of the magnetic lines of force in space due to steady currents flowing on a given curve or a system of curves as mentioned in Section 6 of the present chapter. C. Luehr and the writer considered solution of the problem by computations on an electronic machine in two cases. *1.* In three-dimensional space a current flows on two infinite lines; one of these is the x-axis, the other the line $x = 0$, $z = 1$. *2.* The current flows through a closed curve forming a clover-leaf knot. The problem was to study the qualitative topological features of the field of all the lines of force in space. The general ergodic properties of the fields of lines of force are not understood — they are of practical interest now in magnetohydrodynamics and in astronomy.

The computations were performed on the "Maniac" in Los Alamos. This work is still proceeding, but without a good mechanism of display, the progress in evaluating the properties of the lines of force is slow and the evaluation of the printed data laborious. A visual display would certainly permit one to diminish the number of trials, i.e., the initial points through which one computes the lines of force.

In the study of partial differential equations the advantage of a quick survey of results would, it seems, be even greater. As an example of what we have in mind, the following investigation should be mentioned. A rather deep connection appears to exist between the behavior of solutions of the Hamilton-Jacobi partial differential equation

$$(\partial s/\partial x)^2 + (\partial s/\partial y)^2 + (\partial s/\partial z)^2 + V(x, y, z) = \partial s/\partial t$$

on one hand, and the problem of the random walk, with a variable step proportional to the given function V, described by a diffusion equation. The conjecture, proved so far only in the case of one independent space variable and still open for two or more dimensions, concerns the following problem: Given a function $V(x_1, x_2, \ldots, x_n)$ consider, at time t, the front of the Hamilton-Jacobi wave $S(x^0, x) = t$ originating from any point x^0. This surface is obtainable by the Huyghens construction of envelopes of spheres. On the other hand, consider the crest of the probability function $w(x_1, x_2, \ldots, t)$ corresponding to a random walk from x^0 with a step whose variable length is proportional at the position $(x_1, x_2, \ldots, x_n)$ to the value of the given function V (cf. Everett, Ulam [6]).

By the *crest* of the probability function w we mean the surface $t = f(x, x^0)$ defined by $\partial w/\partial t = 0$. The conjecture states that f and S are simply related; for $n = 1$ one shows $f \sim S^2$ under suitable conditions. Apart from this, we should like to state here the belief that the field of curves or sheaves of surfaces corresponding to the equation $S = t$, where S is the Hamilton-Jacobi function, if suitably displayed on a screen, would help form ideas about the behavior of solutions of a dynamic problem for a great variety of initial conditions all at once.

In problems of hydrodynamics it is again obvious that for qualitative studies, e.g., of the progress of mixing of two fluids, a display of the position of the interface between the two fluids or gases — starting from an unstable configuration — would permit good choices of initial conditions and significant values of parameter. (One of these parameters is the value of the initial acceleration of the lighter fluid into the heavier one.) Here again, in absence of *a priori* knowledge of the character of the motion, it is very hard to code in advance automatic criteria for the selection of such parameters.

It is in the study of games and in the actual playing of games that *synergesis*, i.e., the continuing collaboration between the machine and its operator should prove of immediate value. Recently, some progress has been achieved in coding for electronic computers to play games like checkers (Samuels, IBM Research Lab., on the 704) and chess (Cf. Kister *et al.* [1]). While in the simpler game of checkers the existing code allows the machine to play a very good opening and middle game, in the game of chess the quality of the play, as coded so far, is very rudimentary. No doubt progress will be made during the next few years, but the writer does not believe that chess games of master quality will be achieved in the near future!

An intermediate, and in a way less ambitious, program would be to have codes prepared for assisting a human chess player to explore on a machine certain *special* sequences of moves quite far ahead (i.e., for perhaps 6–8 moves) and then to flash the tentative positions on a screen. This, when coupled with calculations of various evaluation functions generalizing those already conceived would possibly improve the game!

To play a "fair game", the two opponents would have to have at their disposal two identical machines! In other words, it seems that a computing machine could act as a "second" of the player, helping to analyze the positions. To the writer at least, an intelligent program of this sort seems nearer at hand than a completely automatic program enabling the machine to play a high quality game. Going beyond the existing games, it is perhaps permissible to speculate on the invention of entirely new games to be played between two players, each provided with a computer.

All the examples mentioned above deal with methods to facilitate solution of given problems by some general heuristic methods — either of Monte Carlo type, or consisting of specific numerical work. Beyond this it seems that the computing machines will soon be able, at least in elementary cases, to prove theorems in the customary mathematical formalisms. The interesting work of N. Rochester and his associates (Gelernter, Roth, and others) makes it appear plausible that an automatic code will soon be available for *proving* theorems in certain elementary domains (in Euclid's geometry, for example). It will probably take a long time until an analogous program will be realized for more extensive mathematical disciplines. Here again it could be pointed out that an intermediate program of collaboration between a human operator and a machine seems easier of accomplishment! In fact, one may start by practicing this by proving theorems in Euclidean geometry about triangles, etc.; or certainly, if one wanted to automatize proofs in projective geometry, a *display* of the geometric constructions envisaged by a human operator plus automatic searching by the machine for the routine syllogisms might provide a quicker way to find formal proofs.

BIBLIOGRAPHY

ALEXANDER, J. W.

[1] Some Problems in Topology, *Verhandlungen des Internationalen Math. Kong.*, Zürich, 1932, pp. 249–257.

ALEXANDROFF, P. and HOPF, H.

[1] Topologie I, *Die Grundlehren der Math. Wissenschaften*, Vol. XLV, J. Springer, Berlin, 1935.

ANDERSON, R. D.

[1] The Algebraic Simplicity of Certain Groups of Homeomorphisms, to appear in *Proc. Am. Math. Soc.*

ANTOINE, L.

[1] Sur l'homéomorphe de deux figures et leurs voisinages, *J. Math.*, Ser. 8, 4 (1921), pp. 221–325.

ARNOLD, V. I.

[1] *Doklady Akad Nauk U.S.S.R.*, 114, No. 4 (1957).

AUERBACH, H.

[1] Sur un problème de M. Ulam concernant l'équilibre des corps flottants, *Studia Math.*, 7 (1938), pp. 121–142.

BELLMAN, R. and HARRIS, T.

[1] On Age-Dependent Binary Branching Processes, *Ann. of Math.* 55 (2) (1952), pp. 280–295.

BIEBERBACH, L.

[1] Operationsbereiche von Funktionen, *Verhandlungen des Internationalen Math. Kong.*, Zürich, 1932, pp. 162–172.

BING, R. H.

[1] The Cartesian Product of a Certain Nonmanifold and a Line is E. *Bull. Am. Math. Soc.* 64 (1958), pp. 82–84.

BIRKHOFF, G.

[1] *Lattice Theory*, Rev. Ed., Am. Math. Soc. Colloq. Pub. XXV, 1948.

BORSUK, K.

[1] Drei Sätze über die n-dimensionale Euklidische Sphäre, *Fund. Math.*, 20 (1933), pp. 177–190.

BORSUK, K. and ULAM, S. M.

[1] On Symmetric Products of Topological Spaces, *Bull. Am. Math. Soc.*, 37 (1931), pp. 875–882.

[2] Über gewisse Invarianten der ε-Abbildungen, *Math. Ann.*, 108 (1933), pp. 311–318.

BOTT, R.

[1] On Symmetric Products and Steenrod Squares, *Ann. of Math.*, 57 (2) (1953), pp. 579–590.

CANTOR, G.

[1] *Contributions to the Founding of the Theory of Transfinite Numbers*, Dover Publications, New York.

CASHWELL, E. D. and EVERETT, C. J.

[1] *A Practical Manual on the Monte Carlo Method for Random Walk Problems*, LA–2120 Office of Technical Services, U. S. Dept. of Commerce, Washington, D. C.

COHEN, L. W. and ULAM, S. M.

[1] On the Algebra of Systems of Vectors and Some Problems in Kinematics, *Bull. Am. Math. Soc.*, 50 (1944), Abstract 38, p. 61.

EGGLESTON, H. G.

[1] A property of Plane Homeomorphisms, *Fund. Math.*, 42 (1955), pp. 61–74.

ERDÖS, P.

[1] On a Lemma of Littlewood and Offord, *Bull. Am. Math. Soc.*, 51 (1945), pp. 898–902.

ERDÖS, P. and JABOTINSKY, E.

[1] On Sequences of Integers Generated by a Sieving Process, *Koninkl. Nederl. Akad. van Wetenschappen, Amsterdam Proceedings, Ser. A*, 61 (1958), pp. 115–128.

ERDÖS, P. and KAKUTANI, S.

[1] On Non-Denumerable Graphs, *Bull. Am. Math. Soc.* 49 (1943), pp. 457–461.

EVERETT, C. J. and ULAM, S. M.

[1] Projective Algebra I, *Am. J. Math.*, 68 (1946), pp. 77–88.

[2] Multiplicative Systems in Several Variables, I, AECD–2164 (1948); II, AECD–2165 (1948); III, AECD-2532 (1948), Technical Information Branch, Oak Ridge, Tenn.

[3] Multiplicative Systems I, *Proc. Nat. Acad. Sci.*, 34 (1948), pp. 403–405.

[4] On the Problem of Determination of Mathematical Structures by Their Endomorphisms, *Bull. Am. Math. Soc.*, 54 (1948), Abstract 285t, p. 646.

[5] On an Application of a Correspondence Between Matrices Over Real Algebras and Matrices of Positive Real Numbers, *Bull. Am. Math. Soc.*, 56 (1950), p. 63.

[6] Random Walk and the Hamilton-Jacobi Equation, *Bull. Am. Math. Soc.*, 56 (1950), Abstract 97t, p. 63.

FELLER, W.

[1] The Fundamental Limit Theorems in Probability, *Bull. Am. Math. Soc.*, 51 (1945), pp. 800–832.

[2] Diffusion Processes in Genetics, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Univ. of California Press, Berkeley, 1951, pp. 227–246.

FERMI, E., PASTA, J. and ULAM, S. M.

[1] Studies of Nonlinear Problems, LA-1940, Office of Technical Services, U. S. Dept. of Commerce, Washington, D.C.

FOX, R. H.

[1] On a Problem of S. Ulam Concerning Cartesian Products, *Fund. Math.*, 34 (1947), pp. 278–287.

GALE, D. and STEWART, F. M.

[1] *Infinite Games with Perfect Information*, Contributions to the Theory of Games, Vol. II, Princeton, N. J., 1953.

GARDINER, V., LAZARUS, R., METROPOLIS, N. and ULAM, S. M.
[1] On Certain Sequences of Integers Defined by Sieves, *Math. Mag.*, 29 (1956), pp. 117–122.

GÖDEL, K.
[1] Über formal unentscheidbare Sätze der Principia Mathematica und verwandte Systeme I, *Monatshefte Math. Phys.*, 38 (1931), pp. 173–198.
[2] *The Consistency of the Axiom of Choice and the Generalized Continuum Hypothesis with the Axioms of Set Theory*, Ann. of Math. Studies, No. 3, Princeton Univ. Press., Princeton, 1940.

HALMOS, P. R.
[1] Algebraic Logic I, *Compositio Math.*, 12 (1955), pp. 217–249; II, *Fundamenta Math.*, XLIII (1956), pp. 255–325.

HARLOW, F. H. and EVANS, M. W.
[1] The Particle-in-Cell Method for Hydrodynamic Calculations, LA–2139, Office of Technical Services, U. S. Dept. of Commerce, Washington, D. C.

HILBERT, D.
[1] Mathematische Probleme, *Nach. Akad. Wiss. Göttingen*, 1900, pp. 253–297.

HOPF, E.
[1] *Ergodentheorie*, Ergebnisse der Math., Chelsea Publishing Co., New York, 1948.

HUREWICZ, W. and WALLMAN, H.
[1] *Dimension Theory*, Princeton Math. Ser. No. 4, Princeton Univ. Press, Princeton, 1941.

HYERS, D. H.
[1] On the Stability of the Linear Functional Equation, *Proc. Nat. Acad. Sci. U. S.*, 27 (1941), pp. 222–224.

HYERS, D. H. and ULAM, S. M.
[1] On Approximate Isometries, *Bull. Am. Math. Soc.*, 51 (1945), pp. 288–292.
[2] Approximate Isometries of the Space of Continuous Functions, *Ann. of Math.*, 48 (2) (1947), pp. 285–289.
[3] Approximately Convex Functions, *Proc. Am. Math. Soc.*, 3 (1952), pp. 821–828.
[4] On the Stability of Differential Expressions, *Math. Mag.*, 28 (1954), pp. 49–64.

KACZMARZ, S. and TUROWICZ, A.
[1] Sur l'irrationalité des intégrales indéfinies, *Studia Math.*, 8 (1939), pp. 129–134.

KELLY, P. F.
[1] On Isometries of Square Sets, *Bull. Am. Math. Soc.*, 51 (1945), pp. 960–963.

KHINCHIN, A. Y.
[1] *Mathematical Foundations of Statistical Mechanics*, Dover Publications, New York, 1940.
[2] *Three Pearls of Number Theory*, Graylock Press, Rochester, New York, 1952, Chap. I.

KISTER, J., STEIN, P., ULAM, S., WALDEN, W., and WELLS, M.
[1] Experiments in Chess, Jour. Ass'n. Computing Machinery, 4 (1957), pp. 174–177.

KOLMOGOROFF, A. N.

[1] On the Representation of Continuous Functions of Several Variables in the Form of Superposition of Continuous Functions of One Variable and Addition, *Doklady Akad. Nauk S.S.S.R.*, 108, No. 2 (1956).

KURATOWSKI, C.

[1] *Topology I*, Monografje Matematyczne, Warsaw, 1933.

[2] Les ensembles projectifs et l'induction transfinie, *Fund. Math.*, 27 (1936), pp. 269–276.

[3] Les suites transfinies d'ensembles et les ensembles projectifs, *Fund. Math.* 28 (1937), pp. 186–196.

KURATOWSKI, C. and ULAM, S. M.

[1] Quelques propriétés topologiques du produit combinatoire, *Fund. Math.*, 19 (1932), pp. 247–251.

[2] Sur un coefficient lié aux transformations continues d'ensembles, *Fund. Math.*, 20 (1933), pp. 244–253.

LÉVY, P.

[1] Trois théorèmes de calcul des probabilités, *Compt. rend.*, 238 (1954), pp. 2283–2286.

LUSIN, N.

[1] *Leçons sur les Ensembles Analytiques*, Gauthier-Villars, Paris, 1930.

MCKINSEY, J. C. C.

[1] On the Representation of Projective Algebras, *Am. J. Math.*, 70 (1948), pp. 375–384.

MACMAHON, P. A.

[1] *Combinatory Analysis*, Vol. I, II, Cambridge Univ. Press, New York, 1916.

METROPOLIS, N. and ULAM, S. M.

[1] On the Motion of Systems of Mass Points Randomly Distributed on the Infinite Line, *Bull. Am. Math. Soc.*, 55 (1949), Abstract 299t, p. 670.

MOISE, E. E.

[1] Affine Structures in 3-Manifolds, VIII. Invariance of the Knot-Types; Local Tame Imbedding, *Ann. of Math.*, 49 (1954), pp. 159–170 and references to I–VII.

MONTGOMERY, D. and ZIPPIN, L.

[1] *Topological Transformation Groups*, Interscience Publishers, New York–London, 1955.

MORSE, M.

[1] *The Calculus of Variations in the Large*, Am. Math. Soc. Colloq. Pub. XVIII, New York, 1934.

MYCIELSKI, J., SWIERCZKOWSKI, S. and ZIEBA, A.

[1] On Infinite Positional Games, *Bull. acad. polon. sci.*, 4 (1956), pp. 485–488.

MYCIELSKI, J. and ZIEBA, A.

[1] On Infinite Games, *Bull. acad. polon. sci.*, 3 (1955), pp. 133–136.

NETTO E.

[1] *Lehrbuch der Kombinatorik*, 2nd Ed., Chelsea Publishing Co., New York.

Oxtoby, J. C. and Ulam, S. M.
[1] Measure Preserving Homeomorphisms and Metric Transitivity, *Annals of Math.* 42 (1941), pp. 874–920.

Pasta, J. and Ulam, S.
[1] Heuristic Numerical Work in Some Problems of Hydrodynamics, *Math. Tables and Other Aids to Computation*, XIII (1959), pp. 1–12.

Pólya, G. and Szegö, G.
[1] *Aufgabe und Lehrsätze aus der Analysis*, Vol. I, II, Dover Publications, New York.

Pontrjagin, L.
[1] *Topological Groups*, Princeton Math. Ser. No. 2, Princeton Univ. Press, Princeton. 1939.

Rechard, O. W.
[1] Invariant Measures for Many-One Transformations, *Duke Math. J.*, 23 (1956), pp. 477–488.

Ritt, J. F.
[1] *Integration in Finite Terms*, Columbia Univ. Press, New York, 1948.

Roth, K. F.
[1] On Certain Sets of Integers, *J. London Math. Soc.*, 28 (1953), pp. 104–109.

Saks, S.
[1] *Theory of the Integral*, Monografje Matematyczne VII, Stechert and Co., New York, 1937.

Schreier, J.
[1] Über die Drehungsgruppe im Hilbertschen Raum, *Studia Math.*, 5 (1934), pp. 107–110.

Schreier, J. and Ulam, S. M.
[1] Über die Permutationsgruppe der natürlichen Zahlenfolge, *Studia Math.*, IV (1933), pp. 134–141.
[2] Über topologische Abbildungen der euklidischen Sphären, *Fund. Math.*, 23 (1934), pp. 102–118.
[3] Eine Bemerkung über die Gruppe der topologischen Abbildungen der Kreislinie auf sich selbst, *Studia Math.*, 5 (1935), pp. 155–159.
[4] Sur le nombre des générateurs d'un group topologique compact et convexe, *Fund. Math.*, 24 (1935), pp. 302–304.
[5] Über die Automorphismen der Permutationsgruppe der natürlichen Zahlenfolge, *Fund. Math.* 24 (1935), pp. 158–260.

Shapiro, H. N.
[1] Note on a Problem in Number Theory, *Bull. Am. Math. Soc.*, 54 (1948), pp. 890–893.

Sierpinski, W.
[1] *Hypothèse du Continu*, 2nd Ed., Monografje Mathematyczne Tom IV, Warsaw, Chelsea Publishing Co., New York, 1934.
[2] *General Topology*, Univ. of Toronto Press, Toronto, 1952.

Sperner, E.
[1] Ein Satz über Untermengen einer endlichen Menge, *Math. Z.*, 27 (1928), pp. 544–548.

STEIN, P. R. and ULAM, S. M.

[1] A Study of Certain Combinatorial Problems Through Experiments on Computing Machines, *Proceedings of the High Speed Computer Conference*, Louisiana State University, Baton Rouge, 1955.

[2] *Quadratic Transformations*, Part I, Los Alamos Scientific Laboratory Report No. 2305, Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C.

ULAM, S. M.

[1] Zur Masstheorie in der allgemeinen Mengenlehre, *Fund. Math.*, 16 (1930), pp. 141–150.

[2] Infinite Models in Physics, Applied Probability, *Proceedings of Symposia in Applied Mathematics*, Vol. VII, McGraw-Hill Book Co., New York, 1957.

ULAM, S. M. and VON NEUMANN, J.

[1] On the Group of Homeomorphisms of the Surface of the Sphere (abstract), *Bull. Am. Math. Soc.*, 53 (1947), p. 506.

VOLTERRA, V.

[1] *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*, Gauthier-Villars, Paris, 1931.

WALDEN, W.

[1] A Study of Simple Games Through Experiments on Computing Machines, *Contributions to the Theory of Games*, Vol. III, Annals of Math. Studies No. 39, Princeton, 1957.

WEIL, A.

[1] *L'intégration dans les groups topologiques et ses applications*, Actualités Scientifiques et Industrielles 869, Hermann et Cie., Paris, 1938.

WEYL, H.

[1] Mathematics and Logic, *Am. Math. Monthly*, 53 (1946), pp. 2–13.

ZAHORSKI, Z.

[1] Sur l'ensemble des points singuliers d'une fonction d'une variable réelle admettant les dérivés de tous les ordres, *Fund. Math.*, 34 (1947), pp. 183–245; Supplément au mémoire, "Sur l'ensemble . . .", *Fund. Math.*, 36 (1949), pp. 319–320.