It is an intriguing, and nontrivial, exercise to work out the value of λ_2 exactly in this case, in order to determine whether the mixing times $\tau(\varepsilon)$ are closer to the given lower or upper bounds as a function of n.

Let us now return to the proof of Theorem 3.6, establishing the connection between the second-largest eigenvalue and the conductance of a Markov chain. Recall the statement of the Theorem:

Theorem 3.6 *Let $\mathcal{M}$ be a finite, regular, reversible Markov chain and λ_2 the second-largest eigenvalue of its transition matrix. Then:*

(i) $\lambda_2 \leq 1 - \frac{\Phi^2}{2}$,

(ii) $\lambda_2 \geq 1 - 2\Phi$.

Proof. (i) The approach here is to bound Φ in terms of the eigenvalue gap of $\mathcal{M}$, i.e. to show that $\Phi^2/2 \leq 1 - \lambda_2$, from which the claimed result follows.

Thus, consider the eigenvalue $\lambda = \lambda_2$. (The following proof does not in fact depend on this particular choice of eigenvalue $\lambda \neq 1$, but since we are proving an upper bound of the form $\Phi^2/2 \leq 1 - \lambda$, all other eigenvalues yield weaker bounds than λ_2.)

Let e be a left eigenvector $e \neq 0$ such that $eP = \lambda e$. Now e must contain both positive and negative components, since $\sum_i e_i = 0$ as can be seen:

$$\begin{aligned} eP = \lambda e \;\; &\Leftrightarrow \;\; \sum_i e_i p_{ij} = \lambda e_j \quad \forall\, j \\ &\Rightarrow \;\; \sum_j \sum_i e_i p_{ij} = \sum_i e_i \underbrace{\sum_j p_{ij}}_{=1} = \lambda \sum_j e_j \\ &\overset{\lambda \neq 1}{\Rightarrow} \;\; \sum_i e_i = 0. \end{aligned}$$

Define $A = \{i \mid e_i > 0\}$. Assume, without loss of generality, that $\pi(A) \leq 1/2$. (Otherwise we may replace e by $-e$ in the following proof.)

Define further a "π-normalised" version of $e \restriction A$:

$$u_i = \begin{cases} e_i/\pi_i, & \text{if } i \in A \\ 0, & \text{if } i \notin A \end{cases}$$

Without loss of generality we may again assume that the states are indexed so that $u_1 \geq u_2 \geq \ldots \geq u_r > u_{r+1} = \ldots = u_n = 0$, where $r = |A|$.

In the remainder of the proof, the following quantity will be important:

$$D = \frac{\sum_{i<j} w_{ij}(u_i^2 - u_j^2)}{\sum_i \pi_i u_i^2}.$$

We shall prove the following claims:

(a) $\Phi \leq D$,

(b) $D^2/2 \leq 1 - \lambda$,

which suffice to establish our result.

Proof of (a): Denote $A_k = \{1, \ldots, k\}$, for $k = 1, \ldots, r$. The numerator in the definition of D may be expressed in terms of the ergodic flows out of the A_k as follows:

$$\begin{aligned}
\sum_{i<j} w_{ij}(u_i^2 - u_j^2) &= \sum_{i<j} w_{ij} \sum_{i \leq k < j} (u_k^2 - u_{k+1}^2) \\
&= \sum_{k=1}^{r} (u_k^2 - u_{k+1}^2) \sum_{\substack{i \in A_k \\ j \notin A_k}} w_{ij} \\
&= \sum_{k=1}^{r} (u_k^2 - u_{k+1}^2) F_{A_k}.
\end{aligned}$$

Now the capacities of the A_k satisfy $\pi(A_k) \leq \pi(A) \leq 1/2$, so by definition $\Phi_{A_k} \geq \Phi \Rightarrow F_{A_k} \geq \Phi \cdot \pi(A_k)$. Thus,

$$\begin{aligned}
\sum_{i<j} w_{ij}(u_i^2 - u_j^2) &= \sum_{k=1}^{r} (u_k^2 - u_{k+1}^2) F_{A_k} \\
&\geq \Phi \cdot \sum_{k=1}^{r} (u_k^2 - u_{k+1}^2) \pi(A_k) \\
&= \Phi \cdot \sum_{k=1}^{r} (u_k^2 - u_{k+1}^2) \sum_{i=1}^{k} \pi_i \\
&= \Phi \cdot \sum_{i=1}^{r} \pi_i \sum_{k=i}^{r} (u_k^2 - u_{k+1}^2) \\
&= \Phi \cdot \sum_{i \in A} \pi_i u_i^2.
\end{aligned}$$

Hence,

$$\Phi \leq \frac{\sum_{i<j} w_{ij}(u_i^2 - u_j^2)}{\sum_i \pi_i u_i^2} = D.$$

Proof of (b): We introduce one more auxiliary expression:

$$E = \frac{\sum_{i<j} w_{ij}(u_i - u_j)^2}{\sum_i \pi_i u_i^2},$$

and establish that: (b') $D^2 \leq 2E$, (b'') $E \leq 1-\lambda$. This will conclude the proof of Theorem 3.6 (i).

Proof of (b'): Observe first that

$$\sum_{i<j} w_{ij}(u_i + u_j)^2 \leq 2\sum_{i<j} w_{ij}(u_i^2 + u_j^2) \leq 2\sum_{i\in A} \pi_i u_i^2.$$

Then, by the Cauchy-Schwartz inequality:

$$\begin{aligned} D^2 &= \left(\frac{\sum_{i<j} w_{ij}(u_i^2 - u_j^2)}{\sum_i \pi_i u_i^2}\right)^2 \\ &\leq \left(\frac{\sum_{i<j} w_{ij}(u_i + u_j)^2}{\sum_i \pi_i u_i^2}\right)\left(\frac{\sum_{i<j} w_{ij}(u_i - u_j)^2}{\sum_i \pi_i u_i^2}\right) \leq 2E. \end{aligned}$$

Proof of (b''): Denote $Q = I - P$. Then $eQ = (1-\lambda)e$ and thus

$$eQu^T = (1-\lambda)eu^T = (1-\lambda)\sum_{i=1}^{r} \pi_i u_i^2.$$

On the other hand, writing eQu^T out explicitly:

$$\begin{array}{rcl|l}
eQu^T & = & \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{r} q_{ij}e_iu_j & q_{ij} = -p_{ij} = -\dfrac{w_{ij}}{\pi_i}, \quad i \neq j \\
& \geq & \displaystyle\sum_{i=1}^{r}\sum_{j=1}^{r} q_{ij}e_iu_j & q_{ii} = 1 - p_{ii} = \displaystyle\sum_{i\neq j} p_{ij} \\
& = & -\displaystyle\sum_{i\in A}\sum_{\substack{j\in A\\ j\neq i}} w_{ij}u_iu_j + \sum_{i\in A}\sum_{\substack{j\in A\\ j\neq i}} w_{ij}u_i^2 & e_i = \pi_i u_i, \quad i \in A \\
& = & -2\displaystyle\sum_{i<j} w_{ij}u_iu_j + \sum_{i<j} w_{ij}(u_i^2 + u_j^2) & \\
& = & \displaystyle\sum_{i<j} w_{ij}(u_i - u_j)^2. &
\end{array}$$

Thus,

$$E \cdot \sum_i \pi_i u_i^2 = \sum_{i<j} w_{ij}(u_i - u_j)^2 \leq eQu^T = (1-\lambda)\cdot\sum_i \pi_i u_i^2 \;\Rightarrow\; E \leq 1 - \lambda.$$

(ii) Given the stationary distribution vector $\pi \in \mathbb{R}^n$, define an inner product $\langle\cdot,\cdot\rangle_\pi$ in $\mathbb{R}^n$ as:

$$\langle u, v\rangle_\pi = \sum_{i=1}^{n} \pi_i u_i v_i.$$

By (a slight modification of) a standard result (the Courant-Fischer minimax theorem) in matrix theory, and the fact that P is reversible with respect to π, implying $\langle u, Pv\rangle_\pi = \langle Pu, v\rangle_\pi$, one can characterise the eigenvalues of P as:

$$\begin{aligned}
\lambda_1 &= \max\left\{\frac{\langle u, Pu\rangle_\pi}{\langle u, u\rangle_\pi} \mid u \neq 0\right\}, \\
\lambda_2 &= \max\left\{\frac{\langle u, Pu\rangle_\pi}{\langle u, u\rangle_\pi} \mid u \perp \pi, u \neq 0\right\}, \text{ etc.}
\end{aligned}$$

In particular,

$$\lambda_2 \geq \frac{\langle u, Pu\rangle_\pi}{\langle u, u\rangle_\pi} \text{ for any } u \neq 0 \text{ such that } \sum_i \pi_i u_i = 0. \tag{5}$$

Given a set of states $A \subseteq S$, $0 < \pi(A) \leq 1/2$, we shall apply the bound (5) to the vector u defined as:

$$u_i = \begin{cases} \dfrac{1}{\pi(A)}, & \text{if } i \in A \\ -\dfrac{1}{\pi(\bar{A})}, & \text{if } i \in \bar{A} \end{cases}$$

Clearly

$$\sum_i \pi_i u_i = \sum_{i\in A} \frac{\pi_i}{\pi(A)} - \sum_{i\in \bar{A}} \frac{\pi_i}{\pi(\bar{A})} = 1-1=0, \text{ and}$$

$$\langle u,u\rangle_\pi = \sum_i \pi_i u_i^2 = \sum_{i\in A} \frac{\pi_i}{\pi(A)^2} + \sum_{i\in \bar{A}} \frac{\pi_i}{\pi(\bar{A})^2} = \frac{1}{\pi(A)} + \frac{1}{\pi(\bar{A})},$$

so let us compute the value of $\langle u,Pu\rangle_\pi$.

The task can be simplified by representing P as $P = I_n - (I_n - P)$, and first computing $\langle u,(I-P)u\rangle_\pi$:

$$\begin{aligned}
\langle u,(I-P)u\rangle_\pi &= \sum_i \pi_i u_i \sum_j (I-P)_{ij} u_j \\
&= -\sum_i \sum_{j\neq i} \pi_i u_i p_{ij} u_j + \sum_i \sum_{j\neq i} \pi_i u_i p_{ij} u_i \\
&= \sum_i \sum_{j\neq i} \pi_i p_{ij} (u_i^2 - u_i u_j) \\
&= \sum_{i<j} \pi_i p_{ij} (u_i - u_j)^2 \\
&= \sum_{\substack{i\in A \\ j\neq i}} \pi_i p_{ij} \left(\frac{1}{\pi(A)} + \frac{1}{\pi(A)} \right)^2 \\
&= \left(\frac{1}{\pi(A)} + \frac{1}{\pi(A)} \right)^2 F_A.
\end{aligned}$$

Thus,

$$\begin{aligned}
\lambda_2 &\geq \frac{\langle u,Pu\rangle_\pi}{\langle u,u\rangle_\pi} = \frac{1}{\langle u,u\rangle_\pi}\Big(\langle u,u\rangle_\pi - \langle u,(I-P)u\rangle_\pi \Big) \\
&= 1 - \frac{1}{\langle u,u\rangle_\pi} \cdot \langle u,(I-P)u\rangle_\pi \\
&= 1 - \left(\frac{1}{\pi(A)} + \frac{1}{\pi(A)} \right)^{-1} \left(\frac{1}{\pi(A)} + \frac{1}{\pi(A)} \right)^2 \cdot F_A \\
&= 1 - \left(\frac{1}{\pi(A)} + \frac{1}{\pi(A)} \right) \cdot F_A \\
&\geq 1 - \frac{2}{\pi(A)} \cdot F_A = 1 - 2\Phi_A.
\end{aligned}$$

Since the bound (6) holds for any $A \subseteq S$ such that $0 < \pi(A) \leq 1/2$, it follows that it holds also for the conductance

$$\Phi = \min_{0<\pi(A)\leq 1/2} \Phi_A.$$

Thus, we have shown that $\lambda_2 \geq 1 - 2\Phi$, which completes the proof. □

Despite the elegance of the conductance approch, it can be sometimes (often?) difficult to apply in practice – computing graph conductance can be quite difficult. Also the bounds obtained are not necessary the best possible; in particular the square in the upper bound $\lambda_2 \leq 1 - \Phi^2/2$ is unfortunate.

An alternative approch, which is sometimes easier to apply, and can even yield better bounds, is based on the construction of so called "canonical paths" between states of a Markov chain.

Consider again a regular, reversible Markov chain with stationary distribution π, represented as a weighted graph with node set S and edge set $E = \{(i, j) \mid p_{ij} > 0\}$. The weight, or capacity, w_e associated to edge $e = (i, j)$ corresponds to the ergodic flow $\pi_i p_{ij}$ between states i and j.

Specify for each pair of states $k, l \in S$ a *canonical path* γ_{kl} connecting them. The paths should intuitively be chosen as short and as nonoverlapping as possible. (For precise statements, see Theorems 3.9 and 3.11 below.)

Denote $\Gamma = \{\gamma_{kl} \mid k, l \in S\}$ and define the unweighted and weighted *edge loading* induced by Γ on an edge $e \in E$ as:

$$\begin{aligned} \rho_e &= \frac{1}{w_e} \sum_{\gamma_{kl} \ni e} \pi_k \pi_l \\ \bar{\rho}_e &= \frac{1}{w_e} \sum_{\gamma_{kl} \ni e} \pi_k \pi_l |\gamma_{kl}|, \end{aligned}$$

where $|\gamma_{kl}|$ is the length (number of edges) of path γ_{kl}. (Note that here the edges are considered to be *oriented*, so that only paths crossing an edge $e = (i, j)$ in the direction from i to j are counted in determining the loading of e.) The *maximum edge loading* induced by Γ is then:

$$\begin{aligned} \rho &= \rho(\Gamma) = \max_{e \in E} \rho_e \\ \bar{\rho} &= \bar{\rho}(\Gamma) = \max_{e \in E} \bar{\rho}_e. \end{aligned}$$

Theorem 3.9 *For any regular, reversible Markov chain and any choice of canonical paths,*

$$\Phi \geq \frac{1}{2\rho}.$$

Proof. Represent the chain as a weighted graph G, where the weight (capacity) on edge $e = (i, j)$ is defined as:

$$w_{ij} = \pi_i p_{ij} = \pi_j p_{ji}.$$

Every set of states $A \subseteq S$ determines a cut $(A, \bar{A})$ in G, and the conductance of the cut corresponds to its *relative capacity*:

$$\Phi_A = \frac{w(A,\bar{A})}{V_A} = \frac{1}{\pi(A)} \sum_{i \in A, j \in \bar{A}} w_{ij}.$$

Let then A be a set with $0 < \pi(A) \leq \frac{1}{2}$ that minimises Φ_A, and thus has $\Phi_A = \Phi$. Assume some choice of canonical paths $\Gamma = \{\gamma_{ij}\}$, and assign to each path γ_{ij} a "flow" of value $\pi_i \pi_j$. Then the total amount of flow crossing the cut $(A, \bar{A})$ is

$$\sum_{i \in A, j \in \bar{A}} \pi_i \pi_j = \pi(A)\pi(\bar{A}),$$

but the cut edges, i.e. edges crossing the cut, have only total capacity $w(A, \bar{A})$. Thus, some cut edge e must have loading

$$\rho_e = \frac{1}{w_e} \sum_{\gamma_{ij} \ni e} \pi_i \pi_j \geq \frac{\pi(A)\pi(\bar{A})}{w(A,A)} \geq \frac{\pi(A)}{2w(A,A)} = \frac{1}{2\Phi}.$$

The result follows. □

Corollary 3.10 *With notations and assumptions as above,*

$$\lambda_2 \leq 1 - \frac{1}{8\rho^2}.$$

Proof. From Theorems 3.6 and 3.9. □

A more advanced proof yields a tighter result:

Theorem 3.11 *With notations and assumptions as above:*

(i) $\lambda_2 \leq 1 - \frac{1}{\bar{\rho}}$

(ii) $\Delta(t) \leq \frac{(1 - 1/\bar{\rho})^t}{\min_{i \in A} \pi_i}$

(iii) $\tau(\varepsilon) \leq \overline{\rho}\left(\ln\frac{1}{\varepsilon} + \ln\frac{1}{\pi_{min}}\right)$.□

Example 3.2 *Random walk on a ring.* Let us consider again the cyclic random walk of Figure 11. Clearly the stationary distribution is $\pi = [\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}]$, and the ergodic flow on each edge $e = (i, i \pm 1)$ is

$$w_e = \pi_i p_{i,i\pm 1} = \frac{1}{n} \cdot \frac{1}{4} = \frac{1}{4n}.$$

An obvious choice for a canonical path connecting nodes k, l is the shortest one, with length

$$|\gamma_{kl}| = \min\{|l-k|, n-|l-k|\}.$$

It is fairly easy to see that each (oriented) edge is now travelled by 1 canonical path of length 1, 2 of length 2, 3 of length 3, ..., $\frac{n}{2}$ of length $\frac{n}{2}$ (actually the last one is just an upper bound). Thus:

$$\begin{aligned}
\overline{\rho} &= \max_e \frac{1}{w_e} \sum_{\gamma_{kl} \ni e} \pi_k \pi_l |\gamma_{ij}| \leq 4n \sum_{r=1}^{n/2} \frac{1}{n^2} \cdot r^2 \\
&= \frac{4}{n} \cdot \frac{1}{6} \cdot \frac{n}{2} \cdot \left(\frac{n}{2}+1\right) \cdot (n+1) = \frac{1}{6}(n+1)(n+2) \\
\Rightarrow & \\
\tau(\varepsilon) &\leq \tfrac{1}{6}(n+1)(n+2)\left(\ln\, n + \ln\tfrac{1}{\varepsilon}\right) \\
&= \tfrac{1}{6} n^2 \left(\ln\, n + \tfrac{1}{\varepsilon}\right) + O\left(n\left(\ln\, n + \ln\tfrac{1}{\varepsilon}\right)\right).
\end{aligned}$$

Example 3.3 *Sampling permutations.* Let us consider the Markov chain whose states are all possible permutations of $[n] = \{1, 2, \ldots, n\}$, and for any permutations $s, t \in S_n$:

$$p_{st} = \begin{cases} \frac{1}{2}, & \text{if } s = t, \\ \frac{1}{2} \cdot \binom{n}{2}^{-1}, & \text{if } s \text{ can be changed to } t \text{ by transposing two elements,} \\ 0, & \text{otherwise} \end{cases}$$

Thus, e.g. for $n = 3$ we obtain the transition graph in Figure 12.

Clearly, the stationary distribution for this chain is $\pi = \left[\frac{1}{n!}, \frac{1}{n!}, \ldots, \frac{1}{n!}\right]$, and the ergodic flow on each edge $\tau = (s, t)$, with $s \neq t$, $p_{st} > 0$, is:

$$w_\tau = \pi_s p_{st} = \frac{1}{n!} \cdot \frac{1}{2} \cdot \binom{n}{2}^{-1}.$$

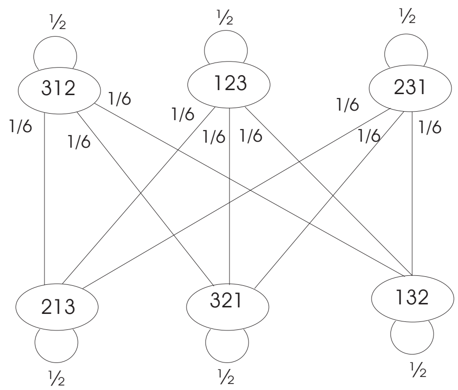

Figure 12: Transition graph for three-element permutations.

A natural canonical path connecting permutation s to permutation t is now obtained as follows:

$$s = s_0 \to s_1 \to s_2 \to \cdots \to s_{n-1} = t.$$

where at each $s_k, s_k(k) = t(k)$. (Thus, each s_k matches t up to element k, $s_k(1\ldots k) = t(1\ldots k)$.)

Thus, e.g. the canonical path connecting $s = (1234)$ to $t = (3142)$ is as follows:

$$(1234) \to \overbrace{(3|214)}^{\omega} \xrightarrow{\tau} \overbrace{(31|24)}^{\omega'} \to (314|2).$$

Now let us denote the set of canonical paths containing a given transition $\tau : \omega \to \omega'$ by $\Gamma(\tau)$. We shall upper bound the size of $\Gamma(t)$ by constructing an injective mapping $\sigma_\tau : \Gamma(\tau) \to S_n$. Obviously, the existence of such a mapping implies that $|\Gamma(\tau)| \leq n!$.

Suppose τ transposes locations $k+1$ and l, $k+1 < l$, of permutation ω. Then for any $\langle s,t\rangle \in \Gamma(\tau)$, define the permutation $z = \sigma_\tau(s,t)$ as follows:

1. Place the elements in $\omega(1\ldots k)$ in the locations they appear in s. (Note that permutation ω is given and fixed as part of τ.)

2. Place the remaining elements in the remaining locations in the order they appear in t.

Thus, for example in the above example case:

$$\sigma_\tau(\langle 1234\rangle, \langle 3142\rangle) \to (\_ \quad \_ \quad 3 \quad \_) \to \underbrace{(1432)}_{z}$$

$$\omega = (3|214), \qquad k = 1$$

Why is this mapping an injection, i.e. how do we recover s and t from a knowledge of τ and $z = \sigma_\tau(s,t)$? The reasoning goes as follows:

1. $t = \omega(1\dots k) +$ "other elements in same order as in z"
2. $s =$ "elements in $\omega(1\dots k)$ at locations indicated in z" + "other elements in locations deducible from the transposition path $s = s_0 \to s_1 \to \cdots \to s_k = \omega$"

This is somewhat tricky, so let us consider an example. Say $\omega = (3 \quad 1|2 \quad 4)$, $k = 2$, $z = (1 \quad 4 \quad 3 \quad 2)$. Then:

1. $t = (3 \quad 1|\_ \quad \_) + (\_ \quad \_|4 \quad 2) = (3 \quad 1|4 \quad 2)$
2.

$$\begin{array}{lllllll}
s & = s_0 & = (1 \;\; _ \;\; 3 \;\; _) & & s_0 & = (1 \;\; _ \;\; 3 \;\; _) \\
 & s_1 & = (3| \;\; _ \;\; _ \;\; _) & \Rightarrow & s_1 & = (3| \;\; 2 \;\; 1 \;\; _) \\
\omega & = s_2 & = (3 \;\; 1| \;\; 2 \;\; 4) & & s_2 & = (3 \;\; 1| \;\; 2 \;\; 4) \\
\hline
\therefore s & = s_0 & = (1 \;\; 2 \;\; 3 \;\; 4) & & s_0 & = (1 \;\; 2 \;\; 3 \;\; 4) \\
 & s_1 & = (3| \;\; 2 \;\; 1 \;\; 4) & \Rightarrow & s_1 & = (3| \;\; 2 \;\; 1 \;\; 4) \\
\omega & = s_2 & = (3 \;\; 1| \;\; 2 \;\; 4) & & s_2 & = (3 \;\; 1| \;\; 2 \;\; 4)
\end{array}$$

Thus, we know that for each transition τ,

$$|\Gamma(\tau)| \leq n!$$

We can now obtain a bound on the unweighted maximum edge loading induced by our collection of canonical paths:

$$\begin{aligned}
\rho &= \max_{\tau \in E} \frac{1}{q_\tau} \sum_{\langle s,t\rangle \in \Gamma(\tau)} \pi_s \pi_t \leq \left(\frac{1}{n!} \cdot \frac{1}{2} \cdot \binom{n}{2}^{-1} \right)^{-1} \cdot n! \cdot \left(\frac{1}{n!} \right)^2 \\
&= 2n! \binom{n}{2} \cdot n! \cdot (\frac{1}{n!})^2 = 2 \cdot \binom{n}{2} = n(n-1).
\end{aligned}$$

By Theorem 3.9, the conductance of this chain is thus $\Phi \geq \frac{1}{2n(n-1)}$, and by Corollary 3.8, its mixing time is thus bounded by

$$\begin{aligned}
\tau_n(\varepsilon) &\leq \frac{2}{\Phi^2} \left(\ln \frac{1}{\varepsilon} + \ln \frac{1}{\pi_{\min}} \right) \leq 2(2n(n-1))^2 \left(\ln \frac{1}{\varepsilon} + \ln n! \right) \\
&= O\left(n^4 \left(n \ln n + \ln \frac{1}{\varepsilon} \right) \right).
\end{aligned}$$