The Kemeny Constant For Finite Homogeneous Ergodic Markov Chains

M. Catral
Department of Mathematics and Statistics
University of Victoria
Victoria, BC
Canada V8W 3R4

S. J. Kirkland
Hamilton Institute
National University of Ireland Maynooth
Maynooth, Co. Kildare
Ireland

M. Neumann
Department of Mathematics
University of Connecticut
Storrs, Connecticut 06269–3009
USA

N.–S. Sze
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hom, Hong Kong

This paper is dedicated to the memory of Professor David Gottlieb 1944–2008

Abstract

A quantity known as the *Kemeny constant*, which is used to measure the expected number of links that a surfer on the World Wide Web, located on a random web page, needs to follow before reaching his/her desired location, coincides with the more well known notion of the *expected time to mixing*, i.e., to reaching stationarity of an ergodic

1

Markov chain. In this paper we present a new formula for the Kemeny constant and we develop several perturbation results for the constant, including conditions under which it is a convex function. Finally, for chains whose transition matrix has a certain directed graph structure we show that the Kemeny constant is dependent only on the common length of the cycles and the total number of vertices and not on the specific transition probabilities of the chain.

1   Introduction

Consider the World Wide Web. One can ask the following question: *If we have a random surfer, who, perhaps having become lost, is located at a random web page, what is the expected number of links that the surfer must follow before reaching his/her desired location?* Levene and Loizou [22] have shown that this expectation is given by a quantity which they call the *Kemeny constant*. This quantity is the same as the quantity known as the *expected time to mixing*, i.e., to reaching stationarity of an ergodic Markov chain which has been explored by Hunter [14] and others. In this paper we shall investigate various properties of the Kemeny constant such as its behavior under perturbations in the transition probabilities of the chain.

The plan of this paper is as follows. In Section 2 we shall present various preliminaries leading to known representations of the Kemeny constant, for example, via the mean first passage times of the chain and its stationary probabilities or via the group generalized inverse $A^{\#}$ of $A = I - T$, where T is the transition matrix of the chain. In Section 3 we shall find a new representation for the Kemeny constant using the inverses of the $(n-1) \times (n-1)$ principal submatrices of A and the diagonal entries of $A^{\#}$.

In Section 4 we shall relate the Kemeny constant to a certain condition number for Markov chains known as $\kappa_6(T)$ in the list of condition numbers in the papers of Cho and Meyer [6, 7]. We mention that condition numbers for Markov chains are used to measure the error in the computation of its stationary distribution vector. In Section 5 we shall consider the effect upon the Kemeny constant of various low rank perturbations of the transition ma-

trix of a chain. We shall also consider conditions under which the Kemeny constant is a convex function of the perturbation.

In the final part, Section 6, of this paper we shall consider the Kemeny constant of a Markov chain whose transition matrix has a certain directed graph structure. For example, if each cycle in the graph has the same length and every pair of cycles intersect, then we shall show that the Kemeny constant is dependent only on the common length of the cycles and the total number of vertices and not on the specific transition probabilities of the chain.

2   Preliminaries Leading To The Kemeny Constant

To introduce the Kemeny constant, we shall require preliminary notions from finite Markov chains. Let $\{X_k\}_{k=0}^{\infty}$ be a finite–state homogeneous ergodic Markov chain on the n states $\mathcal{S}_1, \ldots, \mathcal{S}_n$. For $1 \leq i, j \leq n$, the *mean first passage time* from state $\mathcal{S}_i$ to state $\mathcal{S}_j$, denoted by $m_{i,j}$, is the expected number of time steps to reach state $\mathcal{S}_j$ for the first time when initially the chain is in state $\mathcal{S}_i$. Formally, for $1 \leq i, j \leq n$, $m_{i,j}$ is given by:

$$m_{i,j} \;=\; E(F_{i,j}) \;=\; \sum_{k=1}^{\infty} k Pr(F_{i,j} = k),$$

where $F_{i,j}$ is the random variable representing the smallest number of time–steps for reaching state $\mathcal{S}_j$ for the first time, given that the system was initially in state $\mathcal{S}_i$. That is:

$$F_{i,j} \;=\; \min\{\ell \geq 1 : X_\ell = \mathcal{S}_j | X_0 = \mathcal{S}_i\}.$$

The matrix $M = (m_{i,j})$ is called the *mean first passage matrix* of the chain.

Let $T \in \mathbb{R}^{n,n}$ be the transition matrix of a finite–state homogeneous ergodic Markov chain, so that T is a nonnegative and irreducible matrix. Put $A = I - T$. For $j = 1, \ldots, n$, let A_j denote the $(n-1) \times (n-1)$ principal submatrix of A obtained by deleting its j–th row and column. Then A_j is a nonsingular and diagonally dominant M–matrix. In particular, A_j^{-1} is a nonnegative matrix. Let

$$\overline{M}_j \;=\; [m_{1,j}, \ldots, m_{j-1,j}, m_{j+1,j}, \ldots, m_{n,j}]^T$$

be the j–th column of the mean first passage matrix with the j–th entry deleted. An expression for the mean first passage times in terms of the row sums of A_j^{-1} is as follows (see e.g., [6], [11], [8]):

$$\overline{M}_j \;=\; A_j^{-1}e, \tag{2.1}$$

where $e \in \mathbb{R}^{n-1}$ denotes the vector of all ones.

Next, for $A \in \mathbb{R}^{n,n}$, the *group inverse* of A if it exists, is the unique matrix $A^{\#}$ satisfying the matrix equations $AA^{\#}A = A$, $A^{\#}AA^{\#} = A^{\#}$, and $AA^{\#} = A^{\#}A$. (We note in passing that the group inverse exists if and only if $\text{rank}(A) = \text{rank}(A^2)$.) Let $A = I - T$, where T is the transition matrix belonging to the finite homogeneous ergodic Markov chain mentioned above. In [24], Meyer found the following expression for the mean first passage matrix M in terms of the group inverse $A^{\#}$ and the stationary distribution vector $\pi = [\pi_1, \ldots, \pi_n]^T$ of the chain:

$$M \;=\; (I - A^{\#} + JA_d^{\#}) \begin{bmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_n} \end{bmatrix}, \tag{2.2}$$

where J is the $n \times n$ matrix of all ones and $A_d^{\#}$ is the diagonal matrix whose entries are the corresponding diagonal entries of $A^{\#}$.

Further, in [24], an expression for $A^{\#}$ in terms of A_n^{-1} and the stationary distribution vector is given. In particular, let B be the leading $(n-1)\times(n-1)$ principal submatrix of $A^{\#}$. Then

$$B \;=\; A_n^{-1} + \beta W - A_n^{-1}W - WA_n^{-1}, \tag{2.3}$$

where $\beta := u^T A_n^{-1} e$, $W = eu^T$, and $\pi = [u \; \pi_n]^T$ is the stationary distribution vector of T.

We are now ready to introduce the Kemeny constant. For our n–state homogeneous ergodic Markov chain, several authors have shown that for the expressions

$$\sum_{j=1}^{n} m_{i,j}\pi_j, \quad i = 1, \ldots, n, \tag{2.4}$$

all have a common value independent of i, see e.g., Kemeny and Snell [15], Grinstead and Snell [10], Levene and Loizou [22], and Hunter [14]. Following their analysis of the expected time of exiting a random site, and

entering another desired site, Levene and Loizou call this common value for the chain *Kemeny constant*. For a finite state, homogeneous, and ergodic chain Markov chain with a transition matrix T, we shall denote by $K(T)$ its Kemeny constant.

In [14, Theorem 2.4] Hunter shows that the quantity which we call here the Kemeny constant is given by

$$K(T) \; = \; 1 + \text{trace}(A^{\#}), \qquad (2.5)$$

which can be deduced from (2.2). The fundamental matrix of the chain is defined by $Z = (A + e\pi^T)^{-1}$. Notice that the fundamental matrix Z and the group inverse $A^{\#}$ are related by $Z = A^{\#} + e\pi^T$. Thus,

$$K(T) \; = \; \text{trace}(Z). \qquad (2.6)$$

Let $\lambda_2, \ldots, \lambda_n$ be the eigenvalues of T other than 1. As the nonzero eigenvalues of $A^{\#}$ are the reciprocals of the nonzero eigenvalues of $A = I - T$, we immediately have that:

$$K(T) = 1 + \sum_{i=2}^{n} \frac{1}{1 - \lambda_i}. \qquad (2.7)$$

In the next section we shall develop yet another representation for the Kemeny constant.

3 A New Representation For The Kemeny Constant

In this section we obtain a new representation for the Kemeny constant for the transition matrix T of a finite–state homogeneous ergodic Markov chain.

Proposition 3.1 *Let T be the transition matrix of an n–state ergodic Markov chain and let $A = I - T$. Let $K(T)$ be the Kemeny constant associated with the chain. Then for any $j = 1, \ldots, n$,*

$$K(T) \; = \; \text{trace}\left(A_j^{-1}\right) - \frac{A_{j,j}^{\#}}{\pi_j} + 1, \qquad (3.8)$$

where A_j is the $(n-1) \times (n-1)$ principal submatrix of A obtained by deleting the j–th row and column and $\pi = [\pi_1, \ldots, \pi_n]^T$ is the stationary distribution vector of T.

5

Proof. Without loss of generality we can assume that $j = n$. Then for $i, j = 1, \ldots, n-1$ with $i \neq j$, on using (2.2), we obtain the well–known relation that

$$m_{i,j}\pi_j = A_{j,j}^{\#} - A_{i,j}^{\#}. \tag{3.9}$$

Now using (2.3) we get that:

$$A_{i,j}^{\#} = \left(A_n^{-1}\right)_{i,j} + \beta\pi_j - (A_n^{-1}e)_i\pi_j - \sum_{k=1}^{n-1} \pi_k(A_n^{-1})_{k,j}, \tag{3.10}$$

and

$$A_{j,j}^{\#} = (A_n^{-1})_{j,j} + \beta\pi_j - (A_n^{-1}e)_j\pi_j - \sum_{k=1}^{n-1} \pi_k(A_n^{-1})_{k,j}. \tag{3.11}$$

But then from (3.9)–(3.11) we have that:

$$m_{i,j}\pi_j = (A_n^{-1})_{j,j} - (A_n^{-1})_{i,j} + \pi_j[(A_n^{-1}e)_i - (A_n^{-1}e)_j]$$

and hence, using (2.1), it follows that:

$$m_{i,j}\pi_j = (A_n^{-1})_{j,j} - (A_n^{-1})_{i,j} + \pi_j[m_{in} - m_{jn}], \tag{3.12}$$

for all $i, j = 1, \ldots, n-1$ with $i \neq j$. Summing over $j = 1, \ldots, n-1$ with $j \neq i$ yields that:

$$\sum_{j\neq i} m_{i,j}\pi_j = \operatorname{trace}(A_n^{-1}) - (A_n^{-1}e)_i + m_{i,n}\sum_{j\neq i}\pi_j - \sum_{j\neq i}\pi_j m_{j,n}$$

$$= \operatorname{trace}(A_n^{-1}) - m_{i,n} + m_{i,n}(1 - \pi_i - \pi_n) - \sum_{j\neq i}\pi_j m_{j,n}$$

$$= \operatorname{trace}(A_n^{-1}) - \pi_i m_{i,n} - m_{i,n}\pi_n - \sum_{j\neq i}\pi_j m_{j,n}.$$

Thus,

$$m_{i,n}\pi_n + \sum_{\substack{j = 1 \\ j \neq i}}^{n-1} m_{i,j}\pi_j = \operatorname{trace}(A_n^{-1}) - \pi_i m_{i,n} - \sum_{\substack{j = 1 \\ j \neq i}}^{n-1}\pi_j m_{j,n},$$

or, equivalently,

$$\sum_{\substack{j = 1 \\ j \neq i}}^{n} m_{i,j}\pi_j = \operatorname{trace}(A_n^{-1}) - \sum_{j=1}^{n-1}\pi_j m_{j,n}. \tag{3.13}$$

6

The result follows since the left hand side of (3.13) is equal to $K - 1$ and the relation

$$\sum_{\substack{j\,=\,1 \\ j\,\neq\,k}}^{n} \pi_j m_{j,k} = \frac{A_{k,k}^{\#}}{\pi_k}, \quad k = 1, \ldots, n \tag{3.14}$$

is a consequence of (2.2). $\qquad\square$

As a remark, we relate (3.8) in Proposition 3.1 to group inverses of M–matrices with uniform diagonal entries. These were considered by Kirkland and Neumann in [17]. In one of the results they give a necessary and sufficient condition for the Laplacian of an undirected weighted n–cycle to possess a group inverse with uniform diagonal. A consequence of Proposition 3.1 is the following equivalent condition in terms of the sum of the diagonal entries of the matrices A_j^{-1}, $j = 1, \ldots, n$.

Proposition 3.2 *Let $A = I - T$, where T is an $n \times n$ irreducible doubly stochastic matrix. Then $A^{\#}$ has uniform diagonal entries if and only if* $\mathrm{trace}(A_j^{-1})$ *is constant for $j = 1, \ldots, n$.*

Proof. Since T is doubly stochastic, $\pi_j = \frac{1}{n}$ for $j = 1, \ldots, n$. The result now follows from (3.8). $\qquad\square$

4 Relation of the Kemeny Constant to Condition Numbers for Markov Chains

Let $T \in \mathbb{R}^{n,n}$ and $\hat{T} = T - E$ be transition matrices for two arbitrary, but fixed, finite–state homogeneous ergodic Markov chains and denote their stationary distribution vectors by π and $\hat{\pi}$, respectively. A function $\kappa(\cdot)$ is called a *condition number for Markov chains with respect to the norm pair* (α, β) if for the vector norm $\| \cdot \|_\alpha$ and the matrix norm $\| \cdot \|_\beta$,

$$\|\pi - \hat{\pi}\|_\alpha \leq \kappa(T)\|E\|_\beta.$$

In [14, Theorem 5.1], Hunter shows that $\kappa(T) = K(T) - 1$ is a condition number for Markov chains with respect to the norm pair $(1, \infty)$. Various condition numbers were investigated in the works of Cho and Meyer, see [6, 7], who were able to order $7 - 8$ different condition numbers according to their magnitude. One of the eight condition numbers, due to Seneta [27], is

$$\kappa_6(T) = \frac{1}{2} \max_{1 \leq i,j \leq n} \sum_{s=1}^{n} |A_{i,s}^{\#} - A_{j,s}^{\#}|,$$

which, like $K - 1$, is a condition number with respect to the norm pair $(1, \infty)$. In Kirkland, Neumann, and Sze [19] it has been shown that $\kappa_6(T)$ is optimal with respect to the norm pair $(1, \infty)$ in the sense that there is no other condition number with respect to the pair $(1, \infty)$ whose value at some T is less than $\kappa_6(T)$. We next show the connection between $K(T)$ and $\kappa_6(T)$.

Theorem 4.1 *Suppose that $T \in \mathbb{R}^{n,n}$ is a transition matrix of a finite–state homogeneous ergodic Markov chain with π as its stationary distribution vector. Then*

$$K(T) \ \geq \ \frac{n + \kappa_6(T)}{1 + \kappa_6(T)},$$

with equality holding if and only if

$$T \ = \ -\kappa_6(T)I + (1 + \kappa_6(T))e\pi^T.$$

Furthermore, if $\kappa_6(T) < 1$, then

$$K(T) \ \leq \ \frac{n - \kappa_6(T)}{1 - \kappa_6(T)},$$

with equality holding if and only if

$$T \ = \ \kappa_6(T)I + (1 - \kappa_6(T))e\pi^T.$$

Proof. Label the non–Perron eigenvalues of T as $\lambda_2, \ldots, \lambda_n$, and recall that

$$K(T) \ = \ 1 + \sum_{j=2}^{n} \frac{1}{1 - \lambda_j} \ = \ 1 + \sum_{j=2}^{n} \Re\left(\frac{1}{1 - \lambda_j}\right).$$

Fix an index j between 2 and n. Writing $\lambda_j = x_j + iy_j$, we have

$$\Re\left(\frac{1}{1 - \lambda_j}\right) \ = \ \frac{1 - x_j}{(1 - x_j)^2 + y_j^2} \ \geq \ \frac{1}{1 + \kappa_6(T)},$$

the last inequality follows from the fact that $x_j \geq -\kappa_6(T)$. Thus we find that

$$K(T) \ \geq \ 1 + \frac{n - 1}{1 + \kappa_6(T)} \ = \ \frac{n + \kappa_6(T)}{1 + \kappa_6(T)}.$$

Note also that $K(T) = \frac{n+\kappa_6(T)}{1+\kappa_6(T)}$ if and only if $\lambda_j = -\kappa_6(T), j = 2, \ldots, n$. In that case, it follows from Theorem 2.7 in [18] that, associated with the

eigenvalue $-\kappa_6(T)$, there are $n-1$ linearly independent eigenvectors. Referring to the decomposition of T in terms of orthogonal idempotents, there is a matrix E such that $e\pi^T + E = I$ and $T = e\pi^T - \kappa_6(T)E$. It follows readily that this is equivalent to

$$T = -\kappa_6(T)I + (1 + \kappa_6(T))e\pi^T.$$

Suppose next that $\kappa_6(T) < 1$. Then for each $j = 2, \ldots, n$, we have

$$\Re\left(\frac{1}{1 - \lambda_j}\right) \leq \frac{1}{|1 - \lambda_j|} \leq \frac{1}{1 - |\lambda_j|} \leq \frac{1}{1 - \kappa_6(T)}.$$

Hence we find that

$$K(T) \leq 1 + \frac{n - 1}{1 - \kappa_6(T)} = \frac{n - \kappa_6(T)}{1 - \kappa_6(T)},$$

as desired. Note also that $K(T) = \frac{n - \kappa_6(T)}{1 - \kappa_6(T)}$ if and only if $\lambda_j = \kappa_6(T)$, for $j = 2, \ldots, n$. In that case, as above we find that there is a matrix E such that $e\pi^T + E = I$ and $T = ee^T + \kappa_6(T)E$. This in turn yields that

$$T = \kappa_6(T)I + (1 - \kappa_6(T))e\pi^T.$$

$\square$

5 Perturbation and convexity of the Kemeny constant

In this section, we consider elementary row and column perturbations on a transition matrix of a Markov chain and determine the effects of such perturbations on the Kemeny constant of the chain. In particular, we are interested in determining whether a given perturbation of the transition probabilities will result in a decrease in $K(T)$, the mixing time of the chain.

We will first consider these two types of elementary perturbations:

Type I: $\overline{T} = T + e_r h^T$, where $e_r = [0 \ldots 0 \ \ 1 \ \ 0 \ldots 0]^T$ is the unit vector in $\mathbb{R}^n$ with zeros everywhere except for a 1 in the r–th position, and $h = [h_1, \ldots, h_n]^T$ is any zero sum vector in $\mathbb{R}^n$ such that $\overline{T}$ remains an irreducible transition matrix. Notice that in this perturbation only the r–th

row of T is changed.

Type II: $\overline{T} = T + eh^T$, where $h = [h_1, \ldots, h_n]^T$ is any zero sum vector in $\mathbb{R}^n$ such that $\overline{T}$ remains an irreducible transition matrix.

5.1 Type I perturbations

In order to state our main result concerning the effect of Type I perturbations on the Kemeny constant, we need to recall the so called entrywise *column diagonal dominance property of inverses* of diagonally dominant matrices which is originally due to Metzler [23] and which has been further extended by Fiedler and Ptak in [9]:

Lemma 5.1 (See, for example, [2, Lemma 3.14, page 254]) *Let A be a nonsingular M–matrix of order n whose row sums are all nonnegative; that is, $Ae \geq 0$. Then the entries of A^{-1} satisfy*

$$(A^{-1})_{i,i} \geq (A^{-1})_{k,i}, \quad \text{for all } i, k = 1, \ldots, n.$$

We can now state our first perturbation result:

Theorem 5.2 *Let T be an irreducible and stochastic matrix with stationary distribution vector $\pi = [\pi_1, \ldots, \pi_n]^T$. Let $\overline{T} = T + e_r h^T$, where $h = [h_1, \ldots, h_n]^T$ is any vector in $\mathbb{R}^n$ such that $\overline{T}$ is still irreducible and stochastic, and let $\overline{\pi} = [\overline{\pi}_1, \ldots, \overline{\pi}_n]^T$ be the stationary distribution vector for $\overline{T}$. Put $A = I - T$, $\overline{A} = I - \overline{T}$ and suppose that $M = (m_{i,j})$ and $\overline{M} = (\overline{m}_{i,j})$ are the associated mean first passage matrices of T and $\overline{T}$, respectively. Then*

(i) $\overline{m}_{i,r} = m_{i,r}$, *for all $i \neq r$,*

and

(ii) $\overline{m}_{i,j} \geq m_{i,j} \Longleftrightarrow \overline{\pi}_j \leq \pi_j$, *for all $i, j \neq r$.*

Proof. Without loss of generality, assume that $r = n$. Otherwise, perform a symmetric permutation on the rows and columns of T, interchanging the roles of the indices r and n.

Since the rows of T and $\overline{T}$ coincide except for the n–th row, $\overline{A}_n = A_n$. Using (2.1) we now obtain that:

$$\overline{m}_{i,n} \;=\; (\overline{A}_n^{-1}e)_i \;=\; (A_n^{-1}e)_i \;=\; m_{i,n}, \;\; \text{for } i \neq n.$$

This proves (i). Next, for $i, j = 1, \dots, n-1$, from (3.12), (2.2), and since $(\overline{A}_n^{-1})_{i,j} = (A_n^{-1})_{i,j}$, we have that:

$$\overline{m}_{i,j} - m_{i,j} \;=\; \begin{cases} \left[(A_n^{-1})_{j,j} - (A_n^{-1})_{i,j}\right] \left[\dfrac{1}{\overline{\pi}_j} - \dfrac{1}{\pi_j}\right], & i \neq j \\[2mm] \dfrac{1}{\overline{\pi}_j} - \dfrac{1}{\pi_j}, & i = j. \end{cases} \tag{5.15}$$

Thus, (ii) holds in the case that $i = j$. Finally, suppose that $i \neq j$. Since $Ae = 0$ and A is an irreducible M–matrix, then A_n is a nonsingular diagonally dominant M–matrix. Hence, by Lemma 5.1,

$$(A_n^{-1})_{j,j} \;\geq\; (A_n^{-1})_{i,j}. \tag{5.16}$$

Combining (5.16) and the first part of (5.15) yields that

$$\overline{m}_{i,j} \;\geq\; m_{i,j} \iff \overline{\pi}_j \;\leq \pi_j,$$

and the proof is complete. $\qquad\square$

Notice that if T is perturbed by a Type I perturbation, say in the r–th row, then using Proposition 3.1 and the fact that $\overline{A}_r = A_r$, the Kemeny constants $K = K(T)$ and $\overline{K} = K(\overline{T})$ obey the following relation:

$$\overline{K} - K \;=\; \frac{A_{r,r}^{\#}}{\pi_r} - \frac{\overline{A}_{r,r}^{\#}}{\overline{\pi}_r}. \tag{5.17}$$

This leads us to our second result on Type I perturbations:

Theorem 5.3 *Let T be an irreducible and stochastic matrix with stationary distribution vector $\pi = [\pi_1, \dots, \pi_n]^T$. Let $\overline{T} = T + e_r h^T$ be a Type I perturbation of T and suppose that $K = K(T)$ and $\overline{K} = K(\overline{T})$ are the respective Kemeny constants. Then:*

$$\overline{K} \;\leq\; K \iff \sum_{i \neq r} (\overline{\pi}_i - \pi_i)m_{i,r} \geq 0. \tag{5.18}$$

Proof. By (i) in Theorem 5.2, $\overline{m}_{i,r} = m_{i,r}$, for $i \neq r$. Combining this with (5.17) and (3.14) we get that:

$$\overline{K} - K = \sum_{i \neq r} \pi_i m_{i,r} - \sum_{i \neq r} \overline{\pi}_i \overline{m}_{i,r} = \sum_{i \neq r} (\pi_i - \overline{\pi}_i) m_{i,r},$$

gives (5.18). $\qquad\square$

5.2   Type II perturbations

Since formula (2.7) expresses the Kemeny constant in terms of the eigenvalues other than 1 of T, it is logical to ask to what extent can we perturb T while maintaining control over the change in its eigenvalues? We begin with the following well known perturbation result:

Lemma 5.4 (Brauer [3]) *Let T be an irreducible and stochastic matrix and let h be any vector in $\mathbb{R}^n$. Then the eigenvalues of $T + eh^T$ are given by*

$$1 + h^T e, \ \lambda_2, \ \ldots, \ \lambda_n,$$

where $\lambda_2, \ldots, \lambda_n$ are the eigenvalues of T other than 1.

A direct consequence of the above lemma is that a Type II perturbation leaves unchanged the spectrum of the transition matrix and hence also the Kemeny constant for the chain.

Theorem 5.5 *Let $\overline{T} = T + eh^T$ be a Type II perturbation of T and let $K(T)$ and $K(\overline{T})$ be the respective Kemeny constants. Then $K(\overline{T}) = K(T)$.*

Proof. The result is immediate by the preceding lemma due to the fact that a Type II perturbation requires the entries of h to sum to 0. $\qquad\square$

The equality in (3.14) gives a practical interpretation for the ratios $\frac{A_{j,j}^{\#}}{\pi_j}$, $j = 1, \ldots, n$. Since for a given j, $\sum_{k=1}^{n} \pi_k m_{k,j}$ is the weighted average of the mean first passage times into $\mathcal{S}_j$ from any other state, the expression $\frac{A_{j,j}^{\#}}{\pi_j}$ can be thought of as a measure of such an average mean first passage time into state $\mathcal{S}_j$ from any other state.

In the next proposition it is shown that if the transition matrix is perturbed by uniformly increasing the entries of the j–th column by $\varepsilon > 0$

and then uniformly decreasing the entries of the i–th column by the same amount, then the "average" mean first passage time into state j will not decrease and the "average" mean first passage time into state i will not increase.

Proposition 5.6 *Let $\overline{T} = T + \varepsilon e(e_i - e_j)^T$, where T is a stochastic and irreducible matrix and suppose that $\varepsilon > 0$ is chosen so that $\overline{T}$ remains an irreducible transition matrix. Suppose that $\pi = [\pi_1, \ldots, \pi_n]^T$ and $\overline{\pi} = [\overline{\pi}_1, \ldots, \overline{\pi}_n]^T$ are the respective stationary distribution vectors of T and $\overline{T}$. Then*

$$\frac{\overline{A}_{j,j}^{\#}}{\overline{\pi}_j} \geq \frac{A_{j,j}^{\#}}{\pi_j}, \tag{5.19}$$

and

$$\frac{\overline{A}_{i,i}^{\#}}{\overline{\pi}_i} \leq \frac{A_{i,i}^{\#}}{\pi_i}. \tag{5.20}$$

Proof. We begin by noticing that by the assumptions of the theorem, $\overline{T}$ is obtain from T via a Type II perturbation. Let T_j and $\overline{T}_j$ be the respective $(n-1) \times (n-1)$ principal submatrices of T and $\overline{T}$ obtained by deleting their j–th row and column. Then

$$\overline{T}_j = T_j + \varepsilon e e_i^T$$

and hence, as $\varepsilon > 0$, we can write that:

$$\begin{aligned} \overline{A}_j &= I - \overline{T}_j \\[2mm] &= I - (T_j + \varepsilon e e_i^T) \\[2mm] &= A_j - \varepsilon e e_i^T \\[2mm] &\leq A_j. \end{aligned}$$

Thus $\overline{A}_j$ and A_j are nonsingular M–matrices satisfying $\overline{A}_j \leq A_j$ and from M–matrix theory we have that:

$$\overline{A}_j^{-1} \geq A_j^{-1}. \tag{5.21}$$

Now using Proposition 3.1 we see that:

$$K(\overline{T}) - K(T) = \text{trace}\left(\overline{A}_j^{-1} - A_j^{-1}\right) + \left(\frac{A_{j,j}^{\#}}{\pi_j} - \frac{\overline{A}_{j,j}^{\#}}{\overline{\pi}_j}\right). \tag{5.22}$$

The left-hand side of (5.22) is zero by Theorem 5.5 and thus (5.19) follows from (5.21).

The proof of (5.20) is similar.

$\square$

We list two more types of perturbations on T that lead to a reduction in the mixing time K for the chain. In both cases, the representation of $K(T)$ in (2.7) is used.

Theorem 5.7 *Let T be a symmetric, stochastic and irreducible matrix, and suppose that $\overline{T} = T - E$, where E is a positive semidefinite matrix such that $\overline{T}$ remains stochastic and irreducible. Let $K = K(T)$ and $\overline{K} = K(\overline{T})$ be their associated Kemeny constants. Then $\overline{K} \leq K$.*

Proof. Let λ_i and $\overline{\lambda}_i$, $i = 2, \ldots, n$, be the eigenvalues of T and $\overline{T}$ respectively, that are different from 1. Then by a theorem of Weyl [12, Corollary 4.3.3], $\overline{\lambda}_i \leq \lambda_i$, $i = 2, \ldots, n$. It follows that

$$\sum_{i=2}^{n} \frac{1}{1 - \overline{\lambda}_i} \leq \sum_{i=2}^{n} \frac{1}{1 - \lambda_i},$$

hence $\overline{K} \leq K$ by (2.7). $\square$

Corollary 5.8 *Let $M = (m_{ij})$ and $\overline{M} = (\overline{m}_{ij})$ be the mean first passage matrices associated with T and $\overline{T}$ respectively in the theorem above. Then*

$$\sum_{j=1}^{n} \overline{m}_{ij} \leq \sum_{j=1}^{n} m_{ij}.$$

Proof. First note that since T is symmetric, so is $\overline{T}$. Let $\pi = [\pi_1, \ldots, \pi_n]^T$ and $\overline{\pi} = [\overline{\pi}_1, \ldots, \overline{\pi}_n]^T$ be the respective stationary distribution vectors for T and $\overline{T}$. Since T and $\overline{T}$ are symmetric and stochastic, $\pi = \overline{\pi} = \frac{1}{n}e$ and so $\pi_j = \overline{\pi}_j = \frac{1}{n}$, for $j = 1, \ldots, n$. From Theorem 5.7, $\overline{K} \leq K$ and thus, $\sum_{j=1}^{n} m_{ij}\pi_j = K \leq \overline{K} = \sum_{j=1}^{n} \overline{m}_{ij}\overline{\pi}_j$. The result now follows. $\square$

Let T be a stochastic and irreducible matrix and suppose that $0 \leq \alpha \leq 1$. Suppose further that $v = [v_1, \ldots, v_n]^T$ is a positive probability vector, i.e., $v_j > 0$, for all j and $v^T e = 1$. Let

$$\overline{T} = \alpha T + (1 - \alpha)ev^T. \tag{5.23}$$

Then $\overline{T}$ remains a stochastic and irreducible matrix. Lemma 5.4 is next used to find the eigenvalues of $\overline{T}$ in terms of the eigenvalues of T. First, the eigenvalues of $\overline{T}$ are given in the following lemma and we shall need them in subsequent results:

Lemma 5.9 (Serra–Capizzano [5, Theorem 2.1]) *Let T be stochastic and irreducible, and let $\overline{T}$ be as in (5.23). Let $\sigma(T) = \{1, \lambda_2, \ldots, \lambda_n\}$ be the set of eigenvalues of T. Then the spectrum of eigenvalues of $\overline{T}$ is given by*

$$\sigma(\overline{T}) = \{1, \alpha\lambda_2, \ldots, \alpha\lambda_n\}.$$

Lemma 5.9 leads to the following perturbation results for the Kemeny constant:

Theorem 5.10 *Let T be an $n \times n$ irreducible and stochastic matrix with nonnegative spectrum $\sigma(T) = \{1, \lambda_2, \cdots, \lambda_n\}$. Suppose that $\overline{T} = \alpha T + (1-\alpha)ev^T$, where $0 \leq \alpha \leq 1$, and v is a positive probability vector. Then $K(\overline{T}) \leq K(T)$.*

Proof. By Lemma 5.9, the eigenvalues of $\overline{T}$ are given by $\sigma(\overline{T}) = \{1, \alpha\lambda_2, \cdots, \alpha\lambda_n\}$. Since $0 \leq \alpha \leq 1$ and the λ_i's are nonnegative, $\alpha\lambda_i \leq \lambda_i$, $i = 2, \cdots, n$. It follows that

$$\sum_{i=2}^{n} \frac{1}{1-\alpha\lambda_i} \leq \sum_{i=2}^{n} \frac{1}{1-\lambda_i}$$

and hence $K(\overline{T}) \leq K(T)$ by (2.7). $\qquad\square$

We note the if the assumption that T has nonnegative spectrum is removed in Theorem 5.10, the conclusion $K(\overline{T}) \leq K(T)$ no longer follows.

One can interpret Theorem 5.10 to say that if T_1 is an $n \times n$ irreducible and stochastic matrix with a nonnegative spectrum and $T_2 = ev^T$, with $v \in \mathbb{R}^n$ a positive probability vector, then $K[\alpha T_1 + (1-\alpha)T_2] \leq K(T_1)$ for $0 \leq \alpha \leq 1$. This suggests that we consider the question of convexity of the Kemeny constant. We can prove the following result from which it will follow that, indeed, even under the conditions of Theorem 5.10, we have that:

$$K[\alpha T_1 + (1-\alpha)T_2] \leq \alpha K(T_1) + (1-\alpha)K(T_2), \quad \text{for all } \alpha \in [0,1].$$

Theorem 5.11 *Let T_1 be an $n \times n$ irreducible and stochastic matrix and let $T_2 = ev^T$, where $v \in \mathbb{R}^n$ is a positive probability vector. If*

$$\sigma(T_1) \subseteq [0,1] \cup \{z \in \mathbb{C} : 3\pi/4 \leq \arg(z) \leq 5\pi/4\},$$

then, for $0 \leq \alpha \leq 1$,

$$K[\alpha T_1 + (1-\alpha)T_2] \leq \alpha K(T_1) + (1-\alpha)K(T_2). \qquad (5.24)$$

Proof. Let $\sigma(T_1) = \{1, \lambda_2, \ldots, \lambda_n\}$ satisfy the assumption in the theorem. Now by (2.7),

$$K(T_1) = 1 + \sum_{k=2}^{n} \frac{1}{1 - \lambda_k}. \qquad (5.25)$$

Since rank $T_2 = 1$, (2.7) implies that

$$K(T_2) = n. \qquad (5.26)$$

Next, by Lemma 5.9 and again using (2.7) we can write that

$$K\left([\alpha T_1 + (1-\alpha)T_2]\right) = 1 + \sum_{k=2}^{n} \frac{1}{1 - \alpha\lambda_k},$$

and the left hand side of (5.24) can be written

$$1 + \sum_{k=2}^{n} \left(1 + \frac{\alpha\lambda_k}{1 - \alpha\lambda_k}\right).$$

On the other hand, the right hand side of (5.24) is equal to

$$\alpha\left(1 + \sum_{k=2}^{n} \frac{1}{1 - \lambda_k}\right) + (1-\alpha)n = 1 + \sum_{k=2}^{n} \left(1 + \frac{\alpha\lambda_k}{1 - \lambda_k}\right) \qquad (5.27)$$

by (5.25) and (5.26). Thus, (5.24) holds if we can prove that

$$\sum_{k=2}^{n} \left(\frac{\alpha\lambda_k}{1 - \alpha\lambda_k}\right) \leq \sum_{k=2}^{n} \left(\frac{\alpha\lambda_k}{1 - \lambda_k}\right), \qquad (5.28)$$

for $0 \leq \alpha \leq 1$. Notice that

$$\sum_{k=2}^{n} \left(\frac{\alpha\lambda_k}{1 - \lambda_k}\right) - \sum_{k=2}^{n} \left(\frac{\alpha\lambda_k}{1 - \alpha\lambda_k}\right) = \sum_{k=2}^{n} \alpha\lambda_k \left(\frac{1}{1 - \lambda_k} - \frac{1}{1 - \alpha\lambda_k}\right)$$

$$= \sum_{k=2}^{n} \frac{\alpha(1-\alpha)\lambda_k^2}{(1 - \lambda_k)(1 - \alpha\lambda_k)}$$

is real and so it suffices to show that the real part of each term $\frac{\lambda_k^2}{(1-\lambda_k)(1-\alpha\lambda_k)}$ is nonnegative for all $\alpha \in [0,1]$. It is clearly true if $\lambda_k \in [-1,1)$. Suppose λ_k is a complex number in $\{z \in \mathbb{C} : 3\pi/4 \leq \arg(z) \leq 5\pi/4\}$. Write $\lambda_k = \lambda = x + iy$. Then $-1 < x < 0$ and $x^2 \geq y^2$ and therefore,

$$
\begin{aligned}
\Re\left(\frac{\lambda_k^2}{(1-\lambda_k)(1-\alpha\lambda_k)}\right) &= \Re\left(\frac{\lambda^2(1-\overline{\lambda})(1-\alpha\overline{\lambda})}{|1-\lambda|^2|1-\alpha\lambda|^2}\right) \\[2mm]
&= \Re\left(\frac{\alpha|\lambda|^4 - (1+\alpha)\lambda|\lambda|^2 + \lambda^2}{|1-\lambda|^2|1-\alpha\lambda|^2}\right) \\[2mm]
&= \frac{\alpha|\lambda|^4 - (1+\alpha)x|\lambda|^2 + (x^2-y^2)}{|1-\lambda|^2|1-\alpha\lambda|^2} \geq 0.
\end{aligned}
$$

$\square$

The inequality (5.24) also holds when both T_1 and T_2 are symmetric doubly stochastic matrices.

Theorem 5.12 *Suppose T_1 and T_2 are two $n \times n$ irreducible symmetric doubly stochastic matrices. Then for $0 \leq \alpha \leq 1$,*

$$K[\alpha T_1 + (1-\alpha)T_2] \leq \alpha K(T_1) + (1-\alpha)K(T_2).$$

Proof. We begin by recalling that the map $A \mapsto \mathrm{trace}(A^{-1})$ is strictly convex on the set of the positive definite matrices, see, for example, Horn and Johnson [12, Problem 7.6.2]. Since both $(I - T_1 + J_n)$ and $(I - T_2 + J_n)$, where $J_n = \frac{1}{n}ee^t$, are positive definite,

$$
\begin{aligned}
&\mathrm{trace}([\alpha(I - T_1 + J_n) + (1-\alpha)(I - T_2 + J_n)]^{-1}) \\
&\quad \leq \ \alpha\,\mathrm{trace}[(I - T_1 + J_n)^{-1}] + (1-\alpha)\mathrm{trace}[(I - T_2 + J_n)^{-1}]. \quad (5.29)
\end{aligned}
$$

By (2.6) and the fact that T_1, T_2, and $\alpha T_1 + (1-\alpha)T_2$ have the common stationary distribution vector $\frac{1}{n}e$, the right hand side of (5.29) equals

$$K[\alpha T_1 + (1-\alpha)T_2] = \mathrm{trace}([I - (\alpha T_1 + (1-\alpha)T_2) + J_n]^{-1})$$

while the left hand side of (5.29) becomes $\alpha K(T_1) + (1-\alpha)K(T_2)$. The result therefore follows. $\square$

We mention that there are counter–examples showing that (5.24) does not hold for general irreducible and stochastic T_1 and T_2.

6   The Kemeny constant for certain digraphs

Recall that the *directed graph*, or *digraph*, D of an $n \times n$ matrix $B = (b_{i,j})$, consists of n vertices which can be conveniently labeled $1, \ldots, n$, and directed edges with a directed edge connecting vertex i with vertex j if and only if $b_{i,j} \neq 0$. A *strongly connected digraph* is a digraph in which there is a sequence of directed edges, *a walk*, linking any two vertices. Finally, a cycle in the graph is a sequence of directed edges which starts and ends at the same vertex, and passes through distinct intermediate vertices. Similarly, a path is a sequence of directed edges which passes through distinct intermediate vertices.

In this section we shall investigate the Kemeny constant of Markov chains with an underlying digraph such that the constant is only dependent on the graph structure and not on the actual transition probabilities. For this purpose let D be a strongly connected digraph on n vertices and let Σ_D denote the set of all stochastic matrices of order n whose digraph is contained in D. We begin with the following lemma which finds necessary conditions for the dependency of the Kemeny constant only on the digraph.

Lemma 6.1 *Let D be a strongly connected digraph on n vertices. Suppose that there is γ such that for all irreducible $T \in \Sigma_D$, $K(T) = \gamma$. Then*

(a) *all cycles in D have the same length, say k;*

(b) *every pair of cycles in D intersect in at least one vertex;*

(c) *$(I - T)^{\#}$ exists for all $T \in \Sigma_D$;*

(d) *For all $T \in \Sigma_D$, $K(T) = \frac{2n-k+1}{2}$.*

Proof. (a) Fix a cycle C of D. Let $T \in \Sigma_D$ be a $(0, 1)$–matrix such that the arcs of C correspond to nonzero entries in T and such that C is the unique cycle in the digraph of T, and such that the vertices which are not on C are not on any cycle in the digraph of T. One can construct such a T as follows. For each vertex i not on C, let d_i denote the length of a shortest path from i to a vertex of C, and let p denote the maximum of the d_is. Partition the vertices of D not on C as $S_1, \ldots, S_p$, where $S_j = \{i | d_i = j\}, j = 1, \ldots, p$. Note that for each $i \in S_1$, there is a vertex $m(i)$ on C such that $i \to m(i)$ is an arc of D. Similarly, for each $j = 2, \ldots, p$ and each vertex $i \in S_j$, there is a vertex $m(i) \in S_{j-1}$ such that $i \to m(i)$ is an arc of D. It is readily verified (by induction on p, for example) that subgraph $\hat{D}$ of D formed from the

arcs of C and the arcs $i \to m(i)$ as i ranges over the vertices not on C, has exactly one cycle, namely C. We may then take T to be the $(0,1)$ adjacency matrix of $\hat{D}$.

Let k be the length of the cycle C. Then T can be written as a 2×2 block triangular matrix, with one diagonal block of order k which is equal to A_C, the adjacency matrix of C, and where the other diagonal block is of order $n - k$ and is triangular with all diagonal entries zero. It now follows that $(I - T)^{\#}$ exists, and from the fact (see [16]) that $(I - A_C)^{\#}$ has all diagonal entries equal to $\frac{k-1}{2k}$, we find that that $\operatorname{trace}(I - T)^{\#} = \frac{k-1}{2} + n - k = \frac{2n-k-1}{2}$. Hence, for every $\epsilon > 0$ there is an irreducible $B \in \Sigma_D$ such that $\left|\operatorname{trace}(I - B)^{\#} - \frac{2n-k-1}{2}\right| < \epsilon$; it now follows that $\gamma = \frac{2n-k+1}{2}$. Observe that if D contains a cycle of length $m \neq k$, an analogous argument yields $\gamma = \frac{2n-m+1}{2}$, a contradiction. Thus, all cycles in D have length k.

(b) Suppose that D contains two cycles C_1 and C_2 that do not share any vertices. Let $T \in \Sigma_D$ be a $(0,1)$ matrix such that the arcs of C_1 and C_2 correspond to nonzero entries in T, while the vertices of the digraph not on C_1 and C_2 are not on any cycle. (Such a T can be constructed in a manner analogous to that described in the proof of a) above.) Then the eigenvalues of T consist of the k-th roots of unity, each of algebraic multiplicity two, and 0 of algebraic multiplicity $n - 2k$. Further, there is an irreducible matrix $B \in \Sigma_D$ whose spectrum is arbitrarily close to that of T. For such a B, the spectrum of $I - B$ consists of the algebraically simple eigenvalue 0, a small positive eigenvalue r close to 0, and $n - 2$ other eigenvalues with positive real parts. It follows that there is an irreducible matrix $B \in \Sigma_D$ such that $\operatorname{trace}(I - B)^{\#} \geq \frac{1}{r} > \frac{2n-k-1}{2}$, contrary to our assumption. Hence any two cycles of D must intersect.

(c) Since every pair of cycles in D share a vertex, we see that if $T \in \Sigma_D$ is reducible, then it has a single essential class of indices, so that 1 is an algebraically simple eigenvalue of T. Thus, $(I - T)^{\#}$ exists for each $T \in \Sigma_D$. Finally, (d) follows from continuity. $\square$

We say that a digraph D has *property $S(k)$* if every vertex of D has positive outdegree, all cycles of D intersect, and all cycles of D have length k. Observe that for such a digraph and any $T \in \Sigma_D$, $(I - T)^{\#}$ exists. Furthermore, if D has property $S(k)$, we see that there is a γ such that $K(T) = \gamma$, for all irreducible $T \in \Sigma_D$ if and only if $K(T) = \gamma$, for all $T \in \Sigma_D$.

Recall that the vertices of a strongly connected digraph D that is periodic

with period k can be partitioned into *cyclically transferring classes*, namely into sets of vertices $S_1, \ldots, S_k$, with the property that the only arcs in D are of the form $u \rightarrow v$, for some $u \in S_i$ and $v \in S_{i+1}$, $i = 1, \ldots, k$.

Theorem 6.2 *Let D be a strongly connected digraph on n vertices with property $S(k)$. If there is a γ such that $K(T) = \gamma$, for all $T \in \Sigma_D$, then D has the property that one of its cyclically transferring classes has cardinality 1 or, equivalently, there is a vertex of D that is on every cycle in D.*

Proof. We proceed by induction on the number of arcs in D, and note that the cases that D has just two arcs is readily verified.

Suppose now that for some $m \geq 2$, the result holds for strongly connected digraphs with at most m arcs, and that D has $m+1$ arcs. Let the sizes of the cyclically transferring classes, say $S_1, \ldots, S_k$, be $m_1, \ldots, m_k$, respectively. Note that it is enough to show that $m_i = 1$, for some i. Suppose that D has vertices u_1, u_2, u_3 such that $u_1, u_2 \in S_p$ and $u_3 \in S_{p+1}$, $u_1, u_2 \rightarrow u_3$, and the outdegree of u_1 is at least two, say with $u_1 \rightarrow u_4$, for some $u_4 \neq u_3$. Note that since all cycles in D have length k, no pair of vertices in S_p are on a common cycle. Since $u_1 \rightarrow u_4$ and D is strongly connected, there is a shortest path from u_4 to u_1 in D. Evidently such a path avoids the arc $u_1 \rightarrow u_3$. Also, since $u_2 \rightarrow u_3$, there is a shortest path from u_3 back to u_2 necessarily that avoids the arc $u_1 \rightarrow u_3$. Since all pairs of cycles in D intersect, the cycle using the arc $u_1 \rightarrow u_4$, say C_1, and the cycle using the arc $u_2 \rightarrow u_3$, say C_2, must intersect in a vertex v. Consider the digraph $\hat{D}$ obtained by D by deleting the arc $u_1 \rightarrow u_3$. Since $\hat{D}$ contains a path from u_1 to v and a path from v to u_3, we see that $\hat{D}$ is strongly connected. Note that D and $\hat{D}$ have the same cyclically transferring classes. To see this observe that there is a path from u_3 to u_1 of length $k-1$ in D and hence also in $\hat{D}$, so that there is a walk from u_1 to u_3 in $\hat{D}$ of length $1 \mod k$. Finally, $\hat{D}$ inherits property $S(k)$ from D, and so from the induction hypothesis, some m_i must be 1, as desired.

It remains to consider the case that for each vertex u of D of outdegree two or more, if $u \rightarrow v$, then the indegree of v is one. In particular, for any vertex w of indegree greater than one, if $x \rightarrow w$, then x has outdegree one.

It follows that if $T \in \Sigma_D$ is irreducible, than T can be written as

$$
T = \begin{bmatrix}
0 & T_1 & 0 & \cdots & 0 \\
0 & 0 & T_2 & \cdots & 0 \\
\vdots & & \ddots & \ddots & \vdots \\
0 & & & \ddots & T_{k-1} \\
T_k & 0 & \cdots & 0 & 0
\end{bmatrix}, \tag{6.30}
$$

where each T_i can be permuted to the form

$$
\begin{bmatrix}
y_1^T \\
& \ddots \\
& & y_{l_i}^T \\
& & & \mathbf{1}_{p_1} \\
& & & & \ddots \\
& & & & & \mathbf{1}_{p_{q_i}}
\end{bmatrix},
$$

for some positive vectors y_j, $j = 1, \ldots, l_i$. Suppose that some $p_j \geq 2$. Without loss of generality, we may assume that for some $r \geq 3$, $t_{1r} = t_{2r} = 1$, for some $T \in \Sigma_D$. Write any $T \in \Sigma_D$ as

$$
T = \begin{bmatrix}
0 & 0 & e_{r-2}^T \\
0 & 0 & e_{r-2}^T \\
x & y & \hat{T}
\end{bmatrix}.
$$

Noting that $e_1^T - e_2^T$ is a left null vector for T, it follows that T is similar to the matrix

$$
\begin{bmatrix}
0 & 0 \\
x - y & B
\end{bmatrix}, \quad \text{where} \quad B = \begin{bmatrix}
0 & e_{r-1}^t \\
x + y & T
\end{bmatrix}.
$$

In particular, letting $\tilde{D}$ be the digraph obtained from D by coalescing vertices 1 and 2 into a single vertex labeled a, say, and we see that as T ranges over Σ_D, B ranges over $\Sigma_{\tilde{D}}$. Note that no cycle in D uses both vertices 1 and 2, so each cycle in D corresponds to a unique cycle in $\tilde{D}$. Furthermore, $\tilde{D}$ is strongly connected, and has property $S(k)$. By the induction hypothesis, there is a vertex w of $\tilde{D}$ that is on every cycle in $\tilde{D}$. If $w \neq a$, it follows that every cycle in D goes through w, as desired. On the other hand, if $w = a$, then in D every cycle goes through vertex 1 or vertex 2. But since vertices 1 and 2 have outdegree one with common out-neighbor r, we can find that

every cycle in D goes through the vertex r.

We may now assume that each $p_j = 1$, so that each T_i has the form

$$\begin{bmatrix} y_1^T & & & \\ & \ddots & & \\ & & y_{l_i}^T & \\ & & & I \end{bmatrix}.$$

It follows that $m_i \leq m_{i+1}$, for each $i = 1, \ldots, k$, with $m_{k+1} = m_1$. Hence, $m_1 \leq m_2 \leq \cdots \leq m_k \leq m_1$, so that each of the vectors y_i is in fact 1×1. Thus we find that each $T \in \Sigma_D$ must be a $(0, 1)-$matrix. Thus, each such T is a permutation matrix, and again the conclusion follows. $\square$

Corollary 6.3 *Let D be a strongly connected digraph. There is a γ such that each irreducible matrix $T \in \Sigma_D$ satisfies $K(T) = \gamma$ if and only if D has property $S(k)$ for some k, and there is a vertex of D that lies on every cycle in D.*

Proof. Suppose that D has n vertices, has property $S(k)$ for some $1 \leq k \leq n$, and that there is a vertex that lies on every cycle. If T is an irreducible matrix in Σ_D, then we may write T in cyclic form (6.30). Note that the cyclic products $T_i T_{i+1} \ldots T_k T_1 \ldots T_{i-1}, i = 1, \ldots, k$, share the same non-zero eigenvalues; further, since there is a vertex that lies on every cycle of D, one of the cyclic products above must be the 1×1 matrix [1]. It now follows that the eigenvalues of T consist of the k-th roots of unity (each with multiplicity 1), and 0 with multiplicity $n - k$. As all such T share the same eigenvalues, we find from (2.7) that $K(T)$ is constant as T ranges over the irreducible matrices in Σ_D.

The converse follows from Lemma 6.1 and Theorem 6.3. $\square$

References

[1] Ben–Israel, A., and Greville, T.N.E. (2003). Generalized Inverses: Theory and Applications, 2nd ed., Springer, New York.

[2] Berman, A., and Plemmons, R.J. (1994). Nonnegative Matrices in the Mathematical Sciences, SIAM Publications, Philadelphia.

[3] Brauer, A. (1952). Limits for the characteristic roots of a matrix. IV: Applications to stochastic matrices. Duke Math. J. 19, 75–91.

[4] Campbell, S.L., and Meyer, C.D. (1991). Generalized Inverses of Linear Transformations, Dover Publications, New York.

[5] Serra–Capizzano, S. (2005). Jordan canonical form of the Google matrix: a potential contribution to the Pagerank computation. SIAM J. Matrix Anal. and Appl. 27, 305–312.

[6] Cho, G.E., and Meyer, C.D. (2000). Markov chain sensitivity measured by mean first passage times. Lin. Alg. Appl. 316, 21–28.

[7] Cho, G.E., and Meyer, C.D. (2001). Comparison of perturbation bounds for the stationary distribution of a Markov chain. Lin. Alg. Appl. 335, 137–150.

[8] Dietzenbacher, E. (1990). Perturbations of the Perron vector: Applications to finite Markov chains and demographic population models. Environment and Planning 22, 747–761.

[9] Fiedler, M., and Pták, V. (1967). Diagonally dominant matrices. Czech. Math. J. 17, 420–433.

[10] Grinstead, C.M., and Snell, J.L. (1997). Introduction to Probability, American Mathematical Society, Providence.

[11] Higham, D. (2003). A matrix perturbation view of the small world phenomenon. SIAM J. Matrix Anal. Appl. 25, 429–444.

[12] Horn, R.A., and Johnson, C.R. (1985). Matrix Analysis, Cambridge University Press, New York.

[13] Hunter, J.J. (2005). Stationary distribution and mean first passage times of perturbed Markov chains. Lin. Alg. Appl. 410, 217–243.

[14] Hunter, J.J. (2006). Mixing Times with applications to Perturbed Markov chains. Lin. Alg. Appl. 417, 108–123.

[15] Kemeny, J.G., and Snell, J.L. (1960). Finite Markov Chains, Springer, New York.

[16] Kirkland, S.J. (1995). The group inverse associated with an irreducible periodic nonnegative matrix. SIAM J. Matrix Anal. Appl. 16, 1127–1134.

[17] Kirkland, S.J., and Neumann, M. (1999). On group inverses of M–matrices with uniform diagonal entries. Lin. Alg. Appl. 296, 153–170.

[18] Kirkland, S.J. and Neumann, M. (2009). The case of equality in the Dobrushin–Deutsch–Zenger bound. Lin. Alg. Appl. 431 , 2373-2394.

[19] Kirkland, S.J., Neumann, M., and Sze, N. (2008). On optimal condition numbers for Markov chains. Numer. Math 110, 521–537.

[20] Kirkland, S.J., Neumann, M., and Xu, J. (2007). Transition matrices for well-conditioned Markov chains. Lin. Alg. Appl. 424, 118–131.

[21] Langville, A.N., and Meyer, C.D. (2004). Deeper inside PageRank. Internet Mathematics 1, 335–400.

[22] Levene, M., and Loizou, G. (2002). Kemeny's constant and the random surfer. Amer. Math. Monthly 109, 741–745.

[23] Metzler, L. (1951). A multiple–county theory of income transfer. J. Political Econ. 59, 14–29.

[24] Meyer, C.D. (1975). The role of the group generalized inverse in the theory of finite Markov chains, SIAM Rev. 17, 443–464.

[25] Seneta, E. (1981). Non–Negative Matrices and Markov Chains, Springer, New York.

[26] Seneta, E. (1988). Perturbation of the stationary distribution measured by ergodicity coefficients. Adv. Appl. Prob. 20, 228–230.

[27] Seneta, E. (1991). Sensitivity analysis, ergodicity coefficients, and rank-one updates for finite Markov chains. In Stewart, W.J. (ed.), Numerical Solution of Markov Chains, Marcel Dekker, New York.