THE DUAL BURNSIDE PROCESS

IVAN Z. FENG

Abstract. The Burnside process is a classical Markov chain for sampling uniformly from group orbits. We introduce the *dual Burnside process*, obtained by interchanging the roles of group elements and states. This dual chain has stationary law $\pi(g) \propto |X_g|$, is reversible, and admits a matrix factorization $Q = AB$, $K = BA$ with the classical Burnside kernel K. As a consequence the two chains share all nonzero eigenvalues and have mixing times that differ by at most one step. We further establish universal Doeblin floors, orbit- and conjugacy-class lumpings, and transfer principles between Q and K. We analyze the explicit examples of the value-permutation model $S_k \curvearrowright [k]^n$ and the coordinate-permutation model $S_n \curvearrowright [k]^n$. These results show that the dual chain provides both a conceptual mirror to the classical Burnside process and practical advantages for symmetry-aware Markov Chain Monte Carlo.

1. Introduction

1.1. **Sampling Up to Symmetry.**

Sampling combinatorial objects that carry large symmetry groups is a recurring theme in modern probability, combinatorics, and theoretical computer science. Whenever a finite group G acts on a finite set X, Burnside's lemma[1] tells us that the orbit structure of the action governs key enumerative invariants:

$$z = \frac{1}{|G|} \sum_{g \in G} |X_g|$$

where $z := |X/G|$ is the number of orbits and $X_g = \{x \in X : g \cdot x = x\}$ is the fixed-point set of g.

The Burnside process, first introduced by Jerrum [28], provides a general method for sampling orbits of group actions uniformly at random:

$$x \;\longmapsto\; g \sim \mathrm{Unif}(G_x) \;\longmapsto\; y \sim \mathrm{Unif}(X_g), \qquad \text{move to } y.$$

The chain mixes rapidly in many natural settings ($[k]^n$ with k large [32], contingency tables [12], centralizer abelian (CA) groups[2] [34]), but can also be provably slow [23].

1.2. **The Dual Perspective.**

Reversing the two-stage update of the Burnside chain yields a Markov chain on $G^* := \{g \in G : |X_g| > 0\}$. From a current $g \in G^*$, sample $x \sim \mathrm{Unif}(X_g)$ and then $h \sim \mathrm{Unif}(G_x)$; move to h. Equivalently, with $A(g,x) = \mathbf{1}_{\{x \in X_g\}}/|X_g|,\ B(x,h) = \mathbf{1}_{\{h \in G_x\}}/|G_x|$, the kernel is $Q = AB$, a $|G^*| \times |G^*|$ matrix, whereas the classical Burnside kernel $K = BA$ is $|X| \times |X|$. Because $|G^*|$ is often much smaller than $|X|$, the dual chain provides a more compact and tractable representation of the same symmetry dynamics; moreover, for every $g \in G^*$, $Q(g,g) = Q(g,e)$ (since $X_g \cap X_g = X_g$ and $X_g \cap X_e = X_g$), so diagonal entries are read directly from the e–column (with no analogous simplification for the classical kernel K).

Detailed definitions and structural relations with the classical chain K are developed in Section 3. To the best of our knowledge, this construction has not previously been explored in the literature. The dual inherits reversibility and an explicit stationary law $\pi(g) \propto |X_g|$, and it's deeply connected with behavior of the primal chain. Two special cases motivate the present study:

Date: November 17, 2025.

[1] Often attributed to Cauchy and Frobenius but popularized by Burnside; see the historical discussion in [23].

[2] Here "centralizer abelian (CA) groups" means groups in which the centralizer $C_G(g) = \{h \in G : hg = gh\}$ is abelian for every nontrivial $g \in G$.

Value–permutation model $(G = S_k,\ X = [k]^n)$**:** S_k acts by permuting the *symbols* of the k-ary cube. Equivalently: n labeled (distinguishable) balls in k unlabeled (indistinguishable) boxes; configurations are determined by the set partition of $[n]$ into at most k color classes. For $k \geq n$, Paguyo [32] proves the uniform TV bound $\sup_{u\in[k]^n} \|K^t(u,\cdot) - \pi\|_{\mathrm{TV}} \leq n\Big(1 - \frac{1}{2k}\Big)^t$, hence the mixing time $t_{\mathrm{mix}}(K;\varepsilon) = O\big(k\log(n/\varepsilon)\big)$.

Coordinate-permutation model $(G = S_n,\ X = [k]^n)$**:** S_n acts by permuting the *coordinates* of the k-ary cube. Equivalently: n unlabeled (indistinguishable) balls into k labeled (distinguishable) boxes; configurations are determined by the multiset of symbols (the histogram), equivalently by a weak k-composition $(m_1, \ldots, m_k)$ of n with $\sum_{a=1}^k m_a = n$ and $m_a \geq 0$. For fixed k, Diaconis [10] shows from the all–equal start $\|K^t(x_0,\cdot) - \pi\|_{\mathrm{TV}} \leq (1-c_k)^t$ (for some $c_k > 0$). Uniformly in x, Aldous [1] gives $\|K^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq n(1 - \frac{1}{k})^t$, so $t_{\mathrm{mix}}(K;\varepsilon) = O\big(k\log(n/\varepsilon)\big)$. For the binary case $k = 2$, Diaconis–Zhong [17] further give a closed spectral description, showing from the all–equal start $\frac{1}{4}\Big(\frac{1}{4}\Big)^t \leq \|K^t(x_0,\cdot) - \pi\|_{\mathrm{TV}} \leq 4\Big(\frac{1}{4}\Big)^t$.

1.3. Contributions and Organization.
In this paper, we:

- Develop a unified primal–dual framework for Burnside chains: define the dual process and its universal stationary law; prove the factorizations $Q = AB$, $K = BA$; establish shared nonzero spectra and eigenvector transport; derive one-step TV comparison and mixing-time equivalence; prove model-free Doeblin floors; and relate orbit and conjugacy lumpings, including an auxiliary-variable and block–flip matrix viewpoint (Section 3).
- Work out the *value–permutation* model $S_k \curvearrowright [k]^n$: explicit formulas for $Q(g,h)$ (Stirling, expectation, and coefficient forms), fixed–point–count lumping, n–independent Doeblin bounds via stabilizers, and transfer Paguyo–type TV and spectral bounds from K to Q (Section 4).
- Analyze the *coordinate–permutation* model $S_n \curvearrowright [k]^n$ (Bose–Einstein): orbit–partition formulas for K, joint-orbit formulas and coloring interpretations for Q, universal floors and n–independent bounds for fixed k, and start–specific spectral rates in the binary case via Diaconis–Zhong and Diaconis–Lin–Ram (Section 5).

Acknowledgments

I am grateful to my advisor Jason Fulman for suggesting the study of the dual Burnside process, proposing the closed formulas in the binary coordinate–permutation model, and offering thoughtful feedback throughout. I also thank J. E. Paguyo for valuable discussions and feedback that helped shape this work.

2. Background and Preliminaries

2.1. Notation and Conventions.
Throughout the paper, all sets and groups are finite.

Group actions. A (left) **action** of a finite group G on a finite set X is a map

$$\alpha :\ G \times X \longrightarrow X, \qquad (g,x) \longmapsto g \cdot x$$

satisfying $e \cdot x = x$ and $(g_1g_2)\cdot x = g_1 \cdot (g_2 \cdot x)$ for all $g_1, g_2 \in G,\ x \in X$. We suppress α from the notation and write simply $G \curvearrowright X$.

For $x \in X$ and $g \in G$ we use the standard shorthand

$$O_x := \{\, g\cdot x : g \in G\} \quad (\textbf{orbit} \text{ of } x)$$
$$G_x := \{\, g \in G : g\cdot x = x\} \quad (\textbf{stabilizer} \text{ of } x)$$
$$X_g := \{\, x \in X : g\cdot x = x\} \quad (\textbf{fixed-point set} \text{ of } g)$$

The Orbit-Stabilizer Theorem gives $|O_x| = |G|/|G_x|$.

Burnside's lemma shows that the number of orbits is

$$z = \frac{1}{|G|}\sum_{g\in G}|X_g| = \frac{1}{|G|}\sum_{x\in X}|G_x| \quad \text{(double counting the set } \{(g,x)\in G\times X:\ g\cdot x = x\}).$$

We study two main actions in this paper:

$$\begin{aligned} \textbf{(Value-permutation)} \quad & G = S_k, \quad X = [k]^n, \quad (g\cdot x)_i = g(x_i) \\ \textbf{(Coordinate-permutation)} \quad & G = S_n, \quad X = [k]^n, \quad (g\cdot x)_i = x_{g^{-1}(i)} \end{aligned}$$

Probability and Markov kernels. For a finite set S, the uniform measure is written $\mathrm{Unif}(S)$. A **Markov kernel** (or **transition matrix**) on S is a map $P : S\times S \to [0,1]$ with $\sum_{y\in S} P(x,y) = 1$ (row-summing to 1; i.e., **row–stochastic**). The **total variation (TV) distance** between two probability measures (a.k.a. probability distributions) μ,ν on S is

$$\|\mu-\nu\|_{\mathrm{TV}} \ :=\ \frac{1}{2}\sum_{s\in S}\big|\mu(s)-\nu(s)\big| \ =\ \max_{A\subseteq S}\big|\mu(A)-\nu(A)\big|.$$

Equivalently, $\|\mu-\nu\|_{\mathrm{TV}} = \frac{1}{2}\|\mu-\nu\|_1$, i.e., TV is one half of the ℓ^1 distance on probability measures.

A probability measure π on S is called **stationary** for P if $\pi P = \pi$. **Markov chain Monte Carlo (MCMC)** builds a Markov chain whose stationary distribution is the target π; run the chain long enough and states from the chain provide approximate samples from π.

Write $P_x^t := P^t(x,\cdot)$ for the x-row of the t-step kernel. Define the pointwise distance and **mixing time**

$$d_P(x,t) \ :=\ \|P_x^t - \pi\|_{\mathrm{TV}}, \qquad t_{\mathrm{mix}}(P;x,\varepsilon) \ :=\ \min\{t\ge 0:\ d_P(x,t)\le\varepsilon\}.$$

The worst–case (uniform) versions are

$$d_P(t) \ :=\ \max_{x\in S} d_P(x,t) \ =\ \max_{x\in S}\|P_x^t - \pi\|_{\mathrm{TV}}, \qquad t_{\mathrm{mix}}(P;\varepsilon) \ :=\ \max_{x\in S} t_{\mathrm{mix}}(x,\varepsilon) \ =\ \min\{t\ge 0:\ d_P(t)\le\varepsilon\}.$$

Intuitively, $d_P(t)$ is the worst–case (maximal) distance from stationarity at time t, and $t_{\mathrm{mix}}(P;\varepsilon)$ is the minimal time after which every starting state is ε–close to π. Moreover, we have

$$d_P(t+1) \ \le\ d_P(t),$$

meaning advancing the chain can only move it closer to stationarity (Exercise 4.2 in [30]).

A kernel P on a finite state space S is **reversible** with respect to π (i.e., satisfies **detailed balance**) if

$$\pi(x)P(x,y) = \pi(y)P(y,x) \qquad \forall\, x,y\in S.$$

Intuitively, it means that in stationarity, flow from state x to y equals flow from y to x. Detailed balance implies stationarity:

$$(\pi P)(y) = \sum_x \pi(x)P(x,y) = \sum_x \pi(y)P(y,x) = \pi(y)\sum_x P(y,x) = \pi(y).$$

Equivalently, reversibility is characterized by either of the following equivalent forms:

$$\Pi P = P^\top \Pi, \qquad \Pi := \mathrm{diag}(\pi), \tag{1}$$

$$\langle f, Pg\rangle_\pi = \langle Pf, g\rangle_\pi \quad \text{for all } f,g,\ \langle f,g\rangle_\pi := \sum_{x\in S} f(x)g(x)\pi(x) = f^\top \Pi\, g. \tag{2}$$

From (1) it follows that P is diagonally similar to a symmetric matrix:

$$P \ \sim\ \tilde P := \Pi^{1/2}P\Pi^{-1/2} \quad \text{with} \quad \tilde P^\top = \Pi^{-1/2}P^\top\Pi^{1/2} = \Pi^{-1/2}(\Pi P)\Pi^{-1/2} = \tilde P. \tag{3}$$

From (2), P is self-adjoint on $L^2(\pi)$; by the spectral theorem, its spectrum is real and there exists an orthonormal eigenbasis in $L^2(\pi)$.

Two states x,y **communicate** if each can reach the other; a **communicating class** is a maximal set of pairwise communicating states. A communicating class C is **closed** if $P(x,y) = 0$ for all

$x \in C$ and $y \notin C$ (once entered, it cannot be left). The chain is **irreducible** if every state can be reached from every other state (equivalently, S is a single communicating class).

The **period** $d(x)$ of a state x is

$$d(x) \;:=\; \gcd\{\, n \ge 1 : \; P^n(x,x) > 0 \,\},$$

the greatest common divisor of lengths of closed paths from x back to x with positive probability. A *state* is **aperiodic** if $d(x) = 1$, and the *chain* is **aperiodic** if all states are aperiodic. Intuitively, it means the chain does not return in a fixed rhythm: return times have no common divisor > 1, so it can come back at irregular times. In an *irreducible* chain the period is the same for every state; in particular, if some state has a self-loop $P(x,x) > 0$, then $d(x) = 1$ and the chain is aperiodic.

For reversible P, the eigenvalues of P on $L^2(\pi)$ satisfy

$$1 = \lambda_0 \ge \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_{|S|-1} \ge -1.$$

If P is *irreducible*, then $\lambda_0 = 1$ is a simple eigenvalue (and its eigenspace is spanned by the constant function). If P is *aperiodic* on every closed communicating class, then -1 cannot be an eigenvalue (indeed, $Pf = -f$ would force period 2 on some class). The **spectral gap** is $\gamma := 1 - \lambda_1$, the **absolute spectral gap** is $\gamma_* := 1 - \lambda_*$ with $\lambda_* := \max\{|\lambda| : \lambda \ne 1\} = \max\{\lambda_1, |\lambda_{|S|-1}|\}$, and the **relaxation time** is $t_{\mathrm{rel}} := 1/\gamma_*$. By (3), $\tilde{P} = \Pi^{1/2} P \Pi^{-1/2}$ is *symmetric* with $\mathrm{Spec}(\tilde{P}) = \mathrm{Spec}(P)$, and eigenfunctions f of P correspond to eigenvectors u of $\tilde{P}$ via $u = \Pi^{1/2} f$ (and $f = \Pi^{-1/2} u$). Thus, we can apply Rayleigh–Ritz (Theorem A.21 in [30], applicable to symmetric matrices) to $\tilde{P}$ with the Euclidean inner product $\langle y, z\rangle := y^\top z$, and transport it to P via $y = \Pi^{1/2} f$, giving the $L^2(\pi)$ inner product $\langle f, g\rangle_\pi := \sum_x f(x) g(x) \pi(x)$.

Proposition 2.1 (Rayleigh–Ritz for the symmetric similarity)**.** *Let $\tilde{P} = \Pi^{1/2} P \Pi^{-1/2} = \tilde{P}^\top$ have eigenpairs (λ_k, u_k) with $u_k^\top u_k = 1$ and $1 = \lambda_0 \ge \lambda_1 \ge \cdots \ge \lambda_{n-1}$. Then, for $k = 0, \dots, n-1$,*

$$\lambda_k = \max_{\substack{y \ne 0 \\ y \perp \{u_0, \dots, u_{k-1}\}}} \frac{y^\top \tilde{P} y}{y^\top y} = \min_{\substack{y \ne 0 \\ y \perp \{u_{k+1}, \dots, u_{n-1}\}}} \frac{y^\top \tilde{P} y}{y^\top y}.$$

Corollary 2.2 (Rayleigh–Ritz for P in $L^2(\pi)$)**.** *Let P be reversible w.r.t. π, with eigenpairs (λ_k, e_k) forming an $L^2(\pi)$–orthonormal basis, ordered $1 = \lambda_0 \ge \cdots \ge \lambda_{n-1}$. Then, for $k = 0, \dots, n-1$,*

$$\lambda_k = \max_{\substack{f \ne 0 \\ f \perp_\pi \{e_0, \dots, e_{k-1}\}}} \frac{\langle f, Pf\rangle_\pi}{\langle f, f\rangle_\pi} = \min_{\substack{f \ne 0 \\ f \perp_\pi \{e_{k+1}, \dots, e_{n-1}\}}} \frac{\langle f, Pf\rangle_\pi}{\langle f, f\rangle_\pi},$$

where $\langle f, g\rangle_\pi := \sum_x f(x) g(x) \pi(x)$.

Proof. Set $y := \Pi^{1/2} f$, $u_j := \Pi^{1/2} e_j$. Then

$$\frac{\langle f, Pf\rangle_\pi}{\langle f, f\rangle_\pi} = \frac{f^\top \Pi P f}{f^\top \Pi f} = \frac{y^\top \tilde{P}\, y}{y^\top y}, \qquad f \perp_\pi e_j \iff y \perp u_j.$$

Apply Proposition 2.1 and translate back via these identities. □

Remark 2.3. Obviously, for any eigenpair $Pv = \lambda v$, the Rayleigh quotient equals the eigenvalue:

$$\frac{\langle v, Pv\rangle_\pi}{\langle v, v\rangle_\pi} = \frac{\langle v, \lambda v\rangle_\pi}{\langle v, v\rangle_\pi} = \lambda.$$

Hence, if $\langle f, Pf\rangle_\pi \ge 0$ for all f, then every eigenvalue satisfies $\lambda \ge 0$.

For a reversible kernel P with stationary distribution π and start x, define the (start–specific) χ^2**–distance** at time t by

$$\chi_x^2(t) \;:=\; \sum_{y \in S} \frac{\bigl(P^t(x,y) - \pi(y)\bigr)^2}{\pi(y)} = \sum_{y \in S} \pi(y) \left(\frac{P^t(x,y)}{\pi(y)} - 1\right)^2.$$

By Cauchy–Schwarz,

$$4\,\|P_x^t - \pi\|_{\mathrm{TV}}^2 \;\le\; \chi_x^2(t). \tag{4}$$

If (λ_k, e_k) are the eigenpairs of P in $L^2(\pi)$, with $1 = \lambda_0 \ge \lambda_1 \ge \cdots \ge \lambda_{n-1}$ and $\{e_k\}$ $L^2(\pi)$–orthonormal, then the standard spectral expansion gives

$$\chi_x^2(t) \;=\; \sum_{k=1}^{n-1} \lambda_k^{2t}\, e_k(x)^2. \tag{5}$$

Thus we can bound total variation via the χ^2–distance by (4) and then control $\chi_x^2(t)$ spectrally using (5). (Lemma 12.18 in [30]).

For a reversible, irreducible, aperiodic Markov chain P, we always have $d_P(t) \le C\,\alpha^t$ for some $C > 0$ and $\alpha \in (0,1)$ (Theorem 4.9 in [30]), so $t_{\mathrm{mix}}(P;\varepsilon) \le \left\lceil \frac{\log(C/\varepsilon)}{-\log\alpha} \right\rceil \le \left\lceil \frac{1}{1-\alpha}\,\log\frac{C}{\varepsilon} \right\rceil$. In this case,

$$\lambda_* = \lim_{t\to\infty} d_P(t)^{1/t} \le \lim_{t\to\infty} C^{1/t}\alpha = \alpha. \tag{6}$$

(Corollary 12.7 in [30]). Since $\lambda_1 \le \lambda_*$, this also yields $\lambda_1 \le \alpha$. Conversely, for reversible P, if $\lambda_* \le \beta$ for some $\beta < 1$, then

$$\|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \;\le\; \frac{1}{2}\sqrt{\frac{1}{\pi(x)} - 1}\,\beta^t \;\le\; \frac{1}{2}\sqrt{\frac{1}{\pi_{\min}}}\,\beta^t, \qquad \pi_{\min} := \min_x \pi(x). \tag{7}$$

so we may take $\alpha = \beta$ (e.g., $C = \frac{1}{2}\sqrt{1/\pi_{\min}}$) (Identity (3.10), p. 209 in [4]).

We have the following two facts:

- **Uniqueness (Perron–Frobenius / Fundamental Theorem of Finite Markov Chains).** If P is *irreducible*, then it admits a *unique* stationary distribution π. Equivalently, the eigenvalue $\lambda_0 = 1$ of $P^\top$ is *simple* (algebraic multiplicity 1).
- **Convergence (Ergodic Theorem for Finite Markov Chains).** If P is *irreducible and aperiodic* (a.k.a. **ergodic** for a finite Markov chain), then for every $x \in S$,

  $$\|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \;\xrightarrow[t\to\infty]{}\; 0,$$

  equivalently, $P^t(x,y) \to \pi(y)$ for all $y \in S$. (If P is irreducible but periodic, the Cesàro averages $t^{-1}\sum_{s=0}^{t-1} P^s(x,\cdot)$ still converge to π.)

Re-indexing principle. For any function $\Phi : A \to B$ and any $F : A \to \mathbb{R}$,

$$\sum_{a\in A} F(a) \;=\; \sum_{b\in B}\; \sum_{\substack{a\in A\\ \Phi(a)=b}} F(a). \tag{8}$$

Moreover, if $F(a)$ depends only on $\Phi(a)$, say $F(a) =: G\big(\Phi(a)\big)$, then

$$\sum_{a\in A} G\big(\Phi(a)\big) \;=\; \sum_{b\in B} G(b)\,\big|\Phi^{-1}(b)\big|. \tag{9}$$

In particular, if Φ is a bijection, then $\big|\Phi^{-1}(b)\big| = 1$ here.

This is the discrete version of the pushforward (change of variables) formula and relates to the tower property in probability:

$$\mathbb{E}\big[F(X)\big] \;=\; \mathbb{E}\Big[\mathbb{E}\big[F(X) \mid \Phi(X)\big]\Big].$$

Minorization. Given a Markov kernel P on S, we say P satisfies a (t_0, δ, ν)–**minorization** if there exist $t_0 \in \mathbb{N}$, $\delta > 0$, and a probability measure ν on S such that

$$P^{t_0}(x, A) \;\geq\; \delta\, \nu(A) \qquad \forall\, x \in S,\ \forall\, A \subseteq S.$$

Proposition 2.4 (Rosenthal's bound; [36])**.** *If P satisfies a (t_0, δ, ν)–minorization, then for all $x \in S$ and $t \geq 0$,*

$$\|P_x^t - \pi\|_{\mathrm{TV}} \;\leq\; (1-\delta)^{\lfloor t/t_0 \rfloor}.$$

Consequently,

$$d_P(t) \;\leq\; (1-\delta)^{\lfloor t/t_0 \rfloor} \quad \textit{and} \quad t_{\mathrm{mix}}(P, \varepsilon) \;\leq\; t_0 \left\lceil \frac{\log(1/\varepsilon)}{-\log(1-\delta)} \right\rceil \;\leq\; t_0 \left\lceil \frac{1}{\delta}\, \log\frac{1}{\varepsilon} \right\rceil.$$

Remark 2.5. Since t_{mix} is an *integer*, the ceiling converts the real threshold $\frac{\log(1/\varepsilon)}{-\log(1-\delta)}$ into the minimal integer time that guarantees $\|P_x^t - \pi\|_{\mathrm{TV}} \leq \varepsilon$. Equivalently, we may write

$$t_{\mathrm{mix}}(P; \varepsilon) \;<\; t_0\, \frac{\log(1/\varepsilon)}{-\log(1-\delta)} \;+\; t_0,$$

and for asymptotic statements authors often omit ceilings for readability. Additionally, the ceiling inequality implies

$$t_{\mathrm{mix}}(P; \varepsilon) = O\!\left(t_0\, \frac{1}{\delta}\, \log\frac{1}{\varepsilon}\right).$$

Lumping and lumpability. Let S be a finite set and P a Markov kernel on S. A **lumping (projection)** is given by a partition $\mathcal{B} = \{B_1, \ldots, B_m\}$ (equivalently, an equivalence relation whose classes are the B_i). We say that P is **strongly lumpable** w.r.t. $\mathcal{B}$ if it satisfies **Dynkin's criterion**:

$$\sum_{y \in B_j} P(x, y) \;=\; \sum_{y \in B_j} P(x', y) \qquad \forall\, i, j,\ \forall\, x, x' \in B_i.$$

In that case the **lumped kernel** $\bar{P}$ on $\mathcal{B}$ is *well–defined* by[3]

$$\bar{P}(B_i, B_j) \;:=\; \sum_{y \in B_j} P(x, y) \qquad (x \in B_i). \tag{10}$$

Here "well–defined" means independent of the representative $x \in B_i$. Essentially, strong lumpability means P descends (via *pushforward* along the quotient map $q : X \to X^\sharp$) to a well-defined Markov kernel $\bar{P}$ on the quotient $X^\sharp := X/\!\sim$ (blocks); equivalently, the chain on $X^\sharp$ lifts (via *pullback* q^*) to block–constant evolutions on X. By contrast, **weak lumpability** allows $\bar{P}$ to be well–defined only for some specific initial distributions (e.g. at stationarity); we will use strong lumpability throughout.

Equivalently (under strong lumpability), since the quantity $\sum_{y \in B_j} P(x, y)$ is independent of the representative $x \in B_i$, we may also write the lumped transition as the *block average*

$$\bar{P}(B_i, B_j) \;=\; \frac{1}{|B_i|} \sum_{x \in B_i} \sum_{y \in B_j} P(x, y)\,. \tag{11}$$

In particular, the defining formula shows that the rows of $\bar{P}$ are stochastic (easily checked by summing over all j), and together with $\bar{\pi}(B_i) := \sum_{x \in B_i} \pi(x)$ it yields stationarity of the lumped chain.

We have $\mathrm{Spec}(\bar{P}) \subseteq \mathrm{Spec}(P)$; moreover, if g is a $\bar{P}$–eigenfunction with eigenvalue λ, its lift $g^\flat(x) := g(q(x))$ is a P–eigenfunction with eigenvalue λ. Conversely, if f is a P–eigenfunction

[3] Since B_i only depends on i, we may also write the left-hand side as $\bar{P}(i, j)$. Equivalently, writing blocks as orbit labels, we may denote $\bar{P}([x], [y])$ for $\bar{P}(B_i, B_j)$ when $x \in B_i$, $y \in B_j$.

that is *constant on blocks*, then $f^\sharp([x]) := f(x)$ is a $\bar{P}$–eigenfunction with the same eigenvalue λ. (Lemma 12.9 in [30]).

2.2. **The Classical Burnside Process.** Given the action $G \curvearrowright X$, the **Burnside process** is the Markov chain on the state space X defined by the two-stage auxiliary-variable update:

Algorithm 1 One step of Burnside process from state x

1: Pick $g \in G_x$ uniformly at random (stabilizer draw)
2: Pick $y \in X_g$ uniformly at random (fixed-point draw)
3: Move to y

The transition kernel is

$$K(x,y) \;=\; \sum_{g\in G_x\cap G_y} \frac{1}{|G_x|}\cdot\frac{1}{|X_g|} \;=\; \frac{1}{|G_x|}\sum_{g\in G_x\cap G_y}\frac{1}{|X_g|}.$$

The probability measure $\pi(x) = |G_x|/Z$ where $Z := \sum_{u\in X}|G_u|$ is *stationary* for K. By Burnside's lemma $Z = |G|z$ and orbit-stabilizer $|G| = |O_x||G_x|$, we may further simplify this to $\pi(x) = 1/(z|O_x|)$ where z is the number of orbits. Moreover, K is *reversible* with respect to π:

$$\pi(x)\,K(x,y) = \frac{|G_x|}{Z}\cdot\frac{1}{|G_x|}\sum_{g\in G_x\cap G_y}\frac{1}{|X_g|} = \frac{1}{Z}\sum_{g\in G_x\cap G_y}\frac{1}{|X_g|} = \pi(y)\,K(y,x).$$

Because the identity $e \in G$ belongs to every stabilizer G_x and fixes all of X,

$$K(x,y) \;\ge\; \frac{1}{|G_x|\,|X_e|} \;>\; 0 \qquad \forall\, x,y\in X.$$

Thus K has strictly positive entries, making it *irreducible*; the self-loop probabilities $K(x,x) > 0$ imply *aperiodicity*. Hence π is the *unique* stationary distribution by Perron–Frobenius.

2.3. **Literature Overview.** The Burnside process was introduced by Jerrum [28] in 1993 as a means of sampling combinatorial objects up to symmetry. Jerrum showed rapid mixing (polynomial in the input size) for several groups, including the coordinate-permutation action of S_n on $\{0,1\}^n$. Goldberg–Jerrum [23] later generalized this to any finite alphabet Σ with $|\Sigma| \ge 2$, showing that for every $\delta > 0$, the mixing time $t_{\text{mix}}\big(\frac{1}{3}\big) = \Omega\Big(\exp\big(n^{1/(4+\delta)}\big)\Big)$, meaning the chain on Σ^n under S_n mixes no faster than exponentially slow (torpid).

As well, Aldous–Fill [1] generalized Jerrum's chain to the action of S_n on $[k]^n$ and, via coupling, obtained $O(k\log n)$ mixing time bounds. Diaconis [10] observed connections to Bose–Einstein configurations and used minorization to prove n-independent bounds for fixed or slowly growing k. For $k = 2$, Diaconis–Zhong [17] diagonalized the *orbit–lumped* chain $\bar{K}$ by S_n–orbits: its nontrivial eigenvalues are

$$\lambda_m = \Big(\tbinom{2m}{m}/2^{2m}\Big)^2, \qquad 1 \le m \le \lfloor n/2\rfloor,$$

which yields the sharp two–sided bound, from the all–equal start x_0, $\frac{1}{4}\Big(\frac{1}{4}\Big)^\ell \le \|K^\ell(x_0,\cdot) - \pi\|_{\text{TV}} \le 4\Big(\frac{1}{4}\Big)^\ell$, independently of n. In 2025, Diaconis–Lin–Ram [14] further proved a full spectral description with precise eigenvalues and eigenvectors for the *unlumped* binary Burnside chain K ($k = 2$):

$$\operatorname{Spec}(K) = \{0\} \;\cup\; \{\lambda_m:\; 0 \le m \le \lfloor n/2\rfloor\}, \qquad \lambda_m = \Big(\tbinom{2m}{m}/2^{2m}\Big)^2,$$

with multiplicities $\text{mult}(0) = 2^{\,n-1}$ and $\text{mult}(\lambda_m) = \binom{n}{2m}$. *(Note: $\lambda_0 = 1$.)*

Chen [7] collected several action-agnostic bounds for the classical Burnside chain K: a Doeblin floor yielding the universal ceiling $d_K(n) \le (1 - 1/|G|)^n$ for any finite action $G \curvearrowright X$, an orbit-lumped coupling bound $d_{\bar{K}}(n) \le (1 - 1/|X|)^n$, and standard L^2-to-TV decay via the spectral gap and continuous-time semigroup. More generally, geometric eigenvalue-bound techniques of Diaconis–Stroock [16] and Sinclair–Jerrum [38] remain fundamental tools for bounding relaxation times of these non-local chains.

In 2021, Rahmani briefly introduced the value–permutation model and computed formulas in his thesis [33]. Paguyo [32] further analyzed this process via coupling, showing on $[k]^n$ with $k \ge n$, it is rapidly mixing: $t_{\mathrm{mix}}(\varepsilon) \le \left\lceil 2k \log\bigl(n/\varepsilon\bigr) \right\rceil$, and also provided an n–independent bound in terms of k via a minorization argument.

Other variants include the *commuting chain*, where a group acts on itself by conjugation (so in this case the dual and primal Burnside processes coincide, the stabilizer is the centralizer $G_x = C_G(x) = \{g \in G : gx = xg\}$, and the orbits are the conjugacy classes), studied by Rahmani [34] via coupling and minorization. In 2025, Diaconis–Zhong [18] and Diaconis–Howes [12] also treated this conjugation action: the former combined Burnside sampling with importance sampling to estimate orbit counts (e.g., conjugacy classes in the unitriangular group $U_n(\mathbb{F}_q)$) with provable control, and the latter developed fast (approximately uniform) samplers via lumped/"reflected" Burnside dynamics (e.g., for integer partitions and contingency tables).

Dittmer's thesis [19, Ch. 5] studies the SHM (Split–Hyper–Merge) Markov chain for sampling contingency tables and proves a series of mixing time results. This SHM Markov chain is a special case of the Burnside process for the action of the Young subgroups $H \times K$ on S_n, where $H := S_{\mathbf{a}} \cong S_{a_1} \times \cdots \times S_{a_r}$ and $K := S_{\mathbf{b}} \cong S_{b_1} \times \cdots \times S_{b_c}$ are the Young subgroups and $\mathbf{a} = (a_1, \ldots, a_r)$, $\mathbf{b} = (b_1, \ldots, b_c)$ are compositions of n with $\sum_{i=1}^r a_i = \sum_{j=1}^c b_j = n$, acting by

$$(h, k) \cdot s = h\, s\, k^{-1} \qquad (h \in H,\ k \in K,\ s \in S_n).$$

The set of orbits is $H \backslash S_n / K := \{\, HsK : \ s \in S_n \,\}$; i.e., the orbits are precisely the double cosets, which are in bijection with $r \times c$ contingency tables with fixed margins (a_i) and (b_j). See also [12, §3] for detailed discussion; note that they work with the closely related right–action convention $s \mapsto s_{h,k} = h^{-1} s k$.

Further recent developments include applications to lifted inference in AI by Holtzen et al. [25], uniform generation of unlabeled (Pólya) trees via the S_{n-1}–action on labeled (Cayley) trees $\mathscr{C}_n$ whose orbits are $\mathscr{T}_n$ by Bartholdi–Diaconis [3], a Burnside process on the flag variety for $\mathrm{GL}_n(\mathbb{F}_q)$ by Diaconis and Morton–Ferguson [15], as well as limit profiles and cutoff for the Burnside process on Sylow double cosets by Howes [27], treating the same setup as SHM above with $K = H$ a Sylow p–subgroup of S_n.

3. Main Theoretical Framework

3.1. The Dual Burnside Process.

Definition 3.1 (Dual Burnside process)**.** Given a finite group G acting on a finite set X, the **dual Burnside process** is the Markov chain on $G^* = \{g \in G : |X_g| > 0\}$ with transition kernel:

$$Q(g, h) = \sum_{x \in X_g \cap X_h} \frac{1}{|X_g|} \cdot \frac{1}{|G_x|} = \frac{1}{|X_g|} \sum_{x \in X_g \cap X_h} \frac{1}{|G_x|}.$$

Remark 3.2. We restrict to $G^* = \{g \in G : |X_g| > 0\}$ because the factor $1/|X_g|$ makes $Q(g, \cdot)$ undefined (non-stochastic) when $|X_g| = 0$ (i.e., $X_g = \varnothing$). Since $X_e = X \neq \varnothing$, we have $e \in G^*$, and in fact $g \to e \to h$ with positive probability for all $g, h \in G^*$, so the chain is irreducible (and, with self-loops, aperiodic), as shown below.

For the *primal* kernel K on X, no such restriction is needed; however, we may harmlessly replace the sum over G by a sum over G^*:

$$K(x,y) = \frac{1}{|G_x|} \sum_{g \in G_x \cap G_y} \frac{1}{|X_g|} = \frac{1}{|G_x|} \sum_{g \in (G^*)_x \cap (G^*)_y} \frac{1}{|X_g|},$$

since any g that contributes must satisfy $g \in G_x \cap G_y$, hence $x \in X_g$ and $|X_g| \geq 1$. Thus K is unchanged by restricting the summation to G^*.

The construction follows this algorithm:

Algorithm 2 One step of dual Burnside process from state g

1: Pick $x \in X_g$ uniformly at random (fixed-point set draw)
2: Pick $h \in G_x$ uniformly at random (stabilizer draw)
3: Move to h

Theorem 3.3 (Universal dual stationary law)**.** *Let a finite group G act on a finite set X, and write*

$$z \;=\; \big|X/G\big| \;=\; \frac{1}{|G|} \sum_{g \in G} |X_g|$$

(by Burnside's lemma). Then the probability distribution

$$\pi(g) \;=\; \frac{|X_g|}{z\,|G|}, \qquad g \in G^*,$$

is stationary and reversible for the dual Burnside chain on G^*. Moreover, the dual chain on G^* is irreducible and aperiodic; hence π is the unique stationary distribution.

Proof. **Normalization.** By Burnside's lemma and the earlier definition $Z = \sum_{u \in X} |G_u|$, we also have

$$Z \;=\; \sum_{g \in G} |X_g| \;=\; |G|\, z,$$

so indeed

$$\sum_{g \in G^*} \pi(g) = \frac{1}{Z} \sum_{g \in G^*} |X_g| = \frac{1}{Z} \sum_{g \in G} |X_g| = 1.$$

Stationarity and detailed balance. By definition,

$$\pi(g)\, Q(g,h) = \frac{|X_g|}{|G|\, z} \cdot \frac{1}{|X_g|} \sum_{x \in X_g \cap X_h} \frac{1}{|G_x|} = \frac{1}{|G|\, z} \sum_{x \in X_g \cap X_h} \frac{1}{|G_x|},$$

which is symmetric in g, h. Thus $\pi(g)\, Q(g,h) = \pi(h)\, Q(h,g)$, so π is reversible and hence stationary.

Irreducibility and aperiodicity. For any $g \in G^*$,

$$Q(g,e) = \frac{1}{|X_g|} \sum_{x \in X_g} \frac{1}{|G_x|} > 0, \quad Q(g,g) = \frac{1}{|X_g|} \sum_{x \in X_g} \frac{1}{|G_x|} > 0,$$

and for any $h \in G^*$,

$$Q(e,h) = \frac{1}{|X|} \sum_{x \in X_h} \frac{1}{|G_x|} > 0.$$

Hence every g reaches every h via $g \to e \to h$ with positive probability (irreducible), and each state has a self-loop (aperiodic).

Uniqueness. A finite, irreducible Markov chain has a unique stationary distribution. □

Remark 3.4. The three identities in the "irreducibility and aperiodicity" proof above imply that the *e–column*, the *e–row*, and the *diagonal* of matrix Q are all strictly positive. Indeed, as we see in the proof above, for every $g \in G^*$ we have

$$Q(g,g) = Q(g,e)$$

meaning each diagonal entry equals the corresponding entry in the e–column.

In general, since there are cases where $X_g \cap X_h = \varnothing$, there could exist zero entries in matrix Q (as we shall see in Example 4.2 in the value-permutation model in Section 4). Nonetheless, in certain models (such as the coordinate-permutation model in Section 5), we could also have $Q(g,h) > 0$ for every $g, h \in G^*$ (as we shall see in Corollary 5.14).

On the other hand, for the *primal* Burnside kernel K on X, *every* entry is strictly positive since we always have $e \in G_x \cap G_y$ and $X_e = X$.

Recall for the classical chain

$$\pi(x) \ \propto \ |G_x|.$$

Theorem 3.3 says for any finite group action $G \curvearrowright X$ the dual Burnside chain admits the stationary law

$$\pi(g) \ \propto \ |X_g|.$$

By detailed balance $\big(\pi(g)Q(g,h) = \pi(h)Q(h,g)\big)$ we immediately get the following.

Corollary 3.5 (Reversibility ratio)**.** *For the dual Burnside chain Q with stationary law $\pi(g) \propto |X_g|$, detailed balance yields*

$$\frac{Q(g,h)}{Q(h,g)} \ = \ \frac{\pi(h)}{\pi(g)} \ = \ \frac{|X_h|}{|X_g|} \qquad (g, h \in G^*).$$

In particular,

$$\frac{Q(g,e)}{Q(e,g)} \ = \ \frac{\pi(e)}{\pi(g)} \ = \ \frac{|X|}{|X_g|}.$$

This gives a quick way to populate the matrix: knowing one triangle (or any single column/row) of the matrix determines the other.

3.2. **Primal-Dual Decomposition.**

We write $\mathbf{1}_{\{\cdot\}}$ for indicator functions.

Definition 3.6 (Forward and backward legs)**.** With rows indexed by G^* and X respectively, set

$$A := P_{G\to X}(g,x) \ := \ \frac{\mathbf{1}_{\{x\in X_g\}}}{|X_g|}, \qquad B := P_{X\to G}(x,h) \ := \ \frac{\mathbf{1}_{\{h\in G_x\}}}{|G_x|}.$$

Remark 3.7 (Row–stochasticity)**.**

$$\sum_{x\in X} A(g,x) = \sum_{x} \frac{\mathbf{1}_{\{x\in X_g\}}}{|X_g|} = \frac{|X_g|}{|X_g|} = 1,$$
$$\sum_{h\in G^*} B(x,h) = \sum_{h} \frac{\mathbf{1}_{\{h\in G_x\}}}{|G_x|} = \frac{|G_x|}{|G_x|} = 1.$$

Thus A and B are row–stochastic (rows supported on X_g and G_x, respectively).

Lemma 3.8 (Factorization of the dual and primal kernels)**.** *With A and B defined above, we have*

$$Q = AB \quad \textit{and} \quad K = BA,$$

where

$$Q(g,h) \ = \ \sum_{x\in X_g\cap X_h} \frac{1}{|X_g|\,|G_x|}, \qquad K(x,y) \ = \ \sum_{g\in G_x\cap G_y} \frac{1}{|G_x|\,|X_g|}.$$

Proof. Matrix multiplication gives, for $g, h \in G^*$,

$$(AB)(g,h) = \sum_{x\in X} A(g,x)\, B(x,h) = \sum_{x\in X_g\cap X_h} \frac{1}{|X_g|\,|G_x|},$$

yielding $Q = AB$. The calculation for $K = BA$ is identical with $x, y \in X$ in place of g, h. □

This decomposition establishes the following key result on eigenvalues by Exercise 3.2.11 in Horn–Johnson [26].

Theorem 3.9 (Primal-dual spectral correspondence)**.** *The nonzero spectra of the dual kernel Q and the primal kernel K coincide, with matching algebraic multiplicities:*

$$\mathrm{Spec}_{\neq 0}(Q) = \mathrm{Spec}_{\neq 0}(K).$$

Moreover, the nonsingular parts of the Jordan canonical forms of Q and K are identical.

Let's further study the eigenvectors. For any $\lambda \neq 0$, let

$$E_\lambda(K) := \ker(K - \lambda I) \quad \text{and} \quad E_\lambda(Q) := \ker(Q - \lambda I)$$

denote the (right) eigenspaces. The following results follow from Proposition 1 in Nakatsukasa [31].

Theorem 3.10 (Eigenvectors under the intertwining $Q\,A = A\,K$)**.** *The following holds:*

(a) *If $v \in E_\lambda(K)$, then $Av \in E_\lambda(Q)$ and $Av \neq 0$.*
(b) *If $w \in E_\lambda(Q)$, then $Bw \in E_\lambda(K)$ and $Bw \neq 0$.*
(c) *The maps*

$$A:\ E_\lambda(K) \longrightarrow E_\lambda(Q) \quad \text{and} \quad B:\ E_\lambda(Q) \longrightarrow E_\lambda(K)$$

are mutually inverse up to the scalar λ:

$$A\big(Bw\big) = Qw = \lambda w \quad (w \in E_\lambda(Q)), \qquad B\big(Av\big) = Kv = \lambda v \quad (v \in E_\lambda(K)).$$

Equivalently, A is a linear isomorphism $E_\lambda(K) \xrightarrow{\cong} E_\lambda(Q)$ with inverse $\lambda^{-1}B$, and B is an isomorphism with inverse $\lambda^{-1}A$.

We may also view the results on eigenvalues and eigenvectors above in a complementary and more explicit way: incidence–SVD factorization.

Let $\Gamma \in \{0,1\}^{|G^*|\times|X|}$ be the incidence matrix with $\Gamma(g,x) := \mathbf{1}_{\{g\cdot x = x\}}$, diagonal matrices $D_G(g,g) := |X_g|$, $D_X(x,x) := |G_x|$, and normalized incidence $N := D_G^{-1/2}\,\Gamma\, D_X^{-1/2}$. Then

$$A = D_G^{-1}\,\Gamma, \qquad B = D_X^{-1}\,\Gamma^\top,$$

and

(12)
$$K = BA = D_X^{-1/2}(N^\top N)\, D_X^{1/2} \ \sim\ N^\top N =: \tilde{K}, \qquad Q = AB = D_G^{-1/2}(NN^\top)\, D_G^{1/2} \ \sim\ NN^\top =: \tilde{Q},$$

meaning K and Q are diagonally similar to the symmetric matrices $\tilde{K} = N^\top N$ and $\tilde{Q} = NN^\top$, respectively. Note that with $\Pi_X = c_X D_X$ and $\Pi_G = c_G D_G$, from (1),

$$\Pi_X^{1/2} K \Pi_X^{-1/2} = N^\top N = \tilde{K}, \qquad \Pi_G^{1/2} Q \Pi_G^{-1/2} = NN^\top = \tilde{Q},$$

so the reversibility symmetrizations of K, Q coincide with the normalized-incidence forms.

From this setup, we can easily show:

Lemma 3.11. *Both Burnside kernels K (on X) and Q (on G^*) have nonnegative eigenvalues.*

Proof. Recall that $A \succeq 0$ means *positive semidefinite*: $z^\top A z \ge 0$ for all z. Clearly $\tilde{K} = N^\top N \succeq 0$ and $\tilde{Q} = NN^\top \succeq 0$ because

$$z^\top (N^\top N) z = \|Nz\|_2^2 \ge 0, \qquad y^\top (NN^\top) y = \|N^\top y\|_2^2 \ge 0.$$

Hence all eigenvalues of these Gram matrices are real and nonnegative. By diagonal similarity (12), K and Q share the (real, nonnegative) spectra of $N^\top N$ and $NN^\top$, respectively. □

Remark 3.12 (Rayleigh-form proof)**.** Without this SVD factorization setup, we could still prove this lemma as follows. Since $\pi_K(x) \propto |G_x|$ and $\pi_Q(g) \propto |X_g|$, for any $f : X \to \mathbb{R}$ and $u : G^* \to \mathbb{R}$,

$$\langle f, Kf \rangle_{\pi_K} = \sum_{x,y} f(x) f(y) \pi_K(x) K(x,y) = \frac{1}{Z} \sum_g \frac{1}{|X_g|} \Big(\sum_{x \in X_g} f(x) \Big)^2 \ge 0,$$

$$\langle u, Qu \rangle_{\pi_Q} = \sum_{g,h} u(g) u(h) \pi_Q(g) Q(g,h) = \frac{1}{Z} \sum_x \frac{1}{|G_x|} \Big(\sum_{g:\, x \in X_g} u(g) \Big)^2 \ge 0,$$

where Z is the common normalizing constant:

$$Z \;:=\; \sum_{x \in X} |G_x| \;=\; \sum_{g \in G^*} |X_g| \;=\; |G|\, z, \quad z := |X/G|.$$

That shows that both quadratic forms are nonnegative on $L^2(\pi_K)$ and $L^2(\pi_Q)$. Hence, by Remark 2.3, all eigenvalues of K and Q are real and nonnegative.

Let $\lambda_1(P)$ denote the largest eigenvalue < 1 of a reversible kernel P, and let $\lambda_*(P) := \max\{ |\lambda| : \lambda \in \mathrm{Spec}(P),\ \lambda \ne 1 \}$. Since $\mathrm{Spec}_{\ne 0}(K) = \mathrm{Spec}_{\ne 0}(Q)$ and both spectra are nonnegative, we conclude:

Corollary 3.13 (Gaps and relaxation times coincide)**.** *For both K and Q, the absolute spectral gap equals the usual spectral gap:*

$$\lambda_*(K) = \lambda_1(K) = \lambda_*(Q) = \lambda_1(Q) =: \lambda_1.$$

Consequently,

$$\gamma(K) = \gamma(Q) = \gamma_*(K) = \gamma_*(Q) = 1 - \lambda_1, \qquad t_{\mathrm{rel}}(K) = t_{\mathrm{rel}}(Q) = \frac{1}{1 - \lambda_1}.$$

Remark 3.14. Since Q and K act on spaces of different sizes, the number of zero eigenvalues can differ. From the definitions of A, B, N in terms of Γ and the fact that left/right multiplication by an invertible diagonal matrix does not change rank, we have

$$\mathrm{rank}(A) = \mathrm{rank}(B) = \mathrm{rank}(N) = \mathrm{rank}(\Gamma).$$

Using

$$K = BA = D_X^{-1/2} (N^\top N) D_X^{1/2}, \qquad Q = AB = D_G^{-1/2} (NN^\top) D_G^{1/2},$$

it follows that

$$\mathrm{rank}(K) = \mathrm{rank}(Q) = \mathrm{rank}(N),$$

hence A, B, K, Q, N, Γ all have the same rank and

$$\mathrm{null}(K) = |X| - \mathrm{rank}(N), \qquad \mathrm{null}(Q) = |G^*| - \mathrm{rank}(N).$$

Likewise, to study eigenvectors, suppose $Nv = \sigma u$ and $N^T u = \sigma v$ with $\sigma > 0$; define

$$\phi := D_X^{-1/2} v, \qquad \psi := D_G^{-1/2} u.$$

Then we get

$$\begin{aligned}
A\phi &= D_G^{-1/2} N D_X^{1/2}\phi = D_G^{-1/2} N v = D_G^{-1/2}(\sigma u) = \sigma\,\psi,\\
B\psi &= D_X^{-1/2} N^\top D_G^{1/2}\psi = D_X^{-1/2} N^\top u = D_X^{-1/2}(\sigma v) = \sigma\,\phi,\\
K\phi &= D_X^{-1/2}(N^\top N) D_X^{1/2}\phi = D_X^{-1/2} N^\top N v = D_X^{-1/2} N^\top(\sigma u) = D_X^{-1/2}(\sigma^2 v) = \sigma^2\phi,\\
Q\psi &= D_G^{-1/2}(N N^\top) D_G^{1/2}\psi = D_G^{-1/2} N N^\top u = D_G^{-1/2} N(\sigma v) = D_G^{-1/2}(\sigma^2 u) = \sigma^2\psi,
\end{aligned}$$

so ϕ and ψ are (right) eigenvectors of K and Q with the same eigenvalue σ^2, and A/B transport them with factor σ. Conversely, from any K-eigenpair (ϕ, λ) with $\lambda > 0$, set $v := D_X^{1/2}\phi$ and $u := \lambda^{-1/2} N v$; then $Nv = \sqrt{\lambda}\, u$ and $N^\top u = \sqrt{\lambda}\, v$. Similarly, from any Q-eigenpair (ψ, λ) with $\lambda > 0$, set $u := D_G^{1/2}\psi$ and $v := \lambda^{-1/2} N^\top u$, yielding again $Nv = \sqrt{\lambda}\, u$ and $N^\top u = \sqrt{\lambda}\, v$. This explicitly realizes the eigenvector correspondence of Theorem 3.10 via the singular pairs (u, v) of N.

Finally, it's worth pointing out that we can further define the **block–flip matrix**

$$M \;=\; \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix} \;\in\; \mathbb{R}^{(|G^*|+|X|)\times(|G^*|+|X|)}.$$

A single squaring yields the diagonal decomposition

$$M^2 \;=\; \begin{pmatrix} A\,B & 0 \\ 0 & B\,A \end{pmatrix} \;=\; \begin{pmatrix} Q & 0 \\ 0 & K \end{pmatrix}.$$

Consequently, every *nonzero* eigenvalue λ of $K = B\,A$ (or, equivalently, of $Q = A\,B$) appears in $\mathrm{Spec}(M)$ as a *pair*

$$\{+\sqrt{\lambda},\; -\sqrt{\lambda}\},$$

while the geometric multiplicity of the eigenvalue 0 for M equals the sum of the nullities of K and Q.

Remark 3.15 (Zero eigenvalue multiplicity of M)**.** Indeed, observe

$$\ker M = \Big\{(x, y) :\; Bx = 0,\; Ay = 0\Big\} = \ker B \;\oplus\; \ker A,$$

so $\dim\ker M = \dim\ker B + \dim\ker A$. Since $\mathrm{rank}(A) = \mathrm{rank}(B) = \mathrm{rank}(K) = \mathrm{rank}(Q) = \mathrm{rank}(N)$ (Remark 3.14), we have

$$\dim\ker A = |X| - \mathrm{rank}(N) = \mathrm{null}(K), \qquad \dim\ker B = |G^*| - \mathrm{rank}(N) = \mathrm{null}(Q),$$

hence

$$\mathrm{null}(M) = \mathrm{null}(K) + \mathrm{null}(Q).$$

One benefit of studying this unified matrix M is that it avoids the need to multiply A and B repeatedly when locating eigenvalues of K and Q for any given n and k. Because, by Remark 3.7, every row of A and B is a probability vector, the block matrix $M = \left[\begin{smallmatrix} 0 & A \\ B & 0 \end{smallmatrix}\right]$ is itself a *stochastic kernel* on the disjoint union $E := G^* \sqcup X$. Starting at a group element $g \in G^*$ one step of M moves to $x \in X$ with probability $A(g, x)$; the next step returns to $h \in G^*$ with probability $B(x, h)$. Thus two steps of M restricted to G^* reproduce the dual kernel Q; likewise, two steps restricted to X reproduce the primal kernel K:

$$M^2\big|_{G^*\times G^*} = Q, \qquad M^2\big|_{X\times X} = K.$$

Equivalently, a single transition of Q or K corresponds to two transitions of the alternating walk governed by M, so the latter "moves twice as slowly" in time while combining the dynamics of both chains on a single enlarged state space.

The block kernel M is *irreducible* and *bipartite* (period 2): a unique stationary distribution π exists by irreducibility, but M does not mix since the laws $M^t(u,\cdot)$ oscillate between the two parts.

3.3. **Total Variation and Mixing Time Correspondence.** First, we relate the stationary distributions of K and Q by the decomposition of $K = BA$ and $Q = AB$.

Lemma 3.16 (Transfer of stationarity)**.** *Let π_K be any stationary distribution of the primal kernel $K = BA$ on X. Define*

$$\pi_Q := \pi_K B \in \mathbb{R}^{G^*}.$$

Then π_Q is stationary for Q, i.e., $\pi_Q Q = \pi_Q$, and moreover

$$\pi_K = \pi_Q A.$$

Conversely, if π_Q is stationary for Q, then $\pi_K := \pi_Q A$ is stationary for K and $\pi_K B = \pi_Q$.

Proof. Because $\pi_K K = \pi_K$ and $K = B\,A$,

$$\pi_Q Q \;=\; \pi_K B\,A\,B \;=\; \pi_K (B\,A) B \;=\; \pi_K K\,B \;=\; \pi_K B \;=\; \pi_Q,$$

so π_Q is stationary for Q. Moreover $\pi_Q A = \pi_K B\,A = \pi_K K = \pi_K$.

The reverse implication is identical: starting with $\pi_Q Q = \pi_Q$ and using $Q = A\,B$ we obtain $\pi_K := \pi_Q A$ stationary for K and $\pi_K B = \pi_Q$. □

Thus the stationary measure of either chain can be recovered from the other via the "legs" A and B, providing an alternative proof of stationarity for the dual kernel Q.

Remark 3.17 (Recovering the explicit stationary laws)**.** The transfer identities $\pi_Q = \pi_K B$ and $\pi_K = \pi_Q A$ in Lemma 3.16 recover the universal dual law $\pi_Q(g) \propto |X_g|$ and the classical law $\pi_K(x) \propto |G_x|$ with the correct normalizations.

With $B(x,g) = \mathbf{1}_{\{g\in G_x\}}/|G_x|$ and $\pi_K(x) = |G_x|/Z$ where $Z = \sum_{u\in X} |G_u| = |G|\,z$ (Burnside), we have

$$\pi_Q(g) = (\pi_K B)(g) = \sum_x \pi_K(x) B(x,g) = \sum_{x\in X_g} \frac{|G_x|}{Z}\cdot\frac{1}{|G_x|} = \frac{|X_g|}{|G|\,z}.$$

That is exactly the universal dual law in Theorem 3.3.

Conversely, with $A(g,x) = \mathbf{1}_{\{x\in X_g\}}/|X_g|$ and $\pi_Q(g) = |X_g|/(|G|\,z)$, we get

$$\pi_K(x) = (\pi_Q A)(x) = \sum_g \pi_Q(g) A(g,x) = \sum_{g\in G_x} \frac{|X_g|}{|G|\,z}\cdot\frac{1}{|X_g|} = \frac{|G_x|}{|G|\,z} = \frac{1}{z\,|O_x|},$$

matching the classical formula.

For $t \ge 1$ we have the resolvent identity

$$Q^t = A\,(BA)^{t-1}\,B = A\,K^{t-1}\,B. \tag{13}$$

We derive a pointwise version of the TV comparison between Q and K; the global bound then follows by taking a maximum over $g \in G^*$.

Theorem 3.18 (Pointwise transfer inequality)**.** *For every $g \in G^*$ and every integer $t \ge 1$,*

$$d_Q(g,t) \equiv \|Q^t(g,\cdot) - \pi_Q\|_{\mathrm{TV}} \;\le\; \max_{x\in X_g} \; \|K^{t-1}(x,\cdot) - \pi_K\|_{\mathrm{TV}}.$$

Proof. Recall $A_g = \sum_{x\in X} A_g(x)\,\delta_x$ denotes the g-th row of A, which is supported on X_g; likewise K_x^{t-1} denotes the x-th row of K^{t-1}. The resolvent identity above gives

$$Q^t(g,\cdot) = A_g\,K^{\,t-1}\,B.$$

Using $\pi_Q = \pi_K B$ (Lemma 3.16),

$$Q^t(g,\cdot) - \pi_Q \;=\; (A_g K^{\,t-1} - \pi_K)\,B.$$

Since B is row-stochastic, Exercise 4.2 of [30] (TV contraction) gives $\|\mu B - \nu B\|_{\mathrm{TV}} \le \|\mu - \nu\|_{\mathrm{TV}}$ for distributions μ, ν on X. Taking $\mu =: A_g K^{t-1}$ and $\nu =: \pi_K$,

$$\|Q^t(g,\cdot) - \pi_Q\|_{\mathrm{TV}} = \|(A_g K^{t-1} - \pi_K)B\|_{\mathrm{TV}} \le \|A_g K^{t-1} - \pi_K\|_{\mathrm{TV}}.$$

Because $A_g(x) \ge 0$ and $\sum_{x\in X_g} A_g(x) = \sum_x A_g(x) = 1$,

$$A_g K^{\,t-1} - \pi_K = \sum_{x\in X_g} A_g(x)\, K_x^{t-1} - \sum_{x\in X_g} A_g(x)\,\pi_K = \sum_{x\in X_g} A_g(x)\,(K_x^{t-1} - \pi_K).$$

Total-variation distance is a norm, so Jensen/convexity gives

$$\|A_g K^{\,t-1} - \pi_K\|_{\mathrm{TV}} \;\le\; \sum_{x\in X_g} A_g(x)\,\|K_x^{t-1} - \pi_K\|_{\mathrm{TV}} \;\le\; \max_{x\in X_g} \|K_x^{t-1} - \pi_K\|_{\mathrm{TV}}.$$

where the second inequality is due to the fact that a weighted average of nonnegative numbers is never larger than the largest of them:

$$\sum_x A_g(x)\, a_x \;\le\; \max_x a_x \quad \text{if } A_g(x) \ge 0,\; \sum_x A_g(x) = 1$$

with $a_x =: \|K_x^{t-1} - \pi_K\|_{\mathrm{TV}}$. ☐

Corollary 3.19 (TV one-step comparison)**.** *For all $t \ge 1$,*

$$d_Q(t) \;\le\; d_K(t-1).$$

Proof. Take the maximum over $g \in G^*$ in Theorem 3.18:

$$d_Q(t) = \max_g \|Q^t(g,\cdot) - \pi_Q\|_{\mathrm{TV}} \;\le\; d_K(t-1).$$ ☐

Similarly,

$$K^t \;=\; B\,Q^{\,t-1}\,A,$$

which by the symmetric argument implies for every integer $t \ge 1$:

(14) $$d_K(x,t) \equiv \|K^t(x,\cdot) - \pi_K\|_{\mathrm{TV}} \;\le\; \max_{g\in G_x} \|Q^{t-1}(g,\cdot) - \pi_Q\|_{\mathrm{TV}}, \qquad d_K(t) \;\le\; d_Q(t-1).$$

Because both inequalities hold, the two chains mix essentially at the same speed; each is at most "one step behind" the other.

Theorem 3.20 (Mixing time equivalence)**.** *The mixing times differ by at most one step: for fixed $\varepsilon \in (0,1)$,*

$$\big|t_{\mathrm{mix}}(Q;\varepsilon) - t_{\mathrm{mix}}(K;\varepsilon)\big| \le 1.$$

In particular, we have $t_{\mathrm{mix}}(Q;\varepsilon) = t_{\mathrm{mix}}(K;\varepsilon) + O(1)$ and hence $t_{\mathrm{mix}}(Q;\varepsilon) = \Theta\big(t_{\mathrm{mix}}(K;\varepsilon)\big)$.

Proof. Since $d_Q(t) \le d_K(t-1)$, the moment $d_K(t-1)$ dips to ε, one step later $d_Q(t)$ is also $\le \varepsilon$. Thus $t_{\mathrm{mix}}(Q;\varepsilon) \le t_{\mathrm{mix}}(K;\varepsilon)+1$. Symmetrically, $d_K(t) \le d_Q(t-1)$ gives $t_{\mathrm{mix}}(K;\varepsilon) \le t_{\mathrm{mix}}(Q;\varepsilon) + 1$. ☐

Corollary 3.21 (Transfer from K at x_0 to Q)**.** *Fix $x_0 \in X$, $0 < \alpha < 1$, and constants $0 < C_1 \le C_2 < \infty$ such that, for all $t \ge 0$,*

(15) $$C_1\,\alpha^t \;\le\; d_K(x_0,t) \;=\; \|K^t(x_0,\cdot) - \pi_K\|_{\mathrm{TV}} \;\le\; C_2\,\alpha^t.$$

Then the start–specific mixing time for K satisfies

$$\left\lceil \frac{\log(C_1/\varepsilon)}{-\log\alpha} \right\rceil \;\le\; t_{\mathrm{mix}}(K;x_0,\varepsilon) \;\le\; \left\lceil \frac{\log(C_2/\varepsilon)}{-\log\alpha} \right\rceil.$$

(i) **(Global lower transfer to Q.)** *For all $t \ge 0$,*

$$d_Q(t) \;\ge\; C_1\,\alpha^{\,t+1}.$$

Consequently,

$$t_{\text{mix}}(Q;\varepsilon) \;\geq\; \left\lceil \frac{\log(C_1/\varepsilon)}{-\log\alpha} \right\rceil - 1.$$

(ii) **(Conditional upper transfer to a stabilizer start.)** *Suppose there exists $h_0 \in G_{x_0}$ such that the K–upper bound holds* uniformly *on the fixed–set X_{h_0}, i.e., for all $u \geq 0$ and all $x \in X_{h_0}$,*

$$\|K^{\,u}(x,\cdot) - \pi_K\|_{\text{TV}} \;\leq\; C_2\,\alpha^{\,u}.$$

Then for all $t \geq 1$,

$$d_Q(h_0,t) \;=\; \|Q^{\,t}(h_0,\cdot) - \pi_Q\|_{\text{TV}} \;\leq\; C_2\,\alpha^{\,t-1},$$

and hence

$$t_{\text{mix}}(Q;h_0,\varepsilon) \;\leq\; \left\lceil \frac{\log(C_2/\varepsilon)}{-\log\alpha} \right\rceil + 1.$$

Proof. The bounds for $t_{\text{mix}}(K;x_0,\varepsilon)$ are immediate from the two–sided geometric control on $d_K(x_0,t)$.

For (i), by (14), $d_K(t{+}1) \leq d_Q(t)$. Hence, for all $t \geq 0$, $d_Q(t) \geq d_K(t{+}1) \geq d_K(x_0,t{+}1) \geq C_1\alpha^{t+1}$. The displayed lower bound on $t_{\text{mix}}(Q;\varepsilon)$ follows by solving $C_1\alpha^{t+1} \leq \varepsilon$ for t.

For (ii), fix $h_0 \in G_{x_0}$ satisfying the uniform bound on X_{h_0}. By Theorem 3.18, for $t \geq 1$,

$$\|Q^{\,t}(h_0,\cdot) - \pi_Q\|_{\text{TV}} \;\leq\; \max_{x\in X_{h_0}} \|K^{\,t-1}(x,\cdot) - \pi_K\|_{\text{TV}} \;\leq\; C_2\,\alpha^{\,t-1}.$$

Solving $C_2\alpha^{t-1} \leq \varepsilon$ yields the stated bound on $t_{\text{mix}}(Q;h_0,\varepsilon)$. □

Remark 3.22. The estimate (15) is pointwise in the start and does not bound $d_K(t)$; hence it does not, by itself, give a uniform bound on $d_Q(t)$ via $d_Q(t) \leq d_K(t-1)$. Theorem 3.18 remains valid for the start h_0 considered there.

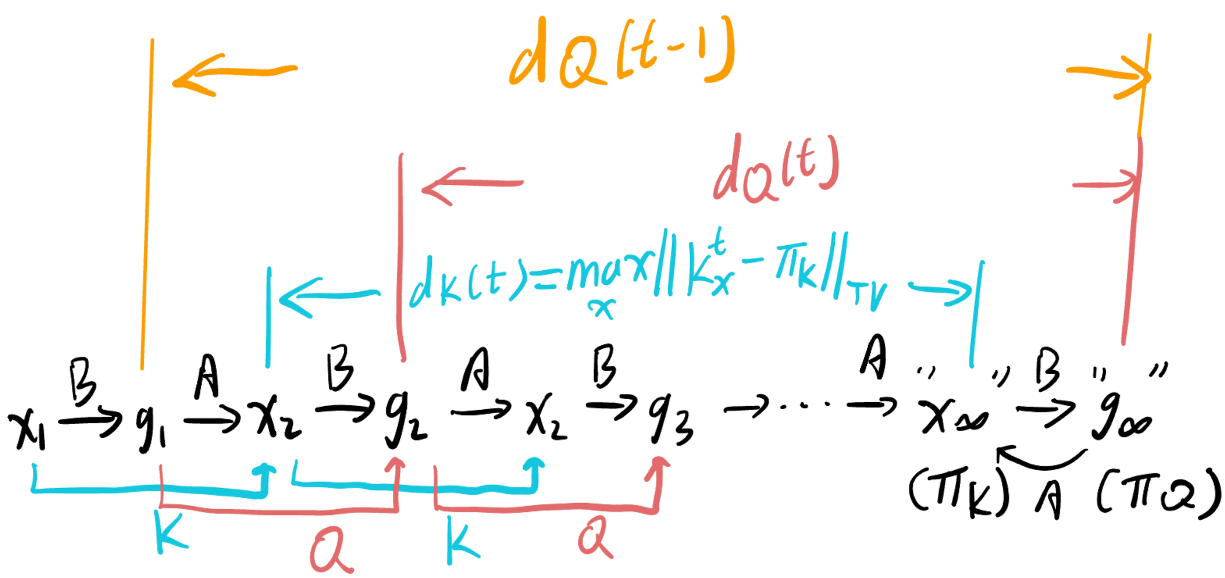

FIGURE 1. Primal–dual alternation via the legs A and B.

Finally, this picture intuitively illustrates the relationship between the primal and dual chains developed in this section. It shows the alternating–leg walk on $E = G^* \sqcup X$: black A-arrows are $G \to X$ steps and black B-arrows are $X \to G$ steps, so the two legs in X give $K = BA$ (blue arrows) and two legs in G^* give $Q = AB$ (red arrows). This yields the one–step TV lag $d_K(t) \leq d_Q(t-1)$

(as easily seen in the picture), and likewise $d_Q(t) \le d_K(t-1)$. It also links the stationary laws by $\pi_Q = \pi_K B$ and $\pi_K = \pi_Q A$.

3.4. **Universal Doeblin Floors (Minorization).** First, Proposition 10 in Section 5 of Chen's thesis [7] states:

Proposition 3.23 (Model-free floor for K; [7]). *For any finite action $G \curvearrowright X$, the Burnside kernel K satisfies*

$$K(x,\cdot) \;\ge\; \frac{1}{|G|}\,\pi(\cdot) \qquad \forall x \in X,$$

hence $d_K(t) \le (1-\frac{1}{|G|})^t$ and $t_{\mathrm{mix}}(K;\varepsilon) \le \left\lceil |G| \log(1/\varepsilon) \right\rceil$.

We now improve this and derive an analogous bound for the dual chain.

Lemma 3.24 (Uniform floors for Q and K). *Let a finite group G act on a finite set X, and set*

$$M \;:=\; \max_{u\in X} |G_u|.$$

Then:

(a) ***(Dual floor).*** *For every $g \in G^*$,*

$$Q(g,\cdot) \;\ge\; \delta_Q\,\delta_e(\cdot), \qquad \delta_Q \;:=\; \frac{1}{M}.$$

(b) ***(Primal floors).*** *For every $x \in X$ and every $y \in X$,*

$$K(x,y) \;\ge\; \frac{1}{|G_x|\,|X|} \;\ge\; \frac{1}{M\,|X|} \;=:\; \underline{\delta}_K.$$

Equivalently (Doeblin form), for every $x \in X$,

$$K(x,\cdot) \;\ge\; \delta_K\ \mathrm{Unif}(X)(\cdot), \qquad \delta_K \;:=\; \frac{1}{M}.$$

Proof. (a) Since $e \in G_x$ for all x and $X_e = X$,

$$Q(g,e) \;=\; \frac{1}{|X_g|}\sum_{x\in X_g}\frac{1}{|G_x|} \;\ge\; \frac{1}{|X_g|}\sum_{x\in X_g}\frac{1}{M} \;=\; \frac{1}{M}.$$

(b) For any x, y, the identity e lies in $G_x \cap G_y$ and $|X_e| = |X|$, so

$$K(x,y) \;\ge\; \frac{1}{|G_x|\,|X|} \;\ge\; \frac{1}{M\,|X|} = \underline{\delta}_K.$$

Summing over $y \in A \subseteq X$ gives the uniform minorization

$$K(x,A) \;\ge\; \frac{|A|}{|G_x|\,|X|} \;\ge\; \frac{|A|}{M\,|X|} = \Big(\frac{1}{M}\Big)\,\mathrm{Unif}(X)(A),$$

i.e., $K(x,\cdot) \ge \delta_K\ \mathrm{Unif}(X)(\cdot)$ with $\delta_K := 1/M$. □

We can then apply Proposition 2.4 to get the following bounds for both chains.

Theorem 3.25 (Rosenthal's bounds). *For all $t \ge 0$ and $0 < \varepsilon < 1$,*

$$d_Q(t) \;\le\; (1-M^{-1})^t, \qquad t_{\mathrm{mix}}(Q;\varepsilon) \;\le\; \left\lceil M\,\log\frac{1}{\varepsilon}\right\rceil,$$

$$d_K(t) \;\le\; (1-M^{-1})^t, \qquad t_{\mathrm{mix}}(K;\varepsilon) \;\le\; \left\lceil M\,\log\frac{1}{\varepsilon}\right\rceil.$$

For the primal chain K, because $M \le |G|$, this is never weaker than Chen's bound and is strict whenever $M < |G|$ (e.g. in $S_k \curvearrowright [k]^n$, $M = (k-1)!$ while $|G| = k!$).

While these one–step Doeblin constants provide uniform ceilings, a minorization for K can be transferred directly to Q (and conversely), though at the cost of two steps. This gives a way to recycle bounds between the two chains:

Theorem 3.26 (Two–step transfer of minorization)**.** *If $K \ge \delta\,\nu$ (row-wise on X for some probability ν), then*

$$Q^2 \;\ge\; \delta\,(\nu B) \quad \text{on } G^*,$$

hence

$$d_Q(t) \;\le\; (1-\delta)^{\lfloor t/2 \rfloor}, \qquad t_{\mathrm{mix}}(Q;\varepsilon) \;\le\; 2\left\lceil \frac{\log(1/\varepsilon)}{-\log(1-\delta)} \right\rceil \;\le\; 2\left\lceil \delta^{-1}\log\frac{1}{\varepsilon} \right\rceil.$$

Proof. Since $K \ge \delta\,\nu$ and B is row-stochastic, $(KB)(x,\cdot) \ge \delta\,(\nu B)(\cdot)$ for all $x \in X$. Since $Q^2 = AKB$,

$$Q^2(g,\cdot) = \sum_{x\in X} A(g,x)\,(KB)(x,\cdot) \;\ge\; \sum_{x\in X} A(g,x)\,\delta\,(\nu B)(\cdot) = \delta\,(\nu B)(\cdot),$$

using $\sum_x A(g,x) = 1$ (rows of A sum to 1). Apply Proposition 2.4 with $t_0 = 2$. □

3.5. **Lumping Principle.** Let P be a Markov kernel on a finite set S with stationary π. Let $\{B_i\}_{i=1}^m$ be a partition of S that satisfies the lumpability condition:

$$\sum_{y\in B_j} P(x,y) = \sum_{y\in B_j} P(x',y) \quad \forall i,j,\ \forall x,x' \in B_i,$$

so the lumped kernel $\bar P(i,j) = \sum_{y\in B_j} P(x,y)$ is *well–defined*. Define $\bar\pi(i) := \sum_{x\in B_i}\pi(x)$; then $\bar\pi$ is stationary for $\bar P$, because

(16)
$$(\bar\pi\bar P)(j) = \sum_i \bar\pi(i)\bar P(i,j) = \sum_i \sum_{x\in B_i} \pi(x) \sum_{y\in B_j} P(x,y) = \sum_{y\in B_j} \underbrace{\sum_{x\in S} \pi(x)P(x,y)}_{=\pi(y)} = \sum_{y\in B_j} \pi(y) = \bar\pi(j).$$

Lemma 3.27 (Reversible lumping; symmetry criterion)**.** *With notation as above:*

(a) ***(Reversibility transfers.)*** *$\bar P$ is reversible w.r.t. $\bar\pi$:*

$$\bar\pi(i)\,\bar P(i,j) \;=\; \bar\pi(j)\,\bar P(j,i) \qquad (1 \le i,j \le m).$$

(b) ***(Irreducibility transfers.)*** *If P is irreducible on S, then the lumped kernel $\bar P$ is irreducible on the quotient $S/\mathcal{B}$.*

(c) ***(Symmetry $\Longleftrightarrow$ uniform pushforward, under irreducibility.)*** *If $\bar P$ is irreducible, then*

$$\bar P(i,j) = \bar P(j,i) \text{ for all } i,j \quad \Longleftrightarrow \quad \bar\pi(1) = \cdots = \bar\pi(m)\,.$$

Proof (sketch). (a) $\bar\pi(i)\bar P(i,j) = \sum_{x\in B_i}\sum_{y\in B_j}\pi(x)P(x,y) = \sum_{x,y}\pi(y)P(y,x) = \bar\pi(j)\bar P(j,i)$. (b) A path in S projects to a path in $S/\mathcal{B}$ with all steps positive. (c) If $\bar\pi$ is uniform, detailed balance gives symmetry. Conversely, symmetry and $\bar P$ irreducible imply $(\bar\pi(i)-\bar\pi(j))\bar P(i,j) = 0$ for all i,j, hence $\bar\pi$ is constant along any path and thus uniform. □

Remark 3.28. Note that in (c), the direction "uniform $\bar\pi \Rightarrow \bar P$ symmetric" does *not* require irreducibility.

Recall $A(g,x) = \mathbf{1}_{\{x\in X_g\}}/|X_g|$ and $B(x,h) = \mathbf{1}_{\{h\in G_x\}}/|G_x|$, so $Q = AB$ on G^* and $K = BA$ on X.

Lemma 3.29 (Subset–aggregation identities for K and Q)**.** *For arbitrary $\Omega \subseteq X$ and $B \subseteq G^*$,*

$$(A\mathbf{1}_\Omega)(g) = \sum_{x\in\Omega} A(g,x) = \frac{|X_g \cap \Omega|}{|X_g|}, \qquad (B\mathbf{1}_B)(x) = \sum_{g\in B} B(x,g) = \frac{|G_x \cap B|}{|G_x|}, \tag{17}$$

$$\sum_{y\in\Omega} K(x,y) = (BA\,\mathbf{1}_\Omega)(x) = \frac{1}{|G_x|}\sum_{h\in G_x} \frac{|X_h\cap\Omega|}{|X_h|}, \tag{18}$$

$$\sum_{h\in B} Q(g,h) = (AB\,\mathbf{1}_B)(g) = \frac{1}{|X_g|}\sum_{u\in X_g} \frac{|G_u\cap B|}{|G_u|}. \tag{19}$$

Proof. For (17), expand the definitions:

$$(A\mathbf{1}_\Omega)(g) = \sum_x A(g,x)\mathbf{1}_\Omega(x) = \frac{1}{|X_g|}\sum_{x\in X_g} \mathbf{1}_{\{x\in\Omega\}} = \frac{|X_g\cap\Omega|}{|X_g|},$$

and similarly $(B\mathbf{1}_B)(x) = |G_x \cap B|/|G_x|$. Then (18) and (19) follow by linearity:

$$(BA\,\mathbf{1}_\Omega)(x) = \sum_h B(x,h)\,(A\mathbf{1}_\Omega)(h) = \frac{1}{|G_x|}\sum_{h\in G_x}\frac{|X_h\cap\Omega|}{|X_h|},$$

$$(AB\,\mathbf{1}_B)(g) = \sum_u A(g,u)\,(B\mathbf{1}_B)(u) = \frac{1}{|X_g|}\sum_{u\in X_g}\frac{|G_u\cap B|}{|G_u|}.$$

□

3.5.1. *TV Preserving Criterion.* We now deduce the precise criterion for when lumping preserves total variation.

Theorem 3.30 (TV preserved by lumping)**.** *Let X be finite and let $\mathcal{B} = \{B_1,\dots,B_m\}$ be a partition of X. For any probability measures μ,ν on X, let $\bar\mu,\bar\nu$ be their pushforwards to $\mathcal{B}$ (i.e., $\bar\mu(B) = \mu(B)$, $\bar\nu(B) = \nu(B)$). Then*

$$\|\bar\mu - \bar\nu\|_{\mathrm{TV}} \;\le\; \|\mu-\nu\|_{\mathrm{TV}},$$

and

$$\|\mu-\nu\|_{\mathrm{TV}} = \|\bar\mu-\bar\nu\|_{\mathrm{TV}} \quad\Longleftrightarrow\quad \text{for every } B\in\mathcal{B},\;\; \mu(x)-\nu(x) \text{ has a constant sign on } B.$$

Equivalently, the set $A^* := \{x\in X : \mu(x)\ge\nu(x)\}$ is a union of blocks of $\mathcal{B}$.

Proof. Write $\Delta(x) := \mu(x)-\nu(x)$. Then

$$\|\mu-\nu\|_{\mathrm{TV}} = \frac12\sum_{x\in X}|\Delta(x)| = \frac12\sum_{B\in\mathcal{B}}\sum_{x\in B}|\Delta(x)|, \qquad \|\bar\mu-\bar\nu\|_{\mathrm{TV}} = \frac12\sum_{B\in\mathcal{B}}\Big|\sum_{x\in B}\Delta(x)\Big|.$$

For each fixed block B we always have the triangle inequality

$$\sum_{x\in B}|\Delta(x)| \;\ge\; \Big|\sum_{x\in B}\Delta(x)\Big|,$$

with equality if and only if all $\{\Delta(x) : x\in B\}$ have the same sign (allowing zeros). Summing over B yields the stated results. The "equivalently" clause follows since $A^* = \{x : \Delta(x)\ge 0\}$ is a union of blocks exactly in that case. □

In particular, applying Theorem 3.30 to the orbit partition ($\mu = P_x^t$, $\nu=\pi$) gives[4]

$$\|P^t(x,\cdot)-\pi\|_{\mathrm{TV}} \;\ge\; \|\bar P^{\,t}([x],\cdot)-\bar\pi\|_{\mathrm{TV}}, \qquad t\ge 0.$$

[4]Recall we may write $[x]$ for the orbit O_x in the definition of $\bar P$.

Hence, for every $\varepsilon \in (0,1)$,

$$t_{\rm mix}(P;x,\varepsilon) \;\ge\; t_{\rm mix}(\bar P;[x],\varepsilon) \quad\text{and}\quad t_{\rm mix}(P;\varepsilon) \;\ge\; t_{\rm mix}(\bar P;\varepsilon),$$

since $d_P(t) \ge d_{\bar P}(t)$ for all $t \ge 0$. When the sign condition holds, lumping *preserves* total variation:

$$\|P^t(x,\cdot)-\pi\|_{\rm TV} \;=\; \|\bar P^{\,t}([x],\cdot)-\bar\pi\|_{\rm TV}, \qquad t\ge 0.$$

As a handy *sufficient condition* for equality in Theorem 3.30, it already suffices that both measures are constant on each block. We record this explicitly:

Corollary 3.31 (Block–constant measures ⇒ TV equality)**.** *Let $\mathcal B=\{B_1,\dots,B_m\}$ be a partition of a finite set X. Suppose μ,ν are* block–constant *on $\mathcal B$, i.e., for each i there exist numbers $\alpha_i,\beta_i\ge 0$ with $\sum_i\alpha_i=\sum_i\beta_i=1$ such that*

$$\mu(x)=\frac{\alpha_i}{|B_i|},\qquad \nu(x)=\frac{\beta_i}{|B_i|}\qquad (x\in B_i).$$

Then

$$\|\mu-\nu\|_{\rm TV} \;=\; \|\bar\mu-\bar\nu\|_{\rm TV} \;=\; \frac12\sum_{i=1}^m|\alpha_i-\beta_i|.$$

In particular, TV is preserved by lumping under $\mathcal B$.

Proof. For $x\in B_i$ we have $\Delta(x)=\mu(x)-\nu(x)=(\alpha_i-\beta_i)/|B_i|$, which has a constant sign on B_i. Hence Theorem 3.30 applies and gives TV equality. The common value is

$$\|\mu-\nu\|_{\rm TV}=\frac12\sum_{i=1}^m\sum_{x\in B_i}\Big|\frac{\alpha_i-\beta_i}{|B_i|}\Big|=\frac12\sum_{i=1}^m|\alpha_i-\beta_i|=\|\bar\mu-\bar\nu\|_{\rm TV}. \qquad \square$$

The next theorem identifies when the chain is lumpable and shows that orbit–invariance of the row propagates forward, making the evolved row block–constant (with respect to the orbit partition). By Corollary 3.31, this block–constancy is a sufficient condition for TV equality, so the total–variation distance matches that of the lumped chain for all $t\ge 1$.

Lemma 3.32 (Orbit–invariance propagates forward)**.** *Let a finite group G act on X, and let P be G–equivariant, i.e., $P(gx,gy)=P(x,y)$ for all $g\in G$ and $x,y\in X$. If for some $x\in X$ the one–step row is orbit–invariant,*

$$P(x,gz)=P(x,z)\qquad(\forall\, z\in X,\ \forall\, g\in G),$$

then for every $t\ge 1$ the t–step row is also orbit–invariant:

$$P^t(x,gz)=P^t(x,z)\qquad(\forall\, z\in X,\ \forall\, g\in G).$$

Proof. Recall our row notation $P_x^t(\cdot):=P^t(x,\cdot)$. We proceed by induction on t. By hypothesis the claim holds for $t=1$. Assume it holds at step t. Using G–equivariance,

$$\begin{aligned}
P_x^{t+1}(gz) &= \sum_{u\in X}P_x^t(u)\,P(u,gz)=\sum_{u\in X}P_x^t(u)\,P(g^{-1}u,z)\\
&= \sum_{v\in X}P_x^t(gv)\,P(v,z)\qquad(\text{set } v=g^{-1}u)\\
&= \sum_{v\in X}P_x^t(v)\,P(v,z)\qquad(\text{inductive hypothesis})\\
&= P_x^{t+1}(z).
\end{aligned}$$

Thus the property propagates to $t+1$, and hence to all $t\ge 1$ by induction. $\square$

Theorem 3.33 (Equivariant lumping and TV comparison)**.** *Let a finite group G act on a finite set S, and let P be a G–equivariant Markov kernel ($P(hs, ht) = P(s, t)$ for all $h \in G$, $s, t \in S$). Let $\mathcal{O}_S := S/G$ be the orbit partition and write $[x]$ for the orbit of x. Define*

$$\bar{P}([s],[t]) \ := \ \sum_{u\in[t]} P(s,u) \qquad (s \in [s]).$$

Then:

(i) ***(Strong lumpability and stationarity.)*** *$\bar{P}$ is well–defined (independent of $s \in [s]$). Moreover, if π is stationary for P, then $\bar{\pi}([s]) := \sum_{u\in[s]} \pi(u)$ is stationary for $\bar{P}$.*

(ii) ***(TV contraction.)*** *For all $t \geq 0$ and $x \in S$,*

$$\|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \ \geq \ \|\bar{P}^{\,t}([x],\cdot) - \bar{\pi}\|_{\mathrm{TV}}.$$

(iii) ***(TV equality under symmetry at the start.)*** *If π is orbit–constant and the row at x is orbit–invariant ($P(x, z) = P(x, gz)$ for all $z \in S$ and $g \in G$), then for all $t \geq 1$,*

$$\|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \ = \ \|\bar{P}^{\,t}([x],\cdot) - \bar{\pi}\|_{\mathrm{TV}}.$$

Proof. (i) Fix $[t] \in \mathcal{O}_S$. If $s' = hs \in [s]$, then

$$\begin{aligned}
\sum_{u\in[t]} P(s',u) &= \sum_{u\in[t]} P(hs,u) && \text{(definition of the sum)}\\
&= \sum_{u\in[t]} P\big(s,h^{-1}u\big) && \text{(G–equivariance of P)}\\
&= \sum_{v\in[t]} P(s,v) && \text{(setwise invariance $h^{-1}[t]=[t]$ since $[t]$ is an orbit).}
\end{aligned}$$

Thus $\bar{P}([s],[t])$ is independent of the chosen $s \in [s]$, so $\bar{P}$ is well–defined.

Moreover, stationarity follows immediately from the general identity (16), applied to the orbit partition $\mathcal{O}_S$ (with $B_i = [s]$).

(ii) As stated above, this is the inequality part of Theorem 3.30 applied to the pushforward on $\mathcal{O}_S$.

(iii) By Lemma 3.32, the row P^t_x is orbit–invariant for all $t \geq 1$:

$$P^t(x,gz) = P^t(x,z) \qquad (\forall\, z \in S,\ \forall\, g \in G),$$

so P^t_x is block–constant on the orbit partition $\mathcal{O}_S$. Since π is orbit–constant by hypothesis, both P^t_x and π are block–constant on $\mathcal{O}_S$. Hence, by Corollary 3.31, total variation is preserved by lumping:

$$\|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} = \|\bar{P}^t_x - \bar{\pi}\|_{\mathrm{TV}} = \|\bar{P}^{\,t}([x],\cdot) - \bar{\pi}\|_{\mathrm{TV}}. \qquad \square$$

3.5.2. *Specializing to Burnside Kernels.* Now, we apply this theorem to the classical and dual kernels. For the K-chain, we take $S = X$ with the left action of G on X (i.e., $x \mapsto gx$). Then $P =: K$ is G-equivariant under this action, so orbits in X lump K.

Lemma 3.34 (Conjugacy/orbit invariances)**.** *Let $G \curvearrowright X$ be a finite group action. For $x \in X$, $g \in G$, $a \in G$, and for any conjugacy class $[h] \subseteq G$ and orbit $[y] \subseteq X$, we have:*

$$u = a\cdot x \implies G_u = a\,G_x\,a^{-1}, \quad |G_u| = |G_x|, \quad |G_u \cap [h]| = |G_x \cap [h]|, \tag{20}$$

$$w = aga^{-1} \implies X_w = a\,X_g, \quad |X_w| = |X_g|, \quad |X_w \cap [y]| = |X_g \cap [y]|. \tag{21}$$

Proof. For (20), if $u = a \cdot x$ then

$$G_u = \{\, t\in G:\ t\cdot u = u \,\} = \{\, t \in G:\ t\cdot(a\cdot x) = a\cdot x \,\} = \{\, t\in G:\ a^{-1}ta\cdot x = x \,\} = a\,G_x\,a^{-1}.$$

Taking cardinalities gives $|G_u| = |G_x|$. Moreover, since conjugation by a maps G_x bijectively onto G_u and preserves the conjugacy class $[h]$, it restricts to a bijection $G_x \cap [h] \to G_u \cap [h]$, so $|G_u \cap [h]| = |G_x \cap [h]|$.

For (21), if $w = aga^{-1}$ then

$$X_w = \{\, y \in X : \ w \cdot y = y \,\} = \{\, y \in X : \ aga^{-1} \cdot y = y \,\} = \{\, y \in X : \ g \cdot (a^{-1} \cdot y) = a^{-1} \cdot y \,\} = a\, X_g.$$

Taking cardinalities gives $|X_w| = |X_g|$. Moreover, since the action of a maps X_g bijectively onto X_w and permutes the orbit $[y]$, it restricts to a bijection $X_g \cap [y] \to X_w \cap [y]$, so $|X_w \cap [y]| = |X_g \cap [y]|$. □

Corollary 3.35 (Orbits lump K and TV comparison)**.** *Let $G \curvearrowright X$ and let K be the Burnside kernel. Let $\mathcal{O} := X/G$ be the set of G–orbits and write $[x]$ for the orbit of x. Define*

$$\bar{K}([x],[y]) \ := \ \sum_{u \in [y]} K(x,u) \qquad (x \in [x]).$$

Then:

(i) ***(Strong lumpability and stationarity.)*** *$\bar{K}$ is well–defined and the pushforward $\bar{\pi}_K([x]) := \sum_{u \in [x]} \pi_K(u)$ is stationary for $\bar{K}$. Moreover, π_K is orbit–constant and $\bar{\pi}_K$ is uniform over orbits:*

$$\pi_K(x) = \frac{1}{z\,|[x]|} \quad (x \in X), \qquad \bar{\pi}_K([x]) = \frac{1}{z} \quad ([x] \in \mathcal{O}).$$

Moreover, $\bar{K}$ is reversible, irreducible, and symmetric:

$$\bar{K}([x],[y]) = \bar{K}([y],[x]) \qquad \textit{for all } [x],[y] \in \mathcal{O}.$$

(ii) ***(TV contraction.)*** *For all $t \geq 0$ and $x \in X$,*

$$\|K^t(x,\cdot) - \pi_K\|_{\mathrm{TV}} \ \geq \ \|\bar{K}^t([x],\cdot) - \bar{\pi}_K\|_{\mathrm{TV}}.$$

(iii) ***(TV equality under symmetry.)*** *If the start x is orbit–invariant ($K(x,z) = K(x,gz)$ for all $z \in X$ and $g \in G$), then for all $t \geq 1$,*

$$\|K^t(x,\cdot) - \pi_K\|_{\mathrm{TV}} \ = \ \|\bar{K}^t([x],\cdot) - \bar{\pi}_K\|_{\mathrm{TV}}.$$

In particular, if G is abelian, every row is orbit–invariant ($K(x,z) = K(x,gz)$ for every $x \in X$); hence TV equality holds for all starts.

Proof. *(i)* K is G–equivariant: $K(hx,hy) = K(x,y)$ for all $h \in G$. Indeed, Lemma 3.34 shows $G_{hx} = hG_xh^{-1}$ and $G_{hy} = hG_yh^{-1}$, and $|X_{hgh^{-1}}| = |X_g|$, so

$$\begin{aligned}
K(hx,hy) &= \frac{1}{|G_{hx}|} \sum_{g \in G_{hx} \cap G_{hy}} \frac{1}{|X_g|} && \text{(definition of } K\text{)} \\
&= \frac{1}{|G_{hx}|} \sum_{g' \in G_x \cap G_y} \frac{1}{|X_{hg'h^{-1}}|} && \text{(change variable } g = hg'h^{-1};\ G_{hx} \cap G_{hy} = h(G_x \cap G_y)h^{-1}\text{)} \\
&= \frac{1}{|G_x|} \sum_{g' \in G_x \cap G_y} \frac{1}{|X_{g'}|} && \text{(conjugation invariances: } |G_{hx}| = |G_x|,\ |X_{hg'h^{-1}}| = |X_{g'}|\text{)} \\
&= K(x,y) && \text{(definition of } K\text{).}
\end{aligned}$$

Thus by Theorem 3.33(i), $\bar{K}$ is well–defined, independent of the chosen $x \in [x]$.

Since $\pi_K(x) = \dfrac{1}{z\,|[x]|}$ and $|[x]| = |[s]|$ for all $x \in [s]$, the lumped stationary is

$$\bar{\pi}_K([s]) = \sum_{x \in [s]} \pi_K(x) = \sum_{x \in [s]} \frac{1}{z\,|[s]|} = \frac{|[s]|}{z\,|[s]|} = \frac{1}{z},$$

so $\bar{\pi}_K$ is uniform over orbits.

By Lemma 3.27 with $P =: K$, reversibility and irreducibility both transfer to $\bar{K}$; with uniform $\bar{\pi}_K$ this yields symmetry of $\bar{K}$.

(ii) This follows directly from Theorem 3.33 (ii) with $P =: K$: pushing forward to the orbit space cannot increase total variation.

(iii) By (i), π_K is orbit–constant and its pushforward $\bar{\pi}_K$ is uniform over orbits, so Theorem 3.33(iii) with $P =: K$ yields the desired TV equality for all $t \geq 1$.

If G is abelian, then $G_{gz} = gG_z g^{-1} = G_z$, so

$$K(x, gz) = \frac{1}{|G_x|} \sum_{h \in G_x \cap G_{gz}} \frac{1}{|X_h|} = \frac{1}{|G_x|} \sum_{h \in G_x \cap G_z} \frac{1}{|X_h|} = K(x, z),$$

i.e., every row is orbit–invariant, so the equality holds for all starts. □

Remark 3.36. Here (iii) recovers Proposition 9 and the statement preceding Corollary 1, Section 5 of Chen [7].

Lumping K by G–orbits is standard (see, e.g., Diaconis–Zhong [17] and Diaconis [10] for coordinate–permuting actions, and Paguyo [32] for value–permuting). Working with the orbit chain $\bar{K}$ simplifies analysis: as shown above, when the start is orbit–invariant (in particular if G is abelian) we have

$$\|K^t(x, \cdot) - \pi_K\|_{\mathrm{TV}} = \|\bar{K}^t([x], \cdot) - \bar{\pi}_K\|_{\mathrm{TV}},$$

so mixing of K is captured exactly by $\bar{K}$.

Example 3.37.

- *Value–permutation model* ($S_k \curvearrowright [k]^n$). If x uses all k symbols, then $G_x = \{e\}$ and the one–step row is flat:
$$K(x, \cdot) \equiv \frac{1}{k^n},$$
hence orbit–invariant. So by Corollary 3.35(iii),
$$\|K^t(x, \cdot) - \pi_K\|_{\mathrm{TV}} = \|\bar{K}^t([x], \cdot) - \bar{\pi}_K\|_{\mathrm{TV}} \qquad (t \geq 1).$$
- *Coordinate–permutation model* ($S_n \curvearrowright [k]^n$). If x is an all–equal word (e.g. 0^n), then $G_x = S_n$, so for any z
$$K(x, z) = \sum_{g \in G_x \cap G_z} \frac{1}{|G_x|\,|X_g|} = \frac{1}{n!} \sum_{g \in G_z} \frac{1}{|X_g|},$$
which, by G–equivariance of K, satisfies $K(x, \sigma z) = K(x, z)$ for all $\sigma \in S_n$; hence it depends only on the S_n–orbit of z. Thus the one–step row $K(x, \cdot)$ is orbit–invariant. So by Corollary 3.35(iii),
$$\|K^t(x, \cdot) - \pi_K\|_{\mathrm{TV}} = \|\bar{K}^t([x], \cdot) - \bar{\pi}_K\|_{\mathrm{TV}} \qquad (t \geq 1).$$
More generally, if the acting group is abelian, every row is orbit–invariant and TV equality holds for all starts (for all $t \geq 1$).

In addition to the model–free Doeblin floor of Proposition 3.23 (with group size $|G|$), Chen [7] (Proposition 11, Section 5) also proves the following result for the lumped chain via a coupling argument:

Proposition 3.38 (Chen's coupling bound; [7])**.** *For every start* $[x]$,

$$d_{\bar{K}}([x], t) \;=\; \|\bar{K}^t([x], \cdot) - \bar{\pi}_K\|_{\mathrm{TV}} \;\leq\; d_{\bar{K}}(t) \;\leq\; \Big(1 - \frac{1}{|X|}\Big)^t, \qquad t_{\mathrm{mix}}\big(\bar{K}; \varepsilon\big) \;\leq\; \Big\lceil |X| \, \log \frac{1}{\varepsilon} \Big\rceil.$$

Moreover, whenever total variation is preserved by orbit lumping at time t (e.g. under the hypothesis of Corollary 3.35 (iii)), the same bound holds for K at that time:

$$\|K^t(x,\cdot)-\pi_K\|_{\mathrm{TV}} \;=\; \|\bar{K}^t([x],\cdot)-\bar{\pi}_K\|_{\mathrm{TV}} \;\le\; \Big(1-\frac{1}{|X|}\Big)^t.$$

By Lemma 3.29 (18) with $\Omega =: [y]$, we obtain the orbit–lumped formula for K:

$$\bar{K}([x],[y]) = \sum_{u\in[y]} K(x,u) = \frac{1}{|G_x|}\sum_{h\in G_x}\frac{|X_h\cap[y]|}{|X_h|} \qquad (x\in[x]). \tag{22}$$

Likewise, for the Q–chain, take $S=G^*$ with the conjugation action $a\cdot g = aga^{-1}$ in Theorem 3.33. Then the following corollary shows that $P =: Q$ is equivariant under conjugation, so conjugacy classes lump Q.

Corollary 3.39 (Conjugacy classes lump Q and TV comparison)**.** *Let $G \curvearrowright G^*$ by conjugation and let Q be the dual kernel. Let $\mathcal{C} := \mathrm{Conj}(G)$ be the set of conjugacy classes in G and write $[g]$ for the conjugacy class of g. Define*

$$\bar{Q}([g],[h]) := \sum_{u\in[h]} Q(g,u) \qquad (g\in[g]).$$

Then:

(i) ***(Strong lumpability and stationarity.)*** *$\bar{Q}$ is well–defined and the pushforward $\bar{\pi}_Q([g]) := \sum_{u\in[g]}\pi_Q(u)$ is stationary for $\bar{Q}$. π_Q is class–constant and*

$$\bar{\pi}_Q([g]) = \frac{|[g]|\,|X_g|}{|G|\,z} \qquad ([g]\in\mathcal{C}).$$

Moreover, $\bar{Q}$ is reversible (with respect to $\bar{\pi}_Q$) and irreducible.

(ii) ***(TV contraction.)*** *For all $t\ge 0$ and $g\in G^*$,*

$$\|Q^t(g,\cdot)-\pi_Q\|_{\mathrm{TV}} \;\ge\; \|\bar{Q}^t([g],\cdot)-\bar{\pi}_Q\|_{\mathrm{TV}}.$$

(iii) ***(TV equality under symmetry.)*** *If the start g has a class–invariant row ($Q(g,h) = Q(g,aha^{-1})$ for all $h\in G^*$ and $a\in G$), then for all $t\ge 1$,*

$$\|Q^t(g,\cdot)-\pi_Q\|_{\mathrm{TV}} \;=\; \|\bar{Q}^t([g],\cdot)-\bar{\pi}_Q\|_{\mathrm{TV}}.$$

In particular, if $g\in Z(G) := \{g\in G : gh = hg \text{ for all } h\in G\}$ (central start), class–invariance ($Q(g,aha^{-1}) = Q(g,h)$ for every $h\in G^*$ and $a\in G$) holds and the TV equality applies to that start; and if G is abelian (conjugacy is trivial), the equality holds tautologically for all starts and times.

Proof. *(i)* Q is conjugacy–equivariant: $Q(aga^{-1},aha^{-1}) = Q(g,h)$ for all $a\in G$. Indeed, by Lemma 3.34,

$$\begin{aligned}
Q(aga^{-1},aha^{-1}) &= \frac{1}{|X_{aga^{-1}}|}\sum_{x\in X_{aga^{-1}}\cap X_{aha^{-1}}}\frac{1}{|G_x|} && \text{(definition of } Q\text{)}\\
&= \frac{1}{|X_{aga^{-1}}|}\sum_{x'\in X_g\cap X_h}\frac{1}{|G_{ax'}|} && \text{(change variable } x = ax';\ X_{aga^{-1}} = aX_g,\ X_{aha^{-1}} = aX_h\text{)}\\
&= \frac{1}{|X_g|}\sum_{x'\in X_g\cap X_h}\frac{1}{|G_{x'}|} && \text{(conjugation invariances: } |X_{aga^{-1}}| = |X_g|,\ |G_{ax'}| = |G_{x'}|\text{)}\\
&= Q(g,h) && \text{(definition of } Q\text{).}
\end{aligned}$$

Thus by Theorem 3.33(i), $\bar{Q}$ is well–defined, independent of the chosen $g\in[g]$.

Since $\pi_Q(g) = \dfrac{|X_g|}{z\,|G|}$ and $|X_g|$ is a class function ($|X_{aga^{-1}}| = |X_g|$, meaning $|X_g|$ is constant on each conjugacy class), π_Q is constant on each class. Thus

$$\bar{\pi}_Q([g]) = \sum_{u\in[g]} \pi_Q(u) = \frac{|[g]|\cdot|X_g|}{z\,|G|},$$

which is the stationary law of the lumped chain.

By Lemma 3.27 with $P =: Q$, reversibility and irreducibility (obtained in Theorem 3.3) both transfer to $\bar{Q}$.

(ii) This follows directly from Theorem 3.33 (ii) with $P =: Q$: pushing forward to conjugacy classes cannot increase total variation.

(iii) By (i), π_Q is class–constant and its pushforward $\bar{\pi}_Q$ is stationary for $\bar{Q}$, so Theorem 3.33 (iii) with $P =: Q$ yields the desired TV equality for all $t \geq 1$ whenever the start g has a class–invariant row.

In particular, for $g \in Z(G)$, since $X_{aha^{-1}} = aX_h$, $X_{aga^{-1}} = X_g$ (central g), and $|G_{ax}| = |G_x|$, the change of variables $x = ay$ bijects $X_g \cap aX_h$ with $X_g \cap X_h$, giving

$$Q(g, aha^{-1}) = \frac{1}{|X_g|}\sum_{x\in X_g\cap aX_h}\frac{1}{|G_x|} = \frac{1}{|X_g|}\sum_{y\in X_g\cap X_h}\frac{1}{|G_y|} = Q(g,h).$$

So the hypothesis holds. If G is abelian, conjugacy is trivial and $\bar{Q} = Q$, so the equality holds tautologically for all starts and times. □

Remark 3.40 (Class invariance at the identity for Q)**.** For any finite action $G \curvearrowright X$,

$$Q(e,h) \;=\; \frac{1}{|X|}\sum_{x\in X_h}\frac{1}{|G_x|},$$

which, by conjugacy–equivariance of Q, satisfies

$$Q(e, aha^{-1}) = Q(aea^{-1}, aha^{-1}) = Q(e,h) \qquad (a \in G).$$

Hence $Q(e,\cdot)$ depends only on the conjugacy class of h; in particular, the one–step row $Q(e,\cdot)$ is class–invariant.

Example 3.41.

- *Value–permutation model* ($S_k \curvearrowright [k]^n$; dual chain Q). If $g = e$, then, by Remark 3.40, $Q(e,\cdot)$ depends only on the conjugacy class of h (indeed only on $f(h) = |\operatorname{Fix}(h)|$, see Section 4), so the one–step row is class-invariant. Hence, by Corollary 3.39(iii),

$$\|Q^t(e,\cdot) - \pi_Q\|_{\mathrm{TV}} = \|\bar{Q}^t([e],\cdot) - \bar{\pi}_Q\|_{\mathrm{TV}} \qquad (t \geq 1).$$

- *Coordinate–permutation model* ($S_n \curvearrowright [k]^n$). 1. If $g = e$, as above, the one–step row $Q(e,\cdot)$ is class–invariant.

  2. If g is an n–cycle, then the one–step row is flat (see Example 5.12 in Section 5):

$$Q(g,\cdot) \;\equiv\; \frac{1}{n!},$$

  hence class–invariant.

  In both cases, by Corollary 3.39(iii),

$$\|Q^t(g,\cdot) - \pi_Q\|_{\mathrm{TV}} = \|\bar{Q}^t([g],\cdot) - \bar{\pi}_Q\|_{\mathrm{TV}} \qquad (t \geq 1).$$

  If G is abelian, conjugacy is trivial and TV equality holds tautologically for all starts and all $t \geq 0$.

By Lemma 3.29 (19) with $B =: [h]$, we obtain the class–lumped formula for Q:

$$\bar{Q}([g],[h]) = \sum_{u\in[h]} Q(g,u) = \frac{1}{|X_g|}\sum_{x\in X_g}\frac{|G_x\cap[h]|}{|G_x|} \qquad (g\in[g]). \tag{23}$$

3.5.3. *Auxiliary-Variable Scheme for Lumped Kernels.* We first point out the **general AV (auxiliary–variable) scheme**: given spaces X (target) and Y (auxiliary), and row–stochastic "legs" $A : X \to Y$ and $B : Y \to X$, set

$$K := BA \;\text{ on } X, \qquad Q := AB \;\text{ on } Y.$$

The Burnside process is the special case with X the object space and Y the auxiliary (e.g., G^*), so all AV results below apply verbatim by taking Y in place of G^*. In particular, the *transfer of stationarity* (Lemma 3.16), the *primal–dual spectral correspondence* and *eigenvector transport* (Theorems 3.9 and 3.10), the *one–step TV lag* and *mixing-time equivalence* (Theorem 3.18, Corollary 3.19, and Theorem 3.20), the *two–step minorization transfer* (Theorem 3.26), and analysis of the *block–flip matrix* M all hold directly in this AV setting. A classical example of the AV paradigm is *Gibbs sampling*; see Diaconis–Khare–Saloff-Coste [13]. Another example is the *twisted* Burnside process introduced by Diaconis–Zhong [17], which allows nonuniform draws and fits the AV paradigm with legs

$$A_v(g,x) = \frac{v(x)\,\mathbf{1}_{\{x\in X_g\}}}{\sum_{u\in X_g} v(u)}, \qquad B_w(x,h) = \frac{w(h)\,\mathbf{1}_{\{h\in G_x\}}}{\sum_{u\in G_x} w(u)}.$$

Let $\mathcal{C} := \mathrm{Conj}(G)$ be the set of conjugacy classes in G and $\mathcal{O} := X/G$ the set of orbits in X. Define the coarse (row–stochastic) legs:

Definition 3.42 (Coarse legs on classes and orbits)**.**

$$\widehat{A} : \mathcal{C}\times\mathcal{O} \to [0,1], \qquad \widehat{A}([g],[x]) \;:=\; \frac{|X_g\cap[x]|}{|X_g|},$$

$$\widehat{B} : \mathcal{O}\times\mathcal{C} \to [0,1], \qquad \widehat{B}([x],[g]) \;:=\; \frac{|G_x\cap[g]|}{|G_x|}.$$

Remark 3.43 (Row–stochasticity of coarse legs)**.** For each $[g]\in\mathcal{C}$,

$$\sum_{[x]\in\mathcal{O}} \widehat{A}([g],[x]) = \frac{1}{|X_g|}\sum_{[x]}|X_g\cap[x]| = \frac{|X_g|}{|X_g|} = 1,$$

and for each $[x]\in\mathcal{O}$,

$$\sum_{[g]\in\mathcal{C}} \widehat{B}([x],[g]) = \frac{1}{|G_x|}\sum_{[g]}|G_x\cap[g]| = \frac{|G_x|}{|G_x|} = 1.$$

Thus $\widehat{A}$ and $\widehat{B}$ are row–stochastic.

Lemma 3.44 (Coarse AB/BA factorization)**.** *We have*

$$\bar{Q} = \widehat{A}\,\widehat{B} \;\text{ on } \mathcal{C}, \qquad \bar{K} = \widehat{B}\,\widehat{A} \;\text{ on } \mathcal{O}.$$

Proof. (a) $\bar{Q} = \widehat{A}\,\widehat{B}$*.* For $[g],[h] \in \mathcal{C}$,

$$\begin{aligned}
(\widehat{A}\widehat{B})([g],[h]) &= \sum_{[x]\in\mathcal{O}} \frac{|X_g \cap [x]|}{|X_g|}\cdot\frac{|G_x\cap[h]|}{|G_x|} \\
&= \frac{1}{|X_g|}\sum_{[x]\in\mathcal{O}}\sum_{u\in[x]} \mathbf{1}_{\{u\in X_g\}}\cdot\frac{|G_x\cap[h]|}{|G_x|} \\
&\overset{(20)}{=} \frac{1}{|X_g|}\sum_{[x]\in\mathcal{O}}\sum_{u\in[x]} \mathbf{1}_{\{u\in X_g\}}\cdot\frac{|G_u\cap[h]|}{|G_u|} \\
&= \frac{1}{|X_g|}\sum_{u\in X} \mathbf{1}_{\{u\in X_g\}}\frac{|G_u\cap[h]|}{|G_u|} = \frac{1}{|X_g|}\sum_{u\in X_g}\frac{|G_u\cap[h]|}{|G_u|} \\
&= \bar{Q}([g],[h]).
\end{aligned}$$

(b) $\bar{K} = \widehat{B}\,\widehat{A}$*.* For $[x],[y] \in \mathcal{O}$,

$$\begin{aligned}
(\widehat{B}\widehat{A})([x],[y]) &= \sum_{[g]\in\mathcal{C}} \frac{|G_x \cap [g]|}{|G_x|}\cdot\frac{|X_g\cap[y]|}{|X_g|} \\
&= \frac{1}{|G_x|}\sum_{[g]\in\mathcal{C}}\sum_{w\in[g]} \mathbf{1}_{\{w\in G_x\}}\cdot\frac{|X_g\cap[y]|}{|X_g|} \\
&\overset{(21)}{=} \frac{1}{|G_x|}\sum_{[g]\in\mathcal{C}}\sum_{w\in[g]} \mathbf{1}_{\{w\in G_x\}}\cdot\frac{|X_w\cap[y]|}{|X_w|} \\
&= \frac{1}{|G_x|}\sum_{w\in G} \mathbf{1}_{\{w\in G_x\}}\frac{|X_w\cap[y]|}{|X_w|} = \frac{1}{|G_x|}\sum_{w\in G_x}\frac{|X_w\cap[y]|}{|X_w|} \\
&= \bar{K}([x],[y]).
\end{aligned}$$

□

That is, the lumped chains $\bar{Q}$ on conjugacy classes and $\bar{K}$ on orbits inherit the same primal–dual factorization pattern as Q and K. Therefore, the *general AV scheme* consequences listed above (stationarity transfer, shared nonzero spectrum/eigenvectors, TV one–step lag with mixing-time equivalence, minorization transfer, and block–flip matrix analysis) carry over verbatim to $(\bar{Q},\bar{K})$.

3.5.4. *Other Lumping Schemes.* Finally, we point out that another useful way to lump Q is by the number of fixed points. Let

$$\mathcal{B} = \{B_s :\ s = 1,2,\dots,|X|\}, \qquad B_s := \{g \in G^* :\ |X_g| = s\}.$$

This partition is *coarser* than conjugacy classes (since $|X_g|$ is a class function), but *strong lumpability is not automatic*: it requires that, for each s,

$$\sum_{h\in B_s} Q(g,h) \ \text{ depends only on } \ |X_g|, \ \text{ not on the finer class of } g.$$

Whenever lumpability holds, the lumped chain is well defined, and TV *contracts* under the push-forward; TV is *not* generally preserved unless the block–constancy criterion in Theorem 3.30 (e.g. Cor. 3.31) applies.

For example, in the *value–permutation* model (Section 4), the partition by $|X_g|$ is *strongly lumpable* for Q, so the lumped chain on $\{B_s\}$ is well defined; for a fixed g, the row $Q(g,\cdot)$ need not be block–constant on the B_s, so TV is *not* preserved in general. We will explore this lumping in detail. On the other hand, in the *coordinate–permutation* model (Section 5), $|X_h| = k^{\,c(h)}$, so

B_s corresponds to fixing $c(h)$; nevertheless this lumping is *not* strongly lumpable in general, because $Q(g,h)$ depends on the joint orbit structure of $\langle g,h\rangle$, not just on $c(h)$. The following is a counterexample.

Example 3.45 (Cycle–count lumping fails for Q in the *coordinate–permutation* model with $k=2,\ n=4$)**.** Let $X=\{0,1\}^4$ and $G=S_4$. Take $g_1=(12)(34)$ and $g_2=(123)$; then $c(g_1)=c(g_2)=2$. Let $\mathcal{B}:=\{h\in S_4:\ c(h)=2\}$ (the union of the 3-cycles and the double transpositions). A direct calculation gives

$$\sum_{h\in\mathcal{B}} Q(g_1,h)=\frac{17}{48}\qquad\text{but}\qquad\sum_{h\in\mathcal{B}} Q(g_2,h)=\frac{19}{48}.$$

Hence the block–sums over $\{h:\ c(h)=2\}$ differ for g_1 and g_2, so the partition of G by the cycle count $c(h)$ (equivalently, by $|X_h|=2^{c(h)}$) is *not* strongly lumpable for Q in this model.

4. The Value-Permutation Model

4.1. **Setting and Basic Properties.** Given integers $n\ge 1$ and $k\ge 1$, we consider $G=S_k$ acting on $X=[k]^n$ by permuting alphabet symbols:

$$(g\cdot x)_i=g(x_i)\quad(g\in S_k,\ x\in[k]^n,\ i\in[n]).$$

Since S_k acts by relabeling symbols, define the **fixed–point set of symbols**

$$\mathrm{Fix}(g):=\{\,a\in[k]:\ g(a)=a\,\},\qquad f(g):=|\,\mathrm{Fix}(g)|=c_1(g),$$

where $c_j(g)$ denotes the number of j–cycles in the cycle decomposition of g. For a word $x\in[k]^n$, define $\mathrm{supp}(x):=\{x_1,\dots,x_n\}$ to be the set of symbols used by x, and the **support size** (number of distinct values)

$$r_x:=\big|\mathrm{supp}(x)\big|=\big|\{x_1,\dots,x_n\}\big|.$$

Key facts.

- $X_g=\mathrm{Fix}(g)^n$ and $|X_g|=f(g)^n$, since $x\in X_g$ if and only if each coordinate $x_i\in\mathrm{Fix}(g)$.
- $G_x\cong S_{k-r_x}$ and $|G_x|=(k-r_x)!$, since the stabilizer of x fixes the used symbols pointwise and permutes the $k-r_x$ unused symbols freely.
- Since the action only relabels symbols, it preserves which coordinates are equal: two words $x,y\in[k]^n$ are in the same orbit if and only if they induce the same set partition of $[n]$ into value–classes, and the orbit count is $z=|X/G|=\displaystyle\sum_{r=0}^{\min\{k,n\}} S(n,r)$ (partitions of $[n]$ into $\le k$ unlabeled blocks), where $S(n,r)$ is the Stirling number of the second kind.
- The dual stationary law is (as detailed in Section 4.3)

$$\pi(g)=\frac{|X_g|}{|G|\,z}=\frac{f(g)^n}{k!\,\sum_{r=0}^{\min\{k,n\}} S(n,r)}.$$

Thus, in this case, the transition matrix reduces to

$$Q(g,h)=\frac{1}{|X_g|}\sum_{x\in X_g\cap X_h}\frac{1}{|G_x|}=\frac{1}{f(g)^n}\sum_{x\in X_g\cap X_h}\frac{1}{(k-r_x)!}.$$

Note that here the state space is $G^*=\{g\in S_k:\ f(g)>0\}=S_k\setminus\{\text{derangements}\}$ with size $|G^*|=k!-!\,k$, where $!\,k=k!\sum_{i=0}^{k}\frac{(-1)^i}{i!}$ is the number of derangements. Notice that as $k\to\infty$ the fraction of derangements in S_k tends to $1/e$ (a constant).

Example 4.1 (Value–permutation model: $k = 5,\ n = 4$)**.** *Action.* S_5 acts on $X = [5]^4$ by permuting symbols.

1. Fixed–point set X_g.
Let $g = (12) \in S_5$. Then

$$\mathrm{Fix}(g) = \{3,4,5\}, \quad X_g = \mathrm{Fix}(g)^{\,4} = \{3,4,5\}^4,$$

so $|X_g| = 3^4 = 81$.

2. Stabilizer G_x.
Take $x = (1,3,3,5) \in [5]^4$. Its set of distinct symbols is $\mathrm{supp}(x) = \{1,3,5\}$, so $r_x = 3$. Thus

$$G_x \cong S_{5-3} = S_2, \quad |G_x| = (5-3)! = 2.$$

Heuristics recap.

(1) For $g = (12)$ to fix a word, every coordinate must lie in $\mathrm{Fix}(g)$, giving $|X_g| = |\,\mathrm{Fix}(g)|^{\,n} = 3^4$.
(2) For a specific word x, only the $k - r_x$ *unused* symbols can be permuted freely: $|G_x| = (k - r_x)!$ (here 2!).

3. Orbits $\Longleftrightarrow$ set partitions (by positions).
The orbit of $x = (1,3,3,5)$ under S_5 corresponds to the partition of the *positions* by equal symbols:

$$\{\{1\},\{2,3\},\{4\}\}.$$

In general, each orbit is determined by the set–partition of $[n]$ into blocks of equal symbols.

4. Box–ball interpretation (twelvefold way).
Place $n = 4$ labeled balls into $k = 5$ labeled boxes ("1", "2", "3", "4", "5") according to the symbols (x_i), i.e., ball i is placed into box x_i. Permuting the box labels (by the S_5–action) erases labels and yields exactly a set–partition of the 4 balls into at most 5 blocks. Hence the number of orbits is

$$z \;=\; \sum_{r=0}^{\min\{5,4\}} S(4,r) \;=\; B_4 \;=\; 15.$$

Example 4.2 (Complete matrix decomposition for $k = 3,\ n = 2$)**.** *States.*

$$G^* = \{e, (12), (13), (23)\}, \qquad X = [3]^2 = \{11,12,13,21,22,23,31,32,33\}.$$

(These orders are used throughout for rows/columns of Q, K, and for A, B.)
Forward leg $A : G^* \to X$ (4×9) and Backward leg $B : X \to G^*$ (9×4)

$$A = \frac{1}{9}\begin{pmatrix} 1&1&1&1&1&1&1&1&1\\ 0&0&0&0&0&0&0&0&9\\ 0&0&0&0&9&0&0&0&0\\ 9&0&0&0&0&0&0&0&0 \end{pmatrix} \qquad B = \frac{1}{2}\begin{pmatrix} 1&0&0&1\\ 2&0&0&0\\ 2&0&0&0\\ 2&0&0&0\\ 1&0&1&0\\ 2&0&0&0\\ 2&0&0&0\\ 2&0&0&0\\ 1&1&0&0 \end{pmatrix}$$

Dual kernel $Q = AB$ (4×4) and Primal kernel $K = BA$ (9×9)

$$Q = \frac{1}{18}\begin{pmatrix} 15 & 1 & 1 & 1 \\ 9 & 9 & 0 & 0 \\ 9 & 0 & 9 & 0 \\ 9 & 0 & 0 & 9 \end{pmatrix} \qquad K = \frac{1}{18}\begin{pmatrix} 10 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 10 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 10 \end{pmatrix}$$

Spectra. $\mathrm{Spec}(Q) = \{1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}\}$, $\mathrm{Spec}(K) = \{1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, 0, 0, 0, 0, 0\}$. The nonzero spectra of Q and K coincide, as predicted by the factorization $Q = AB$, $K = BA$.

4.2. **Canonical Closed Forms for $Q(g, h)$.** In this section we further simplify the transition matrix formula above. Now that we know $X_g = \mathrm{Fix}(g)^n$ and $|X_g| = f(g)^n$, we first observe

$$\begin{aligned} X_g \cap X_h &= [\mathrm{Fix}(g)]^n \cap [\mathrm{Fix}(h)]^n \\ &= \{x \in [k]^n : x_i \in \mathrm{Fix}(g) \text{ and } x_i \in \mathrm{Fix}(h) \text{ for all } i\} \\ &= \{x \in [k]^n : x_i \in \mathrm{Fix}(g) \cap \mathrm{Fix}(h) \text{ for all } i\} \\ &= [\mathrm{Fix}(g) \cap \mathrm{Fix}(h)]^n. \end{aligned} \tag{24}$$

For $n, r \in \mathbb{N}$, recall that $S(n, r)$ denotes the number of partitions of an n–element set into exactly r nonempty blocks. Equivalently, $r!\,S(n, r)$ is the number of surjections $[n] \twoheadrightarrow [r]$. (see Stanley [40], Section 1.9 The Twelvefold Way).

Lemma 4.3 (Count of n-tuples with exactly r distinct values)**.** *Fix integers $j, n \geq 1$ and $1 \leq r \leq \min\{j, n\}$. Then*

$$\big|\{\, x \in [j]^n : \ r_x = r \,\}\big| \;=\; \binom{j}{r} S(n, r)\, r!,$$

where $S(n, r)$ is the Stirling number of the second kind.

Proof. To construct such an x on the left–hand side:

(1) Choose which r values from $[j]$ to use: $\binom{j}{r}$ ways
(2) Partition the n positions into r non-empty blocks (one for each value): $S(n, r)$ ways
(3) Assign the r chosen values to the r blocks: $r!$ ways

Therefore: $|\{x \in [j]^n : r_x = r\}| = \binom{j}{r} S(n, r) r!$. ☐

Note that this formula can be verified using the inclusion-exclusion principle.

Theorem 4.4 (Canonical closed forms for $Q(g, h)$)**.** *Set*

$$a := f(g) = |\,\mathrm{Fix}(g)|, \qquad j := \big|\mathrm{Fix}(g) \cap \mathrm{Fix}(h)\big|.$$

Then the dual Burnside kernel satisfies the following equivalent forms:

$$\text{(Stirling sum)} \qquad Q(g, h) = \frac{1}{a^n} \sum_{r=1}^{\min\{j,n\}} \binom{j}{r} S(n, r) \frac{r!}{(k - r)!}, \tag{25}$$

$$\text{(expectation form)} \qquad Q(g, h) = \left(\frac{j}{a}\right)^n \mathbb{E}\big[(k - R_{j,n})!^{-1}\big], \tag{26}$$

$$\text{(coefficient form)} \qquad Q(g, h) = \frac{n!}{a^n} [u^n]\, [z^k] \; \exp(z) \big(1 + z\,(e^u - 1)\big)^j. \tag{27}$$

Here $R_{j,n}$ denotes the number of distinct symbols used by a uniformly random word in $[j]^n$, i.e.,

$$\mathbb{P}(R_{j,n}=r)=\frac{\binom{j}{r}S(n,r)r!}{j^{\,n}},\qquad r=1,\dots,\min\{j,n\}.$$

Convention. *We interpret $(k-r)!^{-1}=0$ for $r>k$, and for $n\ge 1$ the case $j=0$ yields $Q(g,h)=0$ in all three forms.*

Proof. By (24), $X_g\cap X_h=[j]^n$. Let $\Phi:[j]^n\to\{1,\dots,\min(j,n)\}$ be defined by $\Phi(x):=r_x$. By the re-indexing principle (8) with $F(x):=\dfrac{1}{(k-r_x)!}$,

$$\begin{aligned}
\sum_{x\in[j]^n}F(x)&=\sum_{r=1}^{\min(j,n)}\sum_{\substack{x\in[j]^n\\ \Phi(x)=r}}F(x)\\
&=\sum_{r=1}^{\min(j,n)}\sum_{\substack{x\in[j]^n\\ r_x=r}}\frac{1}{(k-r)!}\\
&=\sum_{r=1}^{\min(j,n)}\frac{1}{(k-r)!}\sum_{\substack{x\in[j]^n\\ r_x=r}}1\\
&=\sum_{r=1}^{\min(j,n)}\frac{1}{(k-r)!}\binom{j}{r}S(n,r)\,r! \qquad\qquad \text{(by Lemma 4.3).}
\end{aligned}$$

Hence (25) follows:

$$Q(g,h)=\frac{1}{a^n}\sum_{x\in[j]^n}\frac{1}{(k-r_x)!}=\sum_{r=1}^{\min(j,n)}\frac{\binom{j}{r}S(n,r)r!}{a^n(k-r)!}.$$

Note that if $j=0$, the intersection is empty so $Q(g,h)=0$; the formula gives an empty sum in this case.

For (26), define $R_{j,n}$ as stated; then

$$\binom{j}{r}S(n,r)r!\;=\;j^{\,n}\mathbb{P}(R_{j,n}=r),$$

so (25) becomes

$$Q(g,h)=\frac{j^{\,n}}{a^{\,n}}\sum_r\mathbb{P}(R_{j,n}=r)\,(k-r)!^{-1}=\big(\frac{j}{a}\big)^n\mathbb{E}[(k-R_{j,n})!^{-1}].$$

For (27), use the coefficient extractions[5]

$$r!\,S(n,r)=n!\,[u^{\,n}]\,(e^u-1)^r,\qquad \frac{1}{(k-r)!}=[z^{\,k-r}]\,e^z.$$

[5]Recall the exponential generating function (EGF) for set partitions into exactly r nonempty blocks is

$$\frac{(e^u-1)^r}{r!}\;=\;\sum_{n\ge r}S(n,r)\,\frac{u^n}{n!}.$$

Insert into (25), interchange the sum with coefficient extractions, and apply the binomial theorem:

$$\begin{aligned} Q(g,h) &= \frac{n!}{a^n}\,[u^n]\sum_{r\ge 0}\binom{j}{r}(e^u-1)^r\,[z^{k-r}]e^z \\ &= \frac{n!}{a^n}\,[u^n]\sum_{r\ge 0}\binom{j}{r}(e^u-1)^r\,[z^k]\big(z^r e^z\big) && \text{(since } [z^{k-r}]e^z=[z^k]z^re^z\text{)} \\ &= \frac{n!}{a^n}\,[u^n][z^k]\,e^z\sum_{r\ge 0}\binom{j}{r}\big(z(e^u-1)\big)^r \\ &= \frac{n!}{a^n}\,[u^n][z^k]\,e^z\,\big(1+z(e^u-1)\big)^j && \text{(binomial theorem).} \end{aligned}$$

This is (27). □

Example 4.5 (Specializing to $g=e$)**.** With $f := |\operatorname{Fix}(h)|$ (so $j=f$ when $g=e$), recall

$$S(n,1)=1,\qquad S(n,2)=2^{n-1}-1,\qquad S(n,3)=\frac{3^n-3\cdot 2^n+3}{6},\qquad S(n,4)=\frac{4^n-4\cdot 3^n+6\cdot 2^n-4}{24}.$$

By (25), if h has exactly one fixed symbol $f=1$, then $j=1$ so

$$Q(e,h)=\frac{1}{k^n}\Big[\binom{1}{1}S(n,1)\frac{1!}{(k-1)!}\Big]=\frac{1}{k^n(k-1)!}.$$

Likewise, for two fixed symbols $f=2$ and $n\ge 2$, then $j=2$ so

$$\begin{aligned} Q(e,h) &= \frac{1}{k^n}\Big[\binom{2}{1}S(n,1)\frac{1!}{(k-1)!}+\binom{2}{2}S(n,2)\frac{2!}{(k-2)!}\Big] \\ &= \frac{1}{k^n}\left[\frac{2}{(k-1)!}+\frac{2\,(2^{n-1}-1)}{(k-2)!}\right]=\frac{2}{k^n}\left(\frac{1}{(k-1)!}+\frac{2^{n-1}-1}{(k-2)!}\right). \end{aligned}$$

4.3. **Stationary Distribution.** We can directly specialize Theorem 3.3 to the value-permutation model and get the following result.

Theorem 4.6 (Dual stationary law)**.** *The dual Burnside chain Q on S_k^* is reversible, irreducible, and aperiodic. Its stationary distribution is unique and given by*

$$\pi(g)=\frac{f(g)^n}{k!\,Z_{k,n}},\qquad g\in S_k^*,$$

where the normalizing constant is

$$Z_{k,n}=\begin{cases} B_n & k\ge n\\ \sum_{d=0}^{k}S(n,d) & k<n\end{cases}$$

and B_n is the n-th Bell number.

Recall that, in the value-permuting case, $\pi_K(x)=(k-r_x)!/(k!\,Z_{k,n})$. In conclusion,

	$k\ge n$	$k<n$
$\pi_K(x)$	$(k-r_x)!/(k!\,B_n)$	$(k-r_x)!/\big(k!\sum_{d\le k}S(n,d)\big)$
$\pi_Q(g)$	$f(g)^n/(k!\,B_n)$	$f(g)^n/\big(k!\sum_{d\le k}S(n,d)\big)$

Lemma 4.7 (Extrema of π_K and π_Q in the value–permutation model)**.** *For $n\ge 1$,*

$$\pi_{K,\max}=\frac{1}{k\,Z_{k,n}},\qquad \pi_{K,\min}=\frac{(k-r_{\max})!}{k!\,Z_{k,n}},\qquad r_{\max}:=\min\{k,n\},$$

attained at words with $r_x=1$ and $r_x=r_{\max}$, respectively.

For $n \geq 1, k \geq 3$,

$$\pi_{Q,\min} = \frac{1}{k!\,Z_{k,n}}, \qquad \pi_{Q,\max} = \frac{k^n}{k!\,Z_{k,n}},$$

attained at permutations with $f(g) = 1$ and $f(g) = k$ (the identity), respectively; for $k = 1, 2$ we have $S_k^* = \{e\}$, so $\pi_{Q,\min} = \pi_{Q,\max} = \pi_Q(e) = 1$ and the dual chain is trivial.[6]

Proof. For K, $r_x \in \{1, \dots, \min\{k, n\}\}$ and $(k - r_x)!$ is strictly decreasing in r_x. Thus $\pi_K(x)$ is maximal at $r_x = 1$ and minimal at $r_x = r_{\max} = \min\{k, n\}$, giving the stated $\pi_{K,\max}$ and $\pi_{K,\min}$.

For Q, $f(g) \in \{1, \dots, k\}$ and $f(g)^n$ is strictly increasing in $f(g)$. If $k \geq 3$,

$$\max_{g\in S_k^*} f(g) = k \quad \text{(attained at } g = e), \qquad \min_{g\in S_k^*} f(g) = 1 \quad \text{(e.g. } g = (1)(2\,3\,\dots\,k) \in S_k^* \text{ for } k \geq 3),$$

giving the stated $\pi_{Q,\min}$ and $\pi_{Q,\max}$. For $k = 1, 2$, the only element with $f(g) > 0$ is e, so $S_k^* = \{e\}$ and $\pi_Q(e) = 1$. □

Remark 4.8 (Stationary–mass ratio)**.** For $k \geq 3$,

$$\frac{\pi_{Q,\max}}{\pi_{Q,\min}} = k^n.$$

We also give a second derivation of the normalization constant via the *cycle–index* of the action: by specializing the cycle–index polynomial $P_k(t_1, \dots, t_k)$ and extracting the relevant coefficient. For cycle–index methods beyond the symmetric groups, see Fulman's work on classical groups [22].

First observe that the dual stationary law is proportional to $|X_g|$ and must sum to 1:

$$\pi_Q(g) \propto |X_g|, \qquad \sum_{h\in G^*} \pi_Q(h) = 1 \;\Rightarrow\; \pi_Q(g) = \frac{|X_g|}{\sum_{h\in G^*} |X_h|}.$$

Recall from [10] (p. 412) that the cycle-index polynomial for the natural action of S_k on $[k]$:

$$P_k(t_1, \dots, t_k) = \frac{1}{k!} \sum_{g\in S_k} t_1^{c_1(g)}\, t_2^{c_2(g)} \cdots t_k^{c_k(g)},$$

where $c_j(g)$ is the number of j–cycles of g. Define

$$F_k(x) := \sum_{g\in S_k} x^{f(g)} = k!\, P_k(x, 1, \dots, 1),$$

where $f(g) = |\operatorname{Fix}(g)| = c_1(g)$ and P_k is the cycle–index polynomial of S_k.

We recall the following classical identity (last equation on p. 412 of [10]).

Proposition 4.9 (Exponential Generating Function; [10])**.**

$$\sum_{k=0}^{\infty} P_k(x_1, \dots, x_k)\, z^k = \exp\Big(\sum_{j=1}^{\infty} \frac{x_j}{j}\, z^j\Big).$$

Theorem 4.10. *With notation as above,*

$$\text{(closed form for } F_k) \qquad F_k(x) = k!\,[z^k]\, \frac{e^{(x-1)z}}{1-z} = k! \sum_{m=0}^{k} \frac{(x-1)^m}{m!}, \tag{28}$$

(29)

$$\text{(Euler operator extraction)} \qquad \sum_{g\in S_k} |X_g| \equiv \sum_{g\in S_k} f(g)^n = \Big(x\frac{d}{dx}\Big)^n F_k(x)\Big|_{x=1} = k! \sum_{m=0}^{k} S(n, m).$$

[6]Recall $S_k^* = S_k \setminus \{\text{derangements}\}$. For $k = 1, 2$, the dual chain has a single state and is stationary at $t = 0$, so $t_{\text{mix}}(\varepsilon) = 0$ for all $\varepsilon \in (0, 1)$.

Proof. Proposition 4.9 with $x_1 = x$ and $x_j = 1$ for $j \geq 2$ yields[7]

$$\sum_{k\geq 0} P_k(x,1,1,\dots)\,z^k = \exp\Bigl(xz + \sum_{j\geq 2}\frac{z^j}{j}\Bigr) = \exp\Bigl(-(\log(1-z)) + (x-1)z\Bigr) = \frac{e^{(x-1)z}}{1-z}.$$

Taking $[z^k]$ and multiplying by $k!$ gives the first form of (28). Expanding $(1-z)^{-1} = \sum_{r\geq 0} z^r$ and $e^{(x-1)z} = \sum_{m\geq 0}\frac{(x-1)^m}{m!}z^m$, the coefficient of z^k is $\sum_{m=0}^{k}\frac{(x-1)^m}{m!}$, proving the second form of (28).

Since $|X_g| = f(g)^n$ for $X = [k]^n$, the first equality in (29) is immediate. Recall that the "degree-counting" Euler operator $x\frac{d}{dx}$ acts on monomials by

$$\Bigl(x\frac{d}{dx}\Bigr)^n x^m \;=\; m^n\,x^m,$$

so

$$\Bigl(x\frac{d}{dx}\Bigr)^n\Bigl(\sum_{g\in S_k} x^{f(g)}\Bigr) = \sum_{g\in S_k} f(g)^n\,x^{f(g)}.$$

Evaluating at $x = 1$ gives the second equality in (29).

Recall the operator identity

$$\bigl(x\frac{d}{dx}\bigr)^n \;=\; \sum_{j=0}^{n} S(n,j)\,x^j\,\frac{d^j}{dx^j}.$$

So

$$\begin{aligned}
(x\frac{d}{dx})^n(x-1)^m\Big|_{x=1} &= \sum_{j=0}^{n} S(n,j)\,x^j\,\frac{d^j}{dx^j}(x-1)^m\Big|_{x=1}\\
&= \sum_{j=0}^{n} S(n,j)\,1^j\,\frac{d^j}{dx^j}(x-1)^m\Big|_{x=1}\\
&= \sum_{j=0}^{n} S(n,j)\,m^{\underline{j}}\,(x-1)^{m-j}\Big|_{x=1}\\
&= \sum_{j=0}^{n} S(n,j)\,m^{\underline{j}}\,\mathbf{1}_{\{j=m\}}\\
&= \mathbf{1}_{\{m\leq n\}}\;S(n,m)\,m^{\underline{m}}\\
&= \mathbf{1}_{\{m\leq n\}}\;S(n,m)\,m!.
\end{aligned}$$

Thus by the second form of (28) we obtain

$$\Bigl(x\frac{d}{dx}\Bigr)^n F_k(x)\Big|_{x=1} = k!\sum_{m=0}^{k}\frac{1}{m!}\,S(n,m)\,m! = k!\sum_{m=0}^{k} S(n,m),$$

the last equality in (29), where we used $S(n,m) = 0$ for $m > n$. ∎

By (29),

$$\pi_Q(g) \equiv \frac{|X_g|}{\sum_h |X_h|} = \frac{f(g)^n}{k!\,\sum_{m=0}^{k} S(n,m)}.$$

[7]Recall the power-series expansion

$$-\log(1-z) \;=\; \sum_{j\geq 1}\frac{z^j}{j} \qquad (|z|<1).$$

In particular, if $k \ge n$, then $\sum_{m=0}^{k} S(n,m) = B_n$ (the nth Bell number), so

$$\pi_Q(g) = \frac{f(g)^n}{k!\, B_n} \qquad (k \ge n).$$

Remark 4.11 (Two quick corollaries)**.** (i) The orbit count is $z = |X/G| = \frac{1}{|G|}\sum_g |X_g| = \frac{1}{k!}\sum_g f(g)^n = \sum_{m=0}^{k} S(n,m)$, so $z = B_n$ when $k \ge n$.

(ii) Since $F_k(x) = \sum_{g\in S_k} x^{f(g)}$, the probability generating function of $f(g)$ under the uniform $g \in S_k$ is

$$G_k(x) \;=\; \mathbb{E}_{g\sim \mathrm{Unif}(S_k)}[\, x^{f(g)}\,] \;=\; \frac{1}{|S_k|}\sum_{g\in S_k} x^{f(g)} \;=\; \frac{1}{k!}\,F_k(x) \;=\; [z^k]\frac{e^{(x-1)z}}{1-z},$$

where the last equality is by the first form of (28). Then, for any $k \ge 2$,[8]

$$\mathbb{E}[f(g)] = G_k'(1) = [z^k]\frac{\partial}{\partial x}\Big(\frac{e^{(x-1)z}}{1-z}\Big)\Big|_{x=1} = [z^k]\frac{z}{1-z} = [z^{k-1}]\frac{1}{1-z} = 1,$$

$$\mathbb{E}\big[f(g)(f(g)-1)\big] = G_k''(1) = [z^k]\frac{\partial^2}{\partial x^2}\Big(\frac{e^{(x-1)z}}{1-z}\Big)\Big|_{x=1} = [z^k]\frac{z^2}{1-z} = [z^{k-2}]\frac{1}{1-z} = 1,$$

and for $k = 0,1$ the same coefficient equalities give the correct edge values. Hence, using $\mathrm{Var}(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \big(\mathbb{E}[X]\big)^2$,

$$\mathrm{Var}(f(g)) = 1 + 1 - 1 = 1 \quad (\text{and } = 0 \text{ for } k = 0,1).$$

4.4. **Mixing and Convergence.** We first recall that $r_x := |\{x_1,\dots,x_n\}|$ is the number of distinct symbols appearing in $x \in X := [k]^n$, so $1 \le r_x \le \min\{k,n\}$. Since $|G_x| = (k-r_x)!$,

$$(k-\min\{n,k\})! \;\le\; |G_x| \;\le\; (k-1)!, \qquad k \ge 2.$$

Equality of the upper bound is attained at $r_x = 1$ (e.g. $x = (1,1,\dots,1)$), hence

$$M := \max_x |G_x| = (k-1)!\,.$$

Remark 4.12. Furthermore, by the fact that the identity $e \in S_k$ lies in every G_u and fixes every v:

$$K(u,v) \;=\; \sum_{g\in G_u\cap G_v} \frac{1}{|G_u|\,|X_g|} \;\ge\; \frac{1}{|G_u|\,|X_e|} \;=\; \frac{1}{|G_u|\,k^n}.$$

Using the upper bound $|G_u| \le (k-1)!$ we obtain a *uniform* lower bound, independent of u, v:

$$\boxed{K(u,v) \;\ge\; \frac{1}{(k-1)!\,k^n} \qquad \forall\, u,v \in [k]^n,\; k \ge 2.}$$

This recovers Lemma 4.4 in Paguyo [32].

Note that the bound is attained whenever $n \ge k-1$ by taking $u = (1,1,\dots,1)$ and $v = (2,3,\dots,k,\text{padding})$; hence, for $n \ge k-1$, $\inf_{u,v} K(u,v) = ((k-1)!\,k^n)^{-1}$, meaning this lower bound cannot be further improved in this region.

To see that the bound can be realized, assume $n \ge k-1$ and set

$$u = (1,1,\dots,1), \qquad v = (2,3,\dots,k,\; \underbrace{2,\dots,2}_{n-(k-1)\text{ times}}\;).$$

[8]Recall $\dfrac{1}{1-z} = \sum_{m\ge 0} z^m$, so the coefficient of z^m is 1 for every $m \ge 0$.

Then $r_u = 1$ and $|G_u| = (k-1)!$. Moreover, every symbol $2, \dots, k$ occurs in v, so the only permutation fixing both u and v is the identity; hence the inequality above is an equality. Consequently $\inf_{u,v} K(u,v) = \big((k-1)!\, k^n\big)^{-1}$, and the infimum is achieved whenever $n \geq k-1$.

Now we collect the standard bounds for the *primal* Burnside chain K in the value–permutation action $S_k \curvearrowright [k]^n$. Each item holds uniformly over starts.

(K1) **Uniform floor via stabilizers.** Lemma 3.24 gives $K(x,\cdot) \geq \delta_K \operatorname{Unif}(X)$, $\delta_K = 1/M$ with $M = \max_x |G_x|$. Here as analyzed above $M = (k-1)!$. Hence, by Theorem 3.25,

$$d_K(t) \;\leq\; \Big(1 - \frac{1}{(k-1)!}\Big)^t \qquad \Rightarrow \qquad t_{\mathrm{mix}}(K;\varepsilon) \;\leq\; \Big\lceil (k-1)!\, \log \frac{1}{\varepsilon} \Big\rceil.$$

This recovers Theorem 4.5 in [32] and also improves the model–free bound $|G| \log(1/\varepsilon) = k! \log(1/\varepsilon)$.

(K2) **Paguyo** [32] **for $k \geq n$.** For $k \geq n$,

$$d_K(t) \;\leq\; n\Big(1 - \frac{1}{2k}\Big)^t, \qquad \Rightarrow \quad t_{\mathrm{mix}}(K;\varepsilon) \;\leq\; \Big\lceil 2k\, \log \frac{n}{\varepsilon} \Big\rceil.$$

This is the main result Theorem 1.1 in [32] obtained via coupling. Equivalently, the relaxation time satisfies $t_{\mathrm{rel}}(K) \leq 2k$ (see also Proposition 4.3 in [32]).

We now turn (K1)–(K2) into consequences for the *dual* chain Q using the one–step TV comparison (Corollary 3.19) and the mixing–time equivalence (Theorem 3.20).

Corollary 4.13 (Global transfer of the stabilizer floor)**.** *From* (K1) *in the value–permutation setting (i.e.,* $M = \max_x |G_x| = (k-1)!$*),*

$$d_Q(t) \;\leq\; \Big(1 - \frac{1}{(k-1)!}\Big)^{t-1}, \qquad t_{\mathrm{mix}}(Q;\varepsilon) \;\leq\; 1 + \Big\lceil (k-1)!\, \log \frac{1}{\varepsilon} \Big\rceil.$$

Proof. By Corollary 3.19, $d_Q(t) \leq d_K(t-1)$. Apply (K1) with $\delta_K = 1/M = 1/(k-1)!$. For mixing times, Theorem 3.20 gives $t_{\mathrm{mix}}(Q;\varepsilon) \leq t_{\mathrm{mix}}(K;\varepsilon) + 1$. ☐

Instead, we could directly apply Theorem 3.25 to Q, which matches Corollary 4.13 up to the immaterial $+1$ step from transfer:

Theorem 4.14 (Direct one–step minorization for Q)**.**

$$d_Q(t) \;\leq\; \Big(1 - \frac{1}{(k-1)!}\Big)^{t}, \qquad t_{\mathrm{mix}}(Q;\varepsilon) \;\leq\; \Big\lceil (k-1)!\, \log \frac{1}{\varepsilon} \Big\rceil.$$

Proof. Independently of K, Lemma 3.24 yields $Q(g,\cdot) \geq \delta_Q\, \delta_e$, $\delta_Q = 1/M = 1/(k-1)!$. Applying Theorem 3.25 with this δ_Q gives the stated bounds. ☐

Theorem 4.15 (Paguyo's Theorem 1.1 transfer to Q)**.** *If $k \geq n$ then for all $t \geq 1$,*

$$d_Q(t) \leq n\Big(1 - \frac{1}{2k}\Big)^{t-1}, \qquad t_{\mathrm{mix}}(Q;\varepsilon) \leq 1 + \Big\lceil 2k \log \frac{n}{\varepsilon} \Big\rceil.$$

Proof. Combine Lemma 3.19 with Paguyo's bound (K2). Then apply Theorem 3.20. ☐

The factor $(1 - \frac{1}{2k})^{t-1} = (2k/(2k-1))(1 - \frac{1}{2k})^t$ shows the dual bound is weaker only by the constant $\frac{2k}{2k-1}$; asymptotically $t^{Q}_{\mathrm{mix}} = O\big(k \ln(n/\varepsilon)\big)$ as in K.

Moreover, the relaxation times agree:

Theorem 4.16 (Paguyo's gap bound (dual form))**.** *For $k \geq n$,*

$$\lambda_1(Q) = \lambda_1(K) \;\leq\; 1 - \frac{1}{2k}.$$

Equivalently, with $\gamma := 1 - \lambda_1$,

$$\gamma(Q) = \gamma(K) \ \geq \ \frac{1}{2k}, \qquad t_{\mathrm{rel}}(Q) = t_{\mathrm{rel}}(K) \ \leq \ 2k.$$

Proof. Follows from Corollary 3.13 and Paguyo's Proposition 4.3 [32]. □

Remark 4.17 (Paguyo's bound via geometric TV)**.** From the uniform TV decay (K2) $d_K(t) \leq n\,(1 - \frac{1}{2k})^t$ and (6) with $\alpha = 1 - \frac{1}{2k}$, we obtain

$$\lambda_1(K) = \lambda_*(K) \ \leq \ 1 - \frac{1}{2k},$$

recovering Paguyo's gap bound directly from the geometric TV inequality.

4.5. **Lumping the Dual Chain Q by Fixed–Point Count.** Paguyo [32, §3–5] shows that lumping the original Burnside kernel K by *set–partition* structure produces a manageable chain $\bar{K}$ on Π_n. As noted in Section 3.5, orbit partitions on X give a natural lumping for K. In the dual setting on $G^* = S_k^* = \{g \in S_k : \ |X_g| > 0\}$, the analog of *orbits* is *conjugacy classes* (Corollary 3.39 and Lemma 3.44). In this section we instead examine a *coarser* lumping: by *number of fixed symbols*. Recall $f(g) = |\operatorname{Fix}(g)|$ and define

$$C_r := \{\, g \in G^* : \ f(g) = r \,\}, \qquad r \in \mathcal{F} := \{1, 2, \dots, k-2, k\}.$$

Note that $C_{k-1} = \varnothing$ since $!\,1 = 0$; see Remark 4.19 below.

Lemma 4.18 (Fixed–point counting formulas)**.** *With notation as above,*

(a) The number of permutations in S_k *with exactly* s *fixed points is*

$$\big|\{g \in S_k : \ f(g) = s\}\big| = \binom{k}{s} \cdot !\,(k-s).$$

(b) Fix $g \in S_k$ *with* $r := |\operatorname{Fix}(g)|$*. For* $0 \leq j \leq \min\{r, s\}$,

$$|\{h \in C_s : \ |\operatorname{Fix}(g) \cap \operatorname{Fix}(h)| = j\}| = \binom{r}{j}\binom{k-r}{s-j} \cdot !\,(k-s).$$

(c) For $x \in [k]^n$*, recall* $|\operatorname{supp}(x)| = r_x$ *where* $\operatorname{supp}(x) = \{x_1, \dots, x_n\}$*. Then*

$$G_x = \{\, h \in S_k : \ \operatorname{supp}(x) \subseteq \operatorname{Fix}(h) \,\}, \qquad |G_x \cap C_s| = \binom{k - r_x}{s - r_x} \cdot !\,(k-s)$$

Proof. (a) Choose the fixed set $S \subseteq [k]$ of size s in $\binom{k}{s}$ ways; on S the permutation is the identity, and on S^c (size $k-s$) it must be a derangement, giving $!\,(k-s)$ possibilities.

(b) Let $F := \operatorname{Fix}(g)$ with $|F| = r$. To obtain $h \in C_s$ with $|\operatorname{Fix}(h) \cap F| = j$, choose the j fixed symbols inside F in $\binom{r}{j}$ ways and the remaining $s-j$ fixed symbols inside F^c in $\binom{k-r}{s-j}$ ways; on the complement (size $k-s$), h must be a derangement, contributing $!\,(k-s)$ possibilities. Multiply.

(c) If $h \in G_x$ then $h \cdot x = x$, so $h(a) = a$ for all $a \in \operatorname{supp}(x)$; hence $\operatorname{supp}(x) \subseteq \operatorname{Fix}(h)$. Conversely, if $\operatorname{supp}(x) \subseteq \operatorname{Fix}(h)$ then $(h \cdot x)_i = h(x_i) = x_i$ for all i, so $h \in G_x$. Thus $G_x = \{h \in S_k : \ \operatorname{supp}(x) \subseteq \operatorname{Fix}(h)\}$.

To count $|G_x \cap C_s|$: the $r_x = |\operatorname{supp}(x)|$ symbols in $\operatorname{supp}(x)$ are forced fixed; as in (b), choose the remaining $s - r_x$ fixed symbols from the complement of size $k - r_x$ in $\binom{k-r_x}{s-r_x}$ ways, and on the $k-s$ nonfixed symbols require a derangement, contributing $!\,(k-s)$. Therefore $|G_x \cap C_s| = \binom{k-r_x}{s-r_x} \cdot !\,(k-s)$, with the usual convention $\binom{n}{m} = 0$ if $m \notin \{0, \dots, n\}$. □

Remark 4.19 (Restriction to S_k^*)**.** Since $S_k^* = \{g \in S_k : \ f(g) > 0\}$, the count in Lemma 4.18(a) remains valid for $s \geq 1$:

$$\big|\{g \in S_k^* : \ f(g) = s\}\big| = \binom{k}{s} \cdot !\,(k-s),$$

while for $s = 0$ the set is empty.

Likewise, for $s \geq 1$ every permutation with exactly s fixed points already lies in $S_k^* = \{g \in S_k : f(g) > 0\}$. Thus $C_s \subseteq S_k^*$, and the formulas in Lemma 4.18 (b)–(c) remain valid if we interpret C_s as a subset of S_k^* rather than of S_k.

Theorem 4.20 (Lumpability by fixed points)**.** *Then the partition $\{C_s : s \in \mathcal{F}\}$ is strongly lumpable for Q, yielding a quotient chain $\bar{Q} = (\bar{Q}(r,s))_{r,s\in\mathcal{F}}$.*

Proof. Recall $Q = AB$ with $Q(g,h) = \sum_{x\in X} A(g,x)B(x,h)$. Fix C_s and write

$$\sum_{h\in C_s} Q(g,h) = \sum_{x\in X} A(g,x) \underbrace{\sum_{h\in C_s} B(x,h)}_{=:u(x)} = \frac{1}{|X_g|}\sum_{x\in X_g} u(x).$$

By Lemma 3.29 (17) and Lemma 4.18,

$$u(x) = \frac{|G_x \cap C_s|}{|G_x|} = \frac{\binom{k-r_x}{s-r_x}\cdot !\,(k-s)}{(k-r_x)!} = \frac{!\,(k-s)}{(k-s)!}\cdot\frac{1}{(s-r_x)!},$$

which depends only on $r_x := |\operatorname{supp}(x)|$ (and s,k), not on g. If $|\operatorname{Fix}(g)| = r$, then $X_g = \operatorname{Fix}(g)^n$ is the set of r–ary words of length n; thus the multiset $\{r_x : x \in X_g\}$ (hence $\{u(x) : x \in X_g\}$) depends only on r (relabel the r symbols). Therefore, for g_1, g_2 with $|\operatorname{Fix}(g_1)| = |\operatorname{Fix}(g_2)|$,

$$\sum_{h\in C_s} Q(g_1,h) = \frac{1}{|X_{g_1}|}\sum_{x\in X_{g_1}} u(x) = \frac{1}{|X_{g_2}|}\sum_{x\in X_{g_2}} u(x) = \sum_{h\in C_s} Q(g_2,h),$$

so strong lumpability holds. □

Example 4.21 (Lumpability check and construction of the lumped chain ($k = 3,\ n = 2$))**.** *Partition by fixed–point count.*

$$C_3 = \{e\} \quad \text{(3 fixed points)}, \qquad C_1 = \{(12),(13),(23)\} \quad \text{(1 fixed point each)}.$$

Original kernel Q (rows/columns ordered $e,(12),(13),(23)$).

$$Q = \begin{array}{c} \\ e \\ (12) \\ (13) \\ (23) \end{array}\!\!\begin{array}{c} \begin{array}{cccc} e & (12) & (13) & (23) \end{array} \\ \begin{pmatrix} \frac{5}{6} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \end{array}, \qquad \operatorname{Spec}(Q) = \left\{1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}\right\}.$$

Lumpability via equal row–sums over blocks.

$$\text{For } x \in C_3: \quad \sum_{y\in C_3} Q(x,y) = \frac{5}{6}, \qquad \sum_{y\in C_1} Q(x,y) = \frac{1}{18}+\frac{1}{18}+\frac{1}{18} = \frac{1}{6}.$$

$$\text{For } x \in C_1: \quad \sum_{y\in C_3} Q(x,y) = \frac{1}{2} \text{ (same for each of (12), (13), (23))}, \qquad \sum_{y\in C_1} Q(x,y) = \frac{1}{2} \text{ (same for each)}.$$

Hence the sums over each target block are *equal within every source block*, so Q is strongly lumpable w.r.t. $\{C_3, C_1\}$.

Lumped entries are exactly these common sums by (10)*.*

$$\bar{Q}(C_3,C_3) = \frac{5}{6}, \quad \bar{Q}(C_3,C_1) = \frac{1}{6}, \quad \bar{Q}(C_1,C_3) = \frac{1}{2}, \quad \bar{Q}(C_1,C_1) = \frac{1}{2}.$$

(Equivalently) block–average formula (11) *gives the same.*

$$\bar{Q}(C_3,C_3) = Q(e,e) = \frac{5}{6}, \qquad \bar{Q}(C_3,C_1) = Q(e,(12)) + Q(e,(13)) + Q(e,(23)) = \frac{1}{6},$$

$$\bar{Q}(C_1, C_3) = \frac{1}{3} \sum_{\sigma \in C_1} Q(\sigma, e) = \frac{1}{3}\left(\frac{1}{2} + \frac{1}{2} + \frac{1}{2}\right) = \frac{1}{2}, \qquad \bar{Q}(C_1, C_1) = \frac{1}{3} \sum_{\sigma \in C_1} \sum_{\tau \in C_1} Q(\sigma, \tau) = \frac{1}{2}.$$

Lumped kernel $\bar{Q}$ on $\{C_3, C_1\}$.

$$\bar{Q} = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad \mathrm{Spec}(\bar{Q}) = \Big\{1, \ \frac{1}{3}\Big\}.$$

(Here the lumped chain retains only a subset of the nonzero eigenvalues of Q.)

Remark 4.22. In this example ($k = 3$), the *fixed–point–count* partition coincides with the *conjugacy* partition on G^*: $C_3 = \{e\}$ and $C_1 = \{(12), (13), (23)\}$ are exactly the two conjugacy classes present in G^*. Hence the lumping here agrees with conjugacy–class lumping.

For $0 \le s \le k$ (since we will only use $s \in \mathcal{F} := \{1, 2, \dots, k-2, k\}$), set

$$\theta_{k,s} \ := \ \frac{!\,(k-s)}{(k-s)!} \qquad \text{(derangement probability on } k-s \text{ symbols)}.$$

For the distinct–value count $R_{r,n}$ of a uniform word in $[r]^n$ (defined in Theorem 4.4) we recall the standard occupancy law

$$\mathbb{P}(R_{r,n} = m) = \frac{\binom{r}{m}\, S(n,m)\, m!}{r^n}, \qquad 1 \le m \le \min\{r, n\}, \tag{30}$$

and adopt the convention $(t!)^{-1} := 0$ for $t < 0$.

We recall that $\pi_Q(g) = \dfrac{f(g)^n}{k!\, Z_{k,n}}$ is the stationary law of Q (Theorem 4.6).

Theorem 4.23 (Lumped stationary distribution)**.** *The lumped stationary distribution $\bar{\pi}$ on $\mathcal{F}$ is*

$$\bar{\pi}(s) = \sum_{g\colon f(g)=s} \pi_Q(g) = \frac{\binom{k}{s} \cdot !\,(k-s)\, s^n}{k!\, Z_{k,n}} = \frac{\theta_{k,s}}{Z_{k,n}} \cdot \frac{s^n}{s!}\,, \qquad s \in \mathcal{F}, \tag{31}$$

where

$$Z_{k,n} = \begin{cases} B_n, & k \ge n, \\ \displaystyle\sum_{d=0}^{k} S(n,d), & k < n. \end{cases}$$

Proof. Pushforward of stationarity gives $\bar{\pi}(s) = \sum_{g\colon f(g)=s} \pi_Q(g)$. By Lemma 4.18(a),

$$\bar{\pi}(s) = \frac{1}{k!\, Z_{k,n}} \sum_{g\colon f(g)=s} f(g)^n = \frac{1}{k!\, Z_{k,n}} \binom{k}{s} \cdot !\,(k-s)\, s^n.$$

Using $\dbinom{k}{s} = \dfrac{k!}{s!\,(k-s)!}$ and $\theta_{k,s} = \dfrac{!\,(k-s)}{(k-s)!}$ yields $\bar{\pi}(s) = \dfrac{\theta_{k,s}}{Z_{k,n}} \cdot \dfrac{s^n}{s!}$. □

Lemma 4.24. *For integers k, r, s, t (with the convention $\binom{n}{m} = 0$ if $m \notin \{0, \dots, n\}$),*

$$\sum_{j \in \mathbb{Z}} \binom{j}{t} \binom{r}{j} \binom{k-r}{s-j} \ = \ \binom{r}{t} \binom{k-t}{s-t}.$$

Proof. By Chu–Vandermonde $\binom{x+a}{n} = \sum_{j=0}^{n} \binom{x}{j}\binom{a}{n-j}$, we have

$$\begin{aligned}
\sum_j \binom{j}{t}\binom{r}{j}\binom{k-r}{s-j} &= \sum_j \Big[\binom{r}{t}\binom{r-t}{j-t}\Big]\binom{k-r}{s-j} && \Big(\text{use } \binom{r}{j}\binom{j}{t} = \binom{r}{t}\binom{r-t}{j-t}\Big) \\
&= \binom{r}{t}\sum_j \binom{r-t}{j-t}\binom{k-r}{s-j} \\
&= \binom{r}{t}\sum_u \binom{r-t}{u}\binom{k-r}{s-t-u} && (\text{set } u := j-t) \\
&= \binom{r}{t}\binom{(r-t)+(k-r)}{s-t} = \binom{r}{t}\binom{k-t}{s-t} && (\text{Chu–Vandermonde}).
\end{aligned}$$

□

Theorem 4.25 (Closed forms for $\bar{Q}$ from Theorem 4.4)**.** *For $r, s \in \mathcal{F}$,*

$$\text{(Stirling sum)} \qquad \bar{Q}(r,s) = \theta_{k,s}\,\frac{1}{r^{\,n}}\sum_{m=1}^{\min\{r,s,n\}}\binom{r}{m} S(n,m)\,\frac{m!}{(s-m)!}, \tag{32}$$

$$\text{(expectation form)} \qquad \bar{Q}(r,s) = \theta_{k,s}\,\mathbb{E}\big[(s-R_{r,n})!^{-1}\big]\,. \tag{33}$$

$$\text{(coefficient form)} \qquad \bar{Q}(r,s) = \theta_{k,s}\,\frac{n!}{r^{\,n}}\,[u^{n}]\,[z^{\,s}]\;e^{z}\left(1+z\,(e^{u}-1)\right)^{r}, \tag{34}$$

Proof. Fix $r, s \in \mathcal{F}$ and choose $g \in C_r$. For $h \in C_s$ define the overlap

$$j(h) := |\operatorname{Fix}(g)\cap\operatorname{Fix}(h)| \in \{0,1,\dots,\min\{r,s\}\}.$$

Using the Stirling–sum form (25) of Theorem 4.4 for $Q(g,h)$,

$$\begin{aligned}
\bar{Q}(r,s) &= \sum_{h\in C_s} Q(g,h) = \sum_{h\in C_s}\frac{1}{r^{\,n}}\sum_{t=1}^{\min\{j(h),\,n\}}\binom{j(h)}{t} S(n,t)\,\frac{t!}{(k-t)!} \\
&= \frac{1}{r^{\,n}}\sum_{t\ge 1} S(n,t)\,\frac{t!}{(k-t)!}\sum_{h\in C_s}\binom{j(h)}{t} \qquad \text{(swap finite sums; terms with } t>n \text{ vanish)} \\
&= \frac{1}{r^{\,n}}\sum_{t\ge 1} S(n,t)\,\frac{t!}{(k-t)!}\sum_{j=0}^{\min\{r,s\}}\binom{j}{t}\,|\{\,h\in C_s:\ j(h)=j\,\}| \\
&\qquad \text{(apply (9) with } A = C_s,\ B=\{0,1,\dots,\min\{r,s\}\}, \\
&\qquad \Phi(h) = j(h) = |\operatorname{Fix}(g)\cap\operatorname{Fix}(h)|,\ G(j) = \binom{j}{t}) \\
&= \frac{1}{r^{\,n}}\sum_{t\ge 1} S(n,t)\,\frac{t!}{(k-t)!}\sum_{j=0}^{\min\{r,s\}}\binom{j}{t}\binom{r}{j}\binom{k-r}{s-j}\cdot !\,(k-s) \qquad \text{(by Lemma 4.18 (b))} \\
&= \frac{1}{r^{\,n}}\sum_{t\ge 1} S(n,t)\,\frac{t!}{(k-t)!}\binom{r}{t}\binom{k-t}{s-t}\cdot !\,(k-s) \qquad \text{(by Lemma 4.24).}
\end{aligned}$$

Since $\binom{k-t}{s-t} = \dfrac{(k-t)!}{(s-t)!\,(k-s)!}$ and $\theta_{k,s} = \dfrac{!\,(k-s)}{(k-s)!}$, we simplify to (32).

For the expectation form (33), use the occupancy law

$$\mathbb{P}(R_{r,n}=m)=\frac{\binom{r}{m}S(n,m)m!}{r^{\,n}},\qquad 1\le m\le \min\{r,n\},$$

and the convention $(t!)^{-1}=0$ for $t<0$, to write

$$\bar{Q}(r,s)=\theta_{k,s}\sum_m \mathbb{P}(R_{r,n}=m)\,\frac{1}{(s-m)!}=\theta_{k,s}\,\mathbb{E}\Big[\frac{1}{(s-R_{r,n})!}\Big].$$

For the coefficient form (34), as in the proof of Theorem 4.4, use the classical identities

$$t!\,S(n,t)=n!\,[u^{\,n}](e^u-1)^t,\qquad \frac{1}{(s-t)!}=[z^{\,s-t}]\,e^z,$$

to rewrite the inner sum as

$$n!\,[u^n]\,[z^{\,s}]\,e^z\sum_{t=0}^{r}\binom{r}{t}\big(z(e^u-1)\big)^t=n!\,[u^n]\,[z^{\,s}]\,e^z\,\big(1+z(e^u-1)\big)^r,$$

and then multiply by $\theta_{k,s}/r^{\,n}$. □

Remark 4.26 (Alternative derivations of the lumped kernel)**.** Besides the proof via Theorem 4.4, we may also derive the closed forms for $\bar{Q}$ in two ways:

(i) Using Lemma 3.29 (19) and Lemma 4.18(c),

$$\bar{Q}(r,s)=\sum_{h\in C_s}Q(g,h)=\frac{1}{|X_g|}\sum_{x\in X_g}\frac{|G_x\cap C_s|}{|G_x|}=\frac{1}{|X_g|}\sum_{x\in X_g}\underbrace{\frac{!\,(k-s)}{(k-s)!}}_{=:\theta_{k,s}}\cdot\frac{1}{(s-r_x)!}=\theta_{k,s}\,\mathbb{E}\Big[\frac{1}{(s-R_{r,n})!}\Big],$$

where $\theta_{k,s}=!\,(k-s)/(k-s)!$ and $R_{r,n}$ is the number of distinct symbols in a uniform word from $[r]^n$ (i.e., $x\sim\mathrm{Unif}(X_g)=\mathrm{Unif}(\mathrm{Fix}(g)^n)$), with the convention $\frac{1}{(s-m)!}:=0$ when $m>s$. This immediately recovers (33).

(ii) From the two–stage description of one Q–step,

$$\bar{Q}(r,s)=\sum_{m=1}^{\min\{r,s,n\}}\underbrace{\frac{\binom{r}{m}\,S(n,m)\,m!}{r^{\,n}}}_{\mathbb{P}(R_{r,n}=m)}\;\underbrace{\frac{\binom{k-m}{s-m}\cdot !\,(k-s)}{(k-m)!}}_{\mathbb{P}(f(h)=s\mid R_{r,n}=m)}. \tag{35}$$

The first factor is the occupancy law (30) for $R_{r,n}$. Conditional on $R_{r,n}=m$, the stabilizer is $G_x\cong S_{k-m}$; a uniform $h\in G_x$ has exactly $s-m$ additional fixed points (choosing which $s-m$ are fixed, deranging the rest) with probability

$$\frac{\binom{k-m}{s-m}\cdot !\,(k-s)}{(k-m)!}=\frac{!\,(k-s)}{(k-s)!}\cdot\frac{1}{(s-m)!}=\theta_{k,s}\,\frac{1}{(s-m)!}.$$

Thus

$$\bar{Q}(r,s)=\theta_{k,s}\,\frac{1}{r^{\,n}}\sum_{m=1}^{\min\{r,s,n\}}\binom{r}{m}\,S(n,m)\,\frac{m!}{(s-m)!},$$

which recovers (32).

5. The Coordinate-Permutation Model

5.1. **Setting and Basic Properties.** Given integers $n \geq 1$ and $k \geq 1$, we consider $G = S_n$ acting on $X = [k]^n$ by permuting *coordinates*:

$$(g \cdot x)_i = x_{g^{-1}(i)} \qquad (g \in S_n,\ x \in [k]^n,\ i \in [n]).$$

For $x = (x_1, \dots, x_n) \in [k]^n$ and each symbol $a \in [k]$, define

$$I_a(x) := \{\, i \in [n] :\ x_i = a \,\}, \qquad m_a(x) := |I_a(x)|, \qquad \mathbf{m}(x) := \big(m_1(x), \dots, m_k(x)\big).$$

Then $\sum_{a=1}^{k} m_a(x) = n$. Here, for $x \in [k]^n$, $I_a(x)$ is the **index set** (or positions) of symbol a, $m_a(x)$ is the **multiplicity** (or count) of symbol a, and $\mathbf{m}(x)$ is the **histogram** (a.k.a. multiplicity vector or composition). For brevity, when x is fixed we suppress (x) and write I_a, m_a, and $\mathbf{m}$. Note that for $k = 2$, it's usually convenient to relabel the alphabet $\{1, 2\}$ as $\{0, 1\}$ (binary), then $m_1(x) = \sum_{i=1}^{n} x_i =: w(x)$ is the **Hamming weight** of x, $m_0(x) = n - w(x)$, and

$$I_1(x) = \{i : x_i = 1\}, \qquad I_0(x) = \{i : x_i = 0\}, \qquad \mathbf{m}(x) = (m_0(x), m_1(x)).$$

We write $r_x := |\{a \in [k] : m_a(x) > 0\}|$ for the **support size** of $\mathbf{m}(x)$ (i.e., number of distinct symbols appearing in x). For $g \in S_n$, define the **total cycle count** by $c(g) := \sum_{j=1}^{n} c_j(g)$, where $c_j(g)$ is the number of j-cycles in the cycle decomposition of g; note that $\sum_{j=1}^{n} j\, c_j(g) = n$.

Key facts.

- $|X_g| = k^{\,c(g)} > 0$, since a word must be constant on each cycle of g, with k choices per cycle. Thus in this model, we have $G^* = \{g :\ |X_g| > 0\} = S_n = G$.
- $G_x \cong S_{m_1(x)} \times \cdots \times S_{m_k(x)}$, and $|G_x| = \prod_{a=1}^{k} m_a(x)!$, since the stabilizer of x may freely permute positions *within* each symbol–class $I_a(x)$ and must preserve different classes setwise.
- Since the action only reorders coordinates, it preserves counts of each symbol: two words $x, y \in [k]^n$ are in the same orbit if and only if they have the same histogram: $\mathbf{m}(x) = \mathbf{m}(y)$, and the orbit count is $z = |X/G| = \left(\!\binom{k}{n}\!\right) = \binom{n+k-1}{k-1}$ (weak k–compositions of n).
- The dual stationary law is (as detailed in Section 5.3)

$$\pi_Q(g) \;=\; \frac{|X_g|}{|G|\, z} = \frac{k^{\,c(g)}}{n!\,\binom{n+k-1}{k-1}}.$$

Thus, in this case, the transition matrix reduces to

$$Q(g,h) \;=\; \frac{1}{|X_g|} \sum_{x \in X_g \cap X_h} \frac{1}{|G_x|} \;=\; \frac{1}{k^{c(g)}} \sum_{x \in X_g \cap X_h} \frac{1}{\prod_{a=1}^{k} m_a(x)!}.$$

Example 5.1 (Coordinate–permutation model: $k = 3$, $n = 4$)**.** *Action.* S_4 acts on $X = \{1, 2, 3\}^4$ by permuting coordinates.

1. Fixed–point set X_g.

Let $g = (12)(34) \in S_4$. A word $x \in \{1, 2, 3\}^4$ is fixed if and only if $x_1 = x_2$ and $x_3 = x_4$. Thus

$$X_g = \{(a, a, b, b) : a, b \in \{1, 2, 3\}\},$$
$$|X_g| = k^{\,c(g)} = 3^{\,2} = 9,$$

since $c(g) = 2$ cycles (blocks $\{1, 2\}$ and $\{3, 4\}$), each with 3 choices.

2. Stabilizer G_x.

Take $x = (2, 2, 1, 3)$. Its symbol counts are

$$\mathbf{m}(x) = (m_1, m_2, m_3) = (1, 2, 1),$$

so equal symbols may be permuted within their position–classes:

$$G_x \cong S_{m_1} \times S_{m_2} \times S_{m_3} \cong S_1 \times S_2 \times S_1,$$

$$|G_x| = \prod_{a=1}^{3} m_a! \;=\; 1! \cdot 2! \cdot 1! \;=\; 2.$$

Heuristics recap.

(1) If g has $c(g)$ cycles, each cycle must be constant, giving $|X_g| = k^{c(g)}$ (here $k = 3$).
(2) For x with histogram $\mathbf{m} = (m_1, m_2, m_3)$, equal symbols can be permuted within their position–classes: $|G_x| = \prod_a m_a!$ (here $1! \cdot 2! \cdot 1!$). In general, for $x \in \{1,2,3\}^n$ with histogram $\mathbf{m} = (m_1, m_2, m_3)$, $|G_x| = \prod_{a=1}^{3} m_a!$.

3. Orbits $\Longleftrightarrow$ histograms (multisets of symbols).

Under S_4, strings are grouped by their histogram (m_1, m_2, m_3) with $m_1 + m_2 + m_3 = n = 4$. For example, $x = (2, 2, 1, 3)$ lies in the orbit with counts $(1, 2, 1)$. In general, each orbit is determined by the unordered multiset of symbols (i.e., the histogram).

4. Bose–Einstein / stars–and–bars interpretation (twelvefold way).

Place $n = 4$ balls into $k = 3$ labeled boxes ("1", "2", "3") according to the coordinates (x_i), i.e., ball i is placed into box x_i. Permuting the ball labels (the S_4–action) erases which ball went where, leaving only the occupancies (m_1, m_2, m_3). The number of orbits equals the number of weak compositions of 4 into 3 parts:

$$z = \binom{4+3-1}{3-1} = \binom{6}{2} = 15.$$

Example 5.2 (Complete matrix decomposition for $k = 2,\ n = 3$ (S_3 acting on $\{0,1\}^3$))**.** *States.*

$$G^* = S_3 = \{e, (12), (13), (23), (123), (132)\}, \qquad X = \{0,1\}^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}.$$

(These orders are used throughout for rows/columns of Q, K, and for A, B.)

Forward leg $A : G^* \to X$ (6×8) and Backward leg $B : X \to G^*$ (8×6)

$$A = \frac{1}{8}\begin{pmatrix} 1&1&1&1&1&1&1&1\\ 2&2&0&0&0&0&2&2\\ 2&0&2&0&0&2&0&2\\ 2&0&0&2&2&0&0&2\\ 4&0&0&0&0&0&0&4\\ 4&0&0&0&0&0&0&4 \end{pmatrix} \qquad B = \frac{1}{6}\begin{pmatrix} 1&1&1&1&1&1\\ 3&3&0&0&0&0\\ 3&0&3&0&0&0\\ 3&0&0&3&0&0\\ 3&0&0&3&0&0\\ 3&0&3&0&0&0\\ 3&3&0&0&0&0\\ 1&1&1&1&1&1 \end{pmatrix}$$

Dual kernel $Q = AB$ (6×6) and Primal kernel $K = BA$ (8×8)

$$Q = \frac{1}{24}\begin{pmatrix} 10&4&4&4&1&1\\ 8&8&2&2&2&2\\ 8&2&8&2&2&2\\ 8&2&2&8&2&2\\ 4&4&4&4&4&4\\ 4&4&4&4&4&4 \end{pmatrix} \qquad K = \frac{1}{16}\begin{pmatrix} 5&1&1&1&1&1&1&5\\ 3&3&1&1&1&1&3&3\\ 3&1&3&1&1&3&1&3\\ 3&1&1&3&3&1&1&3\\ 3&1&1&3&3&1&1&3\\ 3&1&3&1&1&3&1&3\\ 3&3&1&1&1&1&3&3\\ 5&1&1&1&1&1&1&5 \end{pmatrix}$$

Spectra. $\mathrm{Spec}(Q) = \{1, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0, 0\}$, $\mathrm{Spec}(K) = \{1, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0, 0, 0, 0\}$.

Nonzero spectra coincide. (The triple eigenvalue $\frac{1}{4}$ matches the Diaconis–Lin–Ram [14] description.)

5.2. **Canonical Closed Forms for** $Q(g,h)$**.** In this section we further simplify the transition matrix formula above. As in the value-permutation case, our motivation is, now that we know $|X_g| = k^{c(g)}$, we would like to further compute $|X_g \cap X_h|$ for any $g,h \in S_n$.

Definition 5.3 (Joint orbits)**.** For $g,h \in S_n$, let $H := \langle g,h\rangle = \{\, g^{a_1}h^{b_1}g^{a_2}h^{b_2}\cdots g^{a_m}h^{b_m} : m \geq 0,\ a_i,b_i \in \mathbb{Z}\,\}$ (the subgroup generated by g,h), and let $O_1,\dots,O_s$ be the H–orbits on $[n]$ with sizes $b_j := |O_j|$ ($\sum_{j=1}^{s} b_j = n$).

For example, with $n=5$, take

$$g = (1\,2), \qquad h = (3\,4).$$

Then $\langle g,h\rangle$ acts separately on $\{1,2\}$ and on $\{3,4\}$ and fixes 5, so the joint orbits are

$$O_1 = \{1,2\}, \qquad O_2 = \{3,4\}, \qquad O_3 = \{5\},$$

hence $s=3$ and $(b_1,b_2,b_3) = (2,2,1)$.

Definition 5.4 (Common fixed-point set of a subgroup)**.** Let a subgroup $H \leq G$ act on X. Define

$$X_H \;:=\; \{\, x \in X :\ \sigma\cdot x = x \text{ for all } \sigma \in H \,\} \;=\; \bigcap_{\sigma\in H} X_\sigma.$$

Thus X_H is the **common fixed-point set** of H.

In the coordinate-permutation model, for $H = \langle g,h\rangle$, this is equivalent to

$$x \in X_H \quad\Longleftrightarrow\quad x \text{ is constant on each } H\text{-orbit in } [n],$$

i.e.,

$$X_H \;=\; \{\, x \in [k]^n :\ x \text{ is constant on each } O_i \,\},$$

where $\{O_1,\dots,O_s\}$ are the H–orbits on $[n]$.

We establish the following lemma to explicitly represent the set $X_g \cap X_h$ using the definitions above.

Lemma 5.5.

$$X_g \cap X_h = X_H.$$

That is, for $x \in X = [k]^n$, we have $x \in X_g \cap X_h$ if and only if x is constant on each joint orbit O_i of $H := \langle g,h\rangle$. Consequently,

$$|X_g \cap X_h| = k^s,$$

where s is the number of joint orbits.

Proof.

$$\begin{aligned} x \in X_g \cap X_h &\iff g\cdot x = x \text{ and } h\cdot x = x \\ &\overset{(*)}{\iff} \sigma\cdot x = x \text{ for all } \sigma \in \langle g,h\rangle \\ &\iff x_j = x_{\sigma(j)} \text{ for all } \sigma \in \langle g,h\rangle,\ \forall j \\ &\iff x_j = x_{j'} \text{ whenever } j,j' \text{ lie in the same } \langle g,h\rangle\text{-orbit.} \end{aligned}$$

Here $(*)$ holds because in a group action each g acts as a bijection on X, so $g\cdot x = x$ and $h\cdot x = x$ imply $g^{-1}\cdot x = x$ and $h^{-1}\cdot x = x$, and thus every product of $g^{\pm1}$ and $h^{\pm1}$ also fixes x, i.e. every $\sigma \in \langle g,h\rangle$.

Since the action is by coordinate permutation, the last condition says exactly that x is constant on each joint orbit O_i.

For the count, choose independently one of k symbols for each of the s orbits, yielding k^s possibilities. ∎

Next, let's show that a coloring $\phi : [s] \to [k]$ (one color per joint orbit $O_1, \dots, O_s$) determines a *unique* word $x(\phi) \in X_H$, and vice versa. Recall that $[k]^{[s]}$ denotes the set of all maps $\phi : [s] \to [k]$, which we identify with $[k]^s$ by sending each map ϕ to its coordinate vector $(\phi(1), \dots, \phi(s)) \in [k]^s$.

Lemma 5.6 (Orbit-coloring is a bijection)**.** *Let $\{O_1, \dots, O_s\}$ be the H–orbits on $[n]$. Define the map*

$$\Phi : X_H \to [k]^{[s]}, \qquad \Phi(x) = \phi := \big(\phi(1), \dots, \phi(s)\big),$$

by setting, for each $j \in \{1, \dots, s\}$,

$$\phi(j) \;:=\; x_{i_j}, \qquad \text{where } i_j \text{ is any (hence every) coordinate in } O_j.$$

(Equivalently: $\phi(j)$ is the common symbol of x on the block O_j.) Then Φ is a bijection. Its inverse $\Psi : [k]^{[s]} \to X_H$ is given in coordinates by

$$\Psi(\phi) = x(\phi) := (x_1, \dots, x_n) \quad \text{with } x_i := \phi(j) \quad \text{whenever } i \in O_j, \quad j = 1, \dots, s.$$

Proof. Well-defined. If $x \in X_H$ then x is constant on each O_j, so $\phi(j)$ does not depend on the choice of $i_j \in O_j$, and Φ is well-defined. Conversely, for any $\phi : [s] \to [k]$ the definition $x_i(\phi) = \phi(j)$ for $i \in O_j$ makes $x(\phi)$ constant on each O_j, hence $x(\phi) \in X_H$ and Ψ is well-defined.

Mutual inverses. For any $\phi : [s] \to [k]$, by construction $\Phi(\Psi(\phi)) = \phi$. For any $x \in X_H$, choosing $i_j \in O_j$ gives $\Psi(\Phi(x)) = x$. Therefore Φ is bijective with inverse Ψ. □

Just as $m_a(x)$ is the actual count in a specific word x, we define:

Definition 5.7 (Orbit–mass vector: deterministic and random forms)**.** Let $O_1, \dots, O_s$ be the joint orbits of H on $[n]$ with sizes $b_1, \dots, b_s$.

Deterministic (given a coloring). For a coloring $\phi : [s] \to [k]$, define the **orbit–mass counts**

$$M_a(\phi) := \sum_{j:\, \phi(j) = a} b_j = \sum_{j=1}^{s} b_j \, \mathbf{1}_{\{\phi(j) = a\}}, \qquad a = 1, \dots, k.$$

i.e., the total number of coordinates painted color a by ϕ. The **orbit–mass vector** is

$$\mathbf{M}(\phi) := \big(M_1(\phi), \dots, M_k(\phi)\big), \qquad \sum_{a=1}^{k} M_a(\phi) = n.$$

Random (for expectation forms). Let $U_1, \dots, U_s \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}([k])$ and define

$$M_a := \sum_{j=1}^{s} b_j \, \mathbf{1}_{\{U_j = a\}}, \qquad a = 1, \dots, k, \quad \text{so} \quad \sum_{a=1}^{k} M_a = n.$$

Then $(M_1, \dots, M_k)$ has the same distribution as $\mathbf{M}(\phi)$ when ϕ is chosen uniformly from $[k]^{[s]}$.

Remark 5.8 (Orbit–mass equals histogram)**.** For the word $x(\phi)$ uniquely determined by the coloring $\phi : [s] \to [k]$, the usual *multiplicity counts*

$$m_a\big(x(\phi)\big) := \#\{\, i \in [n] : x(\phi)_i = a \,\}$$

agree with the *orbit–mass counts*

$$M_a(\phi) := \sum_{j:\, \phi(j) = a} b_j, \qquad a = 1, \dots, k.$$

Indeed, each orbit O_j with $\phi(j) = a$ contributes b_j coordinates of color a, so $m_a(x(\phi)) = M_a(\phi)$ for every a. Thus

$$\mathbf{m}\big(x(\phi)\big) = \big(m_1(x(\phi)), \dots, m_k(x(\phi))\big) = \big(M_1(\phi), \dots, M_k(\phi)\big) = \mathbf{M}(\phi).$$

Theorem 5.9 (Closed forms for $Q(g,h)$)**.** *With notation as above, the following are equivalent closed forms:*

$$Q(g,h) = \frac{1}{k^{\,c(g)}} \sum_{\phi:[s]\to[k]} \prod_{a=1}^{k} \frac{1}{M_a(\phi)!}, \tag{36 — colorings sum}$$

$$Q(g,h) = \frac{k^{-c(g)}}{n!} \sum_{\phi:[s]\to[k]} \binom{n}{M_1(\phi),\dots,M_k(\phi)}, \tag{37 — multinomial rewrite}$$

$$Q(g,h) = k^{\,s-c(g)}\ \mathbb{E}\Bigg[\prod_{a=1}^{k} \frac{1}{M_a!}\Bigg], \tag{38 — expectation form}$$

$$Q(g,h) = \frac{1}{k^{\,c(g)}}\,[z_1^0\cdots z_k^0]\,\exp\Big(\sum_{a=1}^{k} z_a\Big)\ P^-_{\mathbf{b}}(z_1,\dots,z_k), \tag{39 — coefficient form}$$

$$Q(g,h) = k^{-c(g)}\,\Big[z_1^0\cdots z_k^0\Big]\,\frac{(z_1+\cdots+z_k)^n}{n!}\,P^-_{\mathbf{b}}(z_1,\dots,z_k), \tag{40 — alternative coefficient form}$$

with

$$P^-_{\mathbf{b}}(z_1,\dots,z_k) := \prod_{j=1}^{s}\ \sum_{a=1}^{k} z_a^{-\,b_j} = \prod_{j=1}^{s}\big(z_1^{-\,b_j}+z_2^{-\,b_j}+\cdots+z_k^{-\,b_j}\big).$$

Proof. By Lemma 5.5,

$$Q(g,h) = \frac{1}{k^{c(g)}} \sum_{x\in X_H} \frac{1}{\prod_{a=1}^{k} m_a(x)!}.$$

Let $\Phi : X_H \to [k]^{[s]}$ be the orbit–coloring bijection (Lemma 5.6), and write $\phi := \Phi(x)$ for $x \in X_H$. Then, by Remark 5.8,

$$m_a(x) = M_a(\phi) := \sum_{j:\,\phi(j)=a} b_j, \qquad a = 1,\dots,k.$$

Applying the re–indexing principle (9) with $A := X_H$, $B := [k]^{[s]}$, and

$$F(x)\ :=\ \prod_{a=1}^{k} \frac{1}{m_a(x)!} = \prod_{a=1}^{k} \frac{1}{M_a(\Phi(x))!},$$

which depends only on $\phi = \Phi(x)$, yields (36). Using $\prod_a m_a! = n!/\binom{n}{m_1,\dots,m_k}$ gives (37).

For (38), let $(M_1,\dots,M_k)$ be the random orbit–mass vector from Definition 5.7. Since $(M_1,\dots,M_k)$ has the same law as $\mathbf{M}(\phi) = (M_1(\phi),\dots,M_k(\phi))$ when ϕ is chosen uniformly from $[k]^{[s]}$, we have, for any function F of the totals,

$$\frac{1}{k^s} \sum_{\phi:[s]\to[k]} F\big(M_1(\phi),\dots,M_k(\phi)\big) = \mathbb{E}\big[F(M_1,\dots,M_k)\big].$$

Apply this with $F(\mathbf{M}) = \prod_{a=1}^{k} 1/M_a!$ to get

$$Q(g,h) = \frac{1}{k^{\,c(g)}}\cdot k^s\,\mathbb{E}\Big[\prod_{a=1}^{k}\frac{1}{M_a!}\Big] = k^{\,s-c(g)}\ \mathbb{E}\Bigg[\prod_{a=1}^{k}\frac{1}{M_a!}\Bigg],$$

which is (38).

To deduce the coefficient form, first note that

$$\sum_{\phi:[s]\to[k]} z_1^{-M_1(\phi)}\cdots z_k^{-M_k(\phi)} = \sum_{\phi:[s]\to[k]} \prod_{a=1}^{k} z_a^{-\sum_{j=1}^{s} b_j \mathbf{1}_{\{\phi(j)=a\}}} \quad (\text{definition } M_a(\phi) = \sum_{j:\,\phi(j)=a} b_j)$$

$$= \sum_{\phi:[s]\to[k]} \prod_{j=1}^{s} z_{\phi(j)}^{-b_j} \quad (\text{collect exponents by } j\text{: each } b_j \text{ goes to the chosen color } \phi(j))$$

$$= \prod_{j=1}^{s} \sum_{a=1}^{k} z_a^{-b_j} \quad (\text{distributivity: } \sum_{\phi}\prod_{j} f_j(\phi(j)) = \prod_{j}\sum_{a} f_j(a) \text{ with } f_j(a) = z_a^{-b_j})$$

$$= P_{\mathbf{b}}^{-}(z_1,\dots,z_k). \tag{41}$$

Start from (36),

$$Q(g,h) = \frac{1}{k^{c(g)}} \sum_{\phi:[s]\to[k]} \prod_{a=1}^{k} \left[z_a^{M_a(\phi)}\right] e^{z_a} \quad (\text{basic coeff-of-exp identity } 1/m! = [z^m]e^z)$$

$$= \frac{1}{k^{c(g)}} \sum_{\phi:[s]\to[k]} \left[z_1^{M_1(\phi)}\cdots z_k^{M_k(\phi)}\right] \exp\Big(\sum_{a=1}^{k} z_a\Big) \quad (\text{combine the } k \text{ factors})$$

$$= \frac{1}{k^{c(g)}} \left[z_1^0\cdots z_k^0\right] \exp\Big(\sum_{a=1}^{k} z_a\Big) \sum_{\phi:[s]\to[k]} z_1^{-M_1(\phi)}\cdots z_k^{-M_k(\phi)} \quad ([z^\alpha]F = [z^0]\, F\, z^{-\alpha})$$

$$= \frac{1}{k^{c(g)}} \left[z_1^0\cdots z_k^0\right] \exp\Big(\sum_{a=1}^{k} z_a\Big) P_{\mathbf{b}}^{-}(z_1,\dots,z_k) \quad (\text{by (41) above}),$$

which yields (39).

Alternatively, starting from (37),

$$Q(g,h) = \frac{k^{-c(g)}}{n!} \sum_{\phi:[s]\to[k]} \binom{n}{M_1(\phi),\dots,M_k(\phi)} \quad (\text{multinomial rewrite})$$

$$= \frac{k^{-c(g)}}{n!} \sum_{\phi:[s]\to[k]} \left[z_1^{M_1(\phi)}\cdots z_k^{M_k(\phi)}\right] (z_1+\cdots+z_k)^n \quad (\text{coefficient–extraction identity})$$

$$= \frac{k^{-c(g)}}{n!} \left[z_1^0\cdots z_k^0\right] (z_1+\cdots+z_k)^n \sum_{\phi:[s]\to[k]} z_1^{-M_1(\phi)}\cdots z_k^{-M_k(\phi)} \quad ([z^\alpha]F = [z^0]F\, z^{-\alpha})$$

$$= k^{-c(g)} \left[z_1^0\cdots z_k^0\right] \frac{(z_1+\cdots+z_k)^n}{n!} P_{\mathbf{b}}^{-}(z_1,\dots,z_k) \quad (\text{by (41) above}),$$

which yields (40). □

Remark 5.10. (i) If H is transitive ($s = 1$, $b_1 = n$), then (36) reduces to $Q(g,h) = k^{1-c(g)}/n!$, since each of the k colorings contribute $1/n!$.

(ii) If $h = e$, then the orbits are the cycles of g ($s = c(g)$) and (38) becomes $Q(g,e) = \mathbb{E}\big[1/\prod_{a=1}^{k} M_a!\big]$ for a uniform coloring of the $c(g)$ cycles.

(iii) The last two displays (39) and (40) are equal. Indeed, every monomial of $P^-_{\mathbf{b}}(z)$ has total degree $-n$ (since $\sum_{j=1}^s b_j = n$), and the general identity

$$[z_1^0 \cdots z_k^0]\ \exp\Big(\sum_{a=1}^k z_a\Big)\, F(z) \;=\; \frac{1}{n!}\,[z_1^0 \cdots z_k^0]\ \Big(\sum_{a=1}^k z_a\Big)^n F(z)$$

holds for any F homogeneous of total degree $-n$. Applying this with $F = P^-_{\mathbf{b}}(z)$ yields (39) = (40).

(iv) We can rewrite (39) (and (40)) using positive powers in the product. Just invert the variables $z_a \mapsto z_a^{-1}$ (which preserves the constant term), giving the equivalent forms

$$Q(g,h) = \frac{1}{k^{c(g)}}\,[z_1^0 \cdots z_k^0]\ \exp\Big(\sum_{a=1}^k z_a^{-1}\Big)\, P_{\mathbf{b}}(z_1,\dots,z_k),$$

$$Q(g,h) = \frac{1}{k^{c(g)}}\,[z_1^0 \cdots z_k^0]\ \frac{(z_1^{-1}+\cdots+z_k^{-1})^n}{n!}\, P_{\mathbf{b}}(z_1,\dots,z_k).$$

with

$$P_{\mathbf{b}}(z) := \prod_{j=1}^s \big(z_1^{b_j} + \cdots + z_k^{b_j}\big).$$

In the binary case $k=2$, a coloring $\phi : [s] \to [2]$ is the same as a subset $J \subseteq [s]$ (those orbits colored "1"). Writing $S = \sum_{j\in J} b_j$, the orbit–mass vector is $(M_1, M_2) = (S,\ n-S)$, and this theorem immediately reduces to the following.

Corollary 5.11 (Binary specialization: closed forms for $Q(g,h)$)**.** *For $k=2$, the following are equal:*

$$\begin{aligned}
Q(g,h) &= 2^{-c(g)} \sum_{J\subseteq[s]} \frac{1}{S!\,(n-S)!} = \frac{2^{-c(g)}}{n!} \sum_{J\subseteq[s]} \binom{n}{S} \\
&= 2^{\,s-c(g)}\,\mathbb{E}\left[\frac{1}{S!\,(n-S)!}\right] \\
&= \frac{1}{2^{\,c(g)}}\,[z_1^0 z_2^0]\ \exp(z_1+z_2)\ \prod_{j=1}^s \big(z_1^{-\,b_j} + z_2^{-\,b_j}\big) \\
&= \frac{2^{-c(g)}}{n!}\,[w^0]\,(1+w)^n \prod_{j=1}^s \big(1+w^{-b_j}\big) = \frac{2^{-c(g)}}{n!}\,[w^{\,n}]\,(1+w)^n \prod_{j=1}^s (1+w^{b_j}),
\end{aligned}$$

where the expectation is over independent Bernoulli(1/2) *choices of each orbit.*

Proof. To see the second to last equality (the coefficient form), starting from (39),

$$Q(g,h) = 2^{-c(g)}\,[z_1^0 z_2^0]\ e^{z_1+z_2} \prod_{j=1}^s \big(z_1^{-b_j} + z_2^{-b_j}\big),$$

set $z_1 = wz_2$ so that $[z_1^0 z_2^0] = [w^0 z_2^0]$ and

$$\begin{aligned}
\prod_{j=1}^{s}\bigl(z_1^{-b_j} + z_2^{-b_j}\bigr) &= \prod_{j=1}^{s}\bigl((wz_2)^{-b_j} + z_2^{-b_j}\bigr) && \text{(substitute } z_1 = wz_2\text{)}\\
&= \prod_{j=1}^{s} z_2^{-b_j}\,\bigl(w^{-b_j}+1\bigr) && \text{(factor out } z_2^{-b_j} \text{ in each factor)}\\
&= \left(\prod_{j=1}^{s} z_2^{-b_j}\right)\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(commute products)}\\
&= z_2^{-\sum_{j=1}^{s} b_j}\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(combine the } z_2 \text{ powers)}\\
&= z_2^{-n}\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(since } \textstyle\sum_{j=1}^{s} b_j = n\text{).}
\end{aligned}$$

Then

$$\begin{aligned}
Q(g,h) &= 2^{-c(g)}\,[w^0 z_2^0]\;e^{(w+1)z_2}\,z_2^{-n}\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(after the substitution } z_1 = wz_2\text{)}\\
&= 2^{-c(g)}\,[w^0]\left([z_2^0]\;e^{(w+1)z_2}\,z_2^{-n}\right)\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(separate } [w^0] \text{ and } [z_2^0]\text{)}\\
&= 2^{-c(g)}\,[w^0]\left([z_2^n]\;e^{(w+1)z_2}\right)\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(since } [z_2^0](Fz_2^{-n}) = [z_2^n]F\text{)}\\
&= 2^{-c(g)}\,[w^0]\,\frac{(w+1)^n}{n!}\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(coefficient of } z_2^n \text{ in } e^{(w+1)z_2} \text{ is } (w+1)^n/n!\text{)}\\
&= \frac{2^{-c(g)}}{n!}\,[w^0]\,(1+w)^n\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr).
\end{aligned}$$

To further deduce the last equality,

$$\begin{aligned}
Q(g,h) &= \frac{2^{-c(g)}}{n!}\,[w^0]\,(1+w)^n\prod_{j=1}^{s}\bigl(1+w^{-b_j}\bigr) && \text{(binary constant-term form above)}\\
&= \frac{2^{-c(g)}}{n!}\,[w^0]\,(1+w)^n\prod_{j=1}^{s} w^{-b_j}\bigl(1+w^{b_j}\bigr) && \text{(write } 1+w^{-b_j} = w^{-b_j}(1+w^{b_j})\text{)}\\
&= \frac{2^{-c(g)}}{n!}\,[w^0]\,(1+w)^n\,w^{-n}\prod_{j=1}^{s}\bigl(1+w^{b_j}\bigr) && \text{(combine the } w \text{ powers and use } \textstyle\sum_j b_j = n\text{)}\\
&= \frac{2^{-c(g)}}{n!}\,[w^n]\,(1+w)^n\prod_{j=1}^{s}\bigl(1+w^{b_j}\bigr) && \text{(shift coefficient: } [w^0]F\,w^{-n} = [w^n]F\text{).}
\end{aligned}$$

Finally, note that the expectation form in the statement is the specialization of (38) to $k = 2$. In this case we take

$$S := M_1 = \sum_{j=1}^{s} b_j \, \mathbf{1}_{\{U_j = 1\}},$$

where $U_j \sim \mathrm{Unif}(\{1,2\})$ are i.i.d.; thus $\mathbf{1}_{\{U_j=1\}} \sim \mathrm{Bernoulli}(1/2)$, so the expectation is over independent Bernoulli(1/2) choices of each orbit. □

Example 5.12 (Transitive source gives a flat row)**.** If g is an n–cycle, then $H = \langle g, h\rangle$ is transitive for every $h \in S_n$. By Remark 5.10 (i),

$$Q(g,h) = \frac{k^{\,1-c(g)}}{n!} = \frac{1}{n!} \qquad (\text{since } c(g) = 1).$$

Thus the entire row of Q at g is flat (constant in h).

Example 5.13 (Two joint orbits)**.** Suppose $H = \langle g, h\rangle$ has exactly two orbits O_1, O_2 with sizes $b_1 = a$ and $b_2 = n - a$, where $1 \le a \le n-1$ (by symmetry we may assume $1 \le a \le \lfloor n/2 \rfloor$). Here $s = 2$, so the colorings are maps $\phi : \{1,2\} \to [k]$, and

$$M_c(\phi) = a\, \mathbf{1}_{\{\phi(1)=c\}} + (n-a)\, \mathbf{1}_{\{\phi(2)=c\}} \qquad (c \in [k]).$$

There are two types of colorings ϕ:

- *Both orbits have the same color:* $\phi(1) = \phi(2) = c$. There are k such colorings (one for each $c \in [k]$), and in this case

  $$M_c(\phi) = n, \qquad M_d(\phi) = 0 \;\; (d \ne c),$$

  so

  $$\prod_{a=1}^{k} \frac{1}{M_a(\phi)!} = \frac{1}{n!}.$$

- *The two orbits have distinct colors:* $\phi(1) = c$, $\phi(2) = d$ with $c \ne d$. There are $k(k-1)$ such colorings (ordered pairs (c,d) with $c \ne d$), and in this case

  $$M_c(\phi) = a, \quad M_d(\phi) = n-a, \quad M_e(\phi) = 0 \;\; (e \notin \{c,d\}),$$

  so

  $$\prod_{a=1}^{k} \frac{1}{M_a(\phi)!} = \frac{1}{a!\,(n-a)!}.$$

Plugging these into (36),

$$Q(g,h) = \frac{1}{k^{c(g)}} \sum_{\phi:\{1,2\}\to[k]} \prod_{a=1}^{k} \frac{1}{M_a(\phi)!} = \frac{1}{k^{c(g)}} \left(k \cdot \frac{1}{n!} + k(k-1) \cdot \frac{1}{a!\,(n-a)!} \right).$$

Rewriting,

$$Q(g,h) = \frac{k^{-c(g)}}{n!} \left(k + k(k-1) \binom{n}{a} \right) = \frac{k^{\,1-c(g)}}{n!} \left[1 + (k-1) \binom{n}{a} \right].$$

Corollary 5.14 (Uniform pointwise floor)**.** *For all $g, h \in S_n$,*

$$Q(g,h) \;\ge\; \frac{k^{\,1-c(g)}}{n!},$$

with equality if and only if $\langle g, h\rangle$ is transitive on $[n]$.

Proof. From (36), $Q(g,h) = k^{-c(g)} \sum_{\phi:[s]\to[k]} \prod_{a=1}^{k} 1/M_a(\phi)!$. For each constant coloring $\phi \equiv a$, the term equals $1/n!$; there are k such colorings, yielding the lower bound.

If $s = 1$ (transitive), by Remark 5.10 (i) above, equality holds. If $s \geq 2$, there exist non-constant colorings, and each contributes a strictly positive term (since $M_a(\phi) \geq 0$ and $\sum_a M_a(\phi) = n$), so the sum is strictly larger than $k \cdot (1/n!)$, giving strict inequality. □

Finally, we deduce the closed-form formula for a special transition from e to a single t-cycle.

Theorem 5.15 (Identity to a single t–cycle)**.** *Let $g = e$ and let $h \in S_n$ be a single t–cycle with $2 \leq t \leq n$. Write $m := n - t$. For integers $p > 0$ and $s \geq 0$, set*

$$\kappa_p(s) \;:=\; \sum_{\substack{u_1+\cdots+u_p=s\\ u_i\geq 0}} \prod_{i=1}^{p} \frac{1}{(u_i!)^2}$$

Then for $k \geq 2$,

$$Q(e,h) = k^{1-n} \sum_{j=0}^{m} \binom{m}{j} \frac{(m-j)!}{(t+j)!}\, \kappa_{k-1}(m-j), \tag{42}$$

$$= k^{1-n} \sum_{r=0}^{m} \binom{m}{r} \frac{r!}{(n-r)!}\, \kappa_{k-1}(r), \qquad (r = m-j). \tag{43}$$

Proof. For $g = e$ we have $c(g) = n$. The joint orbits of $H = \langle h \rangle$ on $[n]$ are one block of size t and $m = n - t$ singletons:

$$(b_1,\ldots,b_s) = (t, \underbrace{1,\ldots,1}_{m \text{ times}}), \qquad s = 1 + m.$$

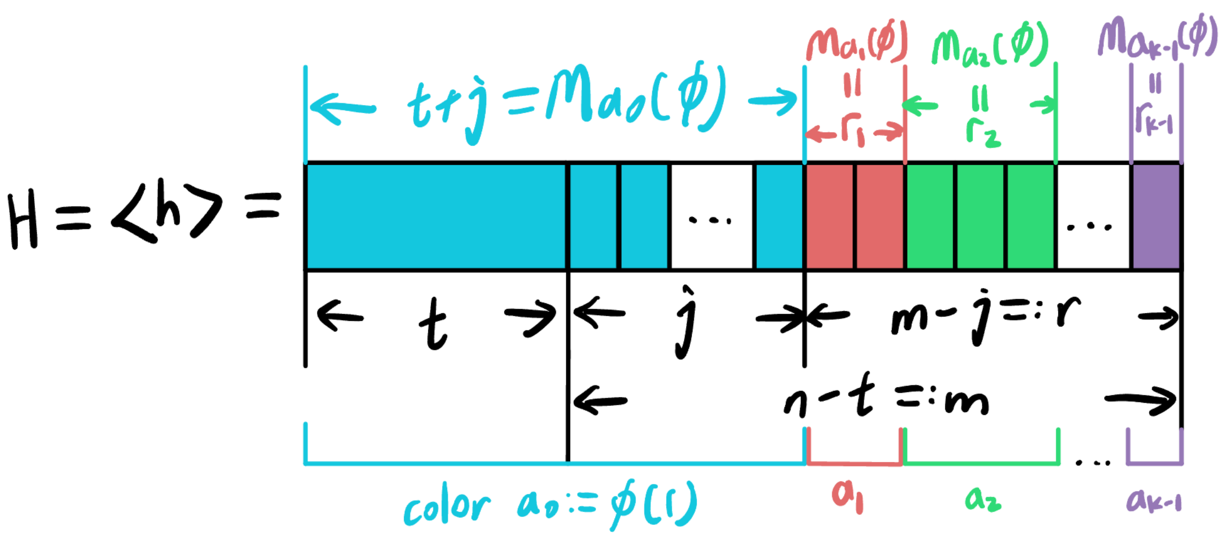

FIGURE 2. Orbits of $\langle h \rangle$ and color counts $M_a(\phi)$.

Start from the colorings form (36):

$$Q(e,h) \;=\; \frac{1}{k^n} \sum_{\phi:[s]\to[k]} \prod_{a=1}^{k} \frac{1}{M_a(\phi)!}.$$

Define the re-indexing map

$$\Phi:\ [k]^{[s]} \to \underbrace{[k]}_{a_0} \times \underbrace{\{0,1,\dots,m\}}_{j} \times \underbrace{\Big\{(r_1,\dots,r_{k-1}) \in \mathbb{N}_0^{k-1} : \textstyle\sum_{\ell=1}^{k-1} r_\ell = r\Big\}}_{r\text{–compositions}},$$

by setting, for a coloring ϕ:

- $a_0 := \phi(1)$ (the color of the t–block);
- $j := \#\{\text{singletons colored } a_0\} \in \{0,\dots,m\}$;
- $r := m-j \in \{0,\dots,m\}$ and $(r_1,\dots,r_{k-1})$ the counts by the remaining $k-1$ colors (so $\sum_{\ell=1}^{k-1} r_\ell = r$).

Recall that $M_a(\phi)$ counts the total number of coordinates painted with color a under the coloring ϕ. Here the color a_0 is used on the t coordinates in the big orbit and on j singletons, so $M_{a_0}(\phi) = t+j$, while each non-a_0 color a_ℓ is used on exactly r_ℓ singletons, so $M_{a_\ell}(\phi) = r_\ell$ for $\ell = 1,\dots,k-1$. Any remaining color a is unused, so $M_a(\phi) = 0$ (no coordinate is painted with color a). Thus

$$\begin{aligned}
\prod_{a=1}^{k} \frac{1}{M_a(\phi)!} &= \Bigg(\prod_{\substack{a\in[k]\\ a=a_0}} \frac{1}{M_a(\phi)!}\Bigg)\Bigg(\prod_{\substack{a\in[k]\\ a\neq a_0,\ \text{used}}} \frac{1}{M_a(\phi)!}\Bigg)\Bigg(\prod_{\substack{a\in[k]\\ \text{unused}}} \frac{1}{M_a(\phi)!}\Bigg)\\
&= \frac{1}{M_{a_0}(\phi)!} \prod_{\ell=1}^{k-1} \frac{1}{M_{a_\ell}(\phi)!} \prod_{\substack{a\in[k]\\ \text{unused}}} \frac{1}{0!} \qquad \text{(fix an order } a_1,\dots,a_{k-1} \text{ of the non-}a_0\text{ colors)}\\
&= \frac{1}{(t+j)!} \prod_{\ell=1}^{k-1} \frac{1}{r_\ell!} \cdot 1 \qquad \text{(since } M_{a_0}(\phi) = t+j,\ M_{a_\ell}(\phi) = r_\ell, \text{ and } 1/0! = 1)\\
&= \frac{1}{(t+j)!} \prod_{\ell=1}^{k-1} \frac{1}{r_\ell!}.
\end{aligned}$$

which depends only on $\Phi(\phi) = (a_0, j, r_\bullet)$, where $r_\bullet := (r_1,\dots,r_{k-1})$ is the $(k-1)$-tuple of counts of the non-a_0 colors. By the re-indexing principle (9),

$$\sum_{\phi:[s]\to[k]} \prod_{a=1}^{k} \frac{1}{M_a(\phi)!} = \sum_{(a_0,j,r_\bullet)} \frac{1}{(t+j)!} \prod_{\ell=1}^{k-1} \frac{1}{r_\ell!} \left|\Phi^{-1}(a_0,j,r_\bullet)\right|.$$

Given $(a_0, j, r_\bullet)$, the fiber size is

$$\left|\Phi^{-1}(a_0,j,r_\bullet)\right| = \binom{m}{j}\binom{r}{r_1,\dots,r_{k-1}} = \binom{m}{j}\frac{r!}{r_1!\cdots r_{k-1}!},$$

since we choose which j singletons take color a_0, then distribute the remaining r singletons to realize $(r_1,\dots,r_{k-1})$. Therefore

$$\sum_{\phi:[s]\to[k]} \prod_{a=1}^{k} \frac{1}{M_a(\phi)!} = \sum_{a_0\in[k]} \sum_{j=0}^{m} \binom{m}{j} \frac{1}{(t+j)!} \sum_{\substack{r_1+\cdots+r_{k-1}=r\\ r=m-j}} \frac{r!}{r_1!\cdots r_{k-1}!} \prod_{\ell=1}^{k-1} \frac{1}{r_\ell!}.$$

The inner sum is

$$\sum_{\substack{r_1+\cdots+r_{k-1}=r\\ r_i\geq 0}} \frac{r!}{r_1!\cdots r_{k-1}!} \prod_{\ell=1}^{k-1} \frac{1}{r_\ell!} = r! \sum_{\substack{r_1+\cdots+r_{k-1}=r\\ r_i\geq 0}} \prod_{\ell=1}^{k-1} \frac{1}{(r_\ell!)^2} = r!\,\kappa_{k-1}(r),$$

by the definition

$$\kappa_{k-1}(r) := \sum_{\substack{u_1+\cdots+u_{k-1}=r \\ u_i\ge 0}} \prod_{i=1}^{k-1} \frac{1}{(u_i!)^2}.$$

Plugging back and summing over $a_0 \in [k]$ yields the factor k, hence

$$Q(e,h) = \frac{1}{k^n}\cdot k \sum_{j=0}^{m} \binom{m}{j} \frac{(m-j)!}{(t+j)!}\,\kappa_{k-1}(m-j) \;=\; k^{\,1-n} \sum_{j=0}^{m} \binom{m}{j} \frac{(m-j)!}{(t+j)!}\,\kappa_{k-1}(m-j),$$

which is (42). Reparametrizing with $r = m - j$ gives (43). □

Remark 5.16 (Two edge cases)**.** (i) If $t = n$ (transitive block, $m = 0$), in (42) the sum has only $j = 0$. Using $\kappa_{k-1}(0) = 1$,

$$Q(e,h) = k^{1-n}\cdot\frac{0!}{n!}\cdot 1 = \frac{k^{1-n}}{n!}.$$

This matches Remark 5.10 (i).

(ii) If $t = n-1$ ($m = 1$), then in (42) the sum has $j = 0, 1$. Use $\kappa_{k-1}(1) = k-1$ and $\kappa_{k-1}(0) = 1$:

$$Q(e,h) = k^{1-n}\left[\binom{1}{0}\frac{1!}{(n-1)!}(k-1) + \binom{1}{1}\frac{0!}{n!}\cdot 1\right] = k^{1-n}\left(\frac{k-1}{(n-1)!} + \frac{1}{n!}\right).$$

For the binary case, if $k = 2$ and $m = n - t$, then $\kappa_1(r) = 1/(r!)^2$, so Theorem 5.15 reduces to the following result.

Corollary 5.17 (Binary specialization: identity to a single t–cycle)**.** *For $k = 2$,*

$$Q(e,h) = \frac{2^{\,1-n}}{n!}\binom{2n-t}{n}.$$

Equivalently, in coefficient form, with $m := n - t$,

$$Q(e,h) = \frac{2^{\,1-n}}{n!}\,[z^{\,m}]\,(1+z)^{\,2n-t}.$$

Proof. With $m := n - t$, by (43),

$$Q(e,h) = 2^{\,1-n}\sum_{r=0}^{m}\binom{m}{r}\frac{1}{r!\,(n-r)!} = \frac{2^{\,1-n}}{n!}\sum_{r=0}^{m}\binom{m}{r}\binom{n}{r} = \frac{2^{\,1-n}}{n!}\sum_{j=0}^{m}\binom{m}{j}\binom{n}{t+j},$$

where the last equality is by the binomial symmetry $\binom{n}{r} = \binom{n}{n-r}$ and the change of variables $j = m - r$ (using $n = t + m$).

By Chu–Vandermonde in the shifted form $\sum_{j=0}^{m}\binom{m}{j}\binom{n}{t+j} = \binom{n+m}{t+m}$, we obtain

$$\sum_{j=0}^{m}\binom{m}{j}\binom{n}{t+j} = \binom{n+m}{t+m} = \binom{2n-t}{n}.$$

Substitute back to get the stated closed form.

The coefficient form follows from the closed form above. □

Example 5.18 ($k = 3$, $n = 4$, $g = e$, $h = (123)$)**.** Let $G = S_4$ act on $X = [3]^4$ by permuting coordinates. Take

$$g = e, \qquad h = (1\,2\,3) \in S_4, \qquad H := \langle g, h\rangle = \langle h\rangle = \{\,e,\ (1\,2\,3),\ (1\,3\,2)\}\ .$$

The H–orbits on [4] are

$$O_1 = \{1,2,3\}, \quad O_2 = \{4\}, \qquad s = 2, \quad (b_1, b_2) = (3,1).$$

- **Common fixed set X_H and the bijection Φ.** By Definition 5.4,

$$X_H = \{x \in [3]^4 : \ x \text{ is constant on each } O_j\} = \{(a,a,a,b) : \ a,b \in [3]\}, \quad |X_H| = 3^2 = 9.$$

In this case $X_g \cap X_h = X_h = X_H$ (since $X_g = X$ for $g = e$), satisfying Lemma 5.5.

As in Lemma 5.6, define $\Phi : X_H \to [3]^{[2]}$ by $\Phi(x) = (\phi(1), \phi(2)) := (x_1, x_4)$, i.e. $\phi(1)$ is the common value on O_1 and $\phi(2)$ the (free) value on O_2. The inverse $\Psi : [3]^{[2]} \to X_H$ is

$$\Psi(u,v) = x(u,v) = (u,u,u,v).$$

Clearly $\Phi \circ \Psi = \mathrm{id}$ and $\Psi \circ \Phi = \mathrm{id}$, so Φ is a bijection.

- **Orbit–mass counts $M_a(\phi)$ and multiplicities $m_a(x)$.** For $\phi = (u,v) \in [3]^2$ and $a \in [3]$,

$$M_a(\phi) = \sum_{j=1}^{2} b_j \, \mathbf{1}_{\{\phi(j)=a\}} = 3\,\mathbf{1}_{\{u=a\}} + 1\,\mathbf{1}_{\{v=a\}}.$$

Obviously, for $x(\phi) = (u,u,u,v)$ we have

$$\mathbf{m}\big(x(\phi)\big) = \big(m_1(x(\phi)), m_2(x(\phi)), m_3(x(\phi))\big) = \big(M_1(\phi), M_2(\phi), M_3(\phi)\big) =: \mathbf{M}(\phi).$$

For example, if $\phi = (u,v) =: (1,3)$, then

$$(M_1, M_2, M_3) = (3, 0, 1),$$

and $x(\phi) = (1,1,1,3)$ has

$$\mathbf{m}\big(x(\phi)\big) = \big(m_1, m_2, m_3\big) = (3,0,1) = \big(M_1(\phi), M_2(\phi), M_3(\phi)\big) =: \mathbf{M}(\phi).$$

- **Colorings sum (36) evaluated.** Here $c(g) = c(e) = n = 4$, so

$$Q(e,h) = \frac{1}{3^4} \sum_{\phi \in [3]^2} \prod_{a=1}^{3} \frac{1}{M_a(\phi)!}.$$

There are 3 constant colorings $(u,v) = (c,c)$, each contributing $\prod_a 1/M_a(\phi)! = 1/4!$, and 6 nonconstant colorings ($u \neq v$), each contributing $\prod_a 1/M_a(\phi)! = 1/(3!\,1!)$. Hence

$$\sum_{\phi} \prod_{a=1}^{3} \frac{1}{M_a(\phi)!} = 3 \cdot \frac{1}{4!} + 6 \cdot \frac{1}{3!\,1!} = \frac{3}{24} + 6 \cdot \frac{1}{6} = \frac{1}{8} + 1 = \frac{9}{8}.$$

Therefore

$$Q(e,h) = 3^{-4} \cdot \frac{9}{8} = \frac{1}{81} \cdot \frac{9}{8} = \boxed{\frac{1}{72}}.$$

- **Multinomial rewrite (37) evaluated.** Equivalently,

$$Q(e,h) = \frac{3^{-4}}{4!} \sum_{\phi \in [3]^2} \binom{4}{M_1(\phi), M_2(\phi), M_3(\phi)}.$$

The 3 constant colorings contribute $\binom{4}{4,0,0} = \binom{4}{0,4,0} = \binom{4}{0,0,4} = 1$ each; the 6 nonconstant colorings contribute $\binom{4}{3,1,0} = \binom{4}{3,0,1} = \binom{4}{1,3,0} = \binom{4}{0,3,1} = \binom{4}{1,0,3} = \binom{4}{0,1,3} = 4$ each. Thus

$$\sum_{\phi} \binom{4}{\mathbf{M}(\phi)} = 3 \cdot 1 + 6 \cdot 4 = 27 \quad \Longrightarrow \quad Q(e,h) = \frac{3^{-4}}{4!} \cdot 27 = \frac{27}{81 \cdot 24} = \boxed{\frac{1}{72}}.$$

- **Expectation form (38) evaluated.** Let $U_1, U_2 \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}([3])$ label the two orbits, and set

$$M_a = 3\,\mathbf{1}_{\{U_1=a\}} + 1\,\mathbf{1}_{\{U_2=a\}}, \quad a = 1,2,3.$$

Then

$$\mathbb{E}\Big[\prod_{a=1}^{3}\frac{1}{M_a!}\Big]=\frac{1}{9}\Big(3\cdot\frac{1}{4!}+6\cdot\frac{1}{3!\,1!}\Big)=\frac{1}{9}\cdot\frac{9}{8}=\frac{1}{8}.$$

Since $s=2$ and $c(g)=4$, we have $k^{\,s-c(g)}=3^{\,2-4}=3^{-2}=1/9$, giving

$$Q(e,h)=\frac{1}{9}\,\mathbb{E}\Big[\prod_{a=1}^{3}\frac{1}{M_a!}\Big]=\frac{1}{9}\cdot\frac{1}{8}=\boxed{\frac{1}{72}}.$$

• **Theorem 5.15 (identity → single t–cycle) verified.** Here $t=3$, $m:=n-t=1$, $k=3$, so

$$\kappa_2(s):=\sum_{u_1+u_2=s}\frac{1}{(u_1!)^2(u_2!)^2}\quad\Rightarrow\quad\kappa_2(0)=1,\quad\kappa_2(1)=\frac{1}{1!^2\,0!^2}+\frac{1}{0!^2\,1!^2}=2.$$

By (42),

$$\begin{aligned}Q(e,h)&=3^{\,1-4}\sum_{j=0}^{1}\binom{1}{j}\frac{(1-j)!}{(3+j)!}\,\kappa_{3-1}(1-j)\\&=3^{\,1-4}\Bigg[\binom{1}{0}\frac{1!}{3!}\,\kappa_2(1)+\binom{1}{1}\frac{0!}{4!}\,\kappa_2(0)\Bigg]=\frac{1}{27}\Big(\frac{1}{6}\cdot 2+\frac{1}{24}\cdot 1\Big)=\frac{1}{27}\cdot\frac{3}{8}=\boxed{\frac{1}{72}}.\end{aligned}$$

5.3. **Stationary Distribution.** We can directly specialize Theorem 3.3 to the coordinate-permutation model and get the following result.

Theorem 5.19 (Dual stationary law)**.** *The dual chain Q on $G=S_n$ (with the coordinate–permutation action on $[k]^n$) is reversible, irreducible, and aperiodic. The unique stationary distribution of Q is*

$$\pi(g)\;=\;\frac{|X_g|}{|G|\,z}\;=\;\frac{k^{\,c(g)}}{n!\,\binom{n+k-1}{k-1}}=\frac{k^{\,c(g)}}{k^{\overline{n}}},$$

where $z=|X/G|=\binom{n+k-1}{k-1}$ is the number of G–orbits on $[k]^n$ and $k^{\overline{n}}:=k(k+1)\cdots(k+n-1)$.

Summary:

Chain	Stationary probability
$\pi_K(x)$	$\dfrac{\prod_{a=1}^{k}m_a(x)!}{n!\,\binom{n+k-1}{k-1}}=\dfrac{1}{\binom{n+k-1}{k-1}\binom{n}{\mathbf{m}(x)}}$
$\pi_Q(g)$	$\dfrac{k^{\,c(g)}}{n!\,\binom{n+k-1}{k-1}}=\dfrac{k^{\,c(g)}}{k^{\overline{n}}}$

Here $m_a(x)=|\{i:\,x_i=a\}|$, $\mathbf{m}(x)$ is its histogram, the multinomial coefficient

$$\binom{n}{\mathbf{m}}:=\binom{n}{m_1,\ldots,m_k}=\frac{n!}{\prod_{i=1}^{k}m_i!},$$

and $c(g)$ is the number of cycles of g.

Binary specialization ($k=2$). Here $m_1(x)=w(x)$ is the Hamming weight, and $m_0(x)=n-w(x)$. Thus

$$\pi_K(x)\;=\;\frac{w(x)!\,\big(n-w(x)\big)!}{n!\,(n+1)}\;=\;\frac{1}{(n+1)\binom{n}{w(x)}},\qquad\pi_Q(g)\;=\;\frac{2^{\,c(g)}}{(n+1)!}.\tag{44}$$

Lemma 5.20 (Extrema of π_K and π_Q in the k-ary coordinate–permutation model)**.** *Writing $n = km + r$ with integers $m \geq 0$ and $0 \leq r < k$, we have*

$$\pi_{K,\max} = \frac{1}{\binom{n+k-1}{k-1}}, \qquad \pi_{K,\min} = \frac{(m!)^{\,k-r}(m+1)!^{\,r}}{\binom{n+k-1}{k-1}\,n!},$$

attained at histograms $(n, 0, \ldots, 0)$ and at balanced histograms with r parts $m+1$ and $k-r$ parts m, respectively.

For Q we have

$$\pi_{Q,\min} = \frac{k}{n!\,\binom{n+k-1}{k-1}}, \qquad \pi_{Q,\max} = \frac{k^n}{n!\,\binom{n+k-1}{k-1}},$$

attained at an n-cycle and the identity permutation, respectively.

Proof. Since $\pi_K(x)$ depends only on $\mathbf{m} = \mathbf{m}(x)$ and there are $\binom{n+k-1}{k-1}$ histograms,

$$\pi_{K,\min} = \frac{1}{\binom{n+k-1}{k-1}\,\max_{\mathbf{m}}\binom{n}{\mathbf{m}}}, \quad \pi_{K,\max} = \frac{1}{\binom{n+k-1}{k-1}\,\min_{\mathbf{m}}\binom{n}{\mathbf{m}}}.$$

Clearly $\min_{\mathbf{m}}\binom{n}{\mathbf{m}} = 1$, attained at $(n, 0, \ldots, 0)$.

For the maximum, take $\mathbf{m} = (\ldots, a, \ldots, b, \ldots)$ with $a \geq b+1$ and set $\mathbf{m}' = (\ldots, a-1, \ldots, b+1, \ldots)$. Then

$$\frac{\binom{n}{\mathbf{m}'}}{\binom{n}{\mathbf{m}}} = \frac{a}{b+1}\begin{cases} > 1, & a \geq b+2, \\ = 1, & a = b+1. \end{cases}$$

Thus any histogram with a gap ≥ 2 between parts is not maximal. Iterating, the maximum occurs when all parts differ by at most 1, i.e. $m_i \in \{m, m+1\}$. Let $t := |\{i : m_i = m+1\}|$; then

$$\sum_{i=1}^{k} m_i = t(m+1) + (k-t)m = km + t = n = km + r,$$

so $t = r$, i.e. exactly r entries equal $m+1$ and $k-r$ entries equal m. For such $\mathbf{m}$,

$$\max_{\mathbf{m}}\binom{n}{\mathbf{m}} = \frac{n!}{(m+1)!^{\,r}\,m!^{\,k-r}},$$

which gives the stated formula for $\pi_{K,\min}$.

For Q, the formula $\pi_Q(g) = k^{c(g)}/(n!\binom{n+k-1}{k-1})$ shows π_Q is minimized when $c(g) = 1$ (an n-cycle) and maximized when $c(g) = n$ (the identity), yielding the formulas for $\pi_{Q,\min}$ and $\pi_{Q,\max}$. □

Remark 5.21 (Stationary–mass ratio)**.** In the k-ary coordinate–permutation model,

$$\frac{\pi_{Q,\max}}{\pi_{Q,\min}} = k^{\,n-1}.$$

Corollary 5.22 (Binary extrema)**.** *In the binary coordinate–permutation model ($k = 2$), we have*

$$\pi_{K,\max} = \frac{1}{n+1}, \qquad \pi_{K,\min} = \frac{1}{(n+1)\binom{n}{\lfloor n/2\rfloor}},$$

and

$$\pi_{Q,\min} = \frac{2}{(n+1)!}, \qquad \pi_{Q,\max} = \frac{2^{\,n}}{(n+1)!}.$$

Proof. Apply Lemma 5.20 with $k = 2$, writing $n = 2m + r$ with $r \in \{0, 1\}$. Then

$$\pi_{K,\max} = \frac{1}{\binom{n+1}{1}} = \frac{1}{n+1}, \qquad \pi_{K,\min} = \frac{(m!)^{2-r}(m+1)!^{\,r}}{\binom{n+1}{1}\,n!} = \frac{(m!)^{2-r}(m+1)!^{\,r}}{(n+1)\,n!}.$$

If $n = 2m$ $(r = 0)$, then

$$\pi_{K,\min} = \frac{(m!)^2}{(n+1)\,n!} = \frac{(m!)^2}{(2m+1)\,(2m)!} = \frac{1}{(2m+1)\binom{2m}{m}} = \frac{1}{(n+1)\binom{n}{n/2}} = \frac{1}{(n+1)\binom{n}{\lfloor n/2\rfloor}}.$$

If $n = 2m+1$ $(r = 1)$, then

$$\pi_{K,\min} = \frac{m!(m+1)!}{(n+1)\,n!} = \frac{m!(m+1)!}{(2m+2)\,(2m+1)!} = \frac{1}{(2m+2)\binom{2m+1}{m}} = \frac{1}{(n+1)\binom{n}{\lfloor n/2\rfloor}}.$$

For Q, the lemma with $k = 2$ gives

$$\pi_{Q,\min} = \frac{2}{n!\binom{n+1}{1}} = \frac{2}{(n+1)!}, \qquad \pi_{Q,\max} = \frac{2^n}{n!\binom{n+1}{1}} = \frac{2^n}{(n+1)!}.$$ □

From Corollary 3.5 (by the detailed balance: $\pi_Q(g)Q(g,h) = \pi_Q(h)Q(h,g)$ with $\pi_Q(\cdot) \propto k^{c(\cdot)}$), we immediately obtain:

Lemma 5.23 (Reversibility ratio)**.** *For $S_n \curvearrowright [k]^n$ and the dual kernel Q, for all $g, h \in S_n$,*

$$\frac{Q(g,h)}{Q(h,g)} \;=\; \frac{\pi_Q(h)}{\pi_Q(g)} \;=\; k^{\,c(h)-c(g)}.$$

Corollary 5.24. *If $g \in S_n$ is a single t–cycle $(2 \le t \le n)$, then*

$$Q(g,g) \;=\; Q(g,e) \;=\; k^{\,t-1}\,Q(e,g).$$

In particular, by Theorem 5.15, with $m := n - t$ and $k \ge 2$,

$$Q(g,g) = Q(g,e) = k^{-m}\sum_{j=0}^{m}\binom{m}{j}\frac{(m-j)!}{(t+j)!}\,\kappa_{k-1}(m-j), \tag{45}$$

$$= k^{-m}\sum_{r=0}^{m}\binom{m}{r}\frac{r!}{(n-r)!}\,\kappa_{k-1}(r), \tag{46}$$

where

$$\kappa_p(s) \;:=\; \sum_{\substack{u_1+\cdots+u_p=s\\ u_i\ge 0}}\;\prod_{i=1}^{p}\frac{1}{(u_i!)^2}.$$

Proof. First, $Q(g,g) = Q(g,e)$ follows from Remark 3.4.

By Lemma 5.23 with $c(e) = n$,

$$Q(g,e) = \frac{\pi_Q(e)}{\pi_Q(g)}\,Q(e,g) = k^{\,n-c(g)}\,Q(e,g).$$

For a single t–cycle, $c(g) = n - t + 1$, hence $n - c(g) = t - 1$, proving $Q(g,e) = k^{t-1}Q(e,g)$. The explicit forms follow by multiplying Theorem 5.15 by k^{t-1}, with $m = n - t$. □

Likewise, Corollary 5.17 specializes as follows.

Corollary 5.25 (Binary specialization for $Q(g,e)$)**.** *Let $k = 2$ and $m := n-t$. If $g \in S_n$ is a single t–cycle, then*

$$Q(g,g) \;=\; Q(g,e) \;=\; \frac{2^{-m}}{n!}\binom{2n-t}{n}.$$

Equivalently, in coefficient form,

$$Q(g,g) \;=\; Q(g,e) \;=\; \frac{2^{-m}}{n!}\,[z^{\,m}]\,(1+z)^{\,2n-t}.$$

5.4. **Mixing and Convergence.** We collect the standard bounds for K in this action; each item holds uniformly over starts unless explicitly marked "start–specific".

(K1) **Uniform floor via stabilizers.** Lemma 3.24 yields $K(x,\cdot) \geq \delta_K\ \mathrm{Unif}(X),\ \delta_K = 1/M$ with $M = \max_x |G_x|$. In the present action $|G_x| = \prod_{a=1}^{k} m_a(x)!$. Since

$$\binom{n}{m_1,\ldots,m_k} = \frac{n!}{\prod_a m_a!} \ \geq\ 1,$$

it follows that $\prod_a m_a! \leq n!$, with equality when one $m_a = n$ and the others are 0 (e.g. x is constant). Hence $M = n!$ and by Theorem 3.25 we get

$$d_K(t)\ \leq\ \Big(1-\frac{1}{n!}\Big)^t \quad\Rightarrow\quad t_{\mathrm{mix}}(K;\varepsilon)\ \leq\ \Big\lceil n!\,\log\frac{1}{\varepsilon}\Big\rceil = \Big\lceil |G|\,\log\frac{1}{\varepsilon}\Big\rceil.$$

So this reproduces the bounds in Proposition 3.23 (independent of k).

(K2) **Aldous** [1] **uniform coupling bound.** For all n,k and all starts $x \in [k]^n$,

$$\|K^t(x,\cdot)-\pi_K\|_{\mathrm{TV}}\ \leq\ n\Big(1-\frac{1}{k}\Big)^t, \qquad \Rightarrow\quad t_{\mathrm{mix}}(K;\varepsilon)\ \leq\ \Big\lceil k\,\log\frac{n}{\varepsilon}\Big\rceil.$$

(K3) **Diaconis** [10] **fixed k (start–specific, n–independent).** For fixed k (or k slowly growing with n), we have the *start–specific* Doeblin/minorization for the all–equal start x_0, namely

$$\|K^t(x_0,\cdot)-\pi_K\|_{\mathrm{TV}}\ \leq\ (1-c_k)^t \quad \text{for some } c_k \in (0,1) \text{ depending only on } k,$$

so $t_{\mathrm{mix}}(K;x_0,\varepsilon)\ \leq\ \Big\lceil c_k^{-1}\log(1/\varepsilon)\Big\rceil$, independent of n.

We now turn (K1)–(K3) into consequences for the *dual* chain Q using the pointwise transfer inequality (Theorem 3.18), the one–step TV comparison (Corollary 3.19), and the mixing–time equivalence (Theorem 3.20).

Corollary 5.26 (Global transfer of the model–free bound)**.**

$$d_Q(t)\ \leq\ \Big(1-\frac{1}{n!}\Big)^{t-1}, \qquad t_{\mathrm{mix}}(Q;\varepsilon)\ \leq\ 1+\Big\lceil n!\,\log\frac{1}{\varepsilon}\Big\rceil.$$

Proof. By Corollary 3.19, $d_Q(t) \leq d_K(t-1)$; apply (K1). For mixing times, Theorem 3.20 gives $t_{\mathrm{mix}}(Q;\varepsilon) \leq t_{\mathrm{mix}}(K;\varepsilon)+1$. □

Instead, we could directly apply Theorem 3.25 to Q, which matches Corollary 5.26 up to the immaterial $+1$ step from transfer:

Theorem 5.27 (Direct one–step minorization for Q)**.**

$$d_Q(t)\ \leq\ \Big(1-\frac{1}{n!}\Big)^{t}, \qquad t_{\mathrm{mix}}(Q;\varepsilon)\ \leq\ \Big\lceil n!\,\log\frac{1}{\varepsilon}\Big\rceil.$$

Proof. Independently of K, Lemma 3.24 yields

$$Q(g,\cdot)\ \geq\ \delta_Q\,\delta_e(\cdot), \qquad \delta_Q = \frac{1}{M} = \frac{1}{n!},$$

hence by Theorem 3.25, we get the results. □

Theorem 5.28 (Uniform Aldous transfer to Q)**.** *For all $n,k \geq 2$ and all $t \geq 1$,*

$$d_Q(t)\ \leq\ n\Big(1-\frac{1}{k}\Big)^{t-1}, \qquad \textit{hence} \qquad t_{\mathrm{mix}}(Q;\varepsilon)\ \leq\ 1\ +\ \Big\lceil k\,\log\frac{n}{\varepsilon}\Big\rceil.$$

Proof. By Corollary 3.19, $d_Q(t) \leq d_K(t-1)$. Apply Aldous's uniform bound (K2): $d_K(t) \leq n(1-\frac{1}{k})^t$ for all t, to obtain the stated TV bound. The mixing-time bound follows from Theorem 3.20: $t_{\mathrm{mix}}(Q;\varepsilon) \leq t_{\mathrm{mix}}(K;\varepsilon)+1 \leq \Big\lceil k\log(n/\varepsilon)\Big\rceil + 1$. □

Theorem 5.29 (Fixed k: start–specific transfer to Q)**.** *If $g \in S_n$ is transitive on coordinates (i.e., an n–cycle, so $X_g = \{a^n : \ a \in [k]\}$), then for all $t \geq 1$,*

$$d_Q(g,t) \ = \ \|Q^t(g,\cdot) - \pi_Q\|_{\mathrm{TV}} \ \leq \ (1-c_k)^{t-1}, \qquad t_{\mathrm{mix}}(Q;g,\varepsilon) \ \leq \ 1 + \left\lceil \frac{1}{c_k} \log \frac{1}{\varepsilon} \right\rceil,$$

with constants independent of n.

Proof. By Theorem 3.18, $\|Q^t(g,\cdot) - \pi_Q\|_{\mathrm{TV}} \leq \max_{x \in X_g} \|K^{t-1}(x,\cdot) - \pi_K\|_{\mathrm{TV}}$. Since X_g consists of constant strings and (K3) holds uniformly over those starts, the stated TV bound follows. The mixing–time bound then follows from Theorem 3.20. □

Finally, for $k = 2$, the (relabeled) binary coordinate–permutation model $S_n \curvearrowright \{0,1\}^n$, Diaconis–Zhong [17] proved the two–sided sharp estimate for the *primal* chain K from the symmetric starts $x_0 \in \{0^n, 1^n\}$:

$$\frac{1}{4}\Big(\frac{1}{4}\Big)^t \ \leq \ d_K(x_0,t) \ = \ \|K^t(x_0,\cdot) - \pi_K\|_{\mathrm{TV}} \ \leq \ 4\Big(\frac{1}{4}\Big)^t, \qquad t \geq 0, \tag{47}$$

which is also *start–specific* (not uniform over all x). Hence $t_{\mathrm{mix}}(K;x_0,\varepsilon) = \min\{t : \|K^t(x_0,\cdot) - \pi_K\|_{\mathrm{TV}} \leq \varepsilon\}$ satisfies

$$\left\lceil \log_4 \frac{1}{\varepsilon} \right\rceil - 1 \ = \ \left\lceil \log_4 \frac{1}{4\varepsilon} \right\rceil \ \leq \ t_{\mathrm{mix}}(K;x_0,\varepsilon) \ \leq \ \left\lceil \log_4 \frac{4}{\varepsilon} \right\rceil \ = \ 1 + \left\lceil \log_4 \frac{1}{\varepsilon} \right\rceil.$$

In particular, $t_{\mathrm{mix}}(K;x_0,\varepsilon) = \lceil \log_4(1/\varepsilon) \rceil + O(1)$ with an additive error at most 1.

Remark 5.30. Diaconis–Zhong [17] first diagonalize the lumped Burnside chain on Hamming weights $\{0,\dots,n\}$ (our $\bar{K}$) by computing all eigenvalues and eigenfunctions (Theorem 1.2 in [17]), and then use the resulting χ^2 bounds to obtain sharp total variation mixing bounds for the original unlumped chain K (Theorem 1.1 in [17]).

As explained in Section 3.5.3, the same lumped/auxiliary-variable scheme applies at the dual level: we could also work with the conjugacy-lumped chain $\bar{Q}$, whose nonzero eigenvalues coincide with those of $\bar{K}$, and then use methods similar to the proof of Theorem 1.2 in [17] to derive mixing bounds for Q.

We now derive consequences for the *dual* chain Q.

Theorem 5.31 (Worst–case lower bound for Q (binary case))**.** *For all $t \geq 0$,*

$$d_Q(t) \ \geq \ \frac{1}{4}\Big(\frac{1}{4}\Big)^{t+1} = 4^{-(t+2)}.$$

Consequently,

$$t_{\mathrm{mix}}(Q;\varepsilon) \ \geq \ \left\lceil \log_4 \frac{1}{\varepsilon} \right\rceil - 2,$$

with absolute constants independent of n.

Proof. From (47), take $C_1 = \alpha = \frac{1}{4}$ in Corollary 3.21 (i). The results follow. □

Theorem 5.32 (Transitive starts for Q (binary case))**.** *If $g \in S_n$ is transitive on coordinates (i.e., a single n–cycle, so $X_g = \{0^n, 1^n\}$), then for all $t \geq 1$,*

$$d_Q(g,t) \ = \ \|Q^t(g,\cdot) - \pi_Q\|_{\mathrm{TV}} \ \leq \ 4\Big(\frac{1}{4}\Big)^{t-1} = 4^{2-t}.$$

Consequently,

$$t_{\mathrm{mix}}(Q;g,\varepsilon) \ \leq \ 2 + \left\lceil \log_4 \frac{1}{\varepsilon} \right\rceil,$$

with absolute constants independent of n.

Proof. Apply Corollary 3.21 (ii) with $x_0 \in \{0^n, 1^n\}$, $C_2 = 4$, $\alpha = \frac{1}{4}$. For an n–cycle g we have $X_g = \{0^n, 1^n\}$, so the uniform K–upper bound on X_g is exactly (47). Hence the results follow. □

Finally, Diaconis–Lin–Ram [14] prove that, for the binary coordinate–permutation model,

$$\lambda_1(\bar{K}) = \lambda_1(K) = \frac{1}{4},$$

so $\frac{1}{4}$ is also the second-largest eigenvalue of the *unlumped* K (and by Corollary 3.13 we may use $\lambda_* = \lambda_1$). Therefore, by (7) with $\beta := \lambda_1 = \frac{1}{4}$, we obtain a uniform decay

$$d_K(t) \le \frac{1}{2\sqrt{\pi_{K,\min}}} \left(\frac{1}{4}\right)^t \qquad (t \ge 0).$$

Recall that Corollary 5.22 shows

$$\pi_{K,\min} = \frac{1}{(n+1)\binom{n}{\lfloor n/2 \rfloor}},$$

so

$$d_K(t) \le \frac{1}{2}\sqrt{(n+1)\binom{n}{\lfloor n/2 \rfloor}} \left(\frac{1}{4}\right)^t \qquad (t \ge 0).$$

Theorem 5.33 (Uniform TV decay transfers to Q (binary case))**.** *For the binary coordinate–permutation model and all $t \ge 1$,*

$$d_Q(t) \le \frac{1}{2\sqrt{\pi_{K,\min}}} \left(\frac{1}{4}\right)^{t-1} = \frac{1}{2}\sqrt{(n+1)\binom{n}{\lfloor n/2 \rfloor}} \left(\frac{1}{4}\right)^{t-1}. \tag{48}$$

Proof. Apply the one–step TV comparison $d_Q(t) \le d_K(t-1)$ (Corollary 3.19) and the bounds above. Substituting $\pi_{K,\min} = \frac{1}{(n+1)\binom{n}{\lfloor n/2 \rfloor}}$ gives the explicit form. □

Remark 5.34 (Direct Q bound and comparison (binary case))**.** By Corollary 3.13 and Diaconis–Lin–Ram [14],

$$\lambda_1(\bar{Q}) = \lambda_1(Q) = \frac{1}{4}.$$

Using Lemma 5.20 with $k = 2$ gives $\pi_{Q,\min} = 2/(n+1)!$. Hence, (7) applied directly to Q yields

$$d_Q(t) \le \frac{1}{2\sqrt{\pi_{Q,\min}}} \left(\frac{1}{4}\right)^t = \frac{1}{2}\sqrt{\frac{(n+1)!}{2}} \left(\frac{1}{4}\right)^t. \tag{49}$$

Comparing with Theorem 5.33 at the same time t, the transferred bound (48) is smaller if and only if $32\binom{n}{\lfloor n/2 \rfloor} \le n!$, which holds for all $n \ge 6$ (strict), while for $1 \le n \le 5$ the direct Q bound (49) is no worse.

6. Conclusion

We introduced and analyzed the *dual* Burnside process, establishing its fundamental properties and connections to the classical chain. They appear to be two different processes on two different state spaces, but the primal–dual factorization reveals that both chains share the same nonzero spectrum and have equivalent mixing times up to a single step. This provides us with a new lens to study the classical Burnside processes and potentially sharpen their mixing-time bounds, since we can transfer them back and forth between the primal and the dual chains. Therefore, the dual perspective provides both theoretical insight and practical advantages for sampling under group symmetry, and opens new avenues for understanding and accelerating symmetry-aware MCMC methods.

Future work includes tightening the bounds through techniques such as lumping, minorization (e.g., [10]), coupling (e.g., [32]), and geometric methods (e.g., [16]). For the value- and coordinate-permutation models, we derived explicit formulas so we may build upon them to find sharper mixing bounds. Additionally, we gave the lumping method for the value-permutation model by the *number of fixed points* (a coarsening of *conjugacy*) in Section 4.5. Likewise, similar to Remark 4.26(i), we may further provide formulas for lumping by *conjugacy classes* in both models based on formulas (22) and (23). This may provide sharper start–specific upper bounds and universal lower bounds, improving existing results.

Furthermore, if we zoom out a bit, the value- and coordinate-permutation models correspond to the following blocks in the twelvefold way (the value–permutation model sits in the Stirling/Bell-number block, while the coordinate–permutation model is the stars–and–bars/Bose–Einstein block):

f: **Domain** $[n] \to$ **Codomain** $[k]$	**Any** f	**Injective** f	**Surjective** f
Distinguishable $\to$ Distinguishable	k^n $[k]^n$ *(Maxwell–Boltzmann)* n-sequence in $[k]$	$k^{\underline{n}} = \dfrac{k!}{(k-n)!}$ n-permutation in $[k]$	$k!\,S(n,k)$ set composition of $[n]$ into k blocks
Distinguishable $\to$ Indistinguishable	$\sum_{j=0}^{k} S(n,j)$ $S_k \curvearrowright [k]^n$ set partition of $[n]$ into $\le k$ blocks	$\begin{cases} 1, & n \le k \\ 0, & n > k \end{cases}$ at most one per block	$S(n,k)$ set partition of $[n]$ into k blocks
Indistinguishable $\to$ Distinguishable	$\binom{n+k-1}{k-1}$ $S_n \curvearrowright [k]^n$ *(Bose–Einstein)* n-multiset of $[k]$	$\binom{k}{n}$ *(Fermi–Dirac)* n-subset of $[k]$	$\binom{n-1}{k-1}$ composition of n into k parts
Indistinguishable $\to$ Indistinguishable	$p_k(n)$ partition of n into $\le k$ parts	$\begin{cases} 1, & n \le k \\ 0, & n > k \end{cases}$ all parts of size 1	$p(n,k)$ partition of n into k parts

where:

- $S(n,k)$ = Stirling number of second kind (partitions into k non-empty subsets)
- $p_k(n)$ = number of partitions of n into at most k parts
- $p(n,k) = p_k(n-k)$ = the number of partitions of n into exactly k parts

Each block corresponds to a canonical combinatorial model as listed above (in teal), and the associated unlabeled configurations are natural targets for MCMC. Therefore, it is natural to study Burnside and dual Burnside processes in the remaining blocks as well, where we expect analogous structure (orbit counts, fixed–point formulas, natural lumpings) and potentially fast symmetry-aware samplers. Since the injective and surjective constraints are G-invariant, the Burnside construction restricts to the corresponding subactions by replacing X_g with $Y_g := X_g \cap Y$. In general, however, $|Y_g|$ depends on the cycle count $c(g)$ (e.g., in the coordinate-permutation model we have $|Y_g| = k!\,S(c(g),k)$ for surjections, with the convention $S(m,k) = 0$ for $m < k$; and for injections only $g = e$ contributes, so $|Y_g| = 0$ for $g \neq e$ while $|Y_e| = |Y| = k(k-1)\cdots(k-n+1)$ when $n \le k$). Consequently, in general, the restricted kernels are not global rescalings of K or Q and require separate analysis.

References

[1] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in progress (2002). `https://www.stat.berkeley.edu/~aldous/RWG/book.pdf`

[2] C. Athanasiadis and P. Diaconis, *Functions of Random Walks on Hyperplane Arrangements*, Adv. Appl. Math. **45** (2010), no. 3, 410–437.

[3] L. Bartholdi and P. Diaconis, *An algorithm for uniform generation of unlabeled trees (Pólya trees), with an extension of Cayley's formula*, arXiv:2411.17613 (2024).

[4] P. Brémaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer, 1999.

[5] W. Burnside, *Theory of Groups of Finite Order*, Cambridge Univ. Press, 1897.

[6] S. Chatterjee and P. Diaconis, A central limit theorem for a new statistic on permutations, *Indian J. Pure Appl. Math.* 48 (2017), 561–573.

[7] W.-K. Chen, *Mixing Times for Burnside Processes*, M.Sc. Thesis, National Chiao Tung University, Hsinchu, Taiwan, July 2006.

[8] N. G. de Bruijn, *Asymptotic Methods in Analysis*, Dover, 1981.

[9] P. Diaconis, *Group Representations in Probability and Statistics*, IMS Lecture Notes—Monograph 11, 1988.

[10] P. Diaconis, Analysis of a Bose–Einstein Markov chain, *Ann. Inst. H. Poincaré Probab. Statist.* 41 (2005), 409–418.

[11] P. Diaconis and J. Fill, Strong stationary times via a new form of duality, *Ann. Probab.* 18 (1990), 1483–1522.

[12] P. Diaconis and M. Howes, Random sampling of partitions and contingency tables: two practical examples of the Burnside process, *Stat. Comput.* (2025), arXiv:2503.02818.

[13] P. Diaconis, K. Khare, and L. Saloff–Coste, *Gibbs sampling, exponential families and orthogonal polynomials*, Ann. Appl. Probab. **18** (2008), no. 4, 1586–1619.

[14] P. Diaconis, A. Lin, and A. Ram, *A curiously slowly mixing Markov chain*, arXiv:2511.01245 (2025).

[15] P. Diaconis and C. Morton–Ferguson, *The Burnside process on the flag variety for* $\mathrm{GL}_n(\mathbb{F}_q)$, arXiv:2510.02285 (2025).

[16] P. Diaconis and D. Stroock, Geometric bounds for eigenvalues of Markov chains, *Ann. Appl. Probab.* 1 (1991), 36–61.

[17] P. Diaconis and C. Zhong, Hahn polynomials and the Burnside process, *J. Algebraic Combin.* 59 (2024), 115–142.

[18] P. Diaconis and C. Zhong, Counting the number of group orbits by marrying the Burnside process with importance sampling, arXiv:2501.11731 [math.PR], 2025.

[19] J. Dittmer, *Counting and Sampling Double Cosets and Applications*, Ph.D. thesis, University of California, Los Angeles, 2023. `https://www.math.ucla.edu/~pak/papers/Dittmer-thesis.pdf`

[20] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed., Wiley, 1968.

[21] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, 2009.

[22] J. Fulman, Cycle indices for the finite classical groups, *J. Group Theory* **2** (1999), 251–289.

[23] L. A. Goldberg and M. Jerrum, The Burnside process converges slowly, *Combin. Probab. Comput.* 11 (2002), 21–34.

[24] L. Grünwald and S. Serafin, Explicit bounds for Bell numbers and their ratios, *J. Math. Anal. Appl.* 540 (2024), 128651.

[25] S. Holtzen, T. Millstein and G. Van den Broeck, Generating and sampling orbits for lifted probabilistic inference, in *Proc. UAI 2020*, PMLR 115 (2020), 985–994.

[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, 2013.

[27] M. Howes, *Limit profiles and cutoff for the Burnside process on Sylow double cosets*, arXiv:2511.01058 (2025).

[28] M. Jerrum, Uniform sampling modulo a group of symmetries using Markov-chain simulation, in *Expanding Graphs*, DIMACS 10, AMS, 1993, 37–47.

[29] R. Koekoek and R. F. Swarttouw, *The Askey Scheme of Hypergeometric Orthogonal Polynomials and its q-Analogue*, Tech. Report 98-17, Delft Univ., 1998.

[30] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times*, 2nd ed., Amer. Math. Soc., 2017.

[31] Y. Nakatsukasa, The low-rank eigenvalue problem, arXiv:1905.11490, 2019.

[32] J. E. Paguyo, Mixing times of a Burnside process Markov chain on set partitions, arXiv:2207.14269 [math.PR], 2022.

[33] J. Rahmani, *Mixing Times for the Commuting Chain*, Ph.D. thesis, University of Southern California, 2021.

[34] J. Rahmani, Mixing times for the commuting chain on CA groups, *J. Theor. Probab.* 35 (2022), 1382–1409.

[35] J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, 1958.

[36] J. S. Rosenthal, Minorization conditions and convergence rates for Markov chain Monte Carlo, *Journal of the American Statistical Association* **90** (1995), no. 430, 558–566.

[37] B. E. Sagan, *The Symmetric Group*, 2nd ed., Grad. Texts in Math. 203, Springer, 2001.

[38] A. Sinclair and M. Jerrum, Approximate counting, uniform generation and rapidly mixing Markov chains, *Inf. Comput.* 82 (1989), 93–133.

[39] A. Sinclair, Improved bounds for mixing rates of Markov chains and multicommodity flow, *Combin. Probab. Comput.* 1 (1992), 351–370.

[40] R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, 2nd ed., Cambridge Univ. Press, 2012.

Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532, USA
Email address: ifeng@usc.edu
URL: https://dornsife.usc.edu/ivan/