Unravelling the (miniature) Rubik's Cube through its Cayley Graph

Daniel Bump[1] and Daniel Auerbach[2]

October 13, 2006

1 Introduction

There are many interesting questions one can ask about a highly symmetric graph. A graph with a transitive automorphism group can be obtained starting with a group G and a small set S of generators. We assume that S is closed under inverses. Then the vertices of this *Cayley graph* are the elements of the group, and elements x and y will be joined by an edge if $x^{-1}y \in S$.

Anyone who has worked with groups knows that they have great individuality. A nonabelian simple group is very different in character than a p-group. A Cayley graph is just a group with additional structure (the edge relation) so one might also expect it to have its own unique individual character. Yet if it is at all large it is a challenge to view it as something more than just an amorphous collection of points with an edge relationship. How can one perceive its true structure?

The Rubik's cube is a popular toy which in reality is nothing but a group with a given set of generators – in other words, its essence is a Cayley graph C_G. Inside the group of the Rubik's cube, one may consider the subgroup with two generators, and the corresponding Cayley graph. For the miniature ($2 \times 2 \times 2$) Rubik's cube, this group of order 29,160 and its Cayley graph (of the same size) are of just the right size to make interesting experiments. It is small enough that a model can be kept in memory in a computer, so that any question can potentially be answered by brute

[1]Department of Mathematics, Building 380, Stanford University, Stanford, CA 94305-2125. **email:** bumpmath.stanford.edu

[2]Google, Inc. **email:** dtauerbachgmail.com

force computation. Yet it is large enough that it may already exhibit unexpected properties of large highly symmetric graphs.

The decision to focus on the two-generator group (generated just by rotations of two adjacent faces) is motivated not only by the desire to obtain a Cayley graph of convenient size, but also by a circumstance described by Singmaster [6] – the two-generator group is a particularly interesting group. It is a wreath product based on the group of permutations of the six vertices that are moved by the two operations, and this permutation group is itself isomorphic to $\mathrm{PGL}_2(\mathbb{F}_5)$, the group of the projective line over the field of five elements; in turn it is isomorphic to S_5.

Cube enthusiasts may note that the two-generator group – of either the standard $(3\times3\times3)$ miniature cube – is the basis of a good puzzle. If one scrambles a cube using just two generators, can one then unscramble it using just those two generators? The fact that this puzzle is a good alternative to the usual cube-scrambling puzzle is a reflection of the remarkable richness of the two generator group.

We will try to bring this Cayley graph C_G to life in this paper. We will see that its diameter is 17. We will consider how the "ball of radius r" consisting of all points of distance r from the origin grows with r. Perhaps surprisingly, when r is near the diameter, the rate of growth of the ball slows dramatically. Although the graph has thousands of vertices, only a few – 18 altogether – are at maximal distance from the origin. We call these vertices the *antipodes* and we will consider the structure of the subgraph formed by these.

We will show that the graph can be used to obtain a presentation of the group, that is, a complete set of relations between the generators. We will apply this information to construct interesting operations for the group of the full $(3 \times 3 \times 3)$ Rubik's cube.

Finally we will turn to questions about the random walk on the graph – how quickly does making random operations from the set of generators scramble the cube? We will consider this question from a couple of different points of view.

Repeatedly making random operations chosen from the set S of generators produces a *Markov process*. How quickly the process converges to randomness is from one point of view a question of computing the eigenvalues of the Markov transition matrix. This in turn reduces to a problem in group representation theory – each irreducible representation of the group will supply some of these eigenvalues. We will compute all of them. On the other hand, we will also study this Markov process brute force computation.

We would like to thank Persi Diaconis and Hua Zhou for interesting conversations and helpful comments. For support, we thank Stanford's VPUE and Gunnar Carlsson (Auerbach) and NSF grant DMS-0354662 (Bump).

2 Introduction to the 2 generator group

Let G denote the group generated by taking clockwise rotations of the right and top faces of the miniature $(2 \times 2 \times 2)$ Rubik's Cube. Denote by R a clockwise rotation of the right face, and let U denote a clockwise rotation of the top face, so that $G = \langle R, U \rangle$. Following Singmaster [6], we also sometimes denote R^{-1} by R', U^{-1} by U'. Let K denote the subgroup of G corresponding to operations that do not change the position of any of the 6 cubes which are generally affected by the group (i.e. K changes only the orientation of these cubes).

Proposition 1. *K is an abelian, normal subgroup of G of order 3^5.*

Proof. Given any operation $g \in G$, and $k \in K$, we clearly have that $gkg^{-1} \in K$, since cube positions are not affected by k. Hence, K is normal in G. Moreover, K is clearly abelian.

To see that the order $|K| \leq 3^5$, we observe that there are 3 possible orientation changes (or twists) for each of the 6 cubes and so the order of K must be $\leq 3^6$. Now label each possible twist by a different number: 0 for no change, 1 for clockwise twist, 2 for counterclockwise twist. It is well-known that the total number of twists must be $\equiv 0$ modulo 3 (Singmaster [6]) and so $|K| \leq 3^5$.

So we need only show that $|K| \geq 3^5$. The operation $RUR'URU^2R'U^2 \in K$ twists three corners clockwise by one twist, and it is easy to see that this operation, together with its conjugates, generates a group containing any operation twisting the corners in such a way that the total number of twists is $\equiv 0$ modulo 3 and so $|K| \geqslant 3^5$. $\square$

Operations in G only affect the locations and orientations of 6 pieces. We label these by the points of the projective line $\mathbb{P}^1(\mathbb{F}_5) = \mathbb{F}_5 \cup \{\infty\}$ over the field with 5 elements (Figure 1).

The group G acts on the six pieces by permuting them and also affecting their orientation. The quotient G/K also acts faithfully as a group of permutations of $\mathbb{P}^1(\mathbb{F}_5)$ (ignoring orientation). Thus it may be regarded as a subgroup of the group S_6 of permutations of the six element set $\mathbb{P}^1(\mathbb{F}_5)$.

A particular group of permutations of $\mathbb{P}^1(\mathbb{F}_5)$ is the group $\mathrm{PGL}(2, \mathbb{F}_5)$ acting by linear fractional transformations. Thus the group $\mathrm{GL}(2, \mathbb{F}_5)$ acts by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : x \longmapsto \frac{ax+b}{cx+d}, \qquad x \in \mathbb{F}_5 \cup \{\infty\},$$

where it is understood that if $x = \infty$ then $\frac{ax+b}{cx+d} = \frac{a}{c}$, while if $cx + d = 0$ then $\frac{ax+b}{cx+d} = \infty$. Since the center Z of $\mathrm{GL}(2, \mathbb{F}_5)$ consisting of scalar matrices acts trivially, this is really an action of $\mathrm{GL}(2, \mathbb{F}_5)/Z = \mathrm{PGL}(2, \mathbb{F}_5)$.

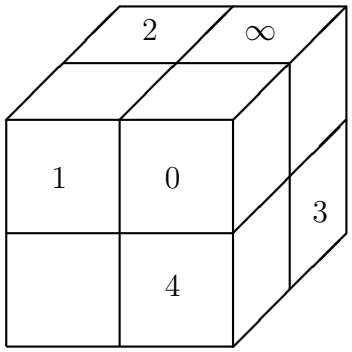


Figure 1: The labeling of the 6 pieces moved by G.

Proposition 2. (Singmaster [6]) *As a permutation group acting on* $\mathbb{P}^1(\mathbb{F}_5)$*, we have* $G/K = \mathrm{PGL}(2, \mathbb{F}_5)$.

Proof. Let us check that generators of G/K are contained in $\mathrm{PGL}(2, \mathbb{F}_5)$. Notice that R, RU generate G, so we need only to show that $R, RU \in \mathrm{PGL}(2, \mathbb{F}_5)$. In keeping with the labeling scheme given and the action described above, we see that R corresponds to the cycle (012∞), while RU corresponds to the cycle (01234). But an easy verification then shows that R and RU have corresponding fractional linear transformations:

$$R = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \in PGL(2, F_5)$$

$$RU = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in PGL(2, F_5)$$

So indeed we see that $G/K \subset \mathrm{PGL}(2, \mathbb{F}_5)$. It is also easy to verify that these two elements generate $\mathrm{PGL}(2, F_5)$. □

Corollary 1. $G/K \cong S_5$*, and* $|G/K| = 5!$.

Proof. The well-known isomorphism $\mathrm{PGL}(2, \mathbb{F}_5) \cong S_5$ may be checked by labeling the 5-Sylow subgroups of S_5 as follows:

$$\begin{array}{lll} \langle(12345)\rangle = \infty, & \langle(12354)\rangle = 0, & \langle(12453)\rangle = 1, \\ \langle(12543)\rangle = 2, & \langle(12534)\rangle = 3, & \langle(12435)\rangle = 4. \end{array}$$

(Here $\langle(12345)\rangle$ denotes the cyclic subgroup generated by the 5-cycle (12345).) The group S_5 acts on $\mathbb{P}^1(\mathbb{F}_5)$ by conjugating its Sylow subgroups, and we claim that

the group of permutations of $\mathbb{P}^1(\mathbb{F}_5)$ thus obtained is $\mathrm{PGL}(2,\mathbb{F}_5)$. To see that it is contained in $\mathrm{PGL}(2,\mathbb{F}_5)$ it is enough to check for generators of S_5 that conjugation induces a linear fractional transformation. For example conjugation by (12345) induces the transformation that fixes ∞ and cyclicly permutes $0 \longrightarrow 1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow 0$, which corresponds to the linear fractional transformation $x \longmapsto x+1$, or the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ in $\mathrm{GL}(2,\mathbb{F}_5)$. Similarly (45) has the effect $0 \longleftrightarrow \infty$, $1 \longleftrightarrow 2$, $3 \longleftrightarrow 4$, which is the linear fractional transformation $x \longmapsto 2/x$, or the matrix $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$ in $\mathrm{GL}(2,\mathbb{F}_5)$. Since S_5 is generated by (12345) and (45), we see that every permutation of the 5-Sylow subgroups thus labeled is in $\mathrm{PGL}(2,\mathbb{F}_5)$ and so we have constructed a homomorphism $S_5 \longrightarrow \mathrm{PGL}(2,\mathbb{F}_5)$. Both groups have the same order 120, and the homomorphism is nontrivial since the only nontrivial normal subgroup of S_5 is A_5, and we have already seen that an element (12345) of A_5 acts nontrivially, so it is an isomorphism. □

Now we show that G is a semidirect product $K \ltimes H$ for a subgroup H. This means that $H \cap K = \{1\}$ and that $G = HK$. (We already know that $K \lhd G$.) This fact – that the two generator group is a semidirect product – is *false* for the larger two generator group of the $3 \times 3 \times 3$ Rubik's cube.

To construct the group H, let us suppose we have a cube in a solved configuration. Suppose further that the F (front) face is red and that the opposite face—the B face—is orange. Consider an operation which leaves only red or orange facelets on both the F and B faces. This element of G is said to *solve the F and B faces mod identification of F and B colors*. Let H denote the set of operations of G which solve the cube mod the identification of F and B colors. It is clear that H is a group.

Proposition 3. *Every element of G/K has a unique representative in H, and so G is the semidirect product $K \rtimes H$.*

Proof. Exactly six cubes move during the operation of the cube. If the locations of these six cubes are given, there is exactly one way for each them to be oriented that solves the F and B faces mod identification of F and B colors. Thus $H \cap K = \{1\}$. What we must show is that if a permutation of the six cubes is attainable within G, then this orientation that solves the cube modulo identification of F and B colors can be achieved within H. Let

$$h_1 = RUR'URU'R'URU^2R'U \in H$$

$$h_2 = URU'RUR'U'RUR^2U'R \in H$$

It is direct to see that these elements are in fact in H, and moreover, that h_1 has the same image as U in G/K and h_2 has the same image as R in G/K. Now consider the subgroup H_1 of H generated by h_1 and h_2. We can write any element of G in the form hk (where $h \in H_1$ and $k \in K$), since up to a twist, the generators of G are contained in H_1. Thus, it is clear that $H_1 = H$ and that $HK = G$, and so we have the desired semi-direct product. □

3 The Cayley graph

As G is a semi-direct product, one can enumerate all elements of G as pairs (k, h) with $k \in K$ and $h \in H$. Given this, the Cayley graph of G is easily modelled using a program in C++, and some facts that are set out in this section were proved using this computer program.

Let C_G denote the *Cayley graph* of G with respect to the set $S = \{R, R', U, U'\}$ of generators. This is the graph whose vertices are the elements of G, and whose edges are the pairs of x and y such that $x^{-1}y \in \{R, R', U, U'\}$. The group G acts transitively on the graph on the left. The graph is kept entirely in memory during the computations.

Proposition 4. *C_G has diameter 17.*

Proof. This was checked by computer. We recursively label the elements of the graph by integers d which will represent the distance from the origin. The identity element is assigned the label $d = 0$, and at the d-th step, all elements that are neighbors of a vertex at distance d that are not already labeled are given the label $d + 1$. After 17 steps, no more unlabeled vertices are found and the algorithm terminates. □

Loops in the Cayley graph correspond to relations between the generators. For example, after applying $R^2U^2R^2U^2R^2U^2$ one returns to the starting point. This means that the graph has a loop, or equivalently, that the "braid relation" $R^2U^2R^2U^2R^2U^2 = 1$ is satisfied in the group.

Lemma 1. *The group G admits a character $\chi : G \longrightarrow \{\pm 1\}$ such that $\chi(x) = -1$ for all $x \in S$. If $x_1 \cdots x_r = 1$ with $x_i \in S$, then r is even.*

Thus every loop in the Cayley graph has even length.

Proof. We have noted that $G/K \cong S_5$, and χ may be taken to be the sign character of S_5 pulled back to G. Applying χ to $x_1 \cdots x_r = 1$ gives $(-1)^r = 1$, so r is even. □

This fact is the basis of a computer algorithm to compute relations in the group, leading to a presentation. Since by Lemma 1 any relation has even length, one can look for situations where there are two distinct paths of equal length that start at the origin and have the same endpoint.

In the course of proving Proposition 4 we used a computer to compute the distance from every vertex to the origin, and we may now reuse this information to implement this idea. Let us a vertex x a *H-vertex* if it has at least two distinct adjacent vertices that are closer to the identity vertex than x (i.e. they have shorter walks to the identity than does x). To find relations, the algorithm that we will now describe goes through the entire graph, identifying H-vertices.

Given the labeling of the vertices by their distance from the origin, the H-vertices are easily identified. For each H-vertex N and each pair of adjacent vertices A and B that are closer to the identity than N, the algorithm produces one relation. Let us describe this relation. Let a be some shortest walk from the identity to N that goes through node A, and let b be some shortest walk that goes from the identity to N going through node B. Then $ab^{-1} = 1$ is clearly a relation, and is the relation generated for the ordered triplet (N, A, B). If there are other H-vertices on either path a or b, then this relation is not uniquely determined, but this does not matter. The important thing is that some such relation exists.

After generating all such relations, the algorithm cuts down on the total number by eliminating unnecessary relations. It does this by considering only a subset of the total number of triplets (N, A, B) that are necessary, by eliminating relations that contain other relations, and by eliminating redundant cycles. To illustrate this last point, let us consider a quick example. If we have two relations $abc = 1$ and $bca = 1$, these two relations are redundant and we can throw one of them out. So we get a presentation of G. There are too many relations to list them all here, but we give 14 of the shortest relations below:

R^4	U^4
$RURU'R'U'R'U'$	$R'U'R'U'R'URURU$
$R'U'R'URURURU'R'U'$	$RU'R'U'R'U'RURU$
$R'URURURU'R'U'$	$RUR'U'R'U'R'URU$
$RURUR'U'R'U'R'U$	$R'U'R'U'RURURU'$
$RURU'R'U'R'U'RU$	$R'U'RURURU'R'U'$
$R'U'R^2URURU^2R'U'$	$R^2U'R'URURUR'U'R$

We have listed all of the relations of length < 12 and two of the relations of length 12.

There is a homomorphism $\phi : G_3 \longrightarrow G$ from the two-generator group G_3 of the $3 \times 3 \times 3$ Rubik's cube, since any operation satisfied by the counterparts of R and

U in G_3 is obviously satisfied in G. The kernel of ϕ is the group of *edge operations* that effect the edges of the full Rubik's cube, but have no effect on the corners. A reasonable strategy for solving the full Rubik's cube is to first put the corners in place, then deal with the edges, so a table of edge operations is a useful thing. As we mentioned in the introduction, a pleasant puzzle is to scramble the cube using the 2 generator group, then restore it using only operations from the 2 generator group, and if one wants to be proficient at this task, it is important to know some edge operations that only use R and U. For example, the operation $RURURU'R'U'R'U'$ is an extremely pleasant edge three-cycle that is easy to use and remember.

Finally, we present a table that partitions all vertices of C_G into their distances from the identity.

r	Number of vertices at distance r	$\frac{\log\ B(r)}{\log(r)}$
0	1	–
1	4	–
2	10	2.32193
3	24	2.46497
4	58	2.6427
5	130	2.84243
6	271	3.02772
7	526	3.19162
8	980	3.33333
9	1750	3.46023
10	2731	3.57449
11	3905	3.6604
12	5229	3.72191
13	5848	3.76469
14	4792	3.77948
15	2375	3.7576
16	508	3.70136
17	18	3.62837

Table 1: The number of vertices at distance r from the origin.

Our next point is speculative. We note the striking fact is that after distance 14, the number of vertices at each distance shrinks dramatically, with only a handful at distance 17. One is tempted to try to visualize the graph as something like a sphere

or other homogenous space of some dimension, since in a sphere of circumference A the area of the set of points at distance between x and $x + \Delta x$ decreases as x approaches $A/2$. But if one accepts this idea, what would be the dimension of the manifold?

The notion of fractal dimension suggests the following considerations. Let $B(r)$ be the number of vertices of distance $< r$ from the identity. In a manifold of dimension d, we would expect the volume of a ball of radius r to satisfy:

$$\log\ B(r) \sim d\ \log(r)$$

for small r. So this suggests that C_G can be thought of as having fractal dimension between 2 and 3. This is not intended to be a rigorous statement, and indeed is a misuse of the concept of dimension. However it seems plausible that on a large scale the graph has some interesting topology that we have not yet been able to discover, and this may be an interesting direction for future work.

Since there are exactly 18 vertices at maximal distance, we wondered if they all were close to each other (which would support the idea that the graph could be thought of as a sphere). We call these 18 vertices the *antipodes*. They are

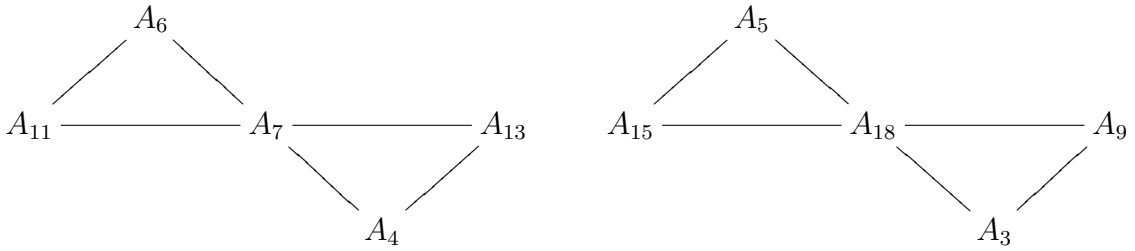
$A_1 = RU'RU'RU'R^2UR'URU'R'UR^2$
$A_2 = R'UR'UR'UR^2UR'U'RUR'UR^2$
$A_3 = U'RU'R^2U'RU'RU'RU'R^2UR^2$
$A_4 = U'R'URU'RUR'URU'RU'RUR^2$
$A_5 = URU'RU'RU'RU'R'UR'U'RUR^2$
$A_6 = UR'UR^2UR'UR'UR'UR^2U'R^2$
$A_7 = R'U^2RU^2RU'RU'RU'RU^2R^2$
$A_8 = RUR^2U'R^2UR'UR'U^2R'UR^2$
$A_9 = U^2RU'R^2U'RU'RU'RUR'UR^2$
$A_{10} = UR'U'RUR'UR'UR'URU'RUR^2$
$A_{11} = U^2RU'RU'RU'RUR'U^2RUR^2$
$A_{12} = U'R'UR'U'RU'RU'RU'R'URU'R^2$
$A_{13} = UR'U^2RUR'UR'UR'UR^2UR^2$
$A_{14} = U^2RU'RUR'UR'UR'U^2R'UR^2$
$A_{15} = U'RU'RU'RU'R^2UR'UR^2UR^2$
$A_{16} = R^2U'RU'R^2UR^2U'RU'RUR^2$
$A_{17} = R'URU'RUR'URU'RU^2R'UR^2$
$A_{18} = UR'UR^2UR'UR'UR'U'R^2UR^2.$

The subgraph of the antipodes can be grouped into 4 clusters. The first two clusters have 4 elements each:

$$A_{12} \text{ —— } A_1 \text{ —— } A_2 \text{ —— } A_{16} \qquad\qquad A_{10} \text{ —— } A_{17} \text{ —— } A_8 \text{ —— } A_{14}$$

Here vertices connected with a horizontal line have distance 6 from each other, and those pairs of vertices that have no lines between them all have distance 8 from each other.

The next two subgraphs are of the form:



Here vertices connected with a horizontal line have distance 6 from each other, and vertices connected with a slanted slash have distance 8 from each other. Moreover, the pairs $\{A_{11}, A_{13}\}$ and $\{A_{15}, A_9\}$ have respective distance 10. The pairs $\{A_6, A_4\}$, $\{A_6, A_{13}\}$, $\{A_{11}, A_4\}$ and $\{A_5, A_3\}$, $\{A_5, A_9\}$, $\{A_{15}, A_3\}$ all have respective distance 12. The smallest distance between any two clusters is 10.

4 The Markov process

The *adjacency matrix* of the Cayley graph of $G = \{g_1, \cdots, g_N\}$ (where $N = |G| =$ 29,160) is the matrix whose rows and columns correspond to the elements of g, with a value 1 in the i, j-th position if g_i and g_j are adjacent in the graph, and otherwise 0. Dividing this matrix by 4 gives a matrix we will denote by M. It is clear that M is a symmetric doubly stochastic matrix. As we will explain, it is the transition matrix of a Markov process. Its eigenvalues will prove important, so we first discuss how to compute them.

If ξ is a function on the group, then we may think of ξ as a vector whose entries are indexed by the group elements, and the i-th entry has value $\xi(g_i)$. Then we may compute the vector $M\xi$, and interpret that as a function on G. Alternatively, we may think of ξ as the entry $\sum \xi(g) \cdot g$ of the group algebra $\mathbb{C}[G]$, and the element of the group algebra corresponding in this way is $\frac{1}{4}(R+R'+U+U')\sum \xi(g)\cdot g$. Hence we may identify vectors with elements of the group algebra, and then application of the matrix M corresponds to left multiplication by the special element $\frac{1}{4}(R+R'+U+U')$ of $\mathbb{C}[G]$. By abuse of notation, we will sometimes write $M = \frac{1}{4}(R + R' + U + U')$ with this understanding.

Lemma 2. *The eigenvalues of M are real. If λ is an eigenvalue of M, so is $-\lambda$. The largest eigenvalue is* 1.

Proof. Since it is symmetric, the eigenvalues are real. If $M\xi = \lambda\xi$, we can multiply the entries in the vector ξ by ± 1 to obtain another vector with eigenvalue $-\lambda$ as

follows. The rows of ξ are indexed by the elements of G multiply the g-th entry by $\chi(g)$, where χ is the character of Lemma 1. By the Perron-Frobenius theorem (Horn and Johnson [2]), the top eigenvalue of M is 1, and it occurs with multiplicity one. $\square$

The eigenvalues of M can be computed using group representation theory. Let $\pi_1, \pi_2, \cdots, \pi_{48}$ be the irreducible representations of G, let $\chi_1, \cdots, \chi_{48}$ be their characters, and let $d_i = \chi_i(1)$ be their degrees. Then the regular representation decomposes as

$$\mathbb{C}[G] \cong \bigoplus_{i=1}^{48} d_i \pi_i.$$

Thus we may find the eigenvalues of M in $\mathbb{C}[G]$ by finding the eigenvalues of M in each of the irreducible representations π_i, then counting each eigenvalue with multiplicity d_i.

Since G is a semi-direct product with abelian kernel K, a theorem of Mackey constructs all irreducible representations by induction from characters of certain subgroups of G. See Mackey [4] Theorem 14.1 and Lang [3], Exercise 7 on page 724. However there is no need to compute these eigenvalues by this method since the computer program GAP [1] is capable of producing the characteristic polynomials of the linear transformations induced by M on each irreducible G-module. The following short program will suffice to produce these after a few hours.

```
G:=Group((1,12,8,6)(3,11,9,5)(2,10,7,4),(10,13,17,8)(7,11,14,18)(15,16,9,12));
r:=(1,12,8,6)(3,11,9,5)(2,10,7,4);
rp:=Inverse(r);
u:= (10,13,17,8)(7,11,14,18)(15,16,9,12);
up:=Inverse(u);
A:=IrreducibleRepresentations(G);
Q:=List([1..48],
     i -> CharacteristicPolynomial((r^A[i]+rp^A[i]+u^A[i]+up^A[i])*1/4);
```

The characteristic polynomials thus produced may then be evaluated using Mathematica. The nonnegative eigenvalues are listed with multiplicity in Table 2. For each $\lambda > 0$ in this table, $-\lambda$ also occurs with the same multiplicity, but these are not listed.

Let $\xi_1, \cdots, \xi_N$ ($N = |G|$ =29,160) be an orthonormal basis of $L^2(G)$ consisting of eigenvectors for M in with respect to the L^2 norm. Let $\lambda_1, \cdots, \lambda_N$ be the eigenvalues, so $M\xi_i = \lambda_i \xi_i$. Such an orthonormal set of eigenvalues exists since the matrix M is

symmetric. By Lemma 2, both 1 and -1 are eigenvalues of M, and both occur with multiplicity one. We order the eigenvalues so that $\lambda_1 = 1$ and $\lambda_2 = -1$. We have

$$\xi_1(g) = 1, \qquad \xi_2(g) = \chi(g).$$

The next eigenvalue $\lambda_3 = 0.964905$ has multiplicity 20 as does $-\lambda_3$, so we can further order the basis such that $\lambda_3 = \lambda_4 = \ldots = \lambda_{22}$ and $\lambda_{23} = \ldots = \lambda_{42} = -\lambda_3$, and $|\lambda_i| < \lambda_3$ when $i > 42$.

Consider the *Markov process* corresponding to the random walk on the Cayley graph. Specifically, let us start with a probability distribution p on G, that is, a function $p : G \to \mathbb{R}^+$ such that $\sum_{g\in G} p(g) = 1$. After we apply a random element of the generating set, we obtain another probability distribution. We may associate with the probability distribution $p : G \to \mathbb{R}^+$ the element $\sum p(g) \cdot g$ of the group algebra $\mathbb{C}[G]$, and after applying a random twist, this element is multiplied by $M = \frac{1}{4}(R + R' + U + U')$.

Let p_0 be the probability distribution corresponding to the solved cube, so that $p_0(1) = 1$, while $p_0(g) = 0$ for $g \neq 1$. Iterating the Markov process by randomly twisting the cube k times gives the distribution $p_k = M^k p_0$. After k operations, the cube will necessarily be at an even or odd distance from the identity, depending on whether k is even or odd. Except for this constraint, p_k should be approximately random if k is large. More precisely, if k is large, we expect $p_k - L_k$ to be small, where

$$L_k(g) = \begin{cases} \frac{2}{|G|} & \text{if } \chi(g) = (-1)^k, \\ 0 & \text{otherwise.} \end{cases}$$

To quantify this expectation, there are various measures of closeness. We will use the L^1 and L^2 norms on G with the total volume normalized to 1. Thus

$$||f||_p = \left(\frac{1}{|G|} \sum_{g\in G} |f(g)|^p \right)^{1/p}.$$

In the literature on random walks on a finite group G, it is customary to measure the distance between two probability distributions p and q by the *total variation distance* (see Diaconis [1] and Saloff-Coste [5]). This is defined by:

$$\|p - q\|_{\text{tv}} = \max_{A\subset G} F(A), \qquad F(A) = \left| \sum_{g\in A} p(g) - q(g) \right|.$$

Proposition 5. *If p and q are probability distributions, we have*

$$\|p - q\|_{\text{tv}} = \frac{|G|}{2} ||p - q||_1. \tag{1}$$

Proof. Let $A_1 = \{g \in G | p(g) > q(g)\}$ and $A_2 = \{g \in G | p(g) < q(g)\}$. Then since p and q are real, it is easy to see that $F(A)$ is maximal when $A = A_1$ or A_2, and since p and q are probability distributions, it is easy to see that $F(A_1) = F(A_2)$. Moreover on A_1 or A_2 we have $|\sum(p(g) - q(g))| = \sum |p(g) - q(g)|$ since all terms have the same sign, so

$$F(A_1) + F(A_2) = \sum_{g \in G} |p(g) - q(g)| = |G| \times ||p - q||_1.$$

□

Proposition 6. *We have*

$$\|p_k - L_k\|_{\text{tv}} \leqslant \frac{|G|}{2^{3/2}} ||p_k - L_k||_2.$$

Proof. In view of Proposition 5, we must show that

$$||p_k - L_k||_1 \leqslant \frac{1}{\sqrt{2}} ||p_k - L_k||_2.$$

This is basically the Cauchy-Schwartz inequality, taking into account that the support of $p_k - L_k$ is contained $\{g|\chi(g) = (-1)^k\}$, a set of order $|G|/2$. To be precise, let $f_k(g) = |p_k(g) - L_k(g)|$. Then

$$||p_k - L_k||_1 = \left\langle f_k, \frac{|G|}{2} L_k \right\rangle \leqslant \|f_k\|_2 \cdot \|(|G|/2)L_k\|_2 = \frac{1}{\sqrt{2}} \|f_k\|_2. \tag{2}$$

□

Proposition 7. *We have*

$$||p_k - L_k||_2 \leqslant \frac{\lambda_3^k}{\sqrt{|G|}},$$

Proof. We can write the initial distribution

$$p_0 = \sum_{i=1}^{N} c_i \xi_i. \tag{3}$$

The first two eigenfunctions are $\xi_1(g) = 1$ and $\xi_2(g) = \chi(g)$. Since p_0 is a probability distribution,

$$1 = \sum_{g \in G} p_0(g) = |G| \langle p_0, \xi_1 \rangle = |G| c_1, \tag{4}$$

so $c_1 = \frac{1}{|G|}$ and $c_1\xi_1$ is the uniform distribution where every state has equal probability. Similarly $c_2 = \frac{1}{|G|}$ since (4) would remain true if we replace ξ_1 by ξ_2, which equals ξ_1 at the origin (which is the only place where $p_0 \neq 0$).

We can choose our orthonormal basis so that ξ_3 is a constant times the projection of p_0 on the λ_3-eigenspace. This means that $c_4 = \ldots = c_{22} = 0$. We can take $\xi_{20+i} = \chi\xi_i$ when $i = 3, 4, \cdots, 22$. Then, since $p_0 = \chi p_0$, we have $c_{23} = c_3$ and $c_{24} = \ldots = c_{42} = 0$. Thus

$$p_0 = c_1\xi_1 + c_2\xi_2 + c_3\xi_3 + c_{23}\xi_{23} + \sum_{i \geqslant 43} c_i\xi_i = \frac{1}{|G|}(\xi_1 + \xi_2) + c_3(\xi_3 + \xi_{23}) + \ldots ,$$

and since $L_k = \frac{1}{|G|}(\xi_1 + (-1)^k\xi_2)$

$$p_k - L_k = c_3\lambda_3^k(\xi_3 + (-1)^k\xi_{23}) + \sum_{i \geqslant 43} c_i\lambda_i^k\xi_i. \tag{5}$$

Thus

$$||p_k - L_k||_2 = ||\sum_{i>2} c_i\xi_i||_2 = \sqrt{\sum_{i>2} |\lambda_i|^{2k}|c_i|^2} \leqslant \lambda_3^k\sqrt{\sum_{i=1}^{|G|} |c_i|^2}, \tag{6}$$

since λ_3 is the largest of the eigenvalues that occur. But by the Plancherel formula

$$\sqrt{\sum_i |c_i|^2} = \|p_0\|_2 = \sqrt{\frac{1}{|G|}\sum_{g\in G} p_0(g)^2} = \frac{1}{\sqrt{|G|}}.$$

□

Theorem 1. *There exist constants $C_1^{\text{odd}}, C_1^{\text{even}}$ and $C_2 > 0$ such that as $k \longrightarrow \infty$*

$$\begin{aligned} \|p_k - L_k\|_{\text{tv}} &\sim \begin{cases} C_{\text{tv}}^{\text{even}}\lambda_3^k & \text{if } k \text{ is even,} \\ C_{\text{tv}}^{\text{odd}}\lambda_3^k & \text{if } k \text{ is odd,} \end{cases} \\ \|p_k - L_k\|_2 &\sim C_2\lambda_3^k. \end{aligned} \tag{7}$$

At the end we will give evidence that $C_{\text{tv}}^{\text{even}} = C_{\text{tv}}^{\text{odd}}$, and conjecture a value for this. This is an interesting empirical observation for which we have no explanation.

Proof. First let k run through the set of all even positive integers, or through all odd positive integers. We will prove the existence of constants C_i^{odd} and C_i^{even} $(i = 1, 2)$ such that (7) is true for k thus restricted. At the end, we will show $C_2^{\text{odd}} = C_2^{\text{even}}$.

It is clear from (5) that $p_k - L_k$ equals $c_3\lambda_3^k(\xi_3 + (-1)^k\xi_{23})$ plus terms that are more rapidly decreasing as $k \longrightarrow \infty$. Thus if $i = 1, 2$

$$\|p_k - L_k\|_i \sim \begin{cases} C_i^{\text{even}}\lambda_3^k & \text{if } k \text{ is even,} \\ C_i^{\text{odd}}\lambda_3^k & \text{if } k \text{ is odd,} \end{cases} \qquad \begin{array}{l} C_i^{\text{even}} = c_3\|\xi_3 + \xi_{23}\|_i\,, \\ C_i^{\text{odd}} = c_3\|\xi_3 - \xi_{23}\|_i\,. \end{array}$$

In view of (1), the statement follows for the total variation distance. We have $C_2^{\text{odd}} = C_2^{\text{even}}$ since ξ_3 and ξ_{23} are orthogonal, so both equal $c_3\sqrt{\|\xi_3\|_2^2 + \|\xi_{23}\|_2^2}$. □

By Propositions 6 and 7 we have

$$C_{\text{tv}}^{\text{odd}}, C_{\text{tv}}^{\text{even}} \leqslant \frac{\sqrt{|G|}}{2^{3/2}} \cong 60.37\cdots\,, \qquad C_2 \leqslant \frac{1}{\sqrt{|G|}} = .005856\cdots\,. \tag{8}$$

How sharp are these bounds? To answer this, we tabulated $\|p_k - L_k\|_{\text{tv}}$ and $\|p_k - L_k\|_2$ in Table 3, simply by iterating the Markov process and computing the values directly. We found that $||p_k - L_k||_{\text{tv}}\ \lambda_3^{-k}$ and $||p_k - L_k||_2\ \lambda_3^{-k}$ tend to limits that are close to square roots of rational numbers that we can identify, and so we conjecture that the actual values are $C_{\text{tv}}^{\text{odd}} = C_{\text{tv}}^{\text{even}} = \sqrt{3} = 1.73205\cdots$ and $C_2 = \frac{\sqrt{40}}{|G|} = 0.000216891\cdots$. Thus the *a priori* bounds (8) are not sharp. Still, they are the best that we know how to obtain by purely theoretical methods.

References

[1] P. Diaconis, *Group representations in probability and statistics.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.

[1] The GAP Group. GAP - Groups, Algorithms, and Programming, Version 4.4. http://www.gap-system.org.

[2] R. Horn and C. Johnson, *Matrix Analysis.* Cambridge University Press, Cambridge 1985.

[3] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[4] G. Mackey, Induced representations of locally compact groups. I. *Ann. of Math.* (2) 55, (1952). 101–139.

[5] Laurent Saloff-Coste. Random walks on finite groups. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 263–346. Springer, Berlin, 2004.

[6] David Singmaster. *Notes on Rubik's magic cube*. Enslow Publishers, Hillside, N.J., fifth edition, 1981.

[7] Hua Zhou. Scrambling a Rubik's cube. *Preprint*, 2005.

1.000000	1	0.964905	20	0.936084	30	0.909367	30	0.908049	30
0.907849	40	0.907045	60	0.906972	60	0.904508	12	0.888077	120
0.886611	30	0.871275	30	0.855006	20	0.847422	20	0.837335	30
0.835757	60	0.835030	60	0.832707	60	0.832706	60	0.832269	30
0.831096	60	0.830034	60	0.829343	48	0.829218	60	0.829145	60
0.827309	60	0.826586	48	0.824754	60	0.820615	60	0.819993	60
0.819279	20	0.819154	30	0.809017	84	0.806201	60	0.806200	60
0.804298	30	0.802612	30	0.800488	20	0.800106	60	0.798320	40
0.795839	60	0.791877	40	0.790766	60	0.790764	60	0.788831	30
0.786276	12	0.785399	30	0.782159	40	0.778619	20	0.775149	20
0.770755	20	0.766137	60	0.766118	40	0.765693	40	0.763270	40
0.758168	40	0.756365	30	0.754269	20	0.752256	60	0.752252	60
0.750000	51	0.736625	60	0.736556	60	0.725562	60	0.721957	30
0.720600	40	0.720196	60	0.720192	60	0.717310	30	0.711052	30
0.710455	60	0.710296	60	0.698521	48	0.688847	60	0.688712	60
0.683013	10	0.673817	60	0.668016	20	0.668013	20	0.667821	20
0.667817	20	0.660964	48	0.660754	60	0.660738	60	0.659446	40
0.658431	30	0.654514	30	0.652425	60	0.641324	60	0.641287	60
0.640388	5	0.635712	30	0.629204	60	0.629181	60	0.624378	60
0.624062	60	0.623212	30	0.619106	60	0.610474	40	0.607727	60
0.606594	60	0.604352	30	0.602651	60	0.601785	60	0.596907	40
0.591803	30	0.591660	30	0.575695	15	0.575694	119	0.575693	30
0.574576	30	0.572932	48	0.567158	60	0.567155	60	0.561738	60
0.552810	30	0.544301	60	0.544279	60	0.542606	60	0.542604	60
0.539565	20	0.537475	60	0.536339	30	0.532916	20	0.520404	20
0.518901	40	0.512703	30	0.511862	60	0.511852	60	0.500000	54
0.492263	120	0.490689	60	0.488542	40	0.487392	60	0.487390	60
0.478352	30	0.457161	30	0.455708	60	0.444737	30	0.444065	60
0.444028	30	0.443226	48	0.438052	120	0.432685	40	0.427291	120
0.427234	40	0.424228	20	0.420769	20	0.420616	30	0.417270	12
0.416355	30	0.415291	30	0.408431	30	0.401645	30	0.400033	60
0.400032	60	0.397816	60	0.397815	60	0.396107	60	0.393558	60
0.393554	60	0.393444	30	0.393070	30	0.392473	60	0.392469	60
0.390388	5	0.381148	60	0.380994	12	0.373564	20	0.368842	60
0.368314	30	0.362423	120	0.361707	40	0.360534	40	0.357762	20
0.356082	30	0.346821	40	0.345492	12	0.344989	20	0.344096	120
0.331079	120	0.325694	164	0.324601	40	0.319624	120	0.319198	30
0.316075	48	0.314854	20	0.309017	84	0.306798	30	0.303628	48
0.301470	20	0.301067	30	0.300165	60	0.297155	120	0.294180	120
0.284554	40	0.283465	40	0.282041	30	0.280577	60	0.276563	40
0.275524	60	0.274968	40	0.268420	30	0.266718	120	0.266447	120
0.266103	120	0.264351	120	0.259595	30	0.250001	10	0.250000	86
0.211505	40	0.198295	30	0.197840	30	0.195990	40	0.192838	120
0.190111	48	0.186409	40	0.183013	10	0.176079	120	0.170410	20
0.169539	30	0.167326	40	0.166980	60	0.157835	120	0.145004	30
0.118314	30	0.117065	48	0.106347	60	0.106303	120	0.104916	30
0.093062	30	0.091040	120	0.089895	120	0.085426	120	0.081867	60
0.076615	30	0.070351	30	0.068591	40	0.060219	30	0.055113	40
0.051983	20	0.044628	60	0.037735	20	0.007424	120	0.000000	4944

Table 2: Nonnegative eigenvalues of M, with multiplicities.

k	$\|\|p_k - L_k\|\|_{\text{tv}}$	$\|\|p_k - L_k\|\|_2$	$\|\|p_k - L_k\|\|_{\text{tv}}\ \lambda_3^{-k}$	$\|\|p_k - L_k\|\|_2\ \lambda_3^{-k}$
5	0.989163	0.00105344	1.18262	0.00125947
10	0.896468	0.00048854	1.28141	0.000698321
15	0.790351	0.000278645	1.35068	0.000476193
20	0.691236	0.000180163	1.41233	0.000368108
25	0.598712	0.0001264	1.46253	0.000308768
30	0.514137	0.0000938488	1.50156	0.000274089
40	0.372838	0.0000575967	1.55646	0.000240445
50	0.267605	0.0000381118	1.59686	0.000227422
60	0.191116	0.0000260214	1.63014	0.000221951
70	0.135807	0.0000179994	1.65578	0.000219452
80	0.0960715	0.0000125222	1.67429	0.00021823
90	0.067769	8.73526×10^{-6}	1.68819	0.000217604
95	0.0568745	7.29994×10^{-6}	1.69389	0.000217414
100	0.047711	6.10191×10^{-6}	1.69888	0.000217275
150	0.00811339	1.02085×10^{-6}	1.72393	0.00021691
175	0.00332912	4.17877×10^{-7}	1.72795	0.000216896
200	0.00136435	1.71063×10^{-7}	1.72987	0.000216892
300	0.0000383523	4.80408×10^{-9}	1.73151	0.000216892
500	3.02499×10^{-8}	3.78898×10^{-12}	1.73159	0.000216892

Table 3: The convergence of the random walk to the uniform distribution.