Sporadic homogeneous structures

by

Gregory Cherlin[1]

*Department of Mathematics, Hill Center, Busch Campus, Rutgers University*
*New Brunswick, NJ 08903*
*Email cherlin@math.rutgers.edu*

ABSTRACT

We discuss Lachlan's classification theory for *finite homogeneous structures* and related problems on finite permutation groups. Lachlan's theory provides a hierarchy of classifications in which structures which are "sporadic" in one context reappear as members of infinite families at later stages. Every finite structure is accounted for at some level in this hierarchy, but for structures associated with familiar primitive permutation groups the combinatorial problem of locating that level precisely can be quite challenging.

> ... *in most categories few objects have the Witt property; those that do are very well behaved indeed.* [As, p. 82]

## Introduction

When classification results are enlivened by the appearance of uninvited guests in the form of "sporadic" objects, those who take an interest in these interlopers may be tempted to account for them in various ways, possibly by viewing them as coming from infinite (perhaps even continuous) families of more general objects which may be natural from some broader point of view. In pure model theory, Lachlan's classification theory for finite homogeneous relational structures provides a relatively well understood illustration (or "toy model", if you will) of this sort of thing. This theory, which will be reviewed below, provides an infinite number of classification theorems of a general character for combinatorial structures with rich automorphism groups, parametrized by certain bounds on the complexity of the structures. Any finite structure will actually appear at some stage in one of these classifications, and may well occur as a sporadic structure initially; in the long run, every sporadic structure winds up belonging to a family parametrized by numerical invariants; at any given stage, only finitely many structures occur as sporadics; and finally, one will never "move beyond" the sporadics: we will always encounter new structures making their appearance as (temporarily) sporadic structures.

---

[1] Supported in part by the Binational Science Foundation BSF 0377215
Key words: classification, sporadic, homogeneity, permutation group, stability

That this sort of thing would occur on a regular basis is only natural, though it may be surprising that this state of affairs would be described by a theorem in a natural and reasonably general setting. The general theorem is actually a consequence of one single finiteness result – the finiteness of the set of sporadic simple groups – and some permutation group theory. (See §2 for an explanation of how we can view structures as a special case of groups, rather than the other way around.) It should be said that whatever the motivation, in working out this theory one does not need to think about the phenomenon of sporadicity as such, and what is really involved is the finiteness theorem given below as a Coordinatization Theorem (§5), which is equivalent to the Bounded-Rank Theorem mentioned briefly in §7.

The notion of homogeneity also leads one to associate a natural measure of complexity $\kappa(G, X)$ with any finite permutation group $(G, X)$, which may be of interest in its own right; this can be studied from a combinatorial point of view without reference to any theoretical background in model theory. This invariant behaves somewhat like a dimension; for example, for a vector space of dimension $n$ (that is, for the group $(\mathrm{GL}(V), V)$) $\kappa$ will be $n + 1$. Determining the complexity of specific permutation groups with precision can be quite challenging, and I have included a selection of open problems in the final section, which can be read independently of the presentation of the general theory, though it will no doubt be helpful to look over the background material in the first three sections, which contain a number of specific examples.

In the last example of §3 we will see how one fairly rich family of examples divides into families and sporadics at each level of analysis, with the number of families and sporadics finite at each stage; but in this case, at least, the number of structures counted as sporadic is exponentially large compared to the number of parametrized families encountered.

Pure model theory at the present time consists largely of ideas connected with Shelah's "classification theory", which attempts to provide a very general theory of classification of structures. This comes in a number of variants, most of which emphasize the classification of infinite structures. For example, when these ideas are specialized to the case of modules over a ring, they involve the classification of pure-injective indecomposable modules, with ideas very closely related to representation theory. It was seen by Lachlan that the model theoretic approach also makes sense for certain broad classes of finite structures: homogeneous structures (for a finite relational language). These results have since been generalized (notably by Hrushovski) to cover reasonably broad classes of finite permutation groups. I will say a little about this as well, mainly in §8.

The point of Lachlan's theory is that it involves an infinite number of related classification problems; for each type of structure, the homogeneous structures of that type can be rather thoroughly classified. In each instance, the structures involved fall naturally into a finite number of families, and within each family the individual objects are parametrized by a finite number of numerical invariants which may be varied independently. It is possible for the number of invariants needed to describe an object to be *zero*, in which case the family degenerates to a single structure – and these are the structures which may be considered *sporadic*. We have said that *any* finite object will eventually be covered by one of these classification schemes; in other words, any structure can be viewed as homogeneous of some type (see the example in §1 and the general considerations of §2). In passing from one classification problem to the next, what generally occurs is that (1) new families arise; (2) in the old families, additional numerical invariants are acquired which may be varied independently. In particular, the sporadic objects from one classification scheme are eventually absorbed into parametrized families. So in this model, sporadic objects can always be understood as part of some larger classification scheme, but sporadicity itself is not escaped.

In the next section I will give a concrete example of all of this: a 1-parameter family of graphs which contains three homogeneous graphs but which enters as a family of homogeneous structures of a slightly more complex type. We will see later that a slight generalization of this provides simple and very uniform families of graphs, some of which will be considered sporadic at more or less every level of Lachlan's hierarchy. The analysis of these examples has been carried out with extraordinary precision by Saracino, building on the rough analysis of [CMS]. The bulk of his results are summarized at the end of §3.

Although Lachlan's theory provides a classification of the homogeneous finite relational structures of any specified type (for definitions, see §2), the literature generally refers to the classification of homogeneous *stable* relational structures, a broader class with a comparatively technical definition, which turns out to consist of the finite ones and their infinite limits. There are good reasons for this generalization: to prove the theorems, it is very convenient to move back and forth between the large finite structures and their infinite limits. Shelah's notion of stability is one of the fundamental concepts of pure model theory; Lachlan realized that in the context of homogeneous relational systems, it is equivalent to "smooth limit of finite" (cf. §4). This has led to a fruitful interaction of model theory and the theory of permutation groups, which involves an interplay between the group theoretic analysis of large finite structures, and the combinatorial analysis of their infinite limits. All of this depends ultimately on the classification of the finite simple groups.

There are a number of expositions of this theory. The theory of (and some major open problems concerning) homogeneous structures in general and finite or stable ones in particular is discussed in Lachlan's ICM lecture [La2]. A detailed account of the classification theory for finite homogeneous structures is given in [KL]; this combines a detailed exposition with some major expositions and an important advance (the form of the Stretching Theorem given in §6 below). It does assume familiarity with the language and point of view of model theory, which to some extent the present article is intended to provide (via examples, the general discussion in §2, and a handful of technical definitions). The subject of homogeneity is also the final topic taken up in the text [DM], which also discusses a number of other topics in permutation group theory which have been important to its users in model theory, such as the O'Nan-Scott Theorem.

In the long run the most general form of this theory would be a structural analysis of large $k$-closed permutation groups of bounded rank (with both $k$ and the bound on rank taken as fixed). This has not been carried out, but an intermediate stage, in which the bound on rank is sharpened to a bound on the number of orbits on 4-tuples, is discussed in [Hr]. This depends on a thorough analysis of the primitive case by group theoretic methods [KLM]; but the analysis of the general case requires a very heavy dose of model theory, and indeed traditional permutation group theory has little to say about the imprimitive case. Technicalities aside, the final form of the theory is quite similar to the theory we will encounter in Lachlan's original case, with the main difference being the appearance in [KLM] of structures conspicuous by their absence in Lachlan's theory: vector spaces with their classical adornments (inner products and quadratic forms), and – in characteristic 2 – also some less classical adornments (see §8).

I thank Saracino for keeping me apprised of his very delicate work, which contains the most subtle analysis of the behavior of the parameter $\kappa$ (§2) for any specific family of primitive permutation groups.

Contents



## §1. The graphs $n^2$.

In the next section we will present the general correspondence between finite permutation groups and finite structures, which are essentially the same thing. Here we look at one example of a family of structures or permutation groups whose properties illustrate some of the issues that turn out to be central in the general theory. This example can be analyzed easily in complete detail, but its natural generalization is still not thoroughly understood (§9, Problem 3).

**Definition**

*The graph $n^d$ has as its vertex set the set of all $d$-sets taken from an $n$-element set, with edges between $d$-tuples $u$ and $v$ if they differ at exactly one vertex.*

Note that the automorphism group of $n^d$ is the wreath product $\mathrm{Sym}\,(n) \wr \mathrm{Sym}\,(d)$; for this reason this graph may be referred to variously as a "power" or a "wreath product action", according as one pays more attention to the graph or permutation group; these two points of view are equivalent for our purposes.

In the present section we consider only the graphs $n^2$, which are also referred to as the *line graphs* $L(K_{n,n})$, as they can be viewed as graphs derived from the complete bipartite graphs $K_{n,n}$ by taking the edges of $K_{n,n}$ as vertices, with two edges adjacent if they have a vertex in common in $K_{n,n}$. For $n \leq 3$ these graphs are *homogeneous* in the following sense: any isomorphism between two induced subgraphs is the restriction of some automorphism. For $n > 3$ they are not homogeneous as graphs, because they then contain two classes of graphs of the form $2 \cdot K_2$ (meaning, two disjoint edges): "parallel" edges and "orthogonal" edges.

For $n = 3$ the graph obtained is considered sporadic (either informally, or as an instance of Lachlan's theory). For $n = 1, 2$ these graphs are not sporadic. For $n = 2$ the graph is complete bipartite graphs and belongs to a family of homogeneous complete multipartite graphs parametrized by two numerical parameters, while for $n = 1$ the graph belongs to the complementary family: the complement of a complete multipartite graph is a disjoint union of complete graphs.

What matters here, though, is the fact that all the graphs $n^2$ do in fact occur as one infinite family of homogeneous structures – but not as *graphs*. We may instead consider these graphs as coming equipped with a 4-place relation of parallelism, $P(v_1, v_2, v_3, v_4)$. This may be defined from the edge relation as follows: if two disjoint edges have the property that the four vertices involved have no further edges between them, the edges are *parallel* if there is no vertex adjacent to all four of the given vertices. In particular the automorphism group of the original graph is also the automorphism group of the enriched structure. However, in the category of graphs-with-parallelism, all of these structures are homogeneous, in the sense that parallelism-preserving isomorphisms between subgraphs are induced by automorphisms.

An important point here is that we consider the two structures on $n^2$ – the graph structure, or the structure of a graph with parallelism – as identical structures, even though they are of different types. We justified this by looking at them as permutation groups; logicians would have expressed this directly, in terms of the structures, by stating that the relations in each structure are definable from the relations in the other structure (definable over $\emptyset$, specifically). Most of the notions coming from model theory can be translated into the language of permutation groups, and it can be technically advantageous to do so on occasion (just as it can be equally advantageous, on other occasions, to make the translation in the other direction). In particular the very rich information contained in the classification of the finite simple groups is largely inaccessible from the structural point of view, though it looms very large indeed in the analysis of the associated permutation groups.

Thus if one traces through Lachlan's classification theory for the case of graphs, the graph $3^2$ will necessarily occur sporadically (while $1^2$ and $2^2$ can easily be absorbed into other infinite families); once the classification is extended to a sufficiently rich class of 4-hypergraphs, the infinite family $n^2$ will occur as an infinite family indexed by the parameter $n$ (this happens, so to speak, automatically, on the basis of general principles).

For the graphs $n^d$ the situation is more complicated and as $d$ increases the complexity of the additional relations which must be considered goes to infinity, as does the number of exceptional cases (with $d$ large relative to $n$) which occur prematurely as sporadic examples. The place of these structures in the hierarchy of homogeneous structures was determined with extraordinary precision by Saracino. We will present this in §3 after setting up our point of view in general in the next section.

If one replaces the natural representation of $\mathrm{Sym}\,(n)$ on $n$ elements by its representation on $k$-sets (so the original representation corresponds to $k = 1$) then the analysis remains very incomplete, though it is possible that Saracino's analysis extends in some reasonable way to $k > 1$.

## §2. Structures and permutation groups

The structures considered in Lachlan's theory will be *relational structures* $\mathcal{X} = (X; \mathbf{R})$ with $\mathbf{R} = (R_1, \ldots, R_r)$ a finite sequence of relations. If the relation $R_i$ is an $n_i$-ary relation, the *signature* of the relational structure $\mathcal{X}$ is the sequence $\tau = (n_1, \ldots, n_r)$. Thus a signature is a finite sequence of natural numbers, and defines a certain type (or category) of relational structure. There are good reasons to add some more data to the signature (for example, one could impose generalized symmetry and irreflexivity properties on each relation) but laying this out in detail would be both tedious and irrelevant to our present purpose.

There is a Galois connection between finite structures and finite permutation groups. We assign to the structure $(\mathcal{X}, \mathbf{R})$ the permutation group $(\mathrm{Aut}\,\mathcal{X}, X)$. In the other direction, given a permutation group $(G, X)$, a relation on $X$ is called *invariant* (specifically, $G$-invariant, if the context is not otherwise clear) if it is invariant under the action of $G$. We may associate to a finite permutation group $(G, X)$ the relational structure $\mathcal{X}$ whose relations are all the $k$-ary invariant relations for $k \leq |X|$.

This establishes a 1-1 correspondence between the image of the connection on both sides: in other words, between the faithful finite permutation groups and some structures, called *canonical structures*, not much encountered in nature, but much encountered in model theory: they are essentially the structures in which every definable relation is given explicitly as part of the structure (apart from the generous cut-off at complexity $k$; there is not, in any case, much use for $(k+1)$-ary relations on a set of size $k$). For many purposes finite structures that correspond to the same permutation group should be identifed; this was illustrated in the previous example by our observation that the parallelism relation is implicit in the graph structure on $n^2$, and hence might as well be added as an ingredient of the structure.

At any given moment one is likely to be working on one side or the other of this Galois connection, but using notions that originate on both sides; so it is useful to build up a glossary giving the meaning of concepts originating on one side in the language of the other. We will give some examples.

As we have said, a relation $R$ on the set $X$ is said to be *invariant* – or alternatively*definable* – in the structure $\mathcal{X} = (X, \mathbf{R})$ if it is $(\mathrm{Aut}\,\mathcal{X})$-invariant. This coincides with *first order definability* without parameters when $X$ is finite – hence the alternative terminology.

A structure is *primitive* if it has no nontrivial invariant equivalence relation – this notion has always been emphasized in permutation group theory, and the same terminology has been adopted by model theorists.

A permutation group is *$k$-closed* if it is the image in the Galois connection of some structure all of whose relations have at most $k$ places, and the *$k$-closure* of a permutation group is the smallest

$k$-closed group containing it. The intrinsic definition runs as follows: the $k$-closure of $(G, X)$ is the set of permutations $\sigma$ with the property that for any $k$-tuple $\mathbf{x}$ in $X$, the image of $\mathbf{x}$ under $\sigma$ lies in the orbit of $\mathbf{x}$ under $G$; and $(G, X)$ is $k$-closed if it equals its $k$-closure. This is an important notion for us, as it recaptures some information about the original presentation of the structure which one might have expected to see washed away by the Galois connection.

This may be refined as follows: the *complexity* of a permutation group $(G, X)$ is the pair $(k, r)$, where $k$ is minimal such that $G$ is $k$-closed, and $r$ is minimal number such that there are $r$ $(\leq k)$-ary invariant relations in a structure $\mathcal{X}$ for which $\operatorname{Aut}\mathcal{X} = G$. The signature of a structure and the complexity of the associated permutation group are closely related; the complexity of a permutation group measures the size of the simplest signature $\sigma$ such that some structure of signature $\sigma$ has the specified group as its automorphism group. In model theory, the signature is taken as given at the outset; in permutation group theoretic terms, this amounts to bounding the complexity.

Another invariant of $k$-closed groups which is relevant here is the number $r_k$ of orbits of $G$ on $X^k$; these orbits are called *$k$-types* and may defined in structural terms as the atoms in the boolean algebra of invariant (= definable) relations. We will have $r \leq r_k$ and in the context of homogeneous structures, $r_k \leq 2^{k2^r}$, so once $k$ is bounded, bounding the complexity or bounding $r_k$ amounts to the same thing.

A finite relational structure is *homogeneous* if every isomorphism between induced substructures is the restriction of an automorphism. We say that a finite relational structure is *$k$-ary* if it is equivalent (or, with some abuse of language: isomorphic) to a homogeneous structure whose relations are $k$-ary. On the permutation group theoretic side, we may translate this to the following condition which is rather tricky to analyze in practice: $(G, X)$ is $k$-ary if for any $n \leq |X|$, if $\mathbf{a}$, $\mathbf{b}$ are two $n$-tuples with the property that all pairs of corresponding subsequences of $\mathbf{a}$, $\mathbf{b}$ of length at most $k$ lie in the same $G$-orbit, then $\mathbf{a}$ and $\mathbf{b}$ lie in the same $G$-orbit. If a group is $k$-ary then it is $k$-closed, but the converse is thoroughly false. The automorphism groups of the graphs $n^2$ in the preceding section are 2-closed by definition, but the main point made about them there was that they are usually *not* binary in our sense (in other words, they are not homogeneous as binary structures); they are 4-ary structures.

Aschbacher refers to homogeneity as the *Witt property* in the passage cited in our epigraph, which continues: "If $X$ is an object with the Witt property and $G$ is its group of automorphisms, then the representation of $G$ on $X$ is usually an excellent tool for studying $G$." Indeed.

The notions of homogeneity and $k$-arity are the starting point for Lachlan's theory, which concerns the class of $k$-ary structures for $k$ fixed, and with a bound on $r_k$ also fixed.

### Notation

*Let $\mathcal{X}$ be a structure. We write $\kappa(\mathcal{X})$ for the degree of homogeneity of $\mathcal{X}$, which is the least $k$ for which $\mathcal{X}$ is $k$-ary.*

While Lachlan's theory provides a good classification of structures with $\kappa(\mathcal{X})$ and $r_{\kappa(\mathcal{X})}$ bounded, and while every finite structure comes into this classification eventually, with $\kappa(X) \leq |\mathcal{X}|$, the theory does not provide any information directly about the point at which a given structure of mathematical interest will occur, that is: how is $\kappa(\mathcal{X})$ computed for interesting structures (or permutation groups) $\mathcal{X}$?

We have now introduced our basic vocabulary, but as we noted in passing above, something odd happens with the notion of isomorphism, and this is worth dwelling on. Two permutation groups $(G,X)$ and $(H,Y)$ will be considered isomorphic if there is a bijection of $X$ with $Y$ carrying $G$ to $H$. This notion will be carried over to structures: $\mathcal{X}$ will considered isomorphic to $\mathcal{Y}$ (via the map $f : X \leftrightarrow Y$) if $(\operatorname{Aut}\mathcal{X}, X)$ is isomorphic to $(\operatorname{Aut}\mathcal{Y}, Y)$ (via the same map). For example: a graph is isomorphic to its complement via the identity map; the graphs $n^2$ are isomorphic to their enrichments by the parallelism relation; a finite set carrying a successor relation is isomorphic to the same set equipped with the induced linear order. Graph theorists may well disagree with the first of these examples, but from our point of view there are two nontrivial relations between vertices, and it is not of much importance which one is *called* the edge relation; in the canonical structure, this amounts to permuting the names by which various relations are known. This cavalier attitude is appropriate in dealing with homogeneity, and more generally with any issues that can be understood at the level of permutation groups. Once one is committed to this notion of isomorphism, one tends to replace the terms "relational structure" and "permutation group" by "permutation structure", and to lose track of which side of the Galois connection one is actually working with at any given moment.

In §1 we encountered an example of the *wreath product* construction. On the permutation group side, one begins in general with two permutation groups $(G,X)$ and $(H,I)$, called, respectively, the *base* group and the *index* group. One then forms the *wreath product* $(G \wr H, X^I)$ which is set-theoretically a power on which $G^I$ acts coordinatewise, while $H$ permutes the coordinates. Thus $G \wr H = G^I \rtimes H$. The same construction applied to a base structure $\mathcal{X}$ and an index structre $\mathcal{I}$ yields a canonical structure $\mathcal{X}^{\mathcal{I}}$, which may be replaced in practice by some other structure with the same automorphism group.

There are two extreme cases: let $d$ represent the permutation group $\operatorname{Sym}(d)$ acting naturally on $\{1, \ldots, d\}$ (as a structure this is a bare set, carrying only the equality relation), and let $\bar{d}$ be the trivial group acting on the same set, which in structural terms is a labeled set whose elements are treated as distinguished constants. The power $\mathcal{X}^{\bar{d}}$ is the traditional power which can be given by lifting each of the relations on $\mathcal{X}$ (including the equality relation) to $d$ relations operating in the $d$ possible "directions". In particular the equality relation lifts to $n$ equivalence relations. The power $\mathcal{X}^d$ is the symmetrized power in which the coordinates may be freely permuted. In particular $n^d$ was presented concretely in §1: it is the wreath product of two sets with no additional structure. These are among the simplest structures occurring in nature, though by no means the simplest structures from the point of view of the computation of $\kappa(\mathcal{X})$.

Some useful permutation group theoretic notions have no apparent analog on the structural side of the Galois connection: notably the *socle*, and more generally the notion of a *normal subgroup*. In particular the structures corresponding to the natural representations of $\operatorname{Sym}(n)$ and $\operatorname{Alt}(n)$

on $\{1,\ldots,n\}$ have virtually nothing in common from the point of view that interests us: the former is degenerate, and the latter has $\kappa$ maximal. On the other hand, as we will see, there are excellent reasons for returning to the structural side of the picture, particularly as we exploit the possibility of taking infinite limits of our structures; though the Galois connection survives in the infinite limit, with some modifications, still its group theoretic side becomes largely useless. To make the connection work well with $X$ infinite, one takes into account the natural topology on $\mathrm{Sym}\,(X)$, and one restricts attention to closed groups having finitely many orbits on $X^k$ for each $k$. For some of this, see [DM] or [Ca2]. In any case we will not actually exploit this connection in infinite structures.

There is one quite reasonable operation which can behave atrociously, regardless of which side one operates on: formation of quotients. This is of some technical importance, and one of the main theorems in the subject is concerned with the structure of such quotients (Proposition 1, §3). In a permutation structure $\mathcal{X}$, the invariant equivalence relations form a lattice with 0 and 1; the structure is primitive if this is the whole lattice. If $E_1$ and $E_2$ are two invariant equivalence relations with $E_1 \geq E_2$ in this lattice, and if $C$ is an $E_1$-class, then the quotient $C/E_2$ is naturally a permutation structure: the group acting on $C/E_2$ is the faithful version of the group induced on $C/E_2$ by the setwise stabilizer in $\mathrm{Aut}\,\mathcal{X}$ of $C$. When $E_1$ covers $E_2$ and $E_2$ is nontrivial on $C$, this quotient is primitive. Some properties are inherited, notably bounds on the number of $k$-types for each $k$. The properties of $k$-closure and $k$-arity behave as badly as one could imagine. In fact:

Every finite structure is a quotient of a finite binary structure.

For example if the automorphism group is transitive, and $(G,X)$ is the corresponding permutation group, then the structure is naturally a quotient of the right regular representation of $G$, which is binary homogeneous (equip the underlying set of $G$ with one binary relation for each element $g$ of $G$, which encodes the action of $g$ by right multiplication).

Another complication: the Jordan-Holder theorem fails badly in this context. Consider for example the structure $\mathcal{X}_n$ consisting of ordered pairs from $\{1,\ldots,n\}$ with distinct entries, on which $\mathrm{Sym}\,(n)$ acts naturally. If $E$ is a nontrivial invariant equivalence relation on $\mathcal{X}_n$, one can analyze the structure as an extension of the quotient $\mathcal{X}_n/E$ by the structure on an $E$-class. If $E$ is the relation of having an identical first coordinate, this means $\mathcal{X}_n$ is treated as an extension of $n$ by $n$. If we consider instead the relation $E'$ of corresponding to the same unordered pair, then $\mathcal{X}_n$ becomes a 2-fold cover of $n^2$, which is a primitive structure; furthermore $\mathcal{X}_n$ is binary homogeneous, and in just one of these two analyses the components are also binary homogeneous. In spite of this, one can still get useful information using induction on the length of such (non-unique) composition series.

The next section contains a variety of examples illustrating the behavior of $\kappa(\mathcal{X})$. This is a digression as far as the general theory is concerned; we return to the main line in §4 with the classification of the finite homogeneous graphs, in which the outlines of a general theory of finite homogeneous structures are very dimly visible – sufficiently visible to Lachlan, in any case, to spark the development of that theory.

## §3. $\kappa(\mathcal{X})$: examples.

Let us write $k(\mathcal{X})$ for the least $k$ such that $\mathrm{Aut}\,(\mathcal{X})$ is $k$-closed, and $\kappa(\mathcal{X})$, as in the previous section, for the least $\kappa$ such that $\mathcal{X}$ has a presentation as a $\kappa$-ary structure; equivalently, this means that every $\mathrm{Aut}\,(\mathcal{X})$-invariant relation is a boolean combination of $\mathrm{Aut}\,(\mathcal{X})$-invariant $\kappa$-place relations. The present section will offer a smørgasbord of examples illustrating the behavior of $\kappa(\mathcal{X})$, and the reader is invited to consult his own appetite.

The computation of both $k(\mathcal{X})$ and $\kappa(\mathcal{X})$ present substantial difficulties, but the basic meaning of the invariant $k(\mathcal{X})$ seems the more accessible. For example, for $k = 2$: a permutation group is 2-closed if it is the automorphism group of the directed graph with colored edges (some symmetric, some asymmetric) obtained by taking the orbits of the group on ordered pairs, and giving each orbit its own color. On the other hand, a permutation group $G$ is *binary* if every $G$-invariant relation is a boolean combination of $G$-invariant binary relations. In practice, this gets decoded as follows: $(G, X)$ is *not* binary if:

> there are two ordered sequences $a_1, \ldots, a_r$ and $b_1, \ldots, b_r$ of points of $X$
> of length $r > 2$, not conjugate under the action of $G$, such that:
> any ordered pair of elements from $\mathbf{a}$ is conjugate to the corresponding pair from $\mathbf{b}$
> under the action of $G$.

Note that if $r > 2$ is minimal with this property, then we will have a stronger condition: any sequence of $r - 1$ elements of $\mathbf{a}$ is conjugate under $G$ to the corresponding subsequence of $\mathbf{b}$. In this case we might as well take $a_i = b_i$ for $i \le r - 1$ here. So we may change the to the following notation: $a_1, \ldots, a_{r-1}, b$ and $a_1, \ldots, a_{r-1}, b'$ are the two sequences, and our condition becomes (for some $r > 2$):

$$(*)\qquad \begin{array}{l} b \text{ and } b' \text{ lie in distinct orbits over } a_1, \ldots, a_{r-1}, \\ \text{but in the same orbit over any } r-2 \text{ of the elements } a_i \end{array}$$

where the orbit "over" a set of points is the orbit under the pointwise stabilizer of that set in $G$. An advantage of the last formulation is that with $r$ fixed, it expresses: $\kappa(\mathcal{X}) \ge r$; so it is no longer tied to the case $\kappa = 2$.

**Example 1**

*If $\mathcal{X}$ is a naked set (i.e. $\mathit{Aut}\,(\mathcal{X}) = \mathit{Sym}\,(X)$) then $\kappa(\mathcal{X}) = 2$.*

This is intuitively obvious and can be read off of $(*)$ directly; the orbit of $b$ over $a_1, \ldots, a_{r-1}$ is determined by its orbit over each of the $a_i$ (which amounts to determining whether $b$ is equal to one of the $a_i$).

**Example 2**

*$\kappa(\mathit{Alt}(n), X) = n - 1$ where $\mathit{Alt}(n)$ acts naturally on $X = \{1, \ldots, n\}$.*

Just look at the sequences $(1, \ldots, n-2, n-1)$ and $(1, \ldots, n-2, n)$. These demonstrate that $\kappa(\mathrm{Alt}(n), X) \geq n-1$. The reverse inequality is equally evident, by the same test. (And $\kappa(G, X) < |X|$ for any permutation group, for the same reason.)

**Example 3**

*Let $\mathcal{V}$ be $(GL(V), V)$ with the natural action. Then $\kappa(\mathcal{V}) = \dim V + 1$ if the base field has more than 2 elements, and is $\dim V$ otherwise.*

The relevant pair of sequences $\mathbf{a}, b$; $\mathbf{a}, b'$ is gotten by taking $\mathbf{a}$ to be a basis, $b = \sum_i a_i$, and $b'$ some other linear combination with all coefficients nonzero, assuming the base field has more than two elements. Otherwise one takes $\mathbf{a}$ to be a basis with one element removed; $b = \sum_i a_i$; and $b'$ is an additional basis element. This provides the relevant lower bound for $\kappa$ in each case, and the upper bound is a triviality; in fact the upper bound $\dim V + 1$ will hold for any group of linear transformations on $V$ (though not necessarily for a group of semilinear transformations).

This is a good example, because $\kappa(\mathcal{X})$ is a measure of complexity which is very like the dimension of an arbitrary finite structure, but taking into account all invariant relations among elements, and not just those provided in the original description of the structure $\mathcal{X}$. The next example continues this line of thought.

**Example 4 [CMS]**

*Let $\left[{n \atop k}\right]$ be the permutation group $Sym(n)$ acting on the set of $k$-sets in a set $X = \{1, \ldots, n\}$, with $n \geq 2k$ (or $n > 2k$ if one insists on primitivity). Then $\kappa(\left[{n \atop k}\right])$ is $[\log_2 k] + 2$.*

We may think of $\left[{n \atop k}\right]$ also as a graph: two $k$-sets may be taken to be adjacent if they are disjoint (e.g., for $n = 5$, $k = 2$: the Petersen graph). This is the same structure (has the same group of automorphisms), so the group is 2-closed. According to the formula, though, $\kappa(\left[{n \atop k}\right]) > 2$ for $k > 1$. For example, $k = 2$: $\kappa(\left[{n \atop k}\right]) = 3$ in this case, and this is illustrated by the triples $\{1, 2\}; \{1, 3\}; \{1, 4\}$ and $\{1, 2\}; \{1, 3\}; \{2, 3\}$. It requires a little more care to check that $\kappa \leq 3$ in this case.

For the general case, one first gives a similar example showing $\kappa(\left[{n \atop k}\right]) \geq [\log_2 k] + 2$. The work comes in the reverse inequality. One notes that the orbit of a sequence of $k$-sets is determined by the cardinalities of the atoms in the boolean algebra they generate (we allow degenerate atoms which are empty; they are really labelled by atoms in the free boolean algebra on the same number of generators). One has to show that these numbers are determined by the corresponding data for the subalgebras generated by $[\log_2 k] + 1$ elements; the main point is just that whenever a set is split into two pieces, one of the pieces is at most half as large as the original.

These particular structures play a fundamental role in the general theory, where they occur as "grassmannian" structures (motivated terminologically by the fact that they are the $q = 1$ versions of grassmannians in vector spaces over $\mathbb{F}_q$).

**Example 5**

*Let $\left[{n \atop k}\right]^\circ$ be the permutation group $Alt(n)$ acting on the set of $k$-sets in a set $X = \{1, \ldots, n\}$,* *with* $n \geq 2k$*. Then* $\kappa(\left[{n \atop k}\right]^\circ) = \max(n-k, n-3)$*, except for* $k = 2$*,* $n = 4$*, where the value is* 3*.*

One gets lower bounds for $\kappa$ from examples, and fortunately the values are high enough that one can get matching upper bounds without getting too badly bogged down. We will give some details, since this is not been documented elsewhere.

One may assume $k \geq 2$, and the case $k = 3$, $n = 6$ is best inspected separately.

*The lower bounds*

For $k = 2$ it suffices to consider the induced action on $\{\{i, n\} : i < n\}$, which is equivalent to $\mathrm{Alt}(n-1)$ in its natural representation. This gives the lower bound $\kappa \geq n-2$ in this case, by example 2. For $n = 4$ one examines the orbits of a specific pair of sequences: $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$ and $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$; the same ones which would be used for the full symmetric group.

For $k \geq 3$ we need an explicit example. We set $a_i = \{i\} \cup \{k, k+1, \ldots, 2k-2\}$ for $i \leq k-2$, and $a_i = \{1, \ldots, k-1\} \cup \{i+2\}$ for $k-1 \leq i \leq n-4$. Taking $b = \{1, \ldots, k-2\} \cup \{n-1\}$ and $b' = \{1, \ldots, k-2\} \cup \{n\}$, we claim that $b$ and $b'$ lie in distinct orbits over $\mathbf{a} = (a_1, \ldots, a_{n-4})$ and in the same orbit over any proper subsequence. To see this one has to compute the pointwise stabilizer in $\mathrm{Alt}(n)$ of $\mathbf{a}$ and of its subsequences; this amounts to computing the boolean algebra generated by one of these subsequences, or at least getting sufficient information about the atoms. The atoms of the algebra generated by $\mathbf{a}$ are $\{i\}$ for $i \leq n-2$ together with $\{n-1, n\}$; the pointwise stabilizer of this algebra in $\mathrm{Alt}(n)$ is trivial. Over proper subsequences one identifies larger atoms easily, and one then sees that $b, b'$ lie in the same orbit for their stabilizers in $\mathrm{Alt}(n)$. The case $k = 3$, $n = 6$ should be treated separately from the general case.

This analysis shows that the value we have given for $\kappa(\left[{n \atop k}\right]^\circ)$ is a valid lower bound.

*The upper bounds*

To get matching upper bounds is more troublesome. One considers any pair of sequences $a_1, \ldots, a_\kappa$, $b_1, \ldots, b_\kappa$ illustrating that $\kappa \geq \kappa(\left[{n \atop k}\right]^\circ)$, and one shows that $\kappa \leq n-2$, and $\kappa \leq n-3$ for $k \geq 3$. Leaving aside the case $k = 3$, $n = 6$, this is done by looking closely at the sequence $a_1, \ldots, a_\kappa$, which has the following property, without loss of generality:

$$(*) \qquad a_i \text{ is not in the boolean algebra generated by } (a_j : j \neq i) \text{ for any } i$$

Indeed, if this fails then $a_i$ is fixed over $(a_j : j \neq i)$ by the full symmetric group, and then Example 4 applies to show $\kappa \leq [\log_2 k] + 2$, a bound which yields $\kappa \leq n-2$ for $k = 2$, apart from the known exception $n = 4$; and the same estimate yields $\kappa \leq n-3$ for $k \geq 3$ and $n \geq 2k$. So we need only consider the case $(*)$.

As $k \geq 2$, the $a_i$ are not all disjoint, so we may suppose $a_1$ meets $a_2$; then the boolean algebra generated by $a_1, a_2$ will have 4 atoms, and applying $(*)$ the boolean algebra generated by the $a_i$ will have at least $\kappa + 2$ atoms, proving $\kappa \leq n-2$ and indicating that the case $\kappa = n-2$ is extreme.

Suppose now that $k \geq 3$. We must eliminate the possibility $\kappa = n-2$. One finds $a_1, a_2, a_3, a_4$ generating 7 atoms, which suffices; in fact one finds $a_1, a_2, a_3$ generating 6 atoms in many cases. A useful observation is that the $a_i$ separate points as there are $n-2$ of them and the algebra they generate must have at least $n$ atoms.

One should dispose first of the case $n = 2k$, $k \geq 4$, leaving the case $n = 6$, $k = 3$ to be dealt with by inspection. The methods are much the same as those used in the main case: $k \geq 3$, $n > 2k$.

Suppose first that one can choose $a_1, a_2$ so that $|a_1 \cap a_2| = k - 1$. As the $a_i$ separate points, we may take $a_3$ splitting $a_1 \cap a_2$. Then $a_i$ cannot split $(a_1 \cup a_2)'$ as otherwise we already have 6 atoms. On the other hand $|(a_1 \cup a_2)'| \geq k$ so $a_i$ cannot contain $(a_1 \cup a_2)'$ either, and $a_3 \subseteq (a_1 \cup a_2)$. Hence $a_i$ consists of $k-2$ elements of $a_1 \cap a_2$ and the two elements of $(a_1 \cup a_2) \setminus a_1 \cap a_2$. In particular this analysis shows that $a_1 \cap a_2$ is not contained in any other $a_j$, as otherwise the same analysis would apply with $a_j$ in place of $a_2$.

Now one uses fully the fact that the $a_i$ separate points. Let $A = a_1 \cup a_2$. We may suppose that $A = \{1, \ldots, k+1\}$, and then after relabeling that $a_i = A - \{i\}$ for $i \leq k$. Then if $a_i$ splits $A'$, we find $a_i \cap A = \emptyset$ as otherwise we can choose $i_1 < i_2 \leq k$ so that $a_{i_1}$, $a_{i_2}$, $a_i$ generate at least 6 atoms. If now follows just by counting that $\kappa \leq n - 3$; in addition to the $k$ elements $a_i$ already determined above, there are at most $n - k - 3$ disjoint from $A$, using $(*)$.

The rest goes similarly and more quickly. Taking $a_1, a_2$ so that $l = |a_1 \cap a_2| > 0$ is minimized, one finds easily that $l = 1$ (distinguishing the cases $l > k/2$, $l \leq k/2$ along the way). Then splitting $(a_1 \cup a_2)'$ and $a_1 \setminus a_2$ by elements $a_3$, $a_4$, one either gets the desired subalgebra on 3 or 4 generators, or in the remaining case one finds $|a_3 \cap a_4| = k - 1$, the case treated at the outset. ∎

We now consider the binary case further. An *affine* group of dimension $d$ is a subgroup of $\mathrm{A\Gamma L}(V)$ containing the translation subgroup $V$, with $V$ $d$-dimensional; it is *strictly linear* if it is a subgroup of $\mathrm{AGL}(V)$. Here $\mathrm{AGL}(V) = V \rtimes \mathrm{GL}(V)$ and $\mathrm{A\Gamma L}(V) = V \rtimes \Gamma\mathrm{L}(V)$.

**Example 6**

*A primitive 1-dimensional strictly linear affine group $G$ is binary if and only if it is cyclic or dihedral; otherwise $\kappa(G) = 3$.*

The structures in the cyclic or dihedral cases are directed or undirected cycles, respectively, of prime order.

Note that the stabilizer of any two points is trivial in this case. This already forces $\kappa(G) \leq 3$. So the only point is to identify the binary cases. Since the group is 1-dimensional, we will denote it $\mathbb{F} \rtimes \mu$ where $\mathbb{F}$ is the additive group of the base field and $\mu$ is a subgroup of the multiplicative group. We assume $|\mu| > 2$ and we show that $G$ is not binary. Consider the triples $(0, -1, g)$ and $(0, -1, g^{-1})$ with $g \in \mu$, $g \neq \pm 1$. Since the stabilizer of two points is trivial and $g \neq g^{-1}$, these triples lie in distinct orbits. However the pairs $(0, g)$ and $(0, g^{-1})$ lie in the same orbit under $\mu$, and the pairs $(-1, g)$ and $(-1, g^{-1})$ lie in the same orbit under $G$ – translate by $+1$ and multiply by $g^{-1}$. Thus we have an explicit violation of binarity.

We will push this a bit further because it will complete the list of currently known primitive binary permutation groups; and the determination of all such would be very welcome.

**Example 7**

*Let $G$ be a primitive 1-dimensional affine group, not strictly linear. Then $\kappa(G) \leq 4$ and $G$ is binary if and only if $G$ has the form $\mathbb{F}_{q^2} \rtimes \mu_{q+1} \rtimes \langle \sigma \rangle$ with $\sigma$ of order 2.*

*Proof*:

For the upper bound, one shows that the permutation group given by the stabilizer of two points is binary. (Any action of a cyclic group is binary.) The typical value is probably 4 and one may be able to identify all the exceptions, but this has not been done. For our purposes, it suffices to identify the binary ones. We leave to the reader the computation that the the groups listed are binary; we will analyse these examples a little more below. The main point is that no other binary examples are to be found in this class.

Suppose $G$ is binary, $G = \mathbb{F} \rtimes H$ with $H \le \mathbb{F}^\times \rtimes \Gamma$, $\Gamma = \operatorname{Aut}\mathbb{F}$. Then the argument given in the linear case applies, but shows only that for $h = a\sigma \in H$, $1^h = 1^{h^{-1}}$, i.e. $a^\sigma = a^{-1}$. Let $\mu$ be the projection of $H$ on $\mathbb{F}^\times$ and let $\Gamma_\circ$ be the projection of $H$ on $\Gamma$; then for $\sigma \in \Gamma_\circ$, we have just seen that $\sigma^2$ fixes $\mu$. However the field generated by $\mu^{\Gamma_\circ}$ is $\mathbb{F}$, by primitivity, so $\sigma^2 = 1$. Thus $|\Gamma_\circ| \le 2$ and, since $G$ is not strictly linear, $\Gamma_\circ = \langle\sigma\rangle$ with $\sigma$ of order 2. Thus $\mathbb{F} = \mathbb{F}_{q^2}$ for some prime power $q$, and $a^\sigma = a^{-1}$ for $a \in \mu$.

Thus $H$ is a subgroup of the desired group $H_1 = \mu_{q+1} \rtimes \langle\sigma\rangle$. Take $a\sigma \in H$. Then $a^\sigma = a^{-1}$ so $a = b/b^\sigma$, some $b$, and conjugating $G$ by $b$ (as an element of $\mathbb{F}^\times$), we may take $\sigma \in H$. Thus $H = \mu_\circ \rtimes \langle\sigma\rangle$ with $\mu_\circ \le \mu_{q+1}$. We will show $\mu_\circ = \mu_{q+1}$.

As $G$ is primitive, $\mu_\circ$ contains some $r \ne \pm 1$. Let $s \in \mu_{q+1}$ be arbitrary. Solve $b^\sigma/b = s$ for $b$. Consider the triples $(0, b, b/(r+1))$ and $(0, bs, bs/(r+1))$. Any pair from the first triple is carried to the corresponding pair of the second triple by one of the maps $\sigma$, $\sigma r$, or $\sigma r^{-1}$. By binarity there is a transformation $g \in G$ taking $(0, b, b/(r+1))$ to $(0, bs, bs/(r+1))$. Then $g \in H$, and since $b^g = bs$ it follows that $g = \sigma$ or $g = s$. As $[b/(r+1)]^g = bs/(r+1)$ we can exclude the former possibility and thus $g = s$. So $s \in G$, for any such $s$. ∎

**Remark**

*Let $\Gamma_q$ be the binary structure corresponding to the binary 1-dimensional affine group $\mathbb{F}_{q^2} \rtimes \mu_{q+1} \rtimes \langle\sigma\rangle$. Then $\Gamma_q$ is a symmetric graph with an edge coloring by $q-1$ colors.*

The symmetry means that we can solve the equation $a^{\sigma r} = -a$ with $r \in \mu_{q+1}$ for any $a$. This just means $-a^\sigma/a \in \mu_{q+1}$, which is the case. This also shows that the orbit over 0 of any point has order $q+1$ and thus there are $q-1$ such orbits.

In particular we have an ordinary uncolored graph only for $q \le 3$. The case $q = 2$ is degenerate and for $q = 3$ the associated graph is the sporadic homogeneous graph $K_3^2$. This is a second way of accounting for this example, quite different from that of §1. In the case $q = 4$ one has three colors, and again this example turns up as one of a small number of sporadics in the classification corresponding to that case, which was carried out long ago by Lachlan. This particular example is also used to show that the Ramsey number $r(3, 3, 3) \ge 17$, as it provides a graph of order 16 with a 3-edge coloring without chromatic triangles (an example given by Andrew Gleason).

We do not know of any other finite primitive binary homogeneous structures, apart from those we have seen: cyclic, dihedral, $\Gamma_q$, and of course $\operatorname{Sym}(n)$ acting naturally.

**Example 8:** $n^d$

Our final example is considerably more subtle. Recall that $n^d$ is the wreath product (acting on a power) of the natural representations of $\mathrm{Sym}(n)$ and $\mathrm{Sym}(d)$. The value of $\kappa(n^d)$ has been worked out by Saracino [Sa]. [CMS] contains estimates of arities of wreath products $\mathcal{X}^{\mathcal{Y}}$ in general, and shows that the upper bound given there provides the exact value for "most" values of $n$ and $d$, specifically: for $n \geq 2[\log_2 d] + 2$.

The upper bound is given in general by

$$\kappa(\mathcal{X}^{\mathcal{Y}}) \leq \kappa(\mathcal{X}) \cdot \kappa(\mathcal{P}(\mathcal{Y}))$$

where $\mathcal{P}(\mathcal{Y})$ is the power set of $\mathcal{Y}$ (with automorphism group $\mathrm{Aut}(\mathcal{Y})$, acting naturally). We have $\kappa(n) = 2$ and $\kappa(\mathcal{P}(\{1, \ldots, d\})) = [\log_2 d] + 1$ [CMS], and thus $\kappa(n^d) \leq 2([\log_2 d] + 1)$; and this turns out to be the exact value for $n \geq 2[\log_2 d] + 2$.

The value of $\kappa(n^d)$ for relatively small values of $n$ follows a more complicated rule, which takes on a distinctly simpler form if we look instead for a formula which computes, in terms of given $n$ and $\kappa$, the minimum value of $d$ for which $\kappa(n^d) \geq \kappa$. For technical reasons it is better to define a very similar function $\delta(\kappa, n)$ as the least value of $d$ such that there are two sequences of length $\kappa$ in $n^d$ of length $\kappa$ which witness that $\kappa(n^d) \geq \kappa$, in the sense of $(*)$ above, and which are not conjugate to sequences occurring in $(n-1)^d$. The most important point here is that we choose to express $d$ in terms of $n$ and $\kappa$, getting a moderately complex set of conditions which can be explicitly but more awkwardly solved for $\kappa$ in terms of $n$ and $d$.

The main formulas that result are:

$$\delta(\kappa, n) = \begin{cases} 2^{\kappa - n/2 - 1} & \text{for } \kappa > n \text{ even} \\ 3 \cdot 2^{r - \frac{n+5}{2}} & \text{for } n \geq 5 \text{ odd and } \kappa > n \text{ even} \\ 9 \cdot 2^{r - \frac{n+7}{2}} & \text{for } \kappa > n \geq 5 \text{ with } \kappa \text{ and } n \text{ odd} \\ 27 \cdot 2^{r - n/2 - 5} & \text{for } \kappa > n \geq 8 \text{ with } \kappa \text{ odd and } n \text{ even} \end{cases}$$

with similar formulas, and some exceptional values, covering the remaining cases.

The method used is to attach a clearcut combinatorial invariant to the orbit of an $r$-tuple in a wreath product. This can be done quite generally, though it is rather messy in general. In the case of $n^d$ the type of an $r$-tuple can be encoded by a multiset consisting whose elements are equivalence relations on the set $\{1, \ldots, r\}$ having at most $n$ classes (and the final, technical part of the definition of $\delta$ corresponds to the condition that at least one of these relations should actually have $n$ classes). The data referred to in condition $(*)$ above would then be two such multisets which coincide whenever a single index $i$ (with $1 \leq i \leq r$) is deleted from all the relations in both multisets). The parameter $\delta$ which is sought is the number of relations occurring in each of these multisets; examples show that the stated values are valid upper bounds, and one must then prove matching lower bounds for the sizes of such multisets. This is what Saracino has done.

All of this yields:

For $n \leq 2[\log_2 d] + 2$,

$$\kappa(n^d) = 2[\log_4 \alpha_n 2^{n/2+1} d] + \epsilon$$

with $\epsilon = 0$ or 1 unless $n = d = 3$,

and with $\alpha_n = 1$ for $n$ even and $4/3\sqrt{2}$ for $n$ odd.

For $n \geq 2[\log_2 d] + 2$,

$$\kappa(n^d) = 2[\log_2 d] + 2$$

Thus a bound on $\kappa$, as specified at the outset in one of Lachlan's classification problems, will pick out: (1) the families $n^d$ for which $n$ is arbitrary and $d < 2^{\kappa/2}$; and (2) the structures $n^d$ for which $n \leq 2[\log_2 d] + 2$ and $2[\log_4 \alpha_n 2^{n/2+1} d] + \epsilon \leq \kappa$, i.e. roughly speaking $n/2 + 1 + [\log_2 d] \leq \kappa$. E.g. for $\kappa \leq 10$, the infinite families are $n^d$ for $d \leq 31$, and the last sporadic structure would be $2^{511}$; so here the number of sporadics is exponentially greater than the number of well-behaved families. (Note that a bound on $r_\kappa$ is not of great importance in this particular context.)

Other examples which one would naturally consider include multiply transitive groups, and the action of $\mathrm{Sym}\,(n)$ or $\mathrm{Alt}(n)$ on partitions of $\{1, \ldots, n\}$ of a fixed type. The multiply transitive representations do not generally present any particular difficulties: the sharply transitive ones of degree $t$ have $\kappa = t + 1$, apart from the natural representation of $\mathrm{Sym}\,(n)$, and in non-sharply transitive cases the value is at least $t + 1$ and not much greater. Actions on partitions are not well understood at all, and to round out this collection one would like to have decent estimates for this case (at least under the full symmetric group) and for corresponding wreath products. A few such problems are discussed further in §9.

## §4. Finite and smoothly approximable homogeneous graphs

We will now present the classification of the finite and countably infinite homogeneous graphs, beginning with the finite ones. It will then be obvious that the list of finite homogeneous graphs can be completed by adjoining their natural infinite limits, and we will need to give a precise definition which describes this completion process in a general setting, in terms of a notion of *smooth approximability*. For homogeneous structures this property is equivalent to one of the fundamental notions of pure model theory: *stability*. There are other infinite homogeneous graphs that fall outside the smoothly approximable scheme; though these have also been classified explicitly, this part of the classification has not been subsumed by any more general theory.

### The finite homogeneous graphs

1. *$m \cdot K_n$: the disjoint sum of $m$ complete graphs, each of order $n$.*

1′. *$(m \cdot K_n)'$: the complementary graphs, complete $m$-partite graphs with parts of order $n$.*

2. *The pentagon $C_5$.*

3. *$3^2 = K_3^2 = L(K_{3,3}) = \Gamma_3$. This has been described at various points above as a wreath product of degenerate structures, the line graph of a complete bipartite graph $K_{3,3}$, or an*

*unusual binary structure associated with a 1-dimensional affine but not strictly linear group over* $\mathbb{F}_3$*.*

Actually in the language of §2, the graphs of types 1 and $1'$ are isomorphic structures; we view antiisomorphisms as a permutation of the language. But this still does not entitle us to view them as isomorphic *graphs*, so one often carries the complementary family along. Each of the sporadic examples (2) and (3) is self-dual. These two graphs can be viewed as degenerate members of families occurring naturally with $\kappa > 2$, in a way which is interesting in the case of $3^2$ but much less so for $C_5$, where the natural family to consider would be one in which the points of $C_5$ are replaced by equivalence classes of arbitrary size. Viewing $C_5$ as part of the family $\{C_n\}$ and $3^2$ as part of the family $\{\Gamma_q\}$ is also very reasonable, but this does not correspond to the hierarchy by degree of homogeneity in the form it to which Lachlan's finiteness theorems apply: in the families $\{C_n\}$ and $\{\Gamma_q\}$, we have $\kappa = 2$, but $r_\kappa$ is unbounded.

One rounds out this list by allowing $m$ and $n$ to become countably infinite in families 1 and $1'$. (However we require them to be countable.) This process needs to be described more intrinsically, and the following definition expresses what we will mean by the *infinite version* of a family of finite structures, in general.

**Definition**

*Let $\mathcal{X}$ be a structure such that $r_k(\mathcal{X})$ is finite for all $k$.*

1 *The structure $\mathcal{X}$ is said to be a* smooth limit *of finite structures if every finite subset $X_\circ$ of the universe $X$ is contained in a set $X_1$ for which the induced structure $\mathcal{X}_1$ is "smoothly embedded" in $\mathcal{X}$, where the latter condition is defined as follows.*

2 $\mathcal{X}_1$ *is* smoothly embedded *in $\mathcal{X}$ if any two finite sequences of elements of $\mathcal{X}_1$ which lie in the same orbit under $Aut(\mathcal{X})$ also lie in the same orbit under $Aut(\mathcal{X}_1)$, where $Aut(\mathcal{X}_1)$ is the group induced on the underlying set $X_1$ by its setwise stabilizer in $Aut(\mathcal{X})$. ($\mathcal{X}_1$ is an induced substructure of $\mathcal{X}$).*

The finiteness assumption on all $r_k$ is harmless in the context of homogeneous structures with $\kappa$ and $r_\kappa$ bounded; any smooth limit of such structures will inherit the same bound on $\kappa$ and $r_\kappa$, and homogeneity then implies that $r_k$ is bounded for all $k$. Smooth approximability is a key condition which is of interest outside the homogeneous context as well; other examples are provided by infinite-dimensional vector spaces over finite fields, which may be decorated with the inner products or quadratic forms which define the various classical groups.

The consideration of these infinite analogs of the finite structures is quite helpful. For example we can reduce list (1) above to the finite smoothly embedded substructures of one structure: $\infty \cdot K_\infty$. It can be convenient to replace infinite families of finite structures by a single infinite limit.

We have mentioned that there are other countably infinite homogeneous infinite graphs, and that these have also been classified. One such is Rado's graph, or "the" random graph, which may be described as follows: if a graph $G$ is constructed by putting in edges randomly and independently, with constant probability $p$ $(0 < p < 1)$, then there is a single graph $G_\infty$ such that with probability

1 the random graph $G$ is isomorphic to $G_\infty$. For each $n$ there is a similar graph $G_n$, called the generic countable graph which contains no clique of order $n$; this cannot be defined probabilistically, but can be defined using topology in place of measure theory. One views the collection $\mathcal{G}_n$ of $K_n$-free graphs as a compact topological space, and one looks for a graph $G_n$ such that the set of graphs in $\mathcal{G}_n$ which are not isomorphic to $G_n$ is topologically meager. There is a better description of these graphs in terms of amalgamation classes (§6) but this will certainly do for the moment.

The full classification of the homogenous graphs is then:

I. The homogeneous graphs which are smoothly approximated by finite graphs;

II. The generic graph omitting the $n$-clique $K_n$, for fixed $n \geq 3$;

II′. The complements of the graphs of type II;

III. The Rado graph (self-dual).

If the Rado graph is an unfamiliar object, one can approach it by considering the rational order $(\mathbb{Q}, <)$ as an analog. Any ordered set can be viewed as a directed graph (in fact a tournament) and one can consider the homogeneous orders. It follows directly from the definition that the only homogeneous orders are: (1) the "order" on one point; and (2) a dense linear order without endpoints, which may be taken to be $(\mathbb{Q}, <)$. The dichotomy occurring in the theory of homogeneous graphs occurs here in an extreme form, as there is one extremely finite example and one extremely infinite example; $(\mathbb{Q}, <)$ has no finite smooth approximation with more than one element. Just as the ordering of $\mathbb{Q}$ can be characterized by its density properties, the Rado graph can be characterized by analogous density properties, stating that any finite subgraph can be extended in all possible ways by the addition of a suitable additional vertex. Peter Winkler and other fans of Arlo Guthrie call this the Alice's restaurant property.

The part of this classification which goes beyond the smoothly approximable case is found in [LW]. Some other classification results of a similar character have been found; the proofs are purely combinatorial, relying on Ramsey's theorem, and are usually long. One may also detect in the case at hand a striking dichotomy, an instance of a more systematic dichotomy in model theory uncovered by Shelah: the stable/unstable distinction. This more technical idea enters heavily into the proofs, and occasionally into the statements, of the main results.

One approach to stability is to define one or more notions of dimension for arbitrary structures, referred to in model theory as ranks – a confusing terminology when used in connection with permutation structures – and to call a structure stable if the rank or ranks used are finite. Following Lachlan, we will use a single notion of rank which is well adapted to the group theoretic viewpoint. We will call this particular notion the *orbit height*.

**Definition**

*Let $\mathcal{X}$ be a homogeneous structure with $\kappa(\mathcal{X})$ and $r_{\kappa(\mathcal{X})}$ finite.*

1. A tree of orbits *of height $n$ in $\mathcal{X}$ is a complete binary branching tree of height $n$, with each vertex labeled by a pair $(A, O)$ with $A$ a finite subset of $X$ and $O$ an orbit of the pointwise stabilizer in $Aut(\mathcal{X})S$ of $A$, such that the sets $A$ increase as one moves along a path in the*

*tree, and the two orbits lying below a given vertex are contained in the orbit attached to that vertex, and are pairwise disjoint.*

2. *The* orbit height *of $\mathcal{X}$ is the maximum value of $n$ (or $\infty$) for which $\mathcal{X}$ has a tree of orbits of height $n$.*

3. The structure $\mathcal{X}$ is stable if its orbit height is finite.

One sees easily that smoothly approximable homogeneous structures with $\kappa$ and $r_\kappa$ finite are stable. The converse is an important structural result. Lachlan's theory consists of one part which provides a structural analysis and finiteness theorem for homogeneous structures in which not only $\kappa$ and $r_\kappa$ are fixed, but a bound on the orbit height is also fixed in advance. To complete the theory one must also bound the height in terms of $\kappa$ and $r_\kappa$. We will look at this issue more closely in the next section.

## §5. The Coordinatization Theorem

Lachlan proposes to consider the following classification problems:

$(P_{\kappa,r})$ Given $\kappa$ and $r$, classify the homogeneous structures $\mathcal{X}$ with $\kappa(\mathcal{X}) \leq \kappa$ and $r_\kappa(\mathcal{X}) \leq r$

It remains to be seen what constitutes a solution; in other words, what constitutes the specification of an infinite family of related structures. There are two approaches to this. For example, if the family in question is $\{m \cdot K_n\}$, disjoint sums of complete graphs of fixed size, we can pass to the limit structure, $\infty \cdot K_\infty$, and define the family in question as the set of finite approximations to the infinite limit – or we can look for the invariants $m, n$ directly. It is reasonable to combine these approaches: identify the limit structures; show that there are finitely many; show that all sufficiently large structures approximate infinite limit structures; and identify the numerical invariants that control isomorphism types within each infinite family. Thus part of the work will take place "at infinity", in a model theoretic context, and part will take place in the "large finite", using combinatorics and permutation group theory. One speaks also of "shrinking" and "stretching"; shrinking an infinite structure to its finite approximations – which is relatively easy – and stretching a large finite structure to its infinite version (which amounts to giving an "explanation" of the finite structure as an approximation to an infinite one). To put the matter more briefly: there are finitely many infinite homogeneous structures of a given type, from which all sufficiently large finite homogeneous ones (and all their infinite limits) arise by a shrinking process. This is the basic finiteness result.

The notion of shrinking is given by smooth embedding: any finite structure which is smoothly embedded in a larger structure can be considered as a shrinking of the larger structure (it might be best at some later stage to exclude some very small smoothly embedded substructures). One can finesse the issue of stretching temporarily by deciding that a stretching of a structure is a structure in which it smoothly embeds, or in other words stretching is the reversal of shrinking; this is reasonable, but it just postpones the question of the existence of infinite stretchings, which is one of the main points: when can a finite structure be interpreted as a "template" for an infinite structure?

In any case, with this terminology, we may state:

**Theorem 1**

*Let $\sigma = (n_1, \ldots, n_r)$ be a type of relational system. There are finitely many homogeneous structures $\Gamma_i$ of type $\sigma$ such that every finite homogenous structure of type $\sigma$ is obtained by shrinking one of the $\Gamma_i$.*

To put some flesh on these bones, it is necessary to look at numerical invariants. These are the analogs of “dimension” in vector spaces, but in the context of relational structures they are of a relatively degenerate type, as illustrated by the family of homogeneous graphs $m \cdot K_n$; here $m$ is the number of classes of an invariant equivalence relation, and $n$ is the size of the classes. This is close to the general case, for Lachlan’s context.

A less trivial illustration is furnished by the graphs $\left[{n \atop k}\right]$. Here there should be a single numerical parameter $n$ ($k$ is fixed, as it can be bounded in terms of $\kappa$). Evidently the parameter $n$ is encoded by $\left[{n \atop k}\right]$ with considerably more subtlety than in the case of $m \cdot K_n$. As mentioned in §3, we will think of $\left[{n \atop k}\right]$ as a grassmannian structure associated with a degenerate geometry. As the examples in §3 may suggest, one would expect the geometries associated to a homogeneous structure to be degenerate (or of bounded size, hence more or less unproblematic). Accordingly we make the following definitions in general.

**Grassmannians and invariants**

1. *A coordinatizing structure will be a structure $\Delta$ carrying an equivalence relation $E$ with finitely many classes, such that the $E$-classes are permuted transitively by $\mathit{Aut}\,\Delta$, and the group fixing the classes setwise is the product of the full symmetric groups on each class.*
2. *A $k$-grassmannian of $\Delta$ is the structure whose points are the subsets of $\Delta$ meeting each class in $k$ points, whose full automorphism group is $\mathit{Aut}\,\Delta$ with the natural action. The invariant attached to a grassmannian is the size of each equivalence class in $\Delta$.*
3. *The invariants attached to a homogeneous structure $\Gamma$ are the invariants attached to all $k$-grassmannian structures which occur as primitive sections of $\Gamma$, with $k$ at most half the size of each equivalence class in $\Gamma$.*

The importance of these invariants can be seen in the following result from [CL].

**Proposition 1 (Coordinatization)**

*For any specified type $\sigma$ of relational structure, there is a bound $m$ such that for every homogeneous relational structure $\Gamma$ of type $\sigma$ and any maximal ($\mathit{Aut}\,\Gamma$)-invariant equivalence relation $E$ on $\Gamma$, one of the following holds:*

1. *$|\Gamma/E| \leq m$; or*
2. *$\Gamma/E$ is a grassmannian of a coordinatizing structure.*

The proof of this relies heavily on permutation group theory, notably the O’Nan-Scott Lemma and the classification of the finite simple groups. Both the model theoretic content and the permutation group theoretic analysis are reviewed in [KL]. Stronger results in a similar vein are found in [KLM] and are essential to the development of theories of broader scope.

This result is the starting point for the study of the "variable" numerical invariants associated with a homogeneous structure. It should be clear enough how one expands or shrinks a grassmannian structure; to do either to a more general homogeneous structure requires more attention. Shrinking can be defined rather directly: one shrinks the coordinate structures and then sees which elements of the structure should be kept.

Stretching, on the other hand, is a problem. Indeed, the essence of the matter is to determine how large a dimension should be in order that it can be stretched freely, and this is a nontrivial question, very much at the heart of determining at what point the sporadic objects run out and the general case is encountered. The main result of [KL] is a clean approach to this problem, which we will indicate very briefly in the next section.

## §6. **Amalgamation**

If $\Gamma$ is a homogeneous structure, then $\mathrm{Sub}(\Gamma)$ denotes the class of finite structures which are isomorphic with induced substructures of $\Gamma$. A countable homogeneous structure $\Gamma$ is determined up to isomorphism by $\mathrm{Sub}(\Gamma)$. Furthermore the relevant classes of finite structures, that is, those which occur as $\mathrm{Sub}\Gamma$ with $\Gamma$ homogeneous, are easily characterized by their intrinsic properties: closure under isomorphism and induced substructure, and the *amalgamation property*, which we now define.

**Definition 2**

1. *An* amalgamation problem *is a triple of structures* $\Gamma_0, \Gamma_1, \Gamma_2$ *together with a pair of embeddings* $\iota_i : \Gamma_0 \longrightarrow \Gamma_i$ *for* $i = 1, 2$. *A* solution *to such a problem is a structure* $\bar{\Gamma}$ *and embeddings* $\bar{\iota}_i : \Gamma_i \longrightarrow \bar{\Gamma}$ *so that* $\bar{\iota}_1\iota_1 = \bar{\iota}_2\iota_2$.
2. *A class* $\mathcal{A}$ *of structures has the* amalgamation property *if any amalgamation problem involving structures in the class has a solution in* $\mathcal{A}$.

To see that $\mathrm{Sub}(\Gamma)$ has the amalgamation property for $\Gamma$ homogeneous, note that we may take $\Gamma_0, \Gamma_1, \Gamma_2$ as substructures of $\Gamma$, and after applying suitable isomorphisms we may assume the embeddings $\iota_i : \Gamma_0 \longrightarrow \Gamma_i$ are *inclusions*; in this case, set $\bar{\Gamma} = \Gamma_1 \cup \Gamma_2$ and let the $\bar{\iota}_i$ be inclusions as well.

Somewhat less evident, but straightforward nonetheless, is the fact that we can construct a homogeneous structure from any amalgamation class closed under isomorphism and induced substructure.

This gives us a quick avenue, in principle, to stretching. Given a large finite homogeneous structure $\Gamma$, let $\mathcal{A}$ be the class of all finite structures which embed "locally" into $\Gamma$ (we will say in a moment what this means). If $\mathcal{A}$ is an amalgamation class, then the corresponding homogeneous structure should play the role of the "stretch" of $\Gamma$. To make this precise we define $\mathrm{Sub}(n, \Gamma)$ as the collection of finite structures such that every substructure of size at most $n$ embeds into $\Gamma$. With

$n$ fixed, $\mathrm{Sub}(n, \Gamma)$ may be taken as the class of structures that embed locally into $\Gamma$. In [KL] the following is proved, as Theorem 9.2:

**Theorem**

*For each fixed type $\sigma$ of relational structure, there is a fixed $N$ such that for every homogeneous finite $\Gamma$ of type $\sigma$, there is some $n \leq N$ for which $\mathrm{Sub}(n, \Gamma)$ is an amalgamation class, and for which the corresponding homogeneous structure is smoothly approximable.*

Obviously this sketch leaves out not only all of the proofs, but a number of highly relevant statements and definitions, particularly relating to the treatment of the numerical invariants in the preceding section and their relation to the theory of stretching. We refer the reader to [KL] for more details and further references, as well as a discussion of effectivity.

## §7. Rank: stability and bounds

We defined the *orbit height* of a permutation group in §4 and we called a homogeneous structure *stable* if its orbit height is finite.

For example, the orbit height of a set on which the full symmetric group acts is 1, since one cannot split an orbit into two orbits of size greater than 1. On the other hand the orbit height of the rational order is infinite, since one can split an infinite interval $(a, b)$ into two such, and repeat. The orbit height is a notion of "dimension" (a measurement of noetherianity). The reader can test this stable/unstable dichotomy on the explicit list of homogeneous graphs given in §4.

The usual definitions of stability in model theory are of a more general character, but reduce to this notion in the homogeneous case. To some extent our initial presentation put the cart before the horse. In the most satisfactory version of Lachlan's theory, the starting point is stability. One proves eventually:

**Proposition 2**

*Within the class of homogeneous structures of a fixed type, the stable ones are those which can be smoothly approximated by finite homogeneous structures of the same type.*

Lachlan's theory gives the classification of all stable homogeneous structures of a fixed type, which includes the classification of the finite ones, but which makes use of the structure of the infinite ones along the way: this class falls into finitely many families, each parametrized by finitely many numerical invariants; and when all invariants are made finite, the resulting structure is finite.

The result on coordinatization (Proposition 1) is actually *equivalent* to the existence of a uniform bound on the orbit heights for stable homogeneous structures of fixed type, though one needs a good dose of model theory to see this. In any case, in all versions of this theory to date one gets the coordinatization result from group theory (at the finite level) and this then allows the introduction of model theoretic techniques to handle the infinite limits, returning, eventually, to the large finite structures. Thus our understanding of the infinite structures requires information coming from their finite approximations, but at the same time our understanding of the finite structures depends on a consideration of their infinite limits.

For unstable homogeneous structures, as we have noted earlier, we have explicit classifications in some nontrivial cases, but no theory.

## §8. Smoothly approximable structures

Lachlan's theory can be extended *mutatis mutandis* to smoothly approximable structures, and if one does so then the supply of geometries is enlarged to include all of the classical geometries (in both affine and projective flavors) which are so visibly absent in the homogeneous case. This is one of the most salient differences. It would be easy enough to carry this aspect along through the theory, but eventually the less trivial geometric structure in these geometries has a certain impact on developments and a more sophisticated dose of model theory comes into play, modeled on standard developments in stability theory. As it happens, the geometries involved are not in fact stable in general, which is one indication that some price is to be paid for the generalization. This is discussed in [Hr, Ch].

Smooth approximability is a somewhat peculiar hypothesis (as a point of departure) from the point of view of model theory, but Lachlan observed that in view of the central role it plays in the theory of stable homogeneous structures and related developments, it is reasonable to work in this category, discarding any more special model theoretic hypotheses. Furthermore dealing with this class amounts to dealing at the finite level with permutation groups with a bound on the number of orbits on 4-tuples; and though one might have preferred to replace 4 by 2 here, this is certainly a natural class to work with.

It is not at all clear *a priori* that the only geometries which are relevant here are classical ones; and strictly speaking, this is not even true, as in characteristic 2 there is another family of geometries falling just outside the classical camp. One can read off the relevant list of geometries from the explicit classification of the large finite primitive structures with a bounded number of orbits on 4-tuples [KLM as modified in Mp]. It is fortunate that permutation group theory is so effective on primitive structures while model theory has good tools for reducing imprimitive structures to primitive ones.

We will mention two of these geometries, to give an indication of the sort of objects that come into play at this level.

**Example 1: polar geometry**

*One has a pair $(V, V^*)$ with $V$ an infinite dimensional vector space over a finite field (of countably infinite dimension) and $V^*$ a dense subspace of its dual. This is the smooth limit of analogous finite dimensional structures in which $V^*$ is the full dual of $V$.*

In model theory we consider the pair $(V, V^*)$ as a single geometry. One of the complications that arises is that in encountering $V$ embedded in a larger structure one cannot easily tell whether it occurs as a subspace with no additional structure, or as half of a polar pair. For example, we might have $V$ together with some grassmannian associated with $V^*$; we would then have to reconstruct $V^*$ from its grassmannian in order to have the second component of the polar pair. None of this is terribly troublesome; it just needs to be dealt with. One of the basic ideas of pure model theory, expressed by the theory of "orthogonality", is that *geometries interact trivially*: either they can be identified (with some deformation), or they are unrelated. If $V$ and $V^*$ are taken to be *separate* geometries, this orthogonality principle is lost; but if such things are not allowed, the principle of orthogonality can be saved.

**Example 2: The quadratic geometries**

There is a family of geometries in characteristic 2 which appears to blend some of the features of affine and projective geometry. It arises because orthogonal groups are contained in symplectic groups in characteristic 2.

The underlying set $Q$ of this geometry in a given (even) dimension is defined as follows: let $V$ be a symplectic space of dimension $2n$, with inner product $( \ , \ )$, and let $Q$ be the set of all quadratic forms $q$ for which we have $q(x+y) = q(x) + q(y) + (x, y)$.

There is a regular action of $V^*$ on $Q$ (or of $V$ on $Q$ after identifying $V$ and $V^*$): $q.\lambda = q + \lambda^2$ for $q \in Q$, $\lambda \in V^*$. In this sense (but only in this sense) the space looks affine. The stabilizer of a point $q$ is of course the associated orthogonal group $O(q)$. The same structure exists in an infinite dimensional version but there is one delicate point: the witt defect of $q \in Q$ is well defined when the dimension is finite, and in passing to a *smooth limit* a formal witt defect function $\omega$ is inherited, though the witt defect itself is no longer very meaningful. (The equivalence relation defined by "same witt defect" can be defined intrinsically from the geometric structure; only the value of the witt defect is lost.) The upshot of all this is that in the infinite limit, one takes $Q$ together with $V^*$ and its regular action, as well as the structure on $V^*$, together with the formal witt defect.

This is the least classical of the geometries that come into play here.

## §9. $\kappa(\mathcal{X})$: problems

The question naturally arises: for one's favorite *primitive* structures, how is $\kappa(\mathcal{X})$ computed in practice? Much of what is understood about this was reviewed in §3. We comment here on cases we do not understand. There are tools for the general analysis of primitive permutation groups, beginning with the O'Nan-Scott lemma and continuing in work of Aschbacher, and among other things, one would like to know how $\kappa$ behaves relative to this. Some general theory for wreath products is found in [CMS] with reasonably satisfying results, some loose and very general, others more detailed under rather specific hypotheses. The following test problem, which remains wide open, is my personal favorite in this area:

**Problem 1**

*Determine the finite primitive binary structures.* ($\kappa(\mathcal{X}) = 2$)

Model theorists might be tempted by an indirect inductive approach involving imprimitive binary structures as well, but this does not seem very promising, and the group theoretic approach may work here. Analogously one might ask for the classification of all sufficiently large primitive structures with a specified bound on $\kappa(\mathcal{X})$. The invariant $\kappa$ may behave well enough – in terms of crude lower bounds – to make an O'Nan-Scott type of analysis feasible. Certainly $\kappa$ is not well behaved under restriction to the socle, as one sees by comparing $\mathrm{Sym}\,(n)$ and $\mathrm{Alt}(n)$, but since only lower bounds are needed, the situation may possibly be manageable.

In §3 we mentioned the conjectured answer in the binary case: oriented and unoriented cycles of prime order; a naked set (with equality); and the peculiar edge-colored graphs associated with primitive homogeneous 1-dimensional affine groups which are not strictly linear, namely those of the form $\Gamma_q = \mathbb{F}_{q^2} \rtimes \mu_{q+1} \rtimes \mathbb{Z}/2\mathbb{Z}$, acting on the base field.

**Problem 2**

*Let $\mathcal{E}_{k,l}$ be the set of partitions of a set of $n = k \cdot l$ elements into $k$ classes of size $l$. Estimate $\kappa(\,\mathrm{Sym}\,(n), \mathcal{E}_{k,l})$, and consider more general partition types, as well as the action of $\mathrm{Alt}(n)$.*

One can get some partial results using estimates for $\kappa$ of a wreath product of actions on $k$-sets, though unfortunately in this connection one also needs the values of $\kappa$ for wreath products falling in the difficult range, such as those handled by Saracino for $k = 1$.

For example we have:

$$\text{With } l_0 < l/2,\ \kappa(\mathcal{E}_{k,l}) \geq \kappa(\mathcal{E}_{2,l}) \geq \kappa(\left[\begin{smallmatrix} l \\ l_0 \end{smallmatrix}\right]^2)$$

For the second inequality, work with the stabilizer of a single point in $\mathcal{E}_{2,l}$. An example: $\kappa(\mathcal{E}_{2,4}) \geq \kappa(4^2) = 4$; the exact value in this case is 5.

One can also show:

$$\kappa(\mathcal{E}_{n,2}) \geq n$$

using a form of the "Möbius band" example of Lachlan which was given at the end of [CMS] in a similar context. This lower bound may possibly be the correct value.

**Problem 3**

*Improve the estimates on $\kappa(\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]^d)$ in the range $n \leq 2k([\log_2 d] + 1)$, getting exact values or at least asymptotically accurate estimates.*

The results of [CMS] and the extraordinarily precise analysis by Saracino in the case $k = 1$ suggest that the best way to approach this is in terms of a function $\delta_k(r, n)$, with the following subtle definition, which will be elucidated momentarily: $\delta_k(r, n)$ is the least $d$ (or $\infty$ if none exists) for which there are two multisets $\mathcal{H}$, $\mathcal{H}'$ of $r$-labeled $k$-uniform hypergraphs on $n$ vertices whose $(r-1)$-restrictions coincide up to isomorphism. Here a multiset is a set with multiplicities; an $r$-labeled $k$-uniform hypergraph on $n$ vertices is a map $\lambda$ from $\{1, \ldots, r\}$ to the $k$-subsets of an $n$-element set; an $A$-restriction of an $r$-labeled hypergraph corresponding to a subset $A$ of $\{1, \ldots, r\}$

is the $A$-labeled hypergraph obtained by restricting $\lambda$ to $A$; and an $(r-1)$-restriction is an $A$-restriction with $|A| = r-1$. Finally, we say that the $A$-restrictions of $\mathcal{H}$ and $\mathcal{H}'$ coincide if there is a bijection between $\mathcal{H}$ and $\mathcal{H}'$ so that corresponding hypergraphs have isomorphic $A$-restrictions, where isomorphisms are permitted to permute the $n$ vertices but the domain $A$ is fixed; and we say that the $(r-1)$-restrictions coincide, if for each $A$ of cardinality $r-1$, the $A$-restrictions coincide. Examples are in order.

For $k = 1$, an $r$-labeled hypergraph is a function from $\{1, \ldots, r\}$ to $\{1, \ldots, n\}$, and the isomorphism types (allowing the action of $\mathrm{Sym}(n)$) are classified by equivalence relations on $\{1, \ldots, r\}$. Thus in this case $\mathcal{H}$ and $\mathcal{H}'$ can be viewed more simply as collections of equivalence relations on $\{1, \ldots, r\}$ with at most $n$ classes. For example for $n = 2$ and $r$ even, we may let $\mathcal{H}$ consist of the partitions of $\{1, \ldots, r\}$ into two pieces of even size (one may be empty); and let $\mathcal{H}'$ consist of the partitions into two pieces of odd size. One may then check directly that the $A$-restrictions, for $|A| = r-1$, consist of all equivalence relations on $A$, each occurring once. This shows $\delta_1(r,2) \leq 2^{r-2}$ and further analysis shows this estimate is exact. For more on $\delta_1$ see §3.

Little is known about $\delta_2$, but some examples may clarify the meaning of the definition.

$\delta_k(r,n) = \delta_{n-k}(r,n)$ for any $k, n$ since there is a canonical idenitification between $k$-sets and their complements. In particular $\delta_2(r,3) = \delta_1(r,3)$. Therefore we will consider the case $n = 4$.

**Example**

$\delta_2(4,4) = 2$. *We give an explicit example. Let $a, b, c, d$ be the four edges of a 4-cycle and define $\mathcal{H}$, $\mathcal{H}'$ as follows (for each $r$-labeled graph we just list the values of $\lambda(1), \ldots, \lambda(4)$ in order:*

*$\mathcal{H}$: (1) $a/b/c/d$; (2) $a/b/a/b$;*

*$\mathcal{H}'$: (1) $a/b/c/b$; (2) $b/a/b/c$.*

This explicit example shows $\delta_2(4,4) \leq 2$; evidently $\delta_2(4,4) > 1$.

**Example**

$\delta_2(5,4) \leq 15$.

Let $\Lambda$ be the set of injective functions from $\{1,2,3,4,5\}$ into $K_4$; these may also be thought of as labelings of $K_4$ minus an edge by distinct labels $1, \ldots, 5$. $\mathrm{Sym}(6)$ and $\mathrm{Sym}(4)$ act naturally on the edges and vertices; $\mathrm{Sym}(4)$ preserves isomorphism types and $\mathrm{Sym}(6)$ acts regularly on $\Lambda$. Thus the isomorphism types represented by $\Lambda$ may be identified with the coset space $\mathrm{Sym}(6)/\mathrm{Sym}(4)$ under a natural embedding of $\mathrm{Sym}(4)$ into $\mathrm{Sym}(6)$ which actually takes $\mathrm{Sym}(4)$ into $\mathrm{Alt}(6)$. In particular this space falls naturally into even and odd types (though the determination of which is which is of course arbitrary). Let $\mathcal{H}$ and $\mathcal{H}'$ be representatives of these two classes. Then $\mathcal{H}$ and $\mathcal{H}'$ each consist of 15 5-labeled graphs on 4 vertices, and we claim that the restrictions of $\mathcal{H}$ and $\mathcal{H}'$ obtained by deleting any one label coincide. For each label $i$ there is a natural bijection between $\mathcal{H}$ and $\mathcal{H}'$ in which each 5-labeled graph is replaced by the corresponding 5-labeled graph in which the label $i$ is moved to the unlabeled edge. In each case the same thing can be accomplished by a transposition in Sym (6), so this switches the even and odd types, and it obviously preserves the $i$-restrictions.

This example shows that $\delta_2(5,4) \leq 15$. Full information on $\delta_2(r,n)$ for a fixed value of $n$ determines the value of $\kappa(\left[\begin{smallmatrix} n \\ 2 \end{smallmatrix}\right]^d)$ for all $d$. For example, our estimate for $\delta_2(5,4)$ suggests that $\kappa(\left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix}\right]^d)$ may be 4 for $2 \leq d \leq 14$ and 5 for $d = 15$, but to pin this down one would need not only the exact value of $\delta_2(5,4)$, but a little more information about $\delta_2(r,4)$ in general.

## References


[As] M. Aschbacher, **The theory of finite groups**, Cambridge Studies in Advanced Mathematics 10, Cambridge University Press, 1986. ix+274 pp.

[Ca1] P. Cameron, *Finite permutation groups and finite simple groups*, Bull. London Math. Society **13** (1981), 1-22.

[Ca2] P. Cameron, **Oligomorphic Permutation Groups**, London Mathematical Society Lecture notes 152, Cambridge University Press, 1990.

[Ch] G. Cherlin, *Large finite structures with few types*, in **Proceedings of NATO ASI**, Fields Institute, Toronto, August 1996, NATO ASI Series C, Kluwer, Dordrecht, 1997 (in press).

[CL] G. Cherlin and A. Lachlan, *Stable finitely homogeneous structures*, TAMS **296** (1986), 815-850.

[CMS] G. Cherlin, G. Martin, and D. Saracino, *Arities of permutation groups: Wreath products and $k$-sets*, J. Combinatorial Theory, Ser. A **74**, 249-286 (1996)

[DM] J. Dixon and B. Mortimer, *Permutation Groups*, Graduate Texts in Mathematics, Springer, New York, 1996.

[Ga] A. Gardiner, *Homogeneous graphs*, J. Comb. Theory Series B **20** (1976), 94-102.

[GK] Ya. Golfand and M. Klin, *On $k$-homogeneous graphs*, in **Algorithmic studies in combinatorics**, 76-85, Nauka, Moscow, 1978.

[Hr] E. Hrushovski, *Finite structures with few types*, in **Finite and Infinite Combinatorics in Sets and Logic**, eds. N. W. Sauer, R. E. Woodrow, and B. Sands, NATO ASI Series C vol. 411, Kluwer, Dordrecht, 1993.

[KLM] W. Kantor, M. Liebeck, and H. Macpherson, *$\aleph_0$-categorical structures smoothly approximated by finite substructures*, Proceedings of the London Mathematical Society, Series 3 **59**, 439-463.

[KL] J. Knight and A. Lachlan, *Shrinking, stretching, and codes for homogeneous structures*, in **Classification Theory**, J. Baldwin ed., LNM 1292, Springer, New York, 1985.

[La1] A. Lachlan, *On countable stable structures which are homogeneous for a finite relational language*, Israel J. of Math. **49** (1984), 69-153.

[La2] A. Lachlan, *Homogeneous structures*, in A. Gleason, ed., **Proceedings of the International Congress of Mathematicians, Berkeley 1986**, v. 1, 314-321, American Mathematical Society, Providence, R.I., 1987.

[LW] A. Lachlan and R. Woodrow, *Countable ultrahomogeneous undirected graphs*, Transactions of the American Mathematical Society **262** (1980), 51-94.

[Mp] H. Macpherson, *Homogeneous and smoothly approximated structures*, in **Proceedings of NATO ASI**, Fields Institute, Toronto, August 1996, NATO ASI Series C, Kluwer, Dordrecht, 1997 (in press).

[Sa] D. Saracino, *The arity of* $n^d$, two preprints, (1996) and manuscript in preparation (1997).

[Sh] J. Sheehan, *Smoothly embeddable subgraphs*, J. London Math. Soc. **9** (1974), 212-218.