Problems around groups, graphs and semigroups

Peter J. Cameron
University of St Andrews

LMS Undergraduate Summer School
Leeds, July 2019

Groups, semigroups, and graphs

Groups and semigroups are algebraic objects. A semigroup is a set with a binary operation satisfying the associative law:

$$a \circ (b \circ c) = (a \circ b) \circ c;$$

a group is a semigroup which also satisfies the identity and inverse laws:

- there is an element e such that $e \circ a = a \circ e = a$ for all a;
- for any a, there exists b such that $a \circ b = b \circ a = e$.

A graph is a combinatorial object: it has a set of vertices, some pairs of which are joined by (undirected) edges.

How strong are these conditions? The table gives the numbers of objects, up to isomorphism.

Order	1	2	3	4	5	6	7	8	9
Groups	1	1	1	2	1	2	1	5	2
Semigroups	1	5	24	188	1915	28634	1627672	3684030417	105978177936292
Graphs	1	2	4	11	34	156	1044	12346	274668

We see that there are very few groups; these axioms are extremely tight. It is a bit surprising how loose the structure of semigroups is; there are many more semigroups than graphs.

Also a bit surprising: there is a formula for the number of graphs (albeit a rather complicated formula involving a sum over partitions of n), but no formula for groups or semigroups. The number of semigroups is only known up to order 10; the number of groups, up to order 2000. The number of groups with orders up to 2000 was announced as a "millennium project". There are about 59.9 billion groups, of which about 59.5 billion have order $2^{10} = 1024$.

There are just two graphs on 2 vertices, namely an edge and a nonedge .

If you have learned any group theory you will know that there is only one group of order 2, the cyclic group (or integers mod 2).

Semigroups of order 2

A semigroup is defined by a multiplication table, for example

$\circ$	a	b
a	a	a
b	a	b

(This one is the semigroup $\{0, 1\}$ with the operation of multiplication.)
There are $2^4 = 16$ possible multiplication tables.

Exercise

- *Show that, up to isomorphism (which means just swapping a and b here) there are* 10 *different* 2×2 *multiplication tables.*
- *Show that five of them are semigroups.*

Graphs from groups

Can we construct one type of object from another, so as to learn something about both types?
One method of forming a graph from a group G is to use a joining rule which depends on the algebraic structure of the group. One examples is the commuting graph: the vertices are the elements of G, and x and y are joined if $xy = yx$.
There are other similar joining rules, which I won't discuss here.
The commuting graph is connected, since any two elements are both joined to the identity (or to any element in the centre of G). So, to make life more interesting, some people remove all elements of the centre from the commuting graph.

Markov chains and random walks

A Markov chain is a system which has a number of states (here always finite) and, at discrete time steps, jumps from one state to another. The probability of jumping from state i to state j is p_{ij}, and the matrix P with (i,j) entry p_{ij} is the transition matrix. Note that its entries are non-negative and its row sums are 1; so its greatest eigenvalue is 1.

An example of a Markov chain is the random walk on a graph (possibly with loops); the states are the vertices, and at each time step the process chooses a random neighbour and moves there.

Theory of Markov chains

A Markov chain is irreducible if it is possible to move from any state to any other; it is aperiodic if the greatest common divisor of the possible return times to states is 1.

Theorem

If a Markov chain is irreducible and aperiodic, then it has a unique limiting probability distribution on the set of states, which it approaches as the number of steps tends to infinity.

A random walk is irreducible if the graph is connected, and aperiodic if the graph is not bipartite (in particular, if it contains loops).

Let P be the transition matrix of an irreducible aperiodic Markov chain. The row sums are all 1, so the all-1 vector is a right eigenvector with eigenvalue 1 (and this is the greatest eigenvalue).

So there is also a left eigenvector with eigenvalue 1, say q. This is a vector with positive entries; if we normalise it to have sum 1, it is the limiting distribution. For if the system is described by this distribution, then after one time step the probability of being in state j is

$$\sum_{i=1}^{n} q_i p_{ij} = q_j,$$

the same as before the step. So this is a stationary state; by the uniqueness it is the limiting state.

A very important question about a Markov chain is how rapidly it approaches its limiting distribution, the so-called mixing time.

Random walk on the commuting graph

It is a remarkable fact that the limiting distribution of the random walk on the commuting graph of a group G is uniform on conjugacy classes of G: that is, the probability of being at an element x is inversely proportional to the size of the conjugacy class containing x.

Exercise

Prove this. (Use the fact, from group theory, that the product of the size of the conjugacy class of x and the number of elements which commute with x is equal to $|G|$, independent of x.)

Problem

What can be said about the rate of convergence of this random walk?

Problem

What can you say about the random walk on the reduced commuting graph?

Sylvester and the six nations

One of the most remarkable facts in mathematics is that the symmetric group S_6 has an outer automorphism (one not induced by conjugation). For no other value, finite or infinite, of n, is this the case.

Here is a very brief sketch, due essentially to Sylvester. Let A be a set of size 6. As well as the six points of A, we consider

- the duads, or pairs of points of A (there are 15 of these);
- the synthemes, or partitions of A into three duads (there are 15 of these);
- the synthematic totals, or partitions of the set of duads into five synthemes (there are 6 of these).

If you are organising a rugby tournament involving six nations, each pair playing once, the duads are the matches, the synthemes the sets of matches that can be played in a single weekend, and the synthematic totals the possible tournament schedules over five weekends.

The outer automorphism of S_6

Any permutation of A induces permutations of the sets of duads, synthemes, and hence of the set X of synthematic totals; so we have an isomorphism from the symmetric group on A to the symmetric group on X which is not induced by a bijection from A to X. This defines an outer automorphism of the symmetric group $\mathrm{Sym}(A)$.

6 is the only number n for which, starting from a set of size n, we can construct another set of size n which is invariant under all permutations of the first set but not naturally bijective with it.

In the language of "general nonsense", 6 is the only number n for which the category of n-element sets and bijections has a non-trivial functor to itself.

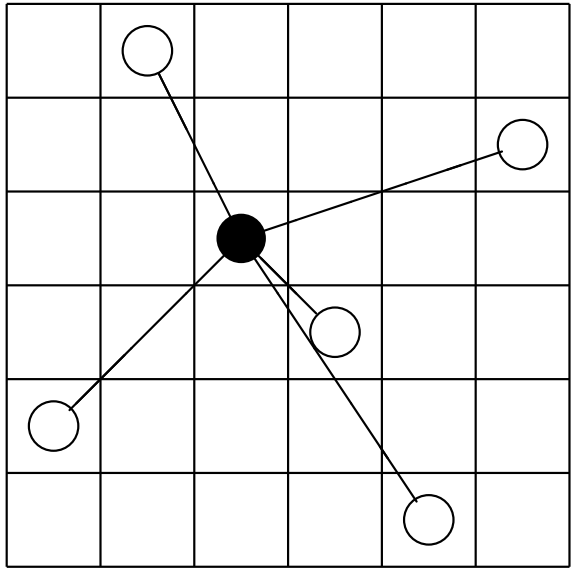
The Sylvester graph

We make a graph on the vertex set $A \times X$ by joining (a, x) to (b, y) if the duad $\{a, b\}$ is a part of the intersection of the synthematic totals x and y (these totals intersect in a unique syntheme). This graph has valency 5, and has remarkable properties.
Now consider the set of 48 subsets of size 6 (called "blocks") consisting of

- the six rows (sets $\{(a, x) : x \in X\}$ for fixed a);
- the six columns (sets $\{(a, x) : a \in A\}$ for fixed x);
- the 36 closed neighbourhoods, consisting of a vertex and its five neighbours – these are called starfish (see next slide).

This collection of subsets forms a block design. It has the property that two points which are joined in the Sylvester graph lie in two blocks, while any other pair of points are contained in a unique block. This design is also remarkable.

Starfish



We define a starfish to consist of a vertex and its neighbours; a galaxy of starfish is the set of six starfish derived from the vertices in a column of the array.

Sylvester designs

We define a Sylvester design to be a block design with 36 points and 48 blocks of size 6 (the points being the vertices of the Sylvester graph) such that two points lie in two blocks if they are adjacent in the Sylvester graph, or one block otherwise. The design described earlier is a Sylvester design, with automorphism group of order 1440 (equal to the automorphism group of S_6). Two others are known, one with 144 automorphisms, and one with only the identity automorphism.

Problem

Find all Sylvester designs.

I suspect that this will require a big computation; but I would be interested in knowing whether, up to isomorphism, the number of such designs is closer to ten or to ten million.

Statistical properties

Statisticians have a variety of optimality criteria for block designs. Very briefly, we form the concurrence matrix of the design, whose (i,j) entry is the number of blocks containing i and j for $i \neq j$, and whose diagonal entries are chosen so that the row and column sums are zero. The most popular criteria maximise functions of the non-zero eigenvalues of this matrix (the harmonic mean, geometric mean, and minimal element in the case of the so-called A, D and E criteria).
If a balanced design or 2-design exists (with all off-diagonal entries equal), then it is optimal. In our case, this would be (more or less) an affine plane of order 6, which is known not to exist!

Problem

Prove that Sylvester designs are optimal in various senses (especially the A criterion defined above).

Semigroups

Now I turn to semigroups.
There are several contenders for the role played by the symmetric group $S(\Omega)$ in group theory:

- the full transformation semigroup $T(\Omega)$ consisting of all maps from Ω into Ω;
- the partial transformation semigroup $P(\Omega)$ consisting of all maps from a subset of Ω into Ω;
- the symmetric inverse semigroup $I(\Omega)$ consisting of all bijections between subsets of Ω.

So there are theories of transformation semigroups, partial transformation semigroups, and inverse semigroups, corresponding to the theory of permutation groups in group theory.
There are also various links between them ...

Endomorphisms and partial isomorphisms

Given a finite group G, let $\mathrm{End}(G)$ be the semigroup of endomorphisms of G, and $\mathrm{PIso}(G)$ the semigroup of partial isomorphisms of G (isomorphisms between subgroups of G). In the notation of the previous slide, $\mathrm{End}(G) \leq P(\Omega)$, while $\mathrm{PIso}(G) \leq I(\Omega)$.

Theorem

If G is abelian, then $|\mathrm{End}(G)| = |\mathrm{PIso}(G)|$.

The proof uses the fact that an abelian group G has a dual G^*, the group of characters of G, such that subgroups of G correspond bijectively to quotients of G^* and *vice versa*.

Exercise

Complete the proof of the theorem!

Problems

Problem

Is there a relation between $\mathrm{End}(G)$ *and* $\mathrm{PIso}(G)$*, for* G *a finite abelian group? (These semigroups are certainly not isomorphic!)*

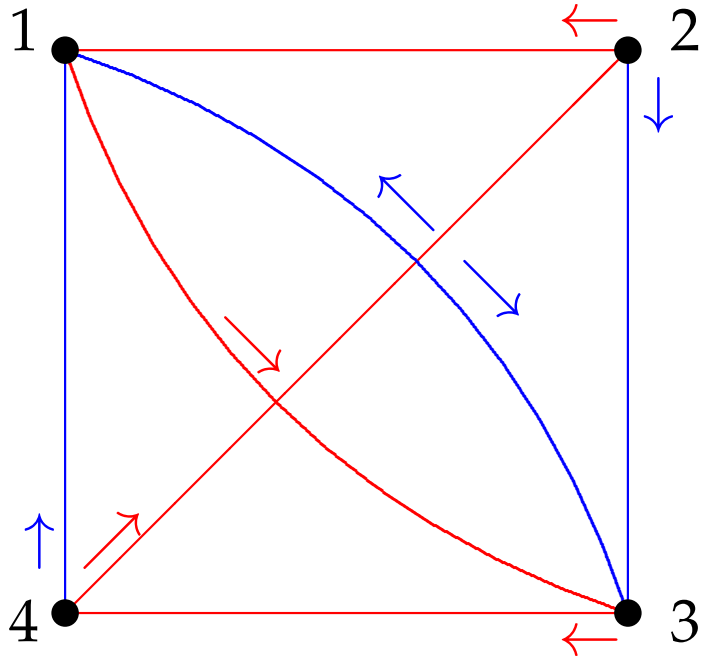
This problem has a linear analogue, where we replace G by a vector space and ask for linear maps. The role of the character group is taken by the dual space.

Problem

Does the converse hold? What can be said about groups where the left hand side is larger (or smaller) than the right hand side?

The dungeon

You are in a dungeon consisting of a number of rooms. Passages are marked with coloured arrows. Each room contains a special door; in one room, the door leads to freedom, but in all the others, to death. You have a map of the dungeon, but you do not know where you are.



You can check that (Blue, Red, Blue) takes you to room 1 no matter where you start.

Automata

The diagram on the last page shows a finite-state deterministic automaton. This is a machine with a finite set of states, and a finite set of transitions, each transition being a map from the set of states to itself. The machine starts in an arbitrary state, and reads a word over an alphabet consisting of labels for the transitions (Red and Blue in the example); each time it reads a letter, it undergoes the corresponding transition.

A reset word is a word with the property that, if the automaton reads this word, it arrives at the same state, independent of its start state. An automaton which possesses a reset word is called synchronizing.

Not every finite automaton has a reset word. For example, if every transition is a permutation, then every word in the transitions evaluates to a permutation. How do we recognise when an automaton is synchronizing?

Automata and transformation semigroups

Combinatorially, an automaton is an edge-coloured digraph with one edge of each colour out of each vertex. Vertices are states, colours are transitions.
Algebraically, if $\Omega = \{1, \ldots, n\}$ is the set of states, then any transition is a map from Ω to itself. Reading a word composes the corresponding maps, so the set of maps corresponding to all words is a transformation semigroup on Ω.
So an automaton is a transformation semigroup with a distinguished generating set.

Graph endomorphisms

Our graphs are simple (no directions, loops, or multiple edges). The clique number $\omega(\Gamma)$ of a graph Γ is the number of vertices in its largest complete subgraph, and the chromatic number $\chi(\Gamma)$ is the smallest number of colours required for a vertex-colouring so that adjacent vertices get different colours.
Let Γ and Δ be graphs. A homomorphism from Γ to Δ is a map f from the vertex set of Γ to that of Δ with the property that, for any edge $\{v, w\}$ of Γ, the image $\{vf, wf\}$ is an edge of Δ.
An endomorphism of Γ is a homomorphism from Γ to itself.

Proposition

- *A homomorphism from K_m to Γ is an embedding of K_m into Γ; such a homomorphism exists if and only if $\omega(\Gamma) \geq m$.*
- *A homomorphism from Γ to K_m is a proper colouring of Γ with m colours; such a homomorphism exists if and only if $\chi(\Gamma) \leq m$.*
- *There are homomorphisms in both directions between Γ and K_m if and only if $\omega(\Gamma) = \chi(\Gamma) = m$.*

The obstruction to synchronization

The endomorphisms of a graph Γ form a transformation semigroup; if Γ is not a null graph, then $\mathrm{End}(\Gamma)$ is not synchronizing, since edges cannot be collapsed.

Theorem

Let S be a transformation semigroup on Ω. Then S fails to be synchronizing if and only if there exists a non-null graph Γ on the vertex set Ω for which $S \leq \mathrm{End}(\Gamma)$. Moreover, we may assume that $\omega(\Gamma) = \chi(\Gamma)$.

Proof.

Given a transformation semigroup S, we define a graph Γ in which x and y are joined if and only if there is no element $s \in S$ with $xs = ys$. Show that $S \leq \mathrm{End}(\Gamma)$, that Γ has equal clique and chromatic number, and that S is synchronizing if and only if Γ is null. □

The probability of synchronization

It is known that, if we choose two random elements of $T(\Omega)$, then with probability tending to 1 as $|\Omega| \to \infty$, they generate a synchronizing semigroup. But the proof is not easy!
Here is a possible approach to a proof of this fact. If two transformations do not generate a synchronizing semigroup, then they must be contained in a maximal non-synchronizing semigroup. This must be the semigroup of endomorphisms of a graph.

Problem

Which graphs have the property that their endomorphism semigroups are maximal non-synchronizing?

If we could solve this problem, we could in principle count the pairs of transformations lying in such a semigroup.

Dixon's theorem

The suggested proof above is based on a theorem of John Dixon, proving a conjecture of Netto.

Theorem

Two random permutations on $\{1, \ldots, n\}$ *generate the symmetric or alternating group of degree* n *with probability tending to* 1 *as* $n \to \infty$.

The proof goes like this. Once we know enough about the maximal subgroups of the symmetric or alternating group, we can do some counting to estimate the number of pairs of permutations which lie together in some maximal subgroup (these are precisely the ones not generating the whole group), and divide by $n!$ to get the required probability.

Other problems on synchronization

Problem

- *The Černý conjecture asserts that if an n-state automaton is synchronizing, it has a reset word of length at most $(n-1)^2$. (This would be best possible, if true.) This conjecture has been open for half a century, and although it looks attractive, it is very hard!*
- *A permutation group cannot be synchronizing as a monoid. So we abuse language by calling a group G synchronizing if the semigroup generated by G and any non-permutation on Ω is synchronizing. The problem is to find all the synchronizing permutation groups. A substantial body of results exists about this.*
- *What about synchronization in the infinite case?*