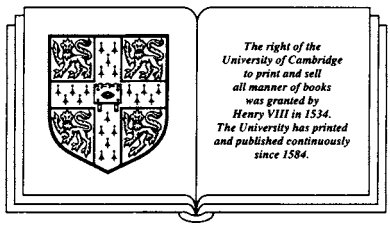
London Mathematical Society Lecture Note Series. 152

Oligomorphic Permutation Groups

Peter J. Cameron
School of Mathematical Sciences, Queen Mary and Westfield College



CAMBRIDGE UNIVERSITY PRESS
Cambridge
New York Port Chester Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011, USA
10, Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1990

Library of Congress cataloguing in publication data available

British Library cataloguing in publication data available

ISBN 0 521 38836 8

Transferred to digital printing 2001

Contents

1. Background

1.1. HISTORY AND NOTATION

In the summer of 1988, a London Mathematical Society symposium was held in Durham on "Model Theory and Groups", organised by Wilfrid Hodges, Otto Kegel and Heter Neumann. This volume of lecture notes is based on the series of lectures I gave at the symposium, but is something more: since no Proceedings of the symposium was published, I have taken the opportunity to incorporate parts of the talks given by other participants, especially David Evans, Udi Hrushovski, Dugald Macpherson, Heter Neumann, Simon Thomas and Boris Zil'ber. (A talk by Richard Kaye revealed new horizons to me which I have not fully assimilated; but Richard's own book should appear soon.) In addition, I have made use of parts of the proceedings of the Oxford–QMC seminar on the same subject which ran weekly in 1987–8 and continues once a term (now as the Oxford–QMW seminar!); contributions by Samson Adeleke, Jacinta Covington, Angus Macintyre and John Truss have been especially valuable to me.

Why model theory and groups? In particular, why the special class of permutation groups considered here?

In the middle 1970s, when my interests were entirely finite, John McDermott asked a question about the relationship between transitivity on ordered and unordered n-tuples for infinite permutation groups. The analogous question, and more besides, had been settled for finite permutation groups by Livingstone and Wagner (1965), with techniques which were largely combinatorial and representation-theoretic, and so not likely to be useful here. McDermott himself had constructed some examples showing that the infinite is very different from the finite.

At that time, "infinite permutation groups" could scarcely be described as a subject. In the Mathematical Reviews classification, permutation groups were explicitly finite.

The only work of substance was Wielandt's Tübingen lecture notes (1959), which was not readily available. Moreover, of the few results which were in the literature, a substantial proportion were by topologists (such as Anderson (1958) and Brown (1959)), and relatively unknown to group theorists.

Another symptom of the situation is illustrated by the following three theorems.

Tits (1952): There is no infinite 4-transitive group in which the stabiliser of 4 points is trivial.

Hall (1954): There is no infinite 4-transitive group in which the stabiliser of 4 points is finite of odd order.

Yoshizawa (1979): There is no infinite 4-transitive group in which the stabiliser of 4 points is finite.

Yoshizawa's theorem is not so much harder than the other two; why, then, the quarter-century gap? In the first two cases, the theorems were regarded as little more than footnotes to the complete determination of finite permutation groups with the same properties (viz. small symmetric, alternating and Mathieu groups). A finite version of Yoshizawa's theorem would be a list of all 4-transitive groups. This was out of reach in the 1950s and 1960s, and by the 1970s it was clear that it would be obtained as a corollary of the classification of finite simple groups; this duly happened in 1980. But the imminence of this classification had also made people think that interesting problems on infinite permutation groups might be waiting.

In addition, there was pressure from outside, especially from model theory. Questions that arise naturally in classification theory and enumeration of models lead to problems about structures with large automorphism groups. Work of Fraïssé and his school (notably Frasnay and Pouzet) leads in the same direction. (See Fraïssé (1986).) Another contribution was from Joyal (1981), who was developing a subject which included "Redfield-Pólya enumeration without groups".

To return to my personal narrative. I was able to answer John McDermott's question and give a classification of permutation groups which are transitive on unordered n-sets for all n. I spoke about this in Oxford, and in the pub afterwards Graham Higman said, "What about groups with finitely many orbits on n-sets for all n? That might be a good topic for a research student." I have researched and studied this topic on-and-off since then, and now I present my thesis.

A permutation group on an infinite set is called oligomorphic if it satisfies the condition of Higman's question (or, equivalently, if it has only finitely many orbits on n-tuples for all n). The main connections between oligomorphic permutation groups and the areas of model theory and combinatorics are provided by two key results which will be described further in Chapter 2, but which can be stated loosely as

follows:

Ryll-Nardzewski's Theorem: A countable (first-order) structure is axiomatisable (that is, characterised, up to isomorphism, as a countable structure, by first-order sentences) if and only if its automorphism group is oligomorphic.

Fraïssé's Theorem: The problem of calculating the numbers of orbits, on n-sets or on n-tuples, of oligomorphic permutation groups is equivalent to that of enumerating the unlabelled or labelled structures in certain classes of finite structures (characterised, more-or-less, by the amalgamation property).

These two results provide the central theme of my lecture notes.

The remainder of this chapter tries to provide a crash course in some of the techniques needed later: first, the two areas principally involved, namely permutation groups and model theory; then, two areas which provide important tools, namely category and measure, and Ramsey theory. The relevant sections can safely be skipped by an expert. Chapter 2 presents the basic properties of oligomorphic permutation groups, and their connection with the theorems of Ryll-Nardzewski and Fraïssé. The third chapter discusses properties of the sequences enumerating orbits on n-sets or n-tuples, especially their growth rates. In Chapter 4 I turn to subgroup theorems explaining how techniques of measure and category, combined with Fraïssé's theorem, allow us to construct various interesting subgroups of closed oligomorphic groups. One of the theorems here has an application to the theory of measurement in mathematical psychology! The final chapter treats some important but miscellaneous topics.

For further reading on the topics of Chapter 1, see Wielandt (1964) for permutation groups, Chang and Keisler (1973) for model theory, Oxtoby (1980) for measure and category, and Graham, Rothschild and Spencer (1980) for Ramsey theory. Other useful books in related areas are those of Fraïssé (1986), Goulden and Jackson (1983) and Shelah (1978).

The exercises are a mixed bag, and should *not* be regarded as routine tests of comprehension. Some are very difficult, and I have given hints, of which the most detailed are outline proofs. Unsolved problems slipped in among the exercises are flagged as such; others are scattered through the text.

A few comments about terminology. As in logic, the natural numbers begin at 0. But I am not totally consistent: I use $\mathbf{N}$ rather than ω to denote the natural numbers, and $\aleph_0$ for their cardinality; and if I want to refer to just two objects, I usually number them 1 and 2 rather than 0 and 1. However, I use the term "ω-sequence" for an

infinite sequence (of order type ω). Also, unlike the logicians, I don't insist that the domain of a structure be non-empty. As in model theory, it is convenient to treat n-tuples flexibly, regarding them as "ordered n-subsets". Thus, if I say that n-tuples $\bar{a} = (a_1, \ldots, a_n)$ and $\bar{b} = (b_1, \ldots, b_n)$ carry isomorphic substructures, I mean that the map $a_i \mapsto b_i$ $(i = 1, \ldots, n)$ is an isomorphism between the induced substructures on $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$.

As usual, $\mathbf{Z}$, $\mathbf{Q}$, and $\mathbf{R}$ are the integers, rationals and real numbers.

1.2. PERMUTATION GROUPS

A *permutation group* G on a set Ω is simply a subgroup of the symmetric group on Ω (the group of all permutations of Ω). However, to allow us to consider a number of permutation groups isomorphic to (or homomorphic images of) a fixed group G, a more general concept is convenient. A *permutation representation* of G on Ω is a homomorphism from G to the symmetric group on Ω. Other terminology is often used: we say that G acts on Ω, or that Ω is a G-set or G-space. The image of the homomorphism is a permutation group, denoted G^Ω, and called the permutation group on Ω *induced* by G.

A permutation representation of G on Ω can be described by a function $\mu : \Omega \times G \to \Omega$, where $\mu(\alpha, g)$ is the image of α under the permutation corresponding to g. This function satisfies

(a) $\mu(\alpha, gh) = \mu(\mu(\alpha, g), h)$;

(b) $\mu(\alpha, 1) = \alpha$.

(These are translations of the closure and identity axioms for a group. Note that the other two group axioms do not have to be translated — composition of permutations is always associative; and the condition derived from the inverses axiom, namely

(c) $\mu(\alpha, g) = \beta \iff \mu(\beta, g^{-1}) = \alpha$,

is a consequence of (a) and (b).) Conversely, given a map μ satisfying (a) and (b), the function carrying g to the permutation $\alpha \mapsto \mu(\alpha, g)$ is a permutation representation of G.

I will from now on suppress the function μ and write αg for the image of α under g. (Notice that I have sneaked in the convention that permutations act on the right!)

Let G act on Ω. Set $\alpha \sim \beta$ if there exists $g \in G$ with $\alpha g = \beta$. This is an equivalence relation on Ω; the reflexive, symmetric and transitive laws correspond naturally to conditions (b), (c) and (a) above. Its equivalence classes are called *orbits*, and G is called *transitive* if it has but one orbit. If a subset Δ of Ω is a union of orbits, then

we have an action of G on Δ. In the case when Δ is a single orbit, the permutation group G^Δ induced on Δ is called a *transitive constituent* of G.

The transitive constituents of a permutation group do not determine it uniquely; but we have:

(1.1) *Any permutation group is a subgroup of the cartesian product of its transitive constituents.*

The cartesian product of $(G_i : i \in I)$ is the set of functions $f : I \to \bigcup G_i$ such that $f(i) \in G_i$ for all $i \in I$; the group operation is componentwise. To each $g \in G$ corresponds the function f_g for which $f_g(i)$ is the restriction of g to the i^{th} orbit: this defines the embedding of G in the cartesian product. In fact, G is a *subcartesian product* of its transitive constituents. (This simply means that it projects onto each factor of the product.)

Note that we have the cartesian product here rather than the (restricted) direct product (which consists of those functions f for which $f(i) = 1$ for all but finitely many $i \in I$). Of course, if there are only finitely many orbits, then the two are indistinguishable.

Given any family $(G_i : i \in I)$ of transitive permutation groups, their cartesian product has a natural action for which the G_i are the transitive constituents. When I refer to the cartesian (or direct) product of permutation groups, this action is intended. There are other actions, which will sometimes be needed; for example, there is an action on the cartesian product of the domains (rather than the disjoint union).

Let G act on Ω. The *stabiliser* G_α of a point $\alpha \in \Omega$ is the set $\{g \in G : \alpha g = \alpha\}$. It is a subgroup of G. Similarly, if $\Delta \subseteq \Omega$, the setwise stabiliser G_Δ of Δ consists of all permutations $g \in G$ which map Δ onto itself; and the pointwise stabiliser $G_{(\Delta)}$ is the set of permutations which fix every point of Δ. Often, $\bar{\alpha}$ will denote an ordered tuple of elements of Ω, and then $G_{\bar{\alpha}}$ will denote the pointwise stabiliser of $\bar{\alpha}$.

Let H be a subgroup of the abstract group G. The *coset space* of H in G is the set of right cosets of H in G; there is an action of G on it given by $(Hx)g = Hxg$ (or, more pedantically, $\mu(Hx, g) = Hxg$). Coset spaces provide "canonical" transitive G-spaces:

(1.2) *If G acts transitively on Ω, then Ω is isomorphic to the coset space of G_α in G, for $\alpha \in \Omega$.*

(A *G-isomorphism* between G-spaces Ω_1, Ω_2 is a bijection θ such that, for all $g \in G$, the diagram

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{\theta} & \Omega_2 \\ \downarrow g & & \downarrow g \\ \Omega_1 & \xrightarrow{\theta} & \Omega_2 \end{array}$$

commutes.)

G acts *regularly* on Ω if it is transitive and the stabiliser of a point is the identity. By (1.2), in this case, Ω is isomorphic to G (on which G acts by right multiplication) — this is called the *right regular representation.* In this situation, G acts *faithfully* on Ω; that is, the map taking elements of G to the corresponding permutations is one-to-one, so that G is isomorphic to its image. This action was used by Cayley to show that every group is isomorphic to a permutation group.

Now suppose that G is transitive on Ω. A *congruence* is a G-invariant equivalence relation on Ω. There are two trivial congruences, namely, equality and the "universal" relation with a single equivalence class. G is said to be *primitive* if there are no other congruences. A transitive group G is primitive if and only if G_α is a maximal subgroup of G. (More generally, the congruences form a lattice isomorphic to the lattice of subgroups lying between G_α and G.)

Once again, if G is not primitive, we can break it down into "smaller" pieces. The reverse construction is the wreath product of permutation groups, defined as follows. Let H and K be permutation groups on Γ and Δ respectively. Take $\Omega = \Gamma \times \Delta$, thought of as a covering space of Δ with fibres bijective with Γ. (See Fig. 1.) Let B (the base group) be the cartesian product of $|\Delta|$ copies of H, one associated with each fibre of Ω (that is, each element of Δ); and let K_1 (the top group) be the permutation group on Ω obtained by letting K permute the fibres according to its given action on Δ. Then the *wreath product* $H \operatorname{Wr} K$ is the (semi-direct) product of B and K_1.

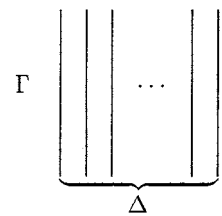


Fig. 1. A covering space

(1.3) *Let G be transitive but imprimitive on Ω. Let Γ be a congruence class; H, the permutation group induced on Γ by its setwise stabiliser in G; Δ, the set of congruence classes; and K, the group induced on Δ by G. Then G can be embedded in a natural way into $H \operatorname{Wr} K$ (as permutation group).*

There is a more general definition of wreath products, which can take account of an arbitrary (partially ordered) set of congruences; but I shall not require this.

Any finite transitive permutation group can be analysed or "broken down" into primitive "components" in finitely many steps (though of course some information is lost). This is not true in general for infinite permutation groups; but it is the case for the groups I shall be considering, those with only finitely many orbits on n-tuples for all n (or indeed, just for $n = 2$). This is because such a group can have only finitely many congruences. (Any congruence, thought of as a set of ordered pairs, is a union of orbits of G in its action on $\Omega \times \Omega$.)

Another important action of the wreath product $H \operatorname{Wr} K$ is the *product action*, on the set of functions $\phi : \Delta \to \Gamma$. (An element f of the base group acts by

$$\phi \cdot f(i) = \phi(i) \cdot f(i),$$

while the top group acts on the arguments of the functions by

$$\phi \cdot k(i) = \phi(ik^{-1}).$$

Exercises

1. For $H \leq G$, the kernel of the action of G on the coset space of H is $\bigcap_{g \in G} g^{-1}Hg$. (This subgroup, called the *core* of H in G, is the largest normal subgroup of G which is contained in H.)

2. The only primitive regular groups are the cyclic groups of (finite) prime order.

3. Let H, K act on Γ, Δ respectively. If $|\Gamma|, |\Delta| > 1$, show that $H \operatorname{Wr} K$ (in its product action) is primitive if and only if H is primitive but not regular on Γ and K is transitive on Δ.

4. Prove the assertion in the text that, if G is transitive on Ω, then the lattice of congruences is isomorphic to the lattice of subgroups between G_α and G. (Consider the stabilisers of the congruence classes containing α.)

1.3. MODEL THEORY

Model theory concerns the relationship between sentences in a formal language and the structures satisfying them. The language appropriate here is that of first-order logic.

The common features of first-order languages are the logical connectives ($\neg$ (*not*) and $\rightarrow$ (*implies*) suffice, though we also use $\vee$ (*or*) and $\wedge$ (*and*)), and quantifiers ($\forall$ (*for all*) and $\exists$ (*there exists*)), punctuation marks (parentheses and comma), and a supply of variables (countably many will be enough). In addition, a language contains symbols for functions, relations and constants, appropriate to the application (the area of mathematics being modelled). Each relation or function symbol is equipped with an *arity* (the number of arguments it takes) as part of its syntax. The language is called *relational* if it contains no function or constant symbols.

For example, for group theory, we could take a binary function (for multiplication), a unary function (for inversion), and a constant (the identity); or we could make do with the first of these alone; or we could define multiplication by a ternary relation R, so that $R(x, y, z)$ is interpreted to mean $xy = z$.

Formulae are defined recursively in a standard way. First, a term is a constant symbol or a variable or a function symbol with the correct number of terms as arguments. An atomic formula is a relation symbol with terms as arguments; a general formula is obtained by combining formulae by means of connectives, or preceding them with quantifiers. It is a sentence if it has no free (unquantified) variables.

A sentence is *universal* ($\forall$), *existential* ($\exists$), or *universal-existential* ($\forall\exists$) if it has the form $(\forall \bar{x})\phi(\bar{x})$, $(\exists \bar{x})\phi(\bar{x})$, or $(\forall \bar{x})(\exists \bar{y})\phi(\bar{x}, \bar{y})$ respectively, where ϕ is quantifier-free.

A structure over a language consists of a set equipped with distinguished constants (i.e. elements of the set), functions and relations corresponding to the symbols in the language (and having the appropriate arities). It is hopefully clear what it means for a sentence ϕ to be satisfied, or valid, in a structure M: we write $M \models \phi$ and say that M is a *model* of ϕ in this case. Similarly for sets of sentences. Thus, a group is a model for the axioms of group theory.

There is a formal deduction system associated with a first-order language. Thias consists of a set of sentences called axioms, and some rules of inference which allow sentences to be derived from others (possibly in the presence of sets of "hypothe-

ses"). In fact, there are several such systems; but all standard ones satisfy *Gödel's completeness theorem*:

(1.4) *A set of sentences has a model if and only if it is consistent (that is, no contradiction can be derived from it).*

From this major result, the two "portals" of model theory are derived.

(1.5) (The compactness theorem.) *A set of sentences has a model if and only if every finite subset has a model*

For, by (1.4), we can replace "has a model" by "is consistent"; and, since proofs in the formal system are finite, if a contradiction could be deduced, then only finitely many hypotheses would be used in the deduction, and this finite set would be inconsistent.

(1.6) (The downward Löwenheim-Skolem theorem.) *A set of sentences over a countable language which has a model has a finite or countable model.*

This comes from the proof of (1.4): the model constructed in that proof is countable. (See the note about equality below.)

Observe that (1.5) and (1.6) contain no reference to the deduction system. Indeed, it will play no further rôle in the discussion.

Equality is obviously important enough to have its own name and conventions. It can be shown, using the compactness theorem, that no set of axioms can force the interpretation of a binary relation to be equality; we can only say that it is an equivalence relation such that "equal" terms can be interchanged in formulae without changing their truth. Then, by factoring out the equivalence relation, we obtain a new structure which satisfies exactly the same sentences as the old one, in which the binary relation really is interpreted as equality. (Such a model is called *normal*.) Because of the importance of equality, it is customary to consider only normal models, and I shall follow this convention. Note, however, that the countable model constructed in the proof of (1.4) may not be normal, and the model obtained by "normalising" it may be finite; hence the possibility of a finite model in (1.6).

The compactness theorem is the source of many results about the limitations to what can be said using first-order sentences. Obviously, we can say everything about a finite structure, simply by listing all the instances and non-instances of relations, etc. But, for example, there is no set Σ of sentences such that all models of Σ are finite but their cardinalities are unbounded. For let ϕ_n be the sentence saying "there

exist at least n points". (See Exercise 1: ϕ_2 is $(\exists x_1)(\exists x_2)(\neg(x_1 = x_2))$.) If such a Σ existed, then Σ together with any finite set of the sentences ϕ_n would have a model, but Σ together with all the ϕ_n would not.

Along the same lines, we have:

(1.7) (The upward Löwenheim-Skolem theorem.) *If a set of sentences has an infinite model, then it has arbitrarily large infinite models.*

To see this, we adjoin to the language a large infinite set of new constant symbols c_i, and to the set of sentences all those of the form $c_i \neq c_j$ (for $i \neq j$). Any finite subset is satisfiable (in the given infinite model), so the whole set is.

We use the term "theory" for "consistent (or satisfiable) set of sentences". There are two opposed points of view here. Some theories, like that of groups, are intended to have many different models; a logical consequence of the theory will be valid in all of them. Others are intended, as far as possible, to describe a single structure. A theory Σ is said to be *complete* if, for every sentence ϕ, either ϕ or $\neg\phi$ is in Σ. This is equivalent to saying that Σ consists of all sentences which hold in some fixed structure M. (We speak of the *theory of* M, written $\mathrm{Th}(M)$.) If M is infinite, then Σ does not determine M, even up to cardinality (by (1.7)). The best we can expect is that M is the only model of Σ of its cardinality. This concept, for countable M, will be very important to us (see §2.5).

Exercises

1. Write down sentences ϕ_n, ψ_n (in a language with equality) such that
   (a) any model of ϕ_n has at least n elements;
   (b) any (normal) model of ψ_n has exactly n elements.

2. Use the compactness theorem to show that a theory (in the language of graphs) having models with arbitrarily large finite diameter has a model with infinite diameter.

3. Show that M is the unique model of $\mathrm{Th}(M)$ (up to isomorphism) if and only if M is finite.

4. Write down a sentence, using equality and one binary relation symbol, all of whose models are infinite. Is this possible with equality alone?

1.4. CATEGORY AND MEASURE

Cantor's celebrated proof of the existence of transcendental numbers went like this: there are so many more complex numbers than algebraic numbers, that there are as many transcendental numbers as complex numbers. The related techniques of Baire category and measure allow refinements on this argument: certain subsets are "small", even though their cardinality is the same as that of the whole set.

Let (X, d) be a complete metric space. A subset of X is dense if its closure is X, i.e. if it meets every open set. A subset is *residual* if it contains the intersection of countably many open dense sets. (Other terms are used: the complement of a residual set is called *meagre*, or *of the first category* — whence a residual set is called *comeagre* — and a set which is not of the first category is *of the second category.*)

(1.8) (The Baire category theorem.) *A residual set in a complete metric space is non-empty.*

Thus, if we can show that the set of elements having some property P is residual, then it follows that some element has the property P. But the interpretation of (1.8) is that a residual set is, in a sense, "large", containing "almost all" of the space; for example, it meets every open dense set. Also, the intersection of countably many residual sets is residual (and hence non-empty).

Metric spaces in these notes always arise in the following way. A point of the space is determined by a countable sequence of choices, and the nearness of two points depends on the initial segment of the choice sequences determining them which agree; we can take

$$d(x, y) = \frac{1}{2^n}$$

if the choice sequences for x and y differ first in the n^{th} term. (The actual value we choose for the distance is not crucial; any decreasing function of n will do. The particular choice is motivated by consideration of Hausdorff measure, see Cameron (1987a).)

For an illustrative example, consider the case where there are just two alternatives for each choice, so that points are represented by an ω-sequence of zeros and ones. (Think of the outcome of countably many coin tosses as determining a point.) Now an open ball consists of all zero-one sequences with a given initial segment. Hence a set S is open if every point of S has an initial segment, all of whose continuations lie in S; and S is dense if every finite sequence has a continuation which lies in S. For

this metric space, the Baire category theorem is easily proved directly (see Exercise 3).

This metric space is homeomorphic to the Cantor ternary set, the set of real numbers in the unit interval whose ternary expansions contain only the digits 0 and 2. (Replace all the ones in a given zero-one sequence by twos, and regard the result as the ternary expansion of a real number.) Note that, if we simply regard the given sequence as the binary expansion of a real number, the resulting map is not a homeomorphism, since it fails to be one-to-one.

The procedure for constructing a measure space is more technical, though in many cases there is an intuitive description. The general technique parallels the construction of Lebesgue measure on $\mathbf{R}$. First we define the measure on a basic class of sets (the open intervals in $\mathbf{R}$); then we extend it by countable additivity to the σ-algebra they generate (the Borel sets in $\mathbf{R}$); then finally (though this is not necessary in many applications) we define inner and outer measure on any subset, and call a set measureable if its inner and outer measure coincide.

In the case of the space of zero-one sequences, the basic sets are those with prescribed initial segment (that is, the open balls), and the measure of the set of sequences with a given initial segment of length n is $1/2^n$. There are two other ways to view this:

(a) We regard a sequence as giving the outcome of infinitely many tosses of a fair coin (so that the outcome of each toss has probability $\frac{1}{2}$, and different tosses are independent). Then we have the standard probability measure.

(b) Identifying a sequence with the binary expansion of a real number in the unit interval, we use Lebesgue measure on the interval. As we noted earlier, this map is not one-to-one; but the failure is not damaging here. (There is a countable, and hence null, set of reals which have two pre-images each.)

The "large" sets in this context are those of measure 1. As with residual sets, they have the property that the intersection of countably many such sets is again of measure 1, and hence non-empty. (This is familiar in the complementary form: the union of countably many null sets is a null set.) Moreover, any non-empty open set has positive measure; so a set of measure 1, like a residual set, is dense.

In the more general case where a point is determined by a sequence of choices which are not restricted to two alternatives each, the assignment of measure to open balls is not so straightforward. The measure of a given open ball has to be split up among its immediate successors. Equal division may not be appropriate. We will see examples later.

Exercises

(All these exercises refer to the space of zero-one sequences.)

1. Prove that, with the metric described in the text, the space of zero-one sequences is a metric space satisfying the *ultrametric inequality*, viz.

$$d(x,z) \leq \max(d(x,y),\ d(y,z)).$$

Show that, in an ultrametric space (one satisfying the ultrametric inequality),
(a) every point in an open ball is its centre;
(b) if two open balls intersect, then one contains the other.

2. Prove that the space of zero-one sequences is a complete metric space, and that its topology is that of pointwise convergence.

3. Show that a residual set is dense and has cardinality $2^{\aleph_0}$.

[*Hint.* Let S be residual and let $S = \bigcap\{X_n : n \in \mathbf{N}\}$, where each X_n is open and dense. To prove that $S \neq \emptyset$, define finite sequences σ_n inductively as follows:
σ_0 is the empty sequence;
if σ_n is defined, choose an infinite continuation s_n of it lying in X_n (possible since X_n is dense) and a finite initial segment σ_{n+1} of s_n all of whose continuations lie in X_n (possible since X_n is open) such that σ_{n+1} is longer than σ_n. The "limit" of the sequences σ_n lies in S.

Now, if U is a given open set, modify the construction by choosing σ_0 so that all continuations of σ_0 lie in U. To demonstrate the cardinality, "code" infinite zero-one sequences into the construction by adding one extra bit to each σ_{n+1}.]

4. A sequence is called *universal* if it contains every finite zero-one sequence as a consecutive subsequence. Show that the set of all universal sequences is residual and has measure 1, i.e. is "large" in both senses.

5.Show that the set of sequences with upper density 1 and lower density 0 is residual.

[The *upper density* of a zero-one sequence s is defined as

$$\limsup_{n\to\infty} d(n)/n,$$

where $d(n)$ is the number of ones among the first n terms of s; the *lower density* is the lim inf of the same quantity.]

Remark. By contrast, the strong law of large numbers asserts that the set of sequences having density 1/2 has measure 1. So, in this instance, the two techniques give conflicting views of the "typical" set. We will see other examples later.

1.5. RAMSEY'S THEOREM

The "pigeonhole principle" asserts that, if the infinite set X is partitioned into finitely many parts, then one at least of these parts is infinite. Ramsey's theorem is a generalisation of this.

(1.9) *Suppose that the set of n-element subsets of the infinite set X is partitioned into finitely many parts. Then there is an infinite subset Y of X, all of whose n-element subsets belong to the same part of the partition.*

I'll prove this for $n = 2$, deducing it from the pigeonhole principle. The general proof is by induction on n, following the same lines, and is outlined in the Exercises.

Let $\{x_0, x_1, \ldots\}$ be an infinite subset of X. We choose an infinite subsequence of (x_i) as follows. Set $y_0 = x_0$. After the $(i-1)^{\text{st}}$ stage, $y_0, \ldots, y_{i-1}$ have been chosen, and there are infinite subsets $Y_1, \ldots, Y_{i-1}$ such that, for $j < i$,

(a) $Y_j \subseteq Y_{j-1}$;

(b) $y_j = \min(Y_j)$;

(c) all edges from y_{j-1} to Y_j (that is, all 2-sets $\{y_{j-1}, z\}$ for $z \in Y_j$) lie in the same part of the partition.

In the i^{th} stage, partition $Y_{i-1} \setminus \{y_{i-1}\}$ so that x lies in the k^{th} part of the partition if and only if $\{y_{i-1}, x\}$ lies in the k^{th} part of the original partition; let Y_i be an infinite part (guaranteed by the pigeonhole principle), and $y_i = \min(Y_i)$.

Now, in the subsequence (y_i), the number m_i of the part of the partition containing a pair $\{y_i, y_j\}$ $(i < j)$ depends only on i, not on j. Another application of the pigeonhole principle yields a subsequence on which m_i is constant. This is the required infinite set.

It is usual to express Ramsey's theorem in the language of colours and colourings. Associating a colour with each part of the given partition, we are provided with a colouring of the n-element subsets of X with finitely many colours. A subset of X is called "monochromatic" if all its n-element subsets have the same colour. Then the theorem asserts the existence of an infinite monochromatic subset of X.

Ramsey's theorem stands at the origin of a flourishing subject, some of whose concerns are quantification of the infinities involved, finite analogues, and similar results for structures other than sets. All I need, however, is a single generalisation of Ramsey's theorem:

(1.10) *Suppose that the n-subsets of an infinite set X are coloured with r colours, all of which are used. Then there are an ordering $c_1, \ldots, c_r$ of the colours, and infinite sets $X_1, \ldots, X_r$, such that X_i contains a set of colour c_i but no set of colour c_j for $j > i$.*

Ramsey's theorem gives us a colour c_1 and an infinite set X_1. We proceed to "rank" the colours; a colour is ranked once we have found an infinite set containing n-sets of that colour and previously ranked colours only. The theorem is proved when all colours have been ranked; and Ramsey's theorem allows us to rank the colour c_1.

Suppose that, at some stage, c is a colour not already ranked, and C an n-set of colour c. The subsets of C are partially ordered by inclusion, and this partial order can be extended to a total order

$$\emptyset = C_0 \prec \ldots \prec C_s = C,$$

where $s = 2^n - 1$. Let Y be an infinite set disjoint from C and containing only sets of ranked colours.

We now proceed through the sequence (C_i), defining infinite sets Y_i, starting with $Y_0 = Y$. At the i^{th} stage, we define a new colouring of the $(n - |C_i|)$-subsets of Y_i, by giving each such set B the colour originally assigned to $B \cup C_i$. There is an infinite set Y_{i+1} with all of its subsets having the same colour. Then all colours of n-sets occurring within $Y_{i+1} \cup C_i$ have been ranked except possibly for the unique colour of n-sets containing C_i; so this colour can be ranked if it hasn't already been. By the time we have worked our way through the entire sequence, the colour of C will have been ranked.

Now just continue this process until all colours are ranked and the theorem is proved.

Exercises

1. Prove Ramsey's theorem.

[*Hint:* The proof is by induction on the size n of the subsets being coloured. The start $n = 1$ of the induction is the pigeonhole principle, and the argument given illustrates

the step from 1 to 2. In general, replace the first (but not the second) application of the pigeonhole principle with the inductive hypothesis for $n-1$.]

2. The finite form of Ramsey's theorem is the following assertion:

Let n, m, r be given positive integers, with $n < m$. Then there is an integer N (depending on n, m, r) with the following property:

If the n-element subsets of an N-element set X are coloured with r colours, then there is an m-element subset Y of X, all of whose n-element subsets have the same colour.

Prove this, by a modification of the argument outlined above.

Harder: Deduce it from the infinite form of Ramsey's theorem by means of the Compactness theorem.

3. Formulate and prove a finite form of (1.10).

4. Show that an infinite sequence of elements of a totally order set contains one of the following: a constant subsequence; a strictly increasing subsequence; a strictly decreasing subsequence.

Using this fact, deduce the Bolzano-Weierstrass theorem (for $\mathbf{R}$) from the "principle of the supremum".