EMBEDDING A MARKOV CHAIN INTO A RANDOM WALK ON A PERMUTATION GROUP

STEVEN N. EVANS

Abstract. Using representation theory, we obtain a necessary and sufficient condition for a discrete–time Markov chain on a finite state space E to be representable as $\Psi_n \Psi_{n-1} \cdots \Psi_1 z$, $n \geq 0$, for any $z \in E$, where the Ψ_i are independent, identically distributed random permutations taking values in some given transitive group of permutations on E. The condition is particularly simple when the group is 2-transitive on E. We also work out the explicit form of our condition for the dihedral group of symmetries of a regular polygon.

1. Introduction

Consider a discrete–time (left) random walk Φ on a transitive group Γ of permutations of a finite set E. That is, $\Phi = (\Phi_n, \mathbb{Q}^\varphi)$ is a Markov chain with state-space Γ such that for $\varphi \in \Gamma$

$$\mathbb{Q}^\varphi\{\Phi_0 = \varphi\} = 1$$

and

$$\mathbb{Q}^\varphi\{\Phi_{n+1} = \psi \mid \Phi_0, \Phi_1, \ldots, \Phi_n\} = Q(\psi \Phi_n^{-1})$$

for some probability distribution Q on Γ. Equivalently, $\Phi_1\Phi_0^{-1}, \Phi_2\Phi_1^{-1}, \ldots$ are i.i.d. under each measure $\mathbb{Q}^\varphi$ with common distribution Q.

For any $z \in E$ we have

$$\mathbb{Q}^\varphi\{\Phi_{n+1} z = y \mid \Phi_0, \ldots, \Phi_n\} = P(\Phi_n z, y),$$

where

$$P(x, y) = \sum_{\psi \in \Gamma : \psi x = y} Q(\psi). \tag{1.1}$$

Thus $(\Phi_n z)_{n \geq 0}$ is a Markov chain with transition matrix P that doesn't depend on z.

We are interested in the extent to which this "projection" of a random walk on Γ onto a Markov chain on E can be reversed. That is, given a Markov chain $X = (X_n, \mathbb{P}^x)$ on E with transition matrix P, when can we "lift" X to find a random walk on Γ with increment distribution Q such that (1.1) holds? The existence of such a lifting allows one to employ the extremely powerful tools, particularly representation theory, that have been used to analyse random walks on groups. A typical and impressive example is [DS87], where the Bernoulli-Laplace diffusion

Date: June 10, 2003.

2000 *Mathematics Subject Classification.* 60J10, 60G50, 60B99, 20B99.

Key words and phrases. doubly stochastic, Markov function, singular value decomposition, generalised inverse, representation, character, 2-transitive group, dihedral group.

Research supported in part by NSF grant DMS-0071468 and a research professorship from the Miller Institute for Basic Research in Science.

model is lifted to a walk on the symmetric group that is bi-invariant under a certain subgroup and the theory of Gelfand pairs is then used to analyse how fast the diffusion model converges to stationarity (see also [Dia88]).

The question of when a lifting exists has a simple answer when Γ is the symmetric group on E (that is, the group of all permutations of E). Firstly, note that if (1.1) holds, then $\sum_{x\in E} P(x,y) = \sum_{\psi\in\Gamma} Q(\psi) = 1$ for each $y \in E$, and hence P is doubly stochastic (this observation holds for an arbitrary group Γ). On the other hand, if P is doubly stochastic, then, by a celebrated result of G. Birkhoff [Bir46, HJ90], P is in the convex hull of the permutation matrices, which is just another way of saying that (1.1) holds.

Even for the symmetric group, the choice of Q is not unique. For example, if $E = \{1,2,3\}$ and $P(x,y) = \frac{1}{3}$ for all x, y, then one possible choices for Q is the probability measure that assigns mass $\frac{1}{6}$ to each possible permutation, and another is the measure that assigns mass $\frac{1}{3}$ to the even permutations and mass 0 to the odd permutations (in the usual cycle notation, the even permutations are $(1)(2)(3)$, $(1,2,3)$, and $(1,3,2)$, while the odd permutations are the transpositions $(1,2)(3)$, $(1,3)(2)$, and $(1)(2,3)$).

In order to describe our results we need to recall a little notation from representation theory. A convenient reference for the facts we need is [Ker99]. Let $\hat{\Gamma}$ denote the collection of irreducible (unitary) representations of Γ. Given $\rho \in \hat{\Gamma}$, write χ_ρ for the character of ρ and d_ρ for the dimension of ρ. The action of Γ on E has an associated representation: each element of Γ is associated with the corresponding $|E| \times |E|$ $\{0,1\}$-valued permutation matrix. This so-called permutation representation decomposes into a direct sum of irreducible representations. Write $\hat{\Gamma}_+$ for the collection of irreducible representations $\rho \in \hat{\Gamma}$ that appear with positive multiplicity $\nu_\rho > 0$ in the decomposition of the permutation representation, and write $\hat{\Gamma}_0$ for the collection of irreducible representations that do not appear. The character of the permutation representation is $N(\varphi) = |\{x : \varphi x = x\}|$ (that is, the number of fixed points of the permutation $\varphi \in \Gamma$). Thus,

$$\begin{aligned}\nu_\rho &= \frac{1}{|\Gamma|} \sum_{\psi\in\Gamma} N(\psi)\chi_\rho(\psi) \\ &= \frac{1}{|\Gamma|} \sum_{x\in E} \sum_{\psi\in\Gamma:\psi x = x} \chi_\rho(\psi),\end{aligned}$$

with $\nu_\rho = 0$ when ρ does not appear in the decomposition of the permutation representation.

Theorem 1.1. *Let $P(x,y)$, $x, y \in E$, be a transition matrix on E. There exists a probability vector Q on Γ such that (1.1) holds if and only if*

$$P(x,y) = \sum_{\rho\in\hat{\Gamma}_+} \sum_{z\in E} \sum_{\varphi\in\Gamma:\varphi x = y} \sum_{\psi\in\Gamma} \frac{d_\rho^2}{|\Gamma|^2 \nu_\rho} \chi_\rho(\psi\varphi^{-1}) P(z,\psi z), \tag{1.2}$$

for all $x, y \in E$, and, for some choice of $h \in \mathbb{R}^\Gamma$,

$$\begin{aligned} R_h(\varphi) := &\sum_{\rho \in \hat{\Gamma}_+} \sum_{z \in E} \sum_{\psi \in \Gamma} \frac{d_\rho^2}{|\Gamma|^2 \nu_\rho} \chi_\rho(\psi\varphi^{-1}) P(z, \psi z) \\ &+ \sum_{\rho \in \hat{\Gamma}_0} \sum_{\psi \in \Gamma} \frac{d_\rho}{|\Gamma|} \chi_\rho(\psi\varphi^{-1}) h(\psi) \\ \geq\ & 0, \end{aligned} \tag{1.3}$$

for all $\varphi \in \Gamma$. Moreover, if (1.2) and (1.3) hold for some $h \in \mathbb{R}^\Gamma$, then the class of probability vectors Q satisfying (1.1) coincides with the class of $R_{h'}$ satisfying (1.3) for some $h' \in \mathbb{R}^\Gamma$ (in particular, all such $R_{h'}$ are automatically probability vectors).

Remark 1.2. If P is such that (1.1) holds for a particular probability vector Q, then it follows from the observations made in the proof of Theorem 1.1 that taking $h = Q$ in (1.3) gives $R_h = Q$.

Remark 1.3. Finding a probability distribution Q that satisfies equation (1.1) involves solving $|E| \times |E|$ equalities and $|\Gamma|$ inequalities in $|\Gamma|$ unknowns, whereas applying Theorem 1.1 involves solving only $|\Gamma|$ inequalties in $|\Gamma|$ unknowns. Although the characters χ_ρ can certainly be complex-valued, it is apparent from the proof of Theorem 1.1 that (1.3) is a system of inequalities of the form $Ax + b \geq 0$, with A a real $|\Gamma| \times |\Gamma|$ matrix and b a real vector of length $|\Gamma|$. Linear programming methods such as the simplex algorithm can be used to solve such inequalities or to ascertain that they are insoluble (see Chapter 6 of [Pad95]). We also remark that there is a Farkas-type "theorem of the alternative" which gives an equivalent condition for (1.3) to hold: namely (1.3) will hold for some h if and only if there is **no** $k \in \mathbb{R}_+^\Gamma$ such that both

$$\sum_{\rho \in \hat{\Gamma}_0} \sum_{\varphi \in \Gamma} \frac{d_\rho}{|\Gamma|} \chi_\rho(\psi\varphi^{-1}) k(\varphi) = 0$$

for all $\psi \in \Gamma$ and

$$\sum_{\varphi \in \Gamma} \sum_{\rho \in \hat{\Gamma}_+} \sum_{z \in E} \sum_{\psi \in \Gamma} \frac{d_\rho^2}{|\Gamma|^2 \nu_\rho} \chi_\rho(\psi\varphi^{-1}) P(z, \psi z) k(\varphi) < 0$$

(see Exercise 6.5(i) of [Pad95]).

Example 1.4. Suppose that $\Gamma = E$ and Γ acts on E via the left regular representation. Of course, in this case it is obvious that a lifting exists if and only if $P(\varphi, \psi)$ is of the form $Q(\psi\varphi^{-1})$, and this Q is the unique lifted increment distribution. It is easy to check that this conclusion follows from Theorem 1.1. In this case $\nu_\rho = d_\rho$ (see Corollary 11.5.4 of [Ker99]) and so

$$\begin{aligned} &\sum_\rho \frac{d_\rho^2}{|\Gamma|^2 \nu_\rho} \sum_{\varphi, \psi \in \Gamma} \mathbf{1}\{y' = \varphi x', y'' = \psi x''\} \chi_\rho(\psi\varphi^{-1}) \\ &= \sum_\rho \frac{d_\rho}{|\Gamma|^2} \chi_\rho(y''(x'')^{-1}(y'(x')^{-1})^{-1}) \\ &= \begin{cases} \frac{1}{|\Gamma|}, & \text{if } y'(x')^{-1} = y''(x'')^{-1}, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Thus conditions (1.2) of Theorem 1.1 holds if and only if $P(x, y)$ only depends on yx^{-1}. We know that there is a lifting in this case. Because $\hat{\Gamma}_0$ is empty, this lifting is unique by Theorem 1.1.

Representation theory is at its most useful for analysing the lifted random walk when the walk has some extra structure. For example, if K is a subgroup of Γ such that (Γ, K) form a Gelfand pair, then the analysis of random walks on Γ that are K-bi-invariant (that is, $Q(k'\varphi k'') = Q(\varphi)$ for $\varphi \in \Gamma$ and $k', k'' \in K$) is particularly simple (see [DS87, Dia88]). A necessary and sufficient condition for a Markov chain on the quotient $E = \Gamma/K$ with transition matrix P to posess a K-bi-invariant lifting is that

$$P(x, y) = P(\varphi x, \varphi y) \text{ for all } x, y \in E \text{ and } \varphi \in \Gamma \tag{1.4}$$

(see Lemma 1 of [DS87], where this result is attributed to Philippe Bougerol).

Another situation in which the representation theoretic analysis of random walks is particularly simple is when the increment distribution is a class function (that is, is constant on conjugacy classes) because the matrix $(\varphi, \psi) \mapsto Q(\psi\varphi^{-1})$ can then be explicitly diagonalised using the characters of Γ – see Chapter 3 of [Dia88]. The following result is clear from Theorem 1.1 and Remark 1.2.

Corollary 1.5. *Let $P(x, y)$, $x, y \in E$, be a transition matrix on E. There exists a probability vector Q on Γ such that (1.1) holds and Q is a class function if and only if condition (1.2) holds, condition (1.3) holds with h a class function, and*

$$\varphi \mapsto \sum_{\rho \in \hat{\Gamma}_+} \sum_{z \in E} \sum_{\psi \in \Gamma} \frac{d_\rho^2}{|\Gamma|^2 \nu_\rho} \chi_\rho(\psi\varphi^{-1}) P(z, \psi z) \tag{1.5}$$

is a class function. Moreover, if these conditions hold for some class function h, then the class of probability vectors Q satisfying (1.1) that are class functions coincides with the class of $R_{h'}$ satisfying (1.3) for some class function h' (in particular, all such $R_{h'}$ are automatically probability vectors).

Remark 1.6. Condition (1.5) is implied by condition (1.4).

Theorem 1.1 takes a considerably simpler form if the group Γ is 2-*transitive* on E; that is, if for any two pairs $(u, v), (x, y) \in E$ with $u \neq v$ and $x \neq y$ there exists a $\varphi \in \Gamma$ with $(u, v) = (\varphi x, \varphi y)$. The group Γ is 2-transitive on E if and only if

$$\frac{1}{|\Gamma|} \sum_{\varphi \in \Gamma} N(\varphi)(N(\varphi) - 1) = 1$$

or, equivalently,

$$\frac{1}{|\Gamma|} \sum_{\varphi \in \Gamma} N(\varphi)^2 = 2$$

(see Corollaries 8.1.2 and 8.1.6 of [Ker99]). For example, the symmetric group of all permutations of E is certainly 2-transitive.

Corollary 1.7. *Suppose that Γ is 2-transitive on E. Let $P(x, y)$, $x, y \in E$, be a transition matrix on E. There exists a probability vector Q on Γ such that (1.1) holds if and only if*

$$\sum_{x \in E} P(x, y) = 1, \tag{1.6}$$

for all $y \in E$, and, for some choice of $h \in \mathbb{R}^\Gamma$,

$$R_h(\varphi) := \frac{1}{|\Gamma|}\left[(|E|-1)\sum_{x\in E} P(x,\varphi x) - |E| + 2\right] + h(\varphi) \\ - \frac{1}{|\Gamma|}\sum_{\psi\in\Gamma}\left[(|E|-1)N(\psi\varphi^{-1}) - |E| + 2\right] h(\psi) \\ \geq 0, \tag{1.7}$$

for all $\varphi \in \Gamma$. Moreover, if (1.6) and (1.7) hold for some $h \in \mathbb{R}^\Gamma$, then the class of probability vectors Q satisfying (1.1) coincides with the class of $R_{h'}$ satisfying (1.7) for some $h' \in \mathbb{R}^\Gamma$ (in particular, all such $R_{h'}$ are automatically probability vectors).

We leave the formulation of an analogue of Corollary 1.5 for 2-transitive groups to the reader.

Example 1.8. Taking h to be a constant in Corollary 1.7, we see that a sufficient condition for equation (1.1) to hold is that equation (1.6) holds and

$$\sum_{x\in E} P(x,\varphi x) \geq \frac{|E|-2}{|E|-1}$$

for all $\varphi \in \Gamma$ (for example, $P(x,y) \geq (|E|-2)/(|E|(|E|-1))$ for all $x, y \in E$ certainly suffices). One can then take

$$Q(\varphi) = \frac{1}{|\Gamma|}\left[(|E|-1)\sum_{x\in E} P(x,\varphi x) - |E| + 2\right].$$

The outline of the rest of the paper is the following. Section 2 briefly recalls some facts about the singular value decomposition of a matrix and its connection to solving linear equations. Section 3 contains a proof of Theorem 1.1 and Section 4 contains a derivation of Corollary 1.7 from Theorem 1.1 as well as an indication of an alternative proof that avoids the use of representation theory. In Section 5 we work out explicitly the objects appearing in Theorem 1.1 for the case where Γ is the group of symmetries of a regular n-gon and E is the corresponding set of vertices.

2. The singular value decomposition and Moore–Penrose inverse

For the sake of completeness and to establish some notation, we recall some facts from linear algebra (see, for example, [Meh77, HJ90]). Let A be an $n \times k$ matrix with rank r (which is also the rank of A^*A and AA^*). The matrix A has the *singular value decomposition*

$$A = ULV^*$$

where:

- $L = \operatorname{diag}(\lambda_1^{1/2}, \ldots, \lambda_r^{1/2})$, with $\lambda_1, \ldots, \lambda_r$ the non-zero eigenvalues of A^*A;
- V is the $k \times r$ matrix with columns consisting of the corresponding orthonormalized eigenvectors (so that $V^*V = I$);
- U is the $n \times r$ matrix given by AVL^{-1} (so that $U^*U = I$ and the columns of U are the orthonormalized eigenvectors of AA^*).

The *Moore–Penrose generalized inverse* of A is the $k \times n$ matrix

$$A^\dagger = V L^{-1} U^*.$$

The linear equation $Ax = b$ has a solution if and only if $AA^\dagger b = b$. Moreover, if $Ax = b$ has a solution, then any solution is of the form

$$x = A^\dagger b + (I - A^\dagger A) z$$

for an arbitrary k-vector z. Note that

$$A^\dagger A = V V^*$$

and

$$A A^\dagger = U U^*,$$

and so $Ax = b$ will have a solution if and only if $UU^* b = b$, in which case a general solution is

$$x = A^\dagger b + (I - V V^*) z$$

for an arbitrary z. The matrix UU^* is the orthogonal projection onto the range of A and the matrix $I - VV^*$ is the orthogonal projection onto the kernel of A.

3. Proof of Theorem 1.1

We can write the equation (1.1) in the form

$$AQ = P,$$

where P is the transition matrix written out as a column vector indexed by $E \times E$ and A is the matrix with rows indexed by $E \times E$ and columns indexed by Γ that is given by

$$A((x,y),\psi) = \begin{cases} 1, & \text{if } y = \psi x, \\ 0, & \text{otherwise.} \end{cases}$$

(Of course, we are seeking solutions Q that have nonnegative entries that sum to 1.)

Now

$$\begin{aligned} A^* A(\varphi,\psi) &= \sum_{x,y} \mathbf{1}\{y = \varphi x\} \mathbf{1}\{y = \psi x\} \\ &= |\{x : \varphi x = \psi x\}| \\ &= N(\psi^{-1}\varphi). \end{aligned}$$

Here, as in the Introduction, N is the character of the permutation representation of Γ that counts the number of fixed points of a permutation. In particular, N is a class function (that is, depends only on the conjugacy class of a permutation).

We now apply a standard procedure to find diagonalize A^*A (see, for example, p48 of [Dia88] for a similar argument).

Consider an irreducible representation ρ of Γ with dimension d_ρ and character χ_ρ. Because N is a class function, the Fourier transform of N at ρ is

$$\begin{aligned}\hat{N}(\rho) &= \sum_{\varphi\in\Gamma} N(\varphi)\rho(\varphi)\\ &= \left(\frac{1}{d_\rho}\sum_{\varphi\in\Gamma} N(\varphi)\chi_\rho(\varphi)\right) I\\ &= \lambda_\rho I,\end{aligned}$$

say (see Lemma 11.5.5 of [Ker99]). Note that

$$\lambda_\rho = \frac{|\Gamma|}{d_\rho}\nu_\rho,$$

where, as in the Introduction,

$$\nu_\rho = \frac{1}{|\Gamma|}\sum_{\varphi\in\Gamma} N(\varphi)\chi_\rho(\varphi)$$

is the multiplicity of the representation ρ in the decomposition of the permutation representation into irreducible components.

Let $(\rho_{ij})_{1\le i,j\le d_\rho}$ be a unitary matrix realization of ρ. We have

$$\begin{aligned}\sum_{\psi\in\Gamma} A^*A(\varphi,\psi)\bar{\rho}_{ji}(\psi) &= \sum_{\psi\in\Gamma} N(\psi^{-1}\varphi)\rho_{ij}(\psi^{-1})\\ &= \sum_{\psi\in\Gamma} N(\psi)\rho_{ij}(\psi\varphi^{-1})\\ &= \sum_{\psi\in\Gamma} N(\psi)\sum_{k=1}^{d_\rho}\rho_{ik}(\psi)\rho_{kj}(\varphi^{-1})\\ &= \sum_{k=1}^{d_\rho}\hat{N}(\rho)_{ik}\bar{\rho}_{jk}(\varphi)\\ &= \lambda_\rho\bar{\rho}_{ji}(\varphi).\end{aligned}$$

Thus

$$v_{\rho,i,j} = \left(\frac{d_\rho}{|\Gamma|}\right)^{1/2}\bar{\rho}_{ji},\ 1\le i,j\le d_\rho,$$

are d_ρ^2 orthonormalized eigenvectors associated with the eigenvalue λ_ρ. Because $\sum_\rho d_\rho^2 = |\Gamma|$ (see Corollary 11.5.4 of [Ker99]), we have found all the eigenvalues of the $|\Gamma|\times|\Gamma|$ matrix A^*A.

The elements appearing in the singular value decomposition of A are thus the following.

- The matrix V has a column $v_{\rho,i,j}$ for each $\rho\in\hat{\Gamma}_+$ (that is, for each ρ appearing in the permutation representation) and each pair $1\le i,j\le d_\rho$.
- The diagonal matrix L has

$$\lambda_\rho^{1/2} = \left(\frac{|\Gamma|\nu_\rho}{d_\rho}\right)^{1/2}$$

appearing d_ρ^2 times on the diagonal.

- The matrix U has columns given by

$$\begin{aligned} u_{\rho,i,j}(x,y) &= \lambda_\rho^{-1/2}(Av_{\rho,i,j})(x,y) \\ &= \lambda_\rho^{-1/2}\sum_{\psi\in\Gamma}\mathbf{1}\{y=\psi x\}\left(\frac{d_\rho}{|\Gamma|}\right)^{1/2}\bar\rho_{ji}(\psi) \\ &= \frac{d_\rho}{|\Gamma|\nu_\rho^{1/2}}\sum_{\psi\in\Gamma}\mathbf{1}\{y=\psi x\}\bar\rho_{ji}(\psi). \end{aligned}$$

The Moore–Penrose inverse of A is thus given by

$$\begin{aligned} A^\dagger(\varphi,(x,y)) &= \sum_{\rho,i,j}\left(\frac{d_\rho}{|\Gamma|}\right)^{1/2}\bar\rho_{ji}(\varphi)\left(\frac{d_\rho}{|\Gamma|\nu_\rho}\right)^{1/2}\frac{d_\rho}{|\Gamma|\nu_\rho^{1/2}}\sum_{\psi\in\Gamma}\mathbf{1}\{y=\psi x\}\rho_{ji}(\psi) \\ &= \sum_{\rho\in\hat\Gamma_+}\frac{d_\rho^2}{|\Gamma|^2\nu_\rho}\sum_{\psi\in\Gamma}\mathbf{1}\{y=\psi x\}\sum_{i,j}\rho_{ji}(\psi)\rho_{ij}(\varphi^{-1}) \\ &= \sum_{\rho\in\hat\Gamma_+}\frac{d_\rho^2}{|\Gamma|^2\nu_\rho}\sum_{\psi\in\Gamma}\mathbf{1}\{y=\psi x\}\chi_\rho(\psi\varphi^{-1}). \end{aligned}$$

Also,

$$\begin{aligned} AA^\dagger((x',y'),(x'',y'')) &= UU^*((x',y'),(x'',y'')) \\ &= \sum_{\rho,i,j}\frac{d_\rho^2}{|\Gamma|^2\nu_\rho}\sum_{\varphi,\psi\in\Gamma}\mathbf{1}\{y'=\varphi x'\}\mathbf{1}\{y''=\psi x''\}\bar\rho_{ji}(\varphi)\rho_{ji}(\psi) \\ &= \sum_{\rho\in\hat\Gamma_+}\frac{d_\rho^2}{|\Gamma|^2\nu_\rho}\sum_{\varphi,\psi\in\Gamma}\mathbf{1}\{(y',y'')=(\varphi x',\psi x'')\}\chi_\rho(\psi\varphi^{-1}) \end{aligned}$$

and

$$\begin{aligned} A^\dagger A(\varphi,\psi) &= VV^*(\varphi,\psi) \\ &= \sum_{\rho,i,j}\frac{d_\rho}{|\Gamma|}\bar\rho_{ji}(\varphi)\rho_{ji}(\psi) \\ &= \sum_{\rho\in\hat\Gamma_+}\frac{d_\rho}{|\Gamma|}\chi_\rho(\psi\varphi^{-1}). \end{aligned}$$

Thus

$$(I-A^\dagger A)(\varphi,\psi)=\sum_{\rho\in\hat\Gamma_0}\frac{d_\rho}{|\Gamma|}\chi_\rho(\psi\varphi^{-1}),$$

because

$$\sum_{\rho\in\hat\Gamma}d_\rho\chi_\rho(\eta)=\begin{cases}|\Gamma|, & \eta=e,\\ 0, & \text{otherwise,}\end{cases}$$

(see Corollary 11.5.4 of [Ker99]).

Theorem 1.1 will now follow if we can show that $\sum_{\varphi\in\Gamma}R_h(\varphi)=1$ for any $h\in\mathbb{R}^\Gamma$. This, however, follows from the two observations:

- $\sum_{\varphi\in\Gamma}\chi_\rho(\psi\varphi^{-1})$ is $|\Gamma|$ if ρ is the trivial one-dimensional representation with character 1 and is 0 for any other $\rho\in\hat\Gamma$ (see Theorem 11.5.3 of [Ker99]);

- the trivial representation appears with multiplicity 1 in the permutation representation (see Lemma 2.1.1 and Theorem 11.5.3 of [Ker99]).

□

4. Proof of Corollary 1.7

The corollary follows directly from Theorem 1.1 and a little algebra once we observe that the permutation representation associated with Γ acting on E decomposes into two irreducible representations. The trivial representation with dimension 1 and character the constant 1 appears with multiplicity 1, and the representation with dimension $|E|-1$ and character $N(\cdot)-1$ also appears with multiplicity 1 (see Exercise 11.5.7 of [Ker99]).

Alternatively, it is interesting to note that it is also possible to prove Corollary 1.7 directly without recourse to the representation theory of Γ. The argument goes as follows.

Let A be as in Section 3. Rather than work with A^*A to find a Moore–Penrose inverse, as we did in Section 3, we will work with AA^*.

By the 2-transitivity of Γ we have

$$\begin{aligned} AA^*((x',y'),(x'',y'')) &= \sum_{\psi\in\Gamma} \mathbf{1}\{y'=\psi(x'), y''=\psi(x'')\} \\ &= \begin{cases} \frac{|\Gamma|}{|E|}, & x'=x'', y'=y'', \\ 0, & x'\neq x'', y'=y'', \\ 0, & x'=x'', y'\neq y'', \\ \frac{|\Gamma|}{|E|(|E|-1)}, & x'\neq x'', y'\neq y''. \end{cases} \end{aligned}$$

Thus, by a suitable indexing of rows and columns, AA^* has the block form

$$\frac{|\Gamma|}{|E|(|E|-1)} \begin{pmatrix} S & T & \dots & T \\ T & S & \dots & T \\ \vdots & \vdots & \ddots & \vdots \\ T & T & \dots & S \end{pmatrix},$$

where $S = (|E|-1)I_{|E|}$ and

$$T = \begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & & 1 \\ \vdots & & \ddots & \\ 1 & & & 0 \end{pmatrix}.$$

Identify E with the cyclic group $\mathcal{C}$ of order $|E|$ by any bijective correspondence. Then $AA^*((x',y'),(x'',y'')) = F((x'',y'')-(x',y'))$, where the function $F:\mathcal{C}\times\mathcal{C}\to\mathbb{R}$ is given by

$$F(x,y) = \begin{cases} \frac{|\Gamma|}{|E|}, & x=0, y=0, \\ \frac{|\Gamma|}{|E|(|E|-1)}, & x\neq 0, y\neq 0, \\ 0, & \text{otherwise}. \end{cases}$$

Because AA^* is a convolution matrix, we can find the eigenvalues and eigenvectors of AA^* using Fourier analysis on $\mathcal{C}\times\mathcal{C}$ in the following manner.

There is an isomorphism between $\mathcal{C}$ and its dual group. Write $\{\theta_a : a \in \mathcal{C}\}$ for the dual group. The characters of $\mathcal{C} \times \mathcal{C}$ are then of the form

$$(x, y) \mapsto \theta_a(x)\theta_b(y), \ a, b \in \mathcal{C}.$$

Then

$$\begin{aligned}
&\sum_{(x'',y'')} AA^*((x',y'),(x'',y''))\theta_a(x'')\theta_b(y'') \\
&= \sum_{(x'',y'')} F((x'',y'') - (x',y'))\theta_a(x'')\theta_b(y'') \\
&= \sum_{(x,y)} F(x,y)\theta_a(x+x')\theta_b(y+y') \\
&= \left[\sum_{(x,y)} F(x,y)\theta_a(x)\theta_b(y)\right]\theta_a(x')\theta_b(y').
\end{aligned}$$

A set of orthonormalized eigenvectors of AA^* is thus $(x, y) \mapsto \frac{1}{|E|}\theta_a(x)\theta_b(y)$ $a, b \in \mathcal{C} \times \mathcal{C}$ (these are all the

eigenvectors because they are linearly independent and there are $|E|^2$ of them), and the corresponding eigenvalues are $\sum_{(x,y)} F(x, y)\theta_a(x)\theta_b(y)$. (This is, of course, is analogous to what we did in in the typically non-commutative setting of Section 3 and is a standard argument: see, for example, [Dav79].)

Observe that

$$\begin{aligned}
\sum_{x\neq 0, y\neq 0} \theta_a(x)\theta_b(y) &= \sum_{x,y} \theta_a(x)\theta_b(y) - \sum_y \theta_b(y) - \sum_x \theta_a(x) + 1 \\
&= [|E|\mathbf{1}\{a = 0\} - 1][|E|\mathbf{1}\{b = 0\} - 1]
\end{aligned}$$

so that

$$\sum_{(x,y)} F(x,y)\theta_a(x)\theta_b(y) = \begin{cases} |\Gamma|, & a = b = 0, \\ \frac{|\Gamma|}{|E|-1}, & a \neq 0, b \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Thus there is a non-zero eigenvalue of AA^* corresponding to each point of

$$\{(0,0)\} \cup (\mathcal{C}\backslash\{0\}) \times (\mathcal{C}\backslash\{0\}) = \mathcal{E}.$$

The ingredients in the singular value decomposition of A are the following.

- The matrix U has a column for each point of $\mathcal{E}$, with the column for (a, b) given by
$$u_{(a,b)}(x,y) = \frac{1}{|E|}\theta_a(x)\theta_b(y), \ x, y \in \mathcal{C}.$$
- The diagonal matrix L has diagonal entries $\lambda_{(a,b)}^{1/2}$, $(a, b) \in \mathcal{E}$, where
$$\lambda_{(0,0)} = |\Gamma|$$
and
$$\lambda_{(a,b)} = \frac{|\Gamma|}{|E| - 1}, \ (a, b) \in \mathcal{E}\backslash\{(0,0)\}.$$

- The matrix $V = A^*UL^{-1}$ has columns $v_{(a,b)}$, $(a,b) \in \mathcal{E}$, with

$$\begin{aligned} v_{(a,b)}(\psi) &= \sum_{x,y} \mathbf{1}\{y = \psi x\} \frac{1}{|E|} \theta_a(x)\theta_b(y)\lambda_{(a,b)}^{-1/2} \\ &= \frac{1}{|E|} \sum_z \theta_a(z)\theta_b(\psi z)\lambda_{(a,b)}^{-1/2}. \end{aligned}$$

It is now a straightforward to compute the matrices $A^\dagger$, $AA^\dagger$ and $A^\dagger A$ and check that one obtains the same objects that one gets using the method of Section 3.

5. An example: the dihedral group

In this section we compute the objects appearing in the statement of Theorem 1.1 in the case where E is the set of vertices of a regular n-gon and Γ is the group of symmetries of the n-gon (that is, Γ is the dihedral group of order $|\Gamma| = 2n$). For simplicity, we will consider the case where n is odd. The case where n is even is similar but a little messier.

A good account of the representation theory of Γ may be found in [Sim96]. The group Γ is the semidirect product of $\mathbb{Z}_n$, the group of integers modulo n, and $\mathbb{Z}_2$, the group of integers modulo 2. It will simplify matters if we think of $\mathbb{Z}_n$ as $\{0, 1, \dots, n-1\}$ and write the group operation as addition, but think of $\mathbb{Z}_2$ as $\{+1, -1\}$ and write the group operation as multiplication. We take $+1 \in \mathbb{Z}_2$ to act on $\mathbb{Z}_n$ as the identity and take $-1 \in \mathbb{Z}_2$ to act on $\mathbb{Z}_n$ via negation (that is, inversion). The group operation is given by $(a,\sigma)(b,\tau) := (a+\sigma b, \sigma\tau)$ for $(a,\sigma),(b,\tau) \in \Gamma$ with $a, b \in \mathbb{Z}_n$ and $\sigma, \tau \in \mathbb{Z}_2$.

The irreducible representations of Γ consist of:

- the (one-dimensional) trivial representation with character the constant function 1,
- the one-dimensional representation arising from the non-trivial representation of $\mathbb{Z}_2$ with character given by $(a,\sigma) \mapsto \sigma$ (where we identify $\pm 1 \in \mathbb{Z}_2$ with $\pm 1 \in \mathbb{R}$),
- $\frac{n-1}{2}$ two-dimensional representations indexed by $\ell = 1, 2, \dots, \frac{n-1}{2}$ with characters

$$\begin{aligned} &(a, +1) \mapsto 2\cos\left(\frac{2\pi a \ell}{n}\right) \\ &(a, -1) \mapsto 0. \end{aligned}$$

The group Γ acts on the set $E = \mathbb{Z}_n$ by $(a,\sigma)x := a + \sigma x$. Hence, the identity element $(0,+1)$ has n fixed points, each element of the form $(a,-1)$ has 1 fixed point, and the remaining group elements are without fixed points. Thus the trivial representation and each of the two-dimensional representations appear in the decomposition of the permutation representation into irreducibles (that is, these representations form the set $\hat{\Gamma}_+$) and the corresponding multiplicities (that is, the numbers ν_ρ) are all 1.

It follows that

$$\begin{aligned} \sum_{\rho \in \hat{\Gamma}_+} \frac{d_\rho^2}{\nu_\rho} \chi_\rho((a,\sigma)) &= 1 + \mathbf{1}\{\sigma = +1\} 8 \sum_{\ell=1}^{\frac{n-1}{2}} \cos\left(\frac{2\pi a \ell}{n}\right) \\ &= 1 + 4\,\mathbf{1}\{\sigma = +1\}[n\mathbf{1}\{a = 0\} - 1]. \end{aligned}$$

Note that for fixed $x, y \in E$, the equation $(a,\sigma)x = y$ has two solutions $(a,\sigma) = (y-x,+1)$ and $(a,\sigma) = (y+x,-1)$. In particular, if (a,σ) solves this equation and (b,τ) solves the equation $(b,\tau)u = v$ for fixed $u, v \in E$, then the possible values of $(b,\tau)(a,\sigma)^{-1}$ are of the form $(v - \tau' u - \tau' y + \sigma'\tau' x, \sigma'\tau')$ where σ' and τ' both range over $\mathbb{Z}_2$. It follows from some straightforward manipulations that the quantity appearing on the right-hand of condition (1.2) is

$$\begin{aligned}
&\frac{1}{4n^2}\sum_{w=0}^{n-1}\sum_{k=0}^{n-1}[(P(w,-x+y+w+k) + P(w,-x-y-w+k) \\
&\qquad + P(w,x-y-w+k) + P(w,x+y+w+k)] \\
&\quad + \frac{(n-1)}{n^2}\sum_{w=0}^{n-1}[(P(w,-x+y+w) + P(w,-x-y-w))] \\
&\quad - \frac{1}{n^2}\sum_{w=0}^{n-1}\sum_{k=1}^{n-1}[P(w,-x+y+w+k) + P(w,-x-y-w+k)] \\
&= \frac{1}{n}\sum_{w=0}^{n-1}[P(w,-x+y+w) + P(w,-x-y-w)] - \frac{1}{n}.
\end{aligned}$$

Similarly, the first term in the quantity on the right-hand side of condition (1.3) is, writing $\varphi = (a,\sigma)$,

$$\begin{aligned}
&\frac{1}{2n} + \frac{n-1}{n^2}\sum_{w=0}^{n-1}P(w,\sigma a + \sigma w) - \frac{1}{n^2}\sum_{w=0}^{n-1}\sum_{k=1}^{n-1}P(w,\sigma a + \sigma w + k) \\
&\quad = \frac{1}{n}\sum_{w=0}^{n-1}P(w,\sigma a + \sigma w) - \frac{1}{2n}.
\end{aligned}$$

Lastly, the second term in the quantity on the right-hand side of condition (1.3) is, again writing $\varphi = (a,\sigma)$,

$$\frac{\sigma}{2n}\sum_{k=0}^{n-1}\left[h((k,+1)) - h((k,-1))\right].$$

A consequence of this last observation is that if Q' and Q'' are two liftings of the same transition matrix P, then there exists a constant c such that $Q'((b,\tau)) = Q''((b,\tau)) + \tau c$ for all $(b,\tau) \in \Gamma$.

Acknowledgment: The author thanks Persi Diaconis, Vaughan Jones, and an anonymous referee for helpful comments.

References

[Bir46] Garrett Birkhoff. Three observations on linear algebra. *Univ. Nac. Tucumán. Revista A.*, 5:147–151, 1946.

[Dav79] Philip J. Davis. *Circulant matrices*. John Wiley & Sons, New York-Chichester-Brisbane, 1979. A Wiley-Interscience Publication, Pure and Applied Mathematics.

[Dia88] Persi Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics, Hayward, CA, 1988.

[DS87] Persi Diaconis and Mehrdad Shahshahani. Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.*, 18(1):208–218, 1987.

[HJ90] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[Ker99] Adalbert Kerber. *Applied finite group actions*. Springer-Verlag, Berlin, second edition, 1999.

[Meh77] M. L. Mehta. *Elements of matrix theory*. Hindustan Publishing Corp., Delhi, 1977.

[Pad95] Manfred Padberg. *Linear optimization and extensions*, volume 12 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1995.

[Sim96] Barry Simon. *Representations of finite and compact groups*, volume 10 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996.

E-mail address: `evans@stat.Berkeley.EDU`

Department of Statistics #3860, University of California at Berkeley, 367 Evans Hall, Berkeley, CA 94720-3860, U.S.A