The University of Warwick

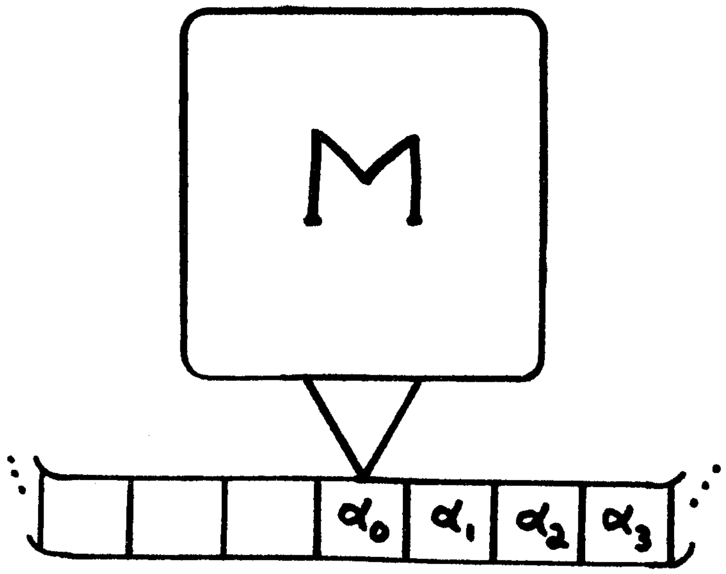
THEORY OF COMPUTATION

REPORT NO.45

INTRODUCTION TO THE BAIRE SPACE

BY

W. W. WADGE



Department of Computer Science
University of Warwick
COVENTRY CV4 7AL
ENGLAND.

October 1982

This report reproduces the chapter 0 of my PhD dissertation, "Reducibility and determinateness on the Baire Space". I have produced it as a Warwick Theory of Computation report because infinite games and the computational approach to topology presented here is, I feel, very relevant to computer science. The basic connection between topology and computability, explained in section E, is as follows: a function from the Baire space to itself is continuous iff it can be computed by a continuously operating numeric 'filter' which has access to a countably infinite database.

It used to be thought that infinite computations (not to mention infinite games) were of little relevance to practical computing, which (it was thought) was inherently finitary. The emergence of the dataflow model of computation (among other factors) has changed all this; computer scientists are now keenly interested in the behaviour and properties of continuously operating (nonterminating) devices. The entire history of the input to, or output from, such a device will normally be an infinite sequence of finite objects. Such a history can be 'coded up' as an infinite sequence of natural numbers; i.e., as an element of the Baire Space. The study of this space could therefore prove to be as important in computer science as it has already proved to be in (say) statistics

Of course I make no claims to have discovered the material presented here. It has been known for many years now that "computability" and "continuity" are closely related. However, I hope that this report will in a small way help to popularise the 'topological' approach to computation. I would like to thank John Addison for introducing me to the 'operational' approach to topology.

Readers interested in the dissertation itself are referred to Theory of Computation report no 44, which is a collection of the more important parts.

Bill Wadge
October 1982

CHAPTER 0

BACKGROUND

0. Introduction

In this chapter we present some simple but vital background of a technical and motivational nature. Most of the material (or at least the technical part) will be familiar to readers already well-versed in modern descriptive set theory.

In section A we present a summary of our mathematical and notational conventions. These are are almost entirely standard.

Section B is a short introduction to the Baire space, its topology, and some closely related spaces.

In section C we give a brief description of the important hierarchies (such as the Borel hierarchy) of subsets of the Baire space.

Section D is an introduction to infinite games of perfect information. We explain the concept of determinateness and the importance of the axiom of determinateness (AD) and its weaker variants.

Finally, in section E we give an informal presentation of the 'algorithmic' interpretation (due to Kleene and Addison) of the topology of the Baire space. We explain, for example, why the clopen sets are 'recursive' and why continuous functions are 'computable'.

A. Notation

The various results (propositions and theorems) announced are to be understood as theorems of Zermelo-Fraenkel set theory (see, for example, Schoenfield (1977)) with the axiom of Dependent Choice. In other words, the statements of the results should be understood as precise specifications of formulas in the language of ZF which are logical consequences of the axioms of ZF without the axiom of Choice but with Dependent Choice. Similarly, the proofs presented should be understood as informal outlines of formal proofs strictly within the language of ZF using only the axioms indicated and the rules of inference of first-order logic. The detail presented should be enough so that in principle at least the purely formal versions of the statements and proofs could be produced (after enormous effort) by any mathematically competent reader with a knowledge of ZF. In practice some of the more obvious proofs may be described only briefly or even omitted, but definitions and results will always be carefully formulated. All this is, of course, accepted practice in mathematical logic.

In ZF the only objects are sets, all of which are built up (by forming sets of sets, sets of sets of sets etc) from the empty set - in other words, from nothing. All the other entities of mathematics - numbers, functions, sequences, relations and so on - must therefore be represented as sets. The reader must understand these representations because explicit use will be made of them. Fortunately many of these representations have become accepted as part of the 'mainstream' of conventional mathematics.

The representation of the natural numbers in ZF is particularly simple. The number 0 is the empty set, the number 1 is the set $\{0\}$ which has 0 as its only element, the number 2 is the set $\{0,1\}$ having 0 and 1 as its only elements, and in general the number n is the set $\{0,1,2,\ldots,n-1\}$ of all smaller numbers. This representation of numbers as sets is due to Von Neumann. It has not yet gained complete acceptance in mathematics as a whole, which is unfortunate, because it is extremely convenient.

One of the advantages of the Von Neumann definition of a number is that it allows us to continue the process of constructing new numbers from old beyond the finite stages. The result is an endless collection of finite and infinite numbers called the <u>ordinals</u>, the finite ordinals being the natural numbers. The first infinite ordinal is ω, the set $\{0,1,2,3,\ldots\}$ of all natural numbers. Next after ω is the ordinal $\omega+1$, the set $\{0,1,2,\ldots\omega\}$. Continuing in this way we construct a second infinite sequence $\omega+2$ (which is the set $\{0,1,2,\ldots,\omega,\omega+1\}$), $\omega+3$, $\omega+4$, ... of ordinals. The collection of all ordinals formed in this way is the ordinal $\omega+\omega$, namely the set $\{0,1,2,\ldots,\omega,\omega+1,\omega+2,\ldots\}$. We then form the ordinals $\omega+\omega+1$, $\omega+\omega+2$, ... then $\omega+\omega+\omega$, and so on.

It is important to realise that the process never stops. No matter how many ordinals we have constructed, there is always a new one, namely the set of all those constructed so far. In particular, there are ordinals of arbitrarily high cardinality. There is no set containing all ordinals for then it itself would be a new ordinal not included. The first uncountable ordinal is Ω, the set of all countable ordinals.

The concepts of "function" and "relation" are, along with those of "set" and "number", among the most important in mathematics. In ZF a relation R between two sets A and B is represented by the set of all ordered pairs (a,b) of elements a and b of A and B respectively such that a is related to b by R. (The ordered pair (a,b) is the set $\{\{a\},\{a,b\}\}$). A function f from A to B is represented as the set of all ordered pairs (a,b) such that $a = f(b)$. Functions in ZF are therefore relations of a special type, namely those which are single valued: in general a relation R is single valued iff $(a,b_1) \in R$ and $(a,b_2) \in R$ implies $b_1 = b_2$.

The domain $Dm(f)$ of a function f the set of all left-hand components of elements of f, i.e. $\{a \in \cup\cup f: (a,b) \in f \text{ for some } b\}$. (We are using conventional mathematical "set builder" notation, which is justified within limits by the comprehension and replacement axioms of ZF. Ingeneral $\cup S$ is the union of all the sets in S, i.e. the set of all x such that $x \in y$ for some y in S). Similarly, the range $Rg(f)$ of a function f is the set $\{b \in \cup\cup f: (a,b) \in f \text{ for some } a\}$ of all right hand components of elements of f. Notice that the empty set $\emptyset$ is a function (because it is a single-valued relation). It is the only function whose domain and range are both empty.

Given two sets A and B, the set ${}^A B$ (read "B-pre-A") is the set of all functions from A to B. We are using prescript notation to avoid confusion with ordinal exponentiation. Thus ω^ω is a countable ordinal (the limit of the sequence ω, ω^2, ω^3, ...) whereas ${}^\omega\omega$ is the set of all functions from ω to ω (${}^{\omega}\omega$ is not an ordinal). The domain of any element of ${}^{A}B$ is A, and the range of an element of ${}^{A}B$ is a subset of B. We might be tempted to refer to B itself as the "codomain" of elements of ${}^{A}B$ but this concept makes no sense given the ZF representation of functions. Anything in ${}^{A}B$ is in ${}^{A}C$ for any superset C of B; there is no uniquely defined codomain.

Notice that ${}^{A}\emptyset$ is $\emptyset$ but that ${}^{\emptyset}B$ is 1; recall that 1 is $\{0\}$ and $\emptyset$ (=0) is the only function (single valued set of ordered pairs) from $\emptyset$ to B.

A function g is said to extend a function f iff the domain of f is a subset of that of g and f and g agree on the smaller domain. In ZF, function g extends function f iff f (as a set of ordered pairs) is a subset of g (as a set of ordered pairs), i.e. if $f \subseteq g$. Two functions f and g are compatible iff they agree on arguments which are in both domains. In ZF functions f and g are compatible iff $f \cup g$ is a function.

Sequences in ZF are (by definition) simply functions whose domains are ordinals. The length $\ln(s)$ of a sequence s is its domain. Thus the sequence $\langle 2,4,6,8\rangle$ is the function $\{(0,2),(1,4),(2,6),(3,8)\}$ and its domain is $\{0,1,2,3\}$, which is of course number 4, the length of the sequence. In general a sequence of length n (n an ordinal) is called an "n-sequence". The above sequence is therefore a 4-sequence.

The elements of the range of a sequence are called the components of the sequence. Since sequences are just functions, the sequence indexing operation (which selects a given component of a sequence) is just function application. For example, component 10 of a sequence s

is simply $s(10)$. The functional notation for indexing is not, however, always the most convenient. We therefore adopt the convention that subscripting denotes application, i.e. that f_x is the same as $f(x)$. This allows us to denote e.g. component 10 of s by s_{10} and so combines conventional mathematical notation with ZF's slightly unusual treatment of sequences. At times, however, the functional notation for indexing is more convenient (say to avoid multiple levels of subscripting) and we will use it as well.

If s is a sequence and n is an ordinal, the function $s|n$ (s restricted to n) is simply the initial segment of s of length n (or s itself if n is not less than the length of s. There is no need for a special initial segment forming operator.

In writing expressions denoting sequences we will use a fairly conventional "sequence-builder" notation which is like set builder notation except that angular parentheses are used instead. Thus $\langle 2,4,6,8 \rangle$ is, as we have already seen, a sequence of length 4, and

$$\langle i^2 \rangle_{i \epsilon \omega}$$

is the ω-sequence of squares of natural numbers. The <u>set</u> $\{i^2\}_{i \epsilon \omega}$ is the range of this sequence. This second form of sequence-builder notation is just a variation of λ-notation; the sequence above is also the value of the λ-expression

$$\lambda i \epsilon \omega \; i^2$$

Our sequence- and set- builder notation is slightly unconventional (in fact, old fashioned) in that the expression which specifies the bound variable and its range (in the above, the expression "$i \epsilon \omega$") appears as a subscript outside the set or sequence brackets, rather than inside, as in the more usual form $\{i^2 : i \epsilon \omega\}$. We prefer our notation because it makes it much clearer that a variable is being bound. In the usual notation there is some ambiguity about exactly which variables are being bound; the expression "$i \epsilon \omega$" looks like a predicate, which it is not. (See also Curry(1958,p86)). When writing set-builder expressions corresponding to comprehension, however, we will use the standard notation (e.g. "$\{i \epsilon \omega : i > 0\}$") because there is no doubt about the bindings.

In denoting sets and sequences we will sometimes use the direct or explicit form (i.e. the form involving no bound variable) together with 'triple dots' to denote an infinite sequence. The sequence of squares, for example, could be expressed as $\langle 0,1,4,9,16,\ldots\rangle$ with the understanding that the first few values given (in this case five) are enough to make the pattern obvious. The triple dot notation can also be used for finite sequences, so that a sequence s of length n can be expressed as $\langle s_0,s_1,s_2,\ldots,s_{n-1}\rangle$.

The 'triple dot' forms are obviously somewhat informal and less precise than the others, but are usually unambiguous, conform to standard mathematical practice, and are often much clearer. There is no requirement, however, to use angle brackets, commas, and triple dots to name a sequence; we can refer to it by its name alone. We can write

the sequence $\langle s_0,s_1,s_2,\ldots s_{n-1}\rangle$

or

the sequence $\langle s_i\rangle_{i \epsilon n}$

but we can also write simply

the sequence s.

There is no need either to underline the name of a sequence, or overline it, or print it in boldface. In our notation even the most complicated objects have the right to a simple name.

Sometimes it is useful to consider a generalization of the notion of sequence in which the indexing set is not required to be an ordinal. These generalized sequences are called "families" and can be thought of as 'labelled' sets (in the same way that sequences can be thought of as 'ordered' sets). It should be apparent, however that a family is simply a function, the domain of which is the indexing set. Sequences were defined in the first place as functions in which the domain was required to be an ordinal; if we drop this requirement, we return to the original concept. Nevertheless, we will on occasion use the term "family" in conjunction with the angle bracket and subscripting notation. Sometimes functions are in fact better thought of as 'labelled' sets rather than as transformations.

Zermelo Fraenkel set theory is a typeless theory in that there is no classification of sets 'built-in' to the syntax as there is in, say, Goedel-Bernays set theory. In practice, however, it is often useful to introduce some notational conventions which help the reader bear in mind the nature of the particular objects denoted. This is especially true in descriptive set theory, where a wide variety of mathematical entities are used.

The simplest objects are the natural numbers, which are the basic "type 0" objects (to use Kleene's terminology). We will generally use the variables "i", "j", "k", "n" and "m" for natural numbers.

Finite sequences of natural numbers are slightly more complicated than natural numbers, but since they are still finite objects they are classified as type 0. We will generally use the variables "s", "t", "u", "v", and "w" to denote elements of Sq (the set of all finite sequences of natural numbers).

At the next level of complexity (type 1, in Kleene's terminology) we have the number theoretic functions, i.e. functions from ω to ω (elements of ${}^{\omega}\omega$). In general we will use the greek letters "α", "β", "γ" and "δ" for elements of the Baire space. Since the objects are sequences, we will also use subscripting and angle brackets in expressions denoting elements of ω. We will occasionally want to refer to subsets of ω (i.e. elements of $\underline{\underline{P}}(\omega)$) but have not adopted any particular convention for naming them.

At level two, we find the sets of number theoretic functions and the second order number theoretic functions, i.e. functions whose arguments and results are ordinary (first order) number theoretic functons. Objects of the first kind, i.e. subsets of the Baire space, will be denoted by uppercase Roman letters such as "A", "B", and "C". The letters "E" and "F" will be reserved for closed subsets (or, more generally $\underline{\underline{\Pi}}^0_\mu$ subsets), and "G" will be reserved for open subsets (or more generally, $\underline{\underline{\Sigma}}^0_\mu$ subsets).

Objects of the second kind, i.e. functions from the Baire space to itself, will be denoted by the letters "f", "g" and "h".

The same conventions will be used for spaces (such as ${}^{\omega}2$) similar to the Baire space.

This work is mainly concerned with classification of subsets of the Baire space. A collection of subsets of the Baire space is a set of type two objects and is therefore of type three. Sets of this type will usually be referred to as "classes" and will be denoted by upper case roman script letters such as "$\underline{\underline{A}}$", "$\underline{\underline{B}}$" and "$\underline{\underline{C}}$". Amongst these the symbols "$\underline{\underline{G}}$" and "$\underline{\underline{F}}$" are reserved for the class of open subsets and the class of closed subsets respectively of the Baire space.

We will be especially interested in classes which (speaking informally) consist of subsets of the Baire space of the same degree of complexity. These classes are called <u>degrees</u> and will be denoted by the lower-case letters "a", "b", "c", "d" and "e". The script letters will be used to denote classes which are not degrees but instead are 'closed downwards' in the sense that every set which is simpler than an element of the class is also in the class.

The study of degrees involves the study of a number of operations on degrees (such as degree addition). Since these operations are applied to type three objects, they themselves are of type four. They will be denoted, however, by fairly ordinary symbols such as "+" and "#". There are a number of other "miscellaneous" sets and operations too varied to be assigned special typefaces or sections of the alphabet. These will be denoted by special names consisting of two roman letters, the first of which may be uppercase (eg "jn" or "Sp").

The only remaining class of special objects is the ordinals (which include representatives of all types). They will be denoted by the greek letters "μ", "ν", "η", "κ", ,"ζ, "υ", "λ" and "ω". The first three will usually denote countable ordinals. The letter "ω" always denotes the set $\{0,1,2,\ldots\}$ of all natural numbers, and "Ω" will always denote the set of countable ordinals.

In applying the conventions just discussed we will usually treat sequences or families of objects of a certain type as being of the same type, even if this is not strictly speaking the case. For example, the script letter "$\underline{\underline{C}}$" might denote either a single subclass of $\underline{\underline{P}}({}^{\omega}\omega)$ or an ω-sequence of such subclasses. This convention is enormously helpful in keeping the symbolism under control and avoiding the notational extravaganzas which result when one insists that sequences must always be referred to using complex expressions such as "$C_0, C_1, \ldots, C_{n-1}$".

The general principle followed in this work is that the definitions and results should make sense on their own, and that the other material is essentially expository and could be omitted. In particular, every definition or result includes a "preamble" stating the exact nature of the objects denoted by all variables used (e.g. "for any subset C of $\underline{\underline{P}}({}^{\omega}\omega)$ and any countable ordinal μ"). An understanding of the notational conventions adopted (i.e. which symbols will be associated with which kinds of objects) is therefore useful for reading the definitions and results but is by no means necessary.

B. The Baire and related spaces

Descriptive set theory began as the study of properties and classifications of subsets of the space $\underline{\underline{R}}$ of real numbers. It was soon realized, however, that the use of $\underline{\underline{R}}$ lead to minor but annoying difficulties. One problem with $\underline{\underline{R}}$ is that it is not homeomorphic to any of its powers, although as far as descriptive set theory is concerned, these spaces have essentially the same properties. (For example, every Borel set is measurable in each of these spaces). Another not unrelated problem with $\underline{\underline{R}}$ is that of representation: the standard decimal representation of real numbers is badly behaved in that numbers very close to each other (such as $1.000\ldots00$ and $0.999\ldots9$) can have completely different expansions.

It was soon realized that these and other difficulties could be avoided simply by omitting the rational numbers, i.e. by working in the space consisting of the set of irrational numbers (say in the interval $(0,1)$) together with the topology induced by that of $\underline{\underline{R}}$. This space (sometimes called the Baire space) is homeomorphic to each of its finite and countable powers. Furthermore every irrational number (in the interval $(0,1)$) can be represented uniquely as a continued fraction of the form

$$1/(1+\alpha_0+1/(1+\alpha_1+1/(1+\alpha_2+1/\ldots)))$$

for some infinite sequence α $(= \langle\alpha_0,\alpha_1,\alpha_2,\ldots\rangle)$ of natural numbers, and this representation is well behaved in that numbers which are very close together will have in common a large initial segment of their respective representations.

The whole approach is further simplified if we drop the 'coding' and work directly with infinite sequences of natural numbers rather than with the irrationals they represent. Therefore in common with almost all modern descriptive set theorists we consider the Baire space to be the set ${}^{\omega}\omega$ of infinite sequences of natural numbers together with the product topology induced when ω is given the discrete topology.

The elements of the Baire space can thus be thought of as paths through a tree the nodes of which are elements of the set Sq of finite sequences of natural numbers. The path determined by an α in ${}^{\omega}\omega$ is the set $\{\alpha|k\}_{k\in\omega}$ of initial segments of α.

For any finite sequence s the <u>interval</u> (<u>of</u> <u>Baire</u>) determined by s (in symbols $[s]$) is the set of all infinite sequences which have s as an initial segment; in terms of the tree, $[s]$ corresponds to the set of all infinite paths which pass through s. The set $\{[s]\}_{s\in Sq}$ is a basis for the Baire topology on ${}^{\omega}\omega$. Thus a subset A of ${}^{\omega}\omega$ is open iff every element α of A is in $[s]$ for some s such that $[s] \subseteq A$. In other words, a set A is open iff for any α in A the fact α is a member of A can be 'deduced' from some finite amount of knowledge about α, i.e. from $\alpha|k$ for some k. For example, the set of all sequences with some occurrence of 0 is open, whereas the set of increasing sequences is not.

The Baire topology is also that induced by the metric d defined by

$$d(\alpha,\beta) = 2^{-n}$$

where n is the least k such that $\alpha(k) \neq \beta(k)$ (if no such k exists, $d(\alpha,\beta)$ is 0). It is complete under this metric. The Baire space is not compact; the open cover $\{[\langle n\rangle]\}_{n\in\omega}$ has no finite subcover.

The Baire space is, as we have indicated, homeomorphic to the set of irrational numbers in the interval $(0,1)$ considered as a subspace of $\underline{\underline{R}}$. The correspondence

$$\alpha \quad \langle---\rangle \quad 1/(1+\alpha_0+1/(1+\alpha_1+1/(\ldots)))$$

is a homeomorphism.

The Baire space and its topology arise quite naturally in various disguises in areas of mathematics other than descriptive set theory. We may, for example, think of a sequence α as representing a formal power series

$$\alpha(0) + \alpha(1)x + \alpha(2)x^2 + \alpha(3)x^3 + \ldots$$

with positive integer (or, by another coding, rational) coefficients. Then the formal arithmetic operations correspond to continuous operations on ${}^{\omega}\omega$. For another example, let θ_0, θ_1, θ_2, ... be an enumeration of the sentences of some first order language, so that we can regard an element α of ${}^{\omega}\omega$ as representing the collection $\{\theta_{\alpha(n)}\}_{n\in\omega}$ of formulas. Then the completeness theorem of first order logic states that set of elements of ${}^{\omega}\omega$ representing consistent sets of formulas is open. Of course these observations are quite simple, but it often happens the very simple topological results can yield nontrivial results in other areas.

The elements of ${}^{\omega}\omega$ can therfore be thought of as 'codings' for countably infinite mathematical objects. However, it may be that some objects like this are not best represented by sequences. For example, a relation or directed graph on ω is a subset of $\omega\times\omega$, i.e. an element of ${}^{\omega\times\omega}2$. The natural (i.e. the product) topology on this space is that obtained by defining an interval a set of the form

$$\{\rho\varepsilon{}^{\omega\times\omega}2 \colon r \subset \rho\}$$

for some function r in ${}^{n\times n}2$ for some n. Then in this topology the set, say, of graphs containing a cycle is open, and the set of linear orders is closed. More generally, an n-ary relation on ω is (essentially) an element of ${}^{\omega\times\omega\times\ldots\times\omega}2$, and and n-ary operation is an element of ${}^{\omega\times\omega\times\ldots\times\omega}\omega$. Thus given any finite first order similarity type we can form, by taking products, a space whose elements are structures of ths type with universe ω. For example, an element of

$$ {}^{\omega\times\omega}\omega \times {}^{\omega}2 \times \omega $$

is a structure with a binary operation, a unary operation and a distinguished element. In a sense, however, these spaces are really only conveniences because they are all homeomorphic to one of

$$ \omega, \quad {}^{\omega}2, \quad \omega \times {}^{\omega}2, \quad {}^{\omega}\omega $$

and each in turn is homeomorphic to a closed subset of the Baire space. The Baire space is thus in a sense universal and this constitutes one more argument for restricting ones attention to it; however, in some contexts these other spaces are very convenient as a means to avoid codings.

C. The classical hierarchies over ${}^{\omega}\omega$

In this section we present a very brief introduction to the classical hierarchies of subclasses of $\underline{P}({}^{\omega}\omega)$.

We have already mentioned that descriptive set theory can be understood as the study of the complexity of subsets of the Baire space. Historically, however, work in the field has almost always been formulated in terms of classifications of subsets of ${}^{\omega}\omega$: typically, one might introduce a new subclass of $\underline{P}({}^{\omega}\omega)$ (such as the Borel sets) and then prove some property (such as measurability) of all sets 'simple' enough to be members of the class in question. Usually, descriptive set theorists are concerned not just with individual classes, but with whole indexed families of classes. These families are usually well ordered by inclusion, and so form hierarchies (in the sense of Addison (1962b)).

It is when the classification approach takes the form of the study of hierarchies that it most clearly seen to be the study of complexity. For given any family of subsets of ${}^{\omega}\omega$ we can define a notion of complexity as follows: a subset A of ${}^{\omega}\omega$ is no more complex than a subset B of ${}^{\omega}\omega$ iff A is a member of every class in the family which contains B. Conversely, given a notion of relative complexity, we can consider the family of all subclasses of $\underline{P}({}^{\omega}\omega)$ which are 'initial' or 'closed downwards' with respect to this notion, i.e. all subclasses which contain all sets no more complex than any of their members. This relationship between classification and complexity is made precise in one of our most important results (namely VF8).

The various hierarchies studied by descriptive set theorists are almost always defined inductively, by taking some class of 'base' sets and a collection of set operations and by defining the various levels of the hierarchy to correspond to the various stages in closing the base class out under the operations. The idea is that the base sets are the simplest, and that in general the complexity of a set is proportional to the number of applications of the given set operations required to obtain the set from base sets. Thus after the 'base sets' themselves the next simplest are those which are the result of one of the set operations applied to base sets; and after them, are those sets which are the result of operations applied to sets from the first two levels; and so on.

The Borel hierarchy was probably the first hierarchy over ${}^{\omega}\omega$ (originally, over the set of reals) to be studied by the classical descriptive set theorists. It is generated by closing out the class $\underline{\underline{G}}$ under the operations of countable union and complementation (relative to ${}^{\omega}\omega$).

For example, at level 3 of this hierarchy is the class $\underline{\underline{G}}^3$ of all sets obtainable from open sets using only three applications of countable union and complementation. The elements of $\underline{\underline{G}}^3$ are those of the form

$$\bigcup_i - \bigcup_j - \bigcup_k - G_{i,j,k}$$

with G an $\omega \times \omega \times \omega$ -ary family of open sets. Simplifying, we see that $\underline{\underline{G}}^3$ consists of those sets of the form

$$\bigcup_i \bigcap_j \bigcup_k F_{i,j,k}$$

with F an $\omega \times \omega \times \omega$ -ary family of closed sets. In classical notation, $\underline{\underline{G}}^3$ is the class $\underline{\underline{F}}_{\sigma\delta\sigma}$ of unions of intersections of unions of closed sets. In modern notaion, it is the class of $\underline{\underline{\Sigma}}_4^0$ sets. The modern notation acknowledges the fact that these sets are those of the form

$$\{\alpha\epsilon^{\omega}\omega\colon \exists i\epsilon\omega\ \forall j\epsilon\omega\ \exists k\epsilon\omega\ \forall l\epsilon\omega\ \alpha\ \epsilon\ D_{i,j,k,l}\}$$

for some ${}^4\omega$ -ary family D of clopen sets. In other words, the $\underline{\underline{\Sigma}}_4^0$ sets are those 'definable' using four blocks of alternating number quantifiers, the first (outermost) block being existential. The class of complements of elements of $\underline{\underline{G}}^3$ is (in our notation) $\underline{\underline{F}}^3$. In classical notation $\underline{\underline{F}}^3$ is $\underline{\underline{G}}_{\sigma\delta\sigma}$ and in modern notation the set of $\underline{\underline{\Pi}}_4^0$ sets (for reasons which should be evident). It is rather unfortunate that three different notations should exist for the same hierarchy but it is unavoidable. The classical notation will not be used in theorems but will be used in informal explanations (at lower levels it is the most suggestive). The modern $\underline{\underline{\Sigma}}$-$\underline{\underline{\Pi}}$ notation will be used in stating results because it is the standard. The third notation, however, has proved to be the most convenient, and will also be used in explanations and results. The trouble with the modern notation is that the numbers assigned at finite levels are too great by one.

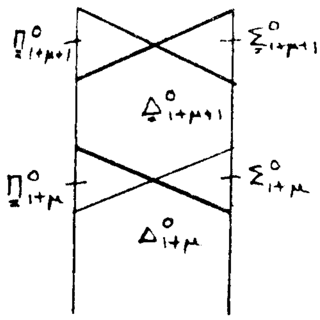
The closing out of the open sets under countable union and complementation requires, of course, Ω many steps. For example the class $\bigsqcup_n \underline{\underline{G}}^n$ (the class of arithmetic sets) is easily seen not to be closed under countable union; if each A_n appears at the n^{th} level but not sooner, we can hardly expect $\bigsqcup_n A_n$ to be in $\underline{\underline{G}}^m$ for any particular m. In general $\underline{\underline{G}}^\mu$ is the class of countable unions of elements of $\{-G\}_{G\epsilon\underline{\underline{G}}^\nu,\nu\epsilon\mu}$. The class of Borel sets is $\bigsqcup_{\mu<\Omega}\underline{\underline{G}}^\mu$ and is easily seen to be closed under the two operations in question.

The $\boldsymbol{\Sigma}$-$\boldsymbol{\Pi}$ notation also extends to give names to the infinite levels but the classical notation does not. To make matters even more complicated, the discrepancy between the numbering of the modern notation and our notation disappears at the infinite levels: in our notation, $\mathbf{G}^{\omega}$ (the collection of countable unions of arithmetic sets) is the class of $\boldsymbol{\Sigma}^0_{\omega}$ sets. The following table may help clear up the confusion.

Modern notation.	Classical notation.	Our notation.
$\boldsymbol{\Sigma}^0_1$, $\boldsymbol{\Pi}^0_1$	$\mathbf{G}$, $\mathbf{F}$	$\mathbf{G}$, $\mathbf{F}$
$\boldsymbol{\Sigma}^0_4$, $\boldsymbol{\Pi}^0_4$	$\mathbf{G}_{\delta\sigma\delta}$, $\mathbf{F}_{\sigma\delta\sigma}$	$\mathbf{G}^3$, $\mathbf{F}^3$
$\boldsymbol{\Sigma}^0_{\omega}$, $\boldsymbol{\Pi}^0_{\omega}$	none	$\mathbf{G}^{\omega}$, $\mathbf{F}^{\omega}$

It is also very helpful to bear in mind the fact that $\mathbf{G}^{\mu}$ is the class of $\boldsymbol{\Sigma}^0_{1+\mu}$ sets, for all μ (finite or infinite) in Ω. We will often be required to use expressions like "$\boldsymbol{\Sigma}^0_{1+\mu}$" in stating our results in modern notation.

The modern notation also gives a name to the class of sets which are both $\boldsymbol{\Sigma}^0_{1+\mu}$ and $\boldsymbol{\Pi}^0_{1+\mu}$; these are the $\boldsymbol{\Delta}^0_{1+\mu}$ sets. (The class of clopen sets is therefore the class of $\boldsymbol{\Delta}^0_1$ sets). The letter "$\boldsymbol{\Delta}$" was chosen because these classes occupy a triangular (or at least pointed area) in the Venn diagram of the $\boldsymbol{\Sigma}^0_{\mu}$ and $\boldsymbol{\Pi}^0_{\mu}$ sets.



A simple cardinality argument shows that not all subsets of ${}^{\omega}\omega$ are Borel (there are $2^{\mathfrak{c}}$ such subsets and only $\mathfrak{c}$ are Borel, $\mathfrak{c}$ being the power of the continuum). Most non Borel sets are of course very complex. There is, however, a class of sets which are 'only just' non Borel. These are the analytic sets, first discovered by the Suslin.

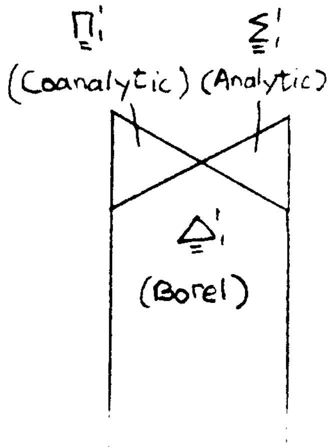
The analytic sets (or, in modern notation, the $\boldsymbol{\Sigma}^1_1$ sets) are those which are projections of closed subsets of the plane; in other words, those of the form

$$\{\alpha \epsilon {}^\omega\omega \colon (\alpha,\beta) \epsilon C \ \text{ for some } \beta\}$$

for some Borel subset C of ${}^\omega\omega \times {}^\omega\omega$. It is not hard to see that the analytic sets are exactly those which are the image of a Borel set under a continuous function, i.e. those of the form $f^*(B)$ for some Borel set B and some continuous function f from ${}^\omega\omega$ to ${}^\omega\omega$. Finally, this class consists of exactly those definable by a simple existential function quantifier; in other words those of the form

$$\{\alpha \epsilon {}^\omega\omega \colon \exists\beta\ \forall n\ R(\alpha,\beta,n)\}$$

for some clopen relation R on ${}^\omega\omega \times {}^\omega\omega \times \omega$. The modern notation is based on this last characterization.



The analytic sets are 'only just' more complicated than the Borel sets in the following sense: a set A is Borel iff both A and -A are analytic. The class of Borel sets therefore occupies the familar triangular area in the Venn diagram of the analytic and coanalytic sets. (In modern notation, the Borel sets are the $\boldsymbol{\Delta}^1_1$ sets).

The collection of analytic and coanalytic sets by no means exhausts the class of Borel sets; together they form only the first level of the projective hierarchy, the hierarchy formed by closing the Borel sets out under the operations of complementation, countable union and projection. For example, at the third level ot this hierarchy (i.e. two levels above the analytic sets) we find the class of $\boldsymbol{\Sigma}^1_3$ sets and its dual, the class of $\boldsymbol{\Pi}^1_3$ sets. The $\boldsymbol{\Sigma}^1_3$ sets are those which are projections of complements of projections of complements of analytic sets. In the older notion, these are the PCPCA sets; the modern notation acknowledges the fact that this class is the collection of all sets of the form

$$\{\alpha \varepsilon {}^{\omega}\omega: \exists\beta\ \forall\gamma\ \exists\delta\ \forall n\ R(\alpha,\beta,\gamma,\delta,n)\}$$

for some clopen subset of ${}^{\omega}\omega \times {}^{\omega}\omega \times {}^{\omega}\omega \times {}^{\omega}\omega \times \omega$. In other words, the class of sets definable using three alternating function quantifiers the first of which is existential. By using countable unions at countable limit ordinals we form the projective hierarchy with Ω levels. The range of this hierarchy is the class of projective sets.

The class of all projective sets is very large (if still less than $\underline{P}({}^{\omega}\omega)$) and its first few levels contain almost all sets used in ordinary analysis. After only the first one or two levels there appear sets which fail to possess the various 'nice' properties enjoyed by all Borel sets. For example, it is consistent with the axioms of set theory that there exists a $\underline{\Delta}^1_2$ set which is not measurable, and an uncountable $\underline{\Pi}^1_1$ set which has no perfect subset. Nevertheless, the projective sets are still simple enough to possess properties not possessed by arbitrary subsets of ${}^{\omega}\omega$. It is thought possible, for example, that it is consistent with ZFC that all projective sets are determinate (determinateness will be discussed in OD).

The projective hierarchy is in a sense the largest and coarsest hierarchy studied by the classical descriptive set theorists. There are, of course, sets which are not projective (assuming even weak forms of AC). We could, for example, study what might be called the higher order hierarchies in which we consider definitions involving higher order quantifiers (the classes of $\underset{=}{\Sigma}^k_\mu$ sets for $k>0$) but these sets are so complex that there are very few general properties shared by them all. Instead, we will follow the classical approach and turn our attention now towards the finer hierarchies which subdivide the Borel hierarchy and give a more precise measure of the complexity of Borel sets.

The first of these subhierarchies was discovered by Hausdorff at the turn of the century, and yields a classification of the class of $\underset{=}{\Delta}^0_2$ sets, i.e. the class of sets which are both $\underset{=}{\Sigma}^0_2$ ($\underset{=}{F}_\sigma$) and $\underset{=}{\Pi}^0_2$ ($\underset{=}{G}_\delta$).

We have already seen that any set which is either $\underset{=}{\Sigma}^0_1$ or $\underset{=}{\Pi}^0_1$ is already both $\underset{=}{\Sigma}^0_2$ and $\underset{=}{\Pi}^0_2$, i.e. that

$$\underline{G}^0 \cup \underline{F}^0 \subseteq \underline{G}^1 \cap \underline{F}^1.$$

It is only natural to ask whether or not there are any $\underset{=}{\Delta}^0_2$ sets which are not open or closed, i.e. (in classical terminology) whether or not there are any sets in the diamond shaped area in the diagram. In fact it is easy to construct such a set: simply take the intersection of two 'appropriate' sets one of which is open and the other of which is closed. The intersection of an open set and a closed set will be $\underset{=}{\Delta}^0_2$ because both the class of $\underset{=}{\Sigma}^0_2$ sets and the class of $\underset{=}{\Pi}^0_2$ sets are closed under intersection; in general, however, such an intersection will be neither open nor closed.

A set which is the intersection of an open set and a closed set is obviously the difference of two closed sets; the class of all such sets (together with its dual) constitutes level 1 of Hausdorff's difference hierarchy over the closed sets. For any n in ω, the class of n-ary differences of closed sets is the class of all sets of the form

$$(F_0-F_1) \cup (F_2-F_3) \cup \ldots \cup (F_{n-1}-F_n)$$

if n is odd, or of the form

$$(F_0-F_1) \cup (F_2-F_3) \cup \cdots \cup F_n$$

if n is even. The sets which are of this form (i.e. which appear in the finite levels of Hausdorff's hierarchy) are exactly the sets which are Boolean combinations of open (and closed) sets.

The finite levels of the hierarchy do not, however, exhaust the class of $\underset{\sim}{\Delta}_2^0$ sets, because (as can be shown) there are $\underset{\sim}{\Delta}_2^0$ sets which cannot be obtained from open and closed sets using only finite unions and intersections. It is necessary to extend the hierarchy through all countable ordinal levels by considering differences of sequences of closed sets of length greater than ω. In so doing it is convenient to consider only sequences F which are antichains (i.e. $F_\nu \supseteq G_\eta$ if $\nu \leq \eta$). For example, if F is a sequence of this type of length $\omega+3$, its difference is

$$(F_0-F_1) \cup (F_2-F_3) \cup \cdots \cup (F_\omega-F_{\omega+1}) \cup (F_{\omega+2}-F_{\omega+3}).$$

Hausdorff proved that the Ω levels of this difference hierarchy exhaust the class of $\underset{\sim}{\Delta}_2^0$ sets.

Hausdorff's result is slightly unsatisfying, however, in the sense that it yields a hierarchy but no construction principle: the levels of the difference hierarchy are not (or do not appear to be) simply the stages in closing the class of open and closed sets out under some ω-ary operations. Nevertheless there is in fact an operation which yields the Hausdorff hierarchy, namely the operation of separated union discovered by Addison. In general a set B is said to be a $\underset{\sim}{\Sigma}_1^0$-separated union of sets in a class $\underset{\sim}{A}$ iff B is the union $\bigsqcup_{i\in\omega} A_i$ of some sequence A

of sets in $\underset{=}{A}$ for which there exists a sequence G of disjoint open sets (the separating sets) such that $A_i \subseteq G_i$ for all i. The existence of the sequence G insures that the components of A are 'far enough apart' to prevent their union becoming too complex.

It will be shown (in IVE) that the class of $\underset{=}{\Delta}_2^0$ sets is that generated by closing the clopen sets out under the operation of separated union. This construction principle can be 'lifted' (using Kuratowski's (α,β)-homeomorphisms) to yield a construcion principle for the class of $\underset{=}{\Delta}_{1+\mu}^0$ sets. In chapter IV we present this generalization in a form which gives a hierarchy for the class of $\underset{=}{\Delta}_\lambda^0$ sets with λ an infinite limit ordinal (thereby solving a long standing open problem posed by Lusin). The generalization involves separated unions in which the separating sets themselves are taken form correspondingly higher levels of the Borel hierarchy.

The last classical hierarchy which we describe covers the one case omitted by the previous results: the class $\underset{=}{F} \cap \underset{=}{G}$ of $\underset{=}{\Delta}_1^0$ (clopen) sets. This was discovered by Barnes in (1966) although Kalmar in (1957) proved essentially the same result in a different setting.

The Barnes-Kalmar result is very simple: the class of $\underset{=}{\Delta}_1^0$ sets is the result of closing $\{\emptyset, {}^\omega\omega\}$ out under the <u>join</u> operation. The join of an ω-sequence A of sets is

$$\{\langle n,\alpha(0),\alpha(1),\alpha(2),\ldots\rangle\}_{\alpha\in A_n}$$

This operation is described more fully in III C.

The proof that the Kalmar hierarchy exhausts the class of $\underset{=}{\Delta}_1^0$ sets is surprisingly simple. Suppose that A is a $\underset{=}{\Delta}_1^0$ set but not in the Kalmar hierarchy. Since A is the join of the details $\langle A_{(\langle n\rangle)}\rangle_{n\in\omega}$, these cannot all be Kalmar sets (i.e. appear in the Kalmar hierarchy). (The detail $C_{(s)}$ of a set C with respect to a finite sequence s is $\{\gamma\in{}^\omega\omega\colon s\gamma \in C\}$).

Therefore, there must be an m in ω such that $A_{(\langle m\rangle)}$ is not Kalmar. In the same way, since $A_{(\langle m\rangle)}$ is the Kalmar union of $\langle A_{(\langle m,n\rangle)}\rangle_{n\epsilon\omega}$, there must be an m' such that $A_{(\langle m,m'\rangle)}$ is not Kalmar. Similarly, there must be an m'' such that $A_{(\langle m,m',m''\rangle)}$ is not Kalmar. In this way we can use DC to construct an infinite sequence α ($= \langle m,m',m'',\ldots\rangle$) such that $A_{(\alpha|n)}$ is not Kalmar for any n. This is, of course, impossible, because A is clopen and so $A_{(\alpha|n)} = \emptyset$ or $A_{(\alpha|n)} = {}^{\omega}\omega$ for large enough n.

The Kalmar hierarchy has some claim to represent the 'last word' in complexity of clopen sets. The hierarchy is not bilateral, so that there are no small diamond shaped areas in the Vens diagram of the hierarchy to fill in with subhierarchies. We cannot, on these grounds alone, rule out the existence of subhierarchies of the Kalmar hierarchy; but it is hard to see how one might distinguish for example between sets in

$$\{\{\alpha\epsilon{}^{\omega}\omega\colon \alpha(0)\ \epsilon\ M\}\}_{M\subseteq\omega}$$

on the basis of 'complexity'. In later chapters we will provide further evidence for considering the levels of the Kalmar hierarchy as representing the finest possible classification of the clopen sets.

The join operation can also be used to give a hierarchy and construction principle for the class of sets A such that both A and -A are differences of open sets. It can be shown that this class is the result of closing the class $\underline{G} \cup \underline{F}$ of open and closed sets out under the operation of Kalmar union. This yields a hierarchy with $\Omega+\Omega$ levels, with the first Ω levels the Kalmar hierarchy of clopen sets and the second Ω levels exhausting the diamond-shaped class

$$(Df_2(\underline{G}) \cap Df_2(\underline{G})^-) - (\underline{G} \cup \underline{F})$$

The analogous result holds at all levels of the difference hierarchy. In general, the class $Df_\mu(\underset{=}{F}) \cap Df_\mu(\underset{=}{F})^-$ is the closure of

$$\bigsqcup_{\nu<\mu}(Df_\nu(\underset{=}{G}) \cap Df_\nu(\underset{=}{G})^-)$$

under the Kalmar union operation (μ a countable ordinal). This yields a hierarchy over the class of $\underset{=}{\Delta}_2^0$ sets with Ω^2 levels. As before, this hierarchy can reasonably be considered the finest one possible.

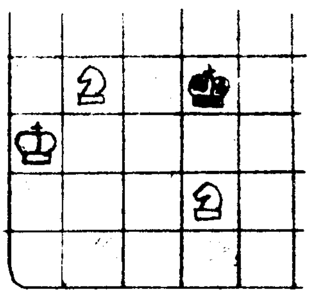
The main goal of this work can be thought of as determining the nature (in particular, the order type) of the 'finest possible' hierarchy over the entire class of Borel sets. In later chapters the Kalmar subhierarchies will be considered as a single unit, so that the hierarchy over the $\underset{=}{\Delta}_2^0$ sets will have order type Ω , not Ω^2. We will show, for example, that the order type of the 'finest' hierarchy over the class of $\underset{=}{\Delta}_3^0$ sets is Ω^Ω (not Ω^3). As another example, we will show that the 'finest' hierarchy over the class of $\underset{=}{\Delta}_\omega^0$ sets has order type $\epsilon_{\Omega+\Omega}$ ($\epsilon_{\Omega+\Omega}$ being the $\Omega+\Omega^{th}$ epsilon number).

D. Infinite games

In this section we present some useful background material about infinite games and the axiom of determinateness.

An infinite game is a game in which a particular 'round' need not terminate so that the outcome of the contest can be determined only by examining the entire history of the contest.

Simple but interesting examples of infinite games can be constructed by extending the standard chessboard infinitely in one or more directions, and by suitably modifying the rules. Consider, for example, the position shown in the diagram (the board extends infinitely in the directions of the dots). White's goal is to checkmate Black's King, and Black's goal is to avoid checkmate. This game is genuinely infinite because Black can win, but cannot achieve certain victory after any finite number of moves, i.e. we cannot in general conclude that Black has won without examining the entire record of the game.



This game nevertheless has a finite aspect in that one of the players (white) cannot win without terminating the game: we might call such a game "half-finite". But it is easy to devise games which are not even half-finite. We could, for example, retain the above board but change the rules so that White's goal is instead to get arbitrarily far away from the Black King, i.e. to play so that no matter how large an integer n is there will be a point in the game after which White's king will never be less than n moves away from the

Black king. Then clearly neither player can ever win in any finite number of moves, and it will always be necessary to look at the entire history of the game to determine the winner.

It is not difficult to give a precise definition of the notion of 'infinite game' provided we restrict ourselves to games in which (i) there are only two players, I and II, who move alternately; (ii) each player on each move has only countably many choices for his next play; and (iii) there are no infinite stages in the game (i.e. all 'rounds' are of length ω). Then the history of a particular round of such a game can be 'coded up' as a pair (α,β) of elements of ${}^{\omega}\omega$, and thus the game itself is completely determined by the subset W of $\omega \times \omega$ consisting of all histories of codes of bouts in which II is the winner. We therefore assume for the sake of simplicity that each player on each move plays a natural number, and define the game to be the set W itself.

(D1) DEFINITION. An (infinite) <u>game</u> is a subset of $\omega \times \omega$.

We can now formalize the notion of a strategy. It is clear that a strategy for II in one of these games is essentially a function from Sq to ω which takes as its argument the finite sequence $\langle\alpha(0),\alpha(1),\ldots,\alpha(n)\rangle$ (for each n) of I's first n moves and gives as its result player II's n^{th} move $\beta(n)$. For our purposes, however, it is more convenient to have a strategy for II yield the entire history $\langle\beta(0),\beta(1),\ldots,\beta(n)\rangle$ of II's moves up to that point. Strategies for player I are defined similarly.

(D2) DEFINITION.

(i) A strategy for II is a monotonic function τ from Sq to Sq such that

$$\ln(\tau(s)) = \ln(s)$$

for every s in Sq;

(ii) a strategy for I is a monotonic function σ from Sq to Sq such that

$$\ln(\sigma(s)) = \ln(s)+1$$

for every s in Sq.

It should be noted that this definition implies that the games we are studying are games of perfect information, i.e. games in which each player has complete knowledge of his opponents moves up to that point.

Now if τ is a strategy for II, we let $\tau^{\sim}$ denote the corresponding function from ${}^{\omega}\omega$ to ${}^{\omega}\omega$ which takes as its argument the entire history of player I's moves and gives as its result the entire history of II's moves. Thus τ is a winning strategy for II for the game W iff $(\alpha,\tau^{\sim}(\alpha))$ is in W for every α; and the notion of a winning strategy for I is similarly defined.

(D3) DEFINITION. For any monotonic function τ from Sq to Sq and any α in ${}^{\omega}\omega$

$$\tau^{\sim}(\alpha) = \bigsqcup_{k\in\omega}\tau(\alpha|k).$$

Note that if τ is a strategy (either for I or for II) then $\tau^{\sim}(\alpha)$ will be in ${}^{\omega}\omega$ for every α.

(D4) DEFINITION. For any game W

(i) a winning strategy for II for W is a strategy τ for II such that $(\alpha,\tau^{\sim}(\alpha)) \in W$ for every α in ${}^{\omega}\omega$;

(ii) a winning strategy for I for W is a strategy σ for I such that $(\sigma^{\sim}(\beta),\beta) \in -W$ for every β in ${}^{\omega}\omega$.

ED

The study of infinite gamesalmost always concerns, in some way or another, the question of determinateness: a game is determinate iff one of the players has a winning strategy (i.e. if the game 'determines' a winner). Since every finite game is determinate, and since also draws are not possible in infinite games (as we have defined them) it might seem plausible to conclude that every infinite game is determinate. This conclusion is, however, not justified.

Now it is certainly true that it cannot be the case that both player I and II have winning strategies for a game W. Given two strategies σ and τ for I and II respectively, we can 'play them off' against each other and form a unique element (α,β) of $\omega \times \omega$ (called by Addison the clash of σ and τ) such that $\alpha = \sigma^{\sim}(\beta)$ and $\beta = \tau^{\sim}(\alpha)$. Then if both σ and τ were winning strategies the clash (α,β) (which is equal to both $(\alpha,\tau^{\sim}(\alpha))$ and $(\sigma^{\sim}(\beta),\beta)$) would have to be in both W and $-W$, impossible. Thus given two strategies for I and II respectively, one of them must be 'superior' to the other.

This argument does not, however, imply that every game is determinate. It may be that given any strategy for player I, player II has a strategy which is superior, but that given any strategy for player II, player I has a strategy which is superior.

In fact it is possible, using the unrestricted axiom of choice, to construct (by 'diagonalizing' over strategies) a game which is not determinate.

Despite these considerations descriptive set theorists have in recent years devoted a great deal of attention to the "axiom of determinateness", which asserts that every infinite game (i.e. subset of ${}^{\omega}\omega \times {}^{\omega}\omega$) is determinate. One of the reasons for this interest is that AD is natural and plausible (unlike most large cardinal axioms) and yet, though this is not obvious, it settles a great many questions (such as the continuum hypothesis) which are known to be independent of the axioms of ZF or ZFC. For example, AD implies that every set of real numbers is measurable, has the property of Baire, and is either countable or contains a perfect subset (see Mycielski (1964)).

The main problem with AD is the fact, mentioned above, that it is inconsistent with AC. This state of affairs can be taken as informal evidence for the 'falseness' of AD and for the incorrectness of our intuitions concerning infinite games; but it can just as easily be taken as (yet more) evidence for the falseness or unreasonableness of the unrestricted axiom of choice. At any rate, a great deal of effort has been expended on trying to resolve this contradiction, either by replacing AC by one of its weaker forms (usually the axiom of Dependent Choice (DC)), or by weakening AD so that it asserts the determinateness of some restricted collection of games, usually a collection (such as the collection of projective subsets of ${}^{\omega}\omega \times {}^{\omega}\omega$) whose elements are in some sense 'definable' or 'constructable'. Naturally, if AD is weakened its consequences may also be weakened. For example, from the assumption that all projective games are determinate we may only be able to conclude that all projective sets are measurable.

One of the most important results concerning determinateness is the theorem of D.A. Martin (1975) that every Borel set is determinate (assuming ZF+DC). Most of the results in this dissertation require Borel Determinateness (BD), and in fact were originally proved 'modulo' BD, before Martin's proof. Martins's result goes a long way towards establishing the 'credibility' of AD, but nevertheless it still cannot be said that the status of the axiom has been finalized: it is still unknown whether or not it is consistent with ZF+DC to assume that every analytic set is determinate (though Martin (1970) has shown that analytic determinateness follows from the existence of a measurable cardinal).

Although we will make heavy use of some results about determinateness (mainly BD), this dissertation is not primarily concerned with AD per se and therefore we will not give more details about AD itself and its relation to higher cardinal axioms, the projective sets and so on. However, to illustrate some of the ideas discussed we conclude this section by presenting a proof that all open games are determinate. This result was first obtained by Gale and Stewart (1953); it can be regarded as the first small step on the way to establishing BD.

(D5) THEOREM. Every open game is determinate.

PROOF. (Outline). Let W be an open subset of ${}^{\omega}\omega \times {}^{\omega}\omega$; we must show that W is determinate. We do this by proving a stronger result, namely that every position for I in W is either a win for I or a win for II. The required result follows easily then from the fact that the position $(\emptyset,\emptyset)$ is a win for one of the players.

We show that all positions for I (i.e. all ordered pairs (s,t) of finite sequences of the same length) are 'determinate' by (i) giving an inductive definition of a class of positions all of which are winning positions for II; then (ii) showing that all the remaining position are winning for I.

The 'base step' in the inductive definition is to take all those positions from which II has 'already' won; that is, the class P_0 of all (s,t) (s and t having the same length) such that $[s] \times [t] \subseteq W$. In other words, P_0 is the class of positions for which II is guaranteed to win no matter <u>how</u> he plays in the rest of the game. Obviously, all elements of P_0 are winning positions for II in W.

The next stage in the induction is to take the class P_1 of all positions from which II can enter P_0 in one move; i.e. the set $\{(s,t) : \forall i \varepsilon \omega\ \exists j \varepsilon \omega\ (si,tj) \ \varepsilon\ P_0\}$. Clearly, all elements of P_1 are winning positions for II in W.

We then proceed to define P_2 as all positions from which II can force entry to P_1 in one move, P_3 as all those from which entry to P_2 can be forced in one move, and so on.

The reader will not be surprised to learn that the construction must be carried on through Ω levels, i.e. we must define P_μ for every countable ordinal μ. The class P_ω, for example, is the set of all positions from which II can force entry into some P_n $(n \ \varepsilon\ \omega)$ in one move. In general, P_μ is the set

$$\{(s,t) : \forall i \varepsilon \omega\ \exists j \varepsilon \omega\ \exists \nu \varepsilon \mu\ (si,tj) \ \varepsilon\ P_\nu\}$$

of all positions from which II can force entry into some P_ν $(\nu \ \varepsilon\ \mu)$ in one move.

We take P_Ω $(=\bigsqcup_{\mu\epsilon\Omega} P_\mu)$ as the desired class of winning positions for II. A simple induction shows that they are indeed winning positions.

Now let (s,t) be an element of $-P_\Omega$. It is not hard to see that Player I can avoid entering P_Ω, at least on the next move. Suppose other wise; then for any move i of Player I, there is a move j of Player II such that the resulting position (si,tj) is in P_{ν_j} for some ν_j. But this in turn implies that (s,t) itself is in P_Ω, because it must be in P_μ where $\mu = \bigsqcup_j \nu_j$. This is of course impossible.

Thus if player I is outside P_Ω he can play from move to move to stay outside P_Ω and can do so indefinitely. It is not hard to see that this constitutes a winning strategy for him. If (α,β) is a final position resulting from the use of such a strategy, then because W is open we cannot have $(\alpha,\beta) \epsilon W$ unless $[\alpha|k] \times [\beta|k] \subseteq W$ for some (in fact all but finitely many) k. In other words, $(\alpha,\beta) \epsilon W$ iff $(\alpha|k,\beta|k) \epsilon P_0$ for some k. This means that II cannot win unless he actually 'enters' W at some finite stage; open games are 'half-finite' in the sense of our earlier discussion. Thus any strategy which avoids P_0 (and therefore any which avoids P_Ω) also avoids W and is a winning strategy for I. $\square$

The proof given here is due to Blackwell, though it has appeared in disguised versions in various contexts. The proof gives a good example of the use of ordinals in game arguments, a technique we will make good use of in later chapters. We can think of the ordinals in the proof as measuring the amount of <u>time</u> it takes II to win from a position in P_Ω. If a position is in P_0, II has already won, no matter how either player might play in the future. If a position is in P_1, it means that II can (if he plays correctly) win in one move. In the same way, if a position is in (say) P_{25}, it means that II (if he plays correctly) can win in at most 25 moves.

This interpretation can also be extended to infinite ordinals. If a position is in P_ω it means that after one move II will be able to 'call' the game, i.e. predict how many more moves he will require to win. If a position is in $P_{\omega+3}$, it means that after 4 moves II will be able to call the game. And if a position is in $P_{\omega+\omega+6}$ it means that after 7 moves II will be able to say how many moves it will take before he will be able to call the game.

The proof given is therefore based on the principle that if II can win the open game W, the time it will take (in general) can be measured by some countable ordinal.

This proof in fact shows that II has a winning strategy for an open game W iff he has a winning strategy for an auxiliary game W' which is the same as W except that (i) on the n^{th} move II also plays a countable ordinal μ_n; (ii) if $\mu_n > 0$ then $\mu_n > \mu_{n+1}$; and (iii) II must enter P_0 before μ_n becomes 0.

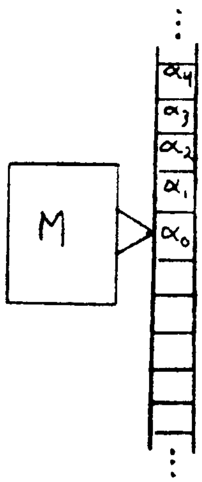
$$\begin{array}{ccc} \mathrm{I} & \mathrm{II} & \\ \alpha_0 & \beta_0 & \mu_0 \\ & & \vee \\ \alpha_1 & \beta_1 & \mu_1 \\ & & \vee \\ \alpha_2 & \beta_2 & \mu_2 \\ \vdots & \vdots & \vdots \end{array}$$

Auxiliary games of this type are very important in many game arguments. The auxiliary ordinal can be considered as a type of clock which imposes a 'time limit' on II.

E. The algorithmic description of the Baire topology

We mentioned that elements of ${}^{\omega}\omega$ can be thought of as codes for countably infinite objects in the same way that elements of ω can be thought of as codes for finite objects, and that descriptive set theory can be thought of as the study of the complexity of sets of infinite objects, in the same way that recursive function theory can be thought of as the study of the complexity of sets of finite objects. Just the same, it might seem unlikely that there would be much connection between, on the one hand, the continuous notions of topology, and on the other the discrete notions of recursive function theory. But in fact the Baire topology and its derived notions can be regarded as natural generalizations of the basic concepts of recursive function theory on ω; in this section we present a very informal view (due to Addison) of the analogy.

We begin by trying to discover which subsets of ${}^{\omega}\omega$ deserve to be called recursive. We know that a subset a of ω is recursive iff there exists a Turing machine which tests for membership in a, i.e. a Turing machine M which, when started on a tape with (a given representation of) a natural number n, eventually halts and prints (say) a "1" if n is in a, otherwise a "0". The fact that a Turing machine tape is infinite and can therefore have an infinite sequence of numbers written on it allows us to carry this definition over directly: we will call a subset A of ${}^{\omega}\omega$ recursive iff there is Turing machine M such that, for any α in ${}^{\omega}\omega$ if M is started on a tape with α



(written on it extending infinitely to the right, with blank squares extending infinitely to the left), then M eventually halts, printing on the last step a '1' if α is in A, otherwise a "0".

For example, the set

$$\{\alpha\varepsilon{}^{\omega}\omega\colon \alpha(2) = 3 \text{ and } \alpha(5) = 0\}$$

is clearly recursive: the machine checks the second and fifth values of α, then gives its answer. The sets $\{\alpha\colon \alpha(2) \leq \alpha(3)\}$ and $\{\alpha\colon \alpha(\alpha(0)) = 5\}$ are also recursive, but not the set $\{\alpha\varepsilon{}^{\omega}\omega\colon \forall n\ \alpha(n) < \alpha(n+1)\}$, because there are an infinite number of comparisons to be checked.

Now suppose that A is a recursive subset of ${}^{\omega}\omega$ such that the Turing machine M tests for membership in A, and that M, having been started with a tape with α written on it, has just halted, and has announced, by writing a "1", that α is in A. A Turing machine computation is finite; therefore M can have examined only a finite number of values of α, say $\alpha(0), \alpha(1), \ldots, \alpha(k-1)$, and on the basis of those values alone concluded that α was in A. This means that any α' which agrees with α for those arguments (i.e. any α' in $[\alpha|k]$) must also be in A, because M's computations when started on a tape with α' on it will be identical. In other words, $\alpha \varepsilon A$ implies that $[\alpha|k] \subseteq A$ for some k - so that A must be open. Similarly, if $\alpha \varepsilon -A$ the machine M will halt and print a "1" after examining only some k values of α, and thus $[\alpha|k] \subseteq -A$. Therefore the complement of A is also open, i.e. A is also closed.

The fact that Turing machine computations are finite therefore implies that every recursive set is both closed and open, i.e. is clopen. It is not true, however, that every clopen set is recursive -

for example, if a is a nonrecursive subset of ω, the set $\{\alpha: \alpha(0) \in a\}$ is clopen but not recursive, because a decision procedure for the latter would yield a decision procedure for a itself. Nevertheless, the set defined above is in a sense 'decidable': to determine whether or not α is in the set, you simply 'examine' $\alpha(0)$, whereas to determine whether or not α is (say) increasing, one must look at an infinite number of values of α.

We can make precise this more general notion of decidability by allowing our Turing machines to have access to a countably infinite "data base", in the form of an extra tape with an element of ${}^{\omega}\omega$. For any subset A of ${}^{\omega}\omega$ and any element δ of ${}^{\omega}\omega$, we say that A is recursive in δ iff there is a two tape Turing machine which decides membership in A (as above), provided it always starts with δ on its extra tape. For example, the set $\{\alpha\in{}^{\omega}\omega: \alpha(0) \in a\}$ is recursive in the characteristic function of a, i.e. in δ where for each n, $\delta(n)$ is 0 if n is in a, otherwise 1.

Computations on the augmented Turing machines are still finite, and so any set recursive in some δ will still be clopen. On the other hand, if A is clopen, let $\langle s_n\rangle_{n\in\omega}$ be some recursive enumeration of Sq, and let δ be a 'code' for A, as follows:

$$\delta(n) = \begin{cases} 0 & \text{if } [s_n] \subseteq -A; \\ 1 & \text{if } [s_n] \subseteq A; \\ 2 & \text{otherwise.} \end{cases}$$

Then it is easy to see that A is recursive in δ to determine whether or not α is in A, the machine examines larger and larger initial segments of α until it finds one which δ codes as a 0 or a 1. The fact that A is open ensures that the computation will always terminate, and we see therefore that the clopen sets are exactly those which are recursive in some element of ${}^{\omega}\omega$

It follows, for example, that the set $\{\alpha: \forall n\ \alpha(n) < \alpha(n+1)\}$ of increasing α cannot be recursive in any δ because it is not open. To use Addison's terminology, deciding membership in a set like $\{\alpha: \alpha(0) \in a\}$ requires mere rote "knowledge" (of a), but deciding whether or not an element of ${}^{\omega}\omega$ is increasing requires genuine "wisdom".

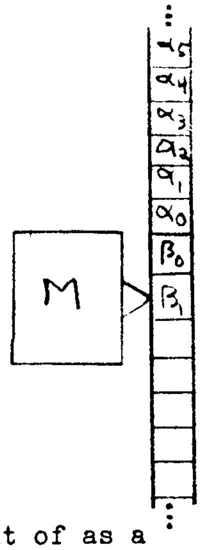
We now consider which subsets of ${}^{\omega}\omega$ deserve to be called recursively enumerable (re). According to the usual definition, a set of natural numbers is re iff there exists a Turing machine which enumerates the set. This definition does not seem to carry over to subsets of ${}^{\omega}\omega$ because it is not clear how a Turing machine can 'enumerate' a possibly uncountably infinite set. There is, however, an alternate definition of the class of re subsets of ω: a subset a of ω is re iff there exists a Turing machine which accepts a, i.e. iff there is a machine such that, when started with a natural number n written on its tape, eventually prints a "1" and halts iff n is in a.

This alternate definition carries over easily, and so we say that a subset A of ${}^{\omega}\omega$ is re iff there is a Turing machine M such that for any α in ${}^{\omega}\omega$ if M is started with α written on its tape, then M eventually halts and writes a "1" iff α is in A. As before, if A is re and α is in A, then will accept α after having examined only some finite initial segment $\alpha(0), \alpha(1), \ldots, \alpha(k-1)$ of α, and so every element of $[\alpha|k]$ will also be accepted. Thus for any α, if $\alpha \in A$ then $[\alpha|k] \subseteq A$ for some k, and so A is open.

The fact that Turing machine computations are finite therefore implies that every re set is open. It is easy to see that the converse is not true; if a is not re, neither is $\{\alpha\epsilon {}^{\omega}\omega\colon \alpha(0) \in a\}$ (even though it is recursive in some δ). Therefore, for any subset A of ${}^{\omega}\omega$ and element δ of ${}^{\omega}\omega$, we say that A is re in δ iff A is the set accepted (as above) by some two tape Turing machine which is always started with the "data base" δ on its second tape. Every set re in some δ is clearly still open; and conversely, if A is open and δ is some code for the set $\{s\epsilon Sq\colon [s] \subseteq A\}$, then it is not hard to see that A is re in δ Thus the class of subsets of ${}^{\omega}\omega$ which are re in some element of ${}^{\omega}\omega$ is exactly the class of open subsets of ${}^{\omega}\omega$

Finally, we consider which functions from ${}^{\omega}\omega$ to ${}^{\omega}\omega$ deserve to be called recursive. According to the usual definition, a function f from ω to ω is recursive iff there is a Turing machine which computes f, i.e. when started with a natural number n written on its tape, it eventually prints $f(n)$ on the tape and halts. This definition cannot be carried over directly because although a machine to 'compute' a function f from ${}^{\omega}\omega$ to ${}^{\omega}\omega$ can be presented with an argument α in ${}^{\omega}\omega$ it cannot print out the entire result β $(= f(\alpha))$ before halting. The machine, however, cannot examine all of α at once, it can only examine α value by value. It is therefore reasonable to require only that the machine print out the values of β in the same way, one by one.

We therefore define a function f from ${}^{\omega}\omega$ to ${}^{\omega}\omega$ to be recursive iff there exists a Turing machine M such that when started with a tape with α written on it, computes for a while, then prints $f(\alpha)(0)$ (in some special format, say, to distinguish it from intermediate results), computes for a while, then prints $f(\alpha)(1)$, then later $f(\alpha)(2)$, and so on, forever, without halting.



The machine which computes f may also be thought of as a continuously operating process or factory, with values of α being fed one by one in one end, and values of β $(= f(\alpha))$ being produced one by one out the other end. Note, however, that the values of β may be turned out at a different rate (perhaps slower) than the rate at which values of α are being fed in, and the factory may require storage because computing each value of β may require access to any of the values of α already fed in.

The machine M which computes a recursive function f from ${}^{\omega}\omega$ to ${}^{\omega}\omega$ operates without halting, but nevertheless any finite number of output values will be produced after only a finite number of steps. Therefore, for any input α, if M computes m values of $f(\alpha)$ to be the finite sequence s $(= \langle s_0, s_1, \ldots, s_{m-1}\rangle)$, it can have examined only a finite number $\alpha(0)$, $\alpha(1)$, ..., $\alpha(k-1)$ of values of α before printing out the first m values of $f(\alpha)$, and so any α' which agrees with α on its first k values (i.e. any α' in $[\alpha|k]$) will also have s as the first m values of $f(\alpha')$. The set of α for which

$f(\alpha)$ begins with s is $f^{-1}([s])$; we have therefore shown that $\alpha \in f^{-1}([s])$ implies that $[\alpha|k] \subseteq f^{-1}([s])$ for some k, i.e. we have shown that $f^{-1}([s])$ is open. Since the inverse image of every interval is open, the inverse image of every open set is open and so f is <u>continuous</u>.

We have therefore shown that every recursive function is continuous. The converse is of course, not true, but if we define the notion of "recursive" by allowing, as before, machines with 'data bases' it is easy to see first, that every function recursive in some δ is continuous, and second, if f is continuous, then f is recursive in a δ which codes the set $\{(s,t): [s] \subseteq f^{-1}([t])\}$. Thus the recursive functions are exactly those which are recursive in some element of ${}^{\omega}\omega$.

So we have a well defined analogy between topological notions and computability notions, in which clopen sets correspond to recursive sets, open sets correspond to re sets and continuous functions correspond to recursive functions. This analogy can in fact be carried much further, so that for example, the Borel sets correspond to hyperarithmetic sets, the analytic sets correspond to the $\underset{\sim}{\Sigma}^1_1$ sets and Baire measurable functions correspond to hyperarithmetic functions.

The significance of the ordinal indices of the levels of the Kalmar hierarchy can be explained in terms of Turing machines. The Kalmar indices can be understood as 'bounds' on the number of values which the machine which decides membership in A has to read in before halting. We equip our Turing machines with an extra 'counter' register capable of holding a countable ordinal. The contents of the register must be decreased every time a component of α is read in, and the machine must halt when the ordinal stored becomes 0. It can be shown that a set A is in level μ of the Kalmar hierarchy iff there is a Turing machine M with a counter which is able to decide membership in A provided its counter initially contains μ.