A STRUCTURED CONDITION NUMBER FOR KEMENY'S CONSTANT*

JANE BREEN† AND STEVE KIRKLAND‡

Abstract. Kemeny's constant is an interesting and useful quantifier describing the global average behaviour of a Markov chain. In this article, we examine the sensitivity of Kemeny's constant to perturbations in the transition probabilities. That is, we consider the problem of generating a condition number for Kemeny's constant, to give an indication of the size of the change in its value relative to the size of the perturbation. We provide a structured condition number and determine some illuminating upper and lower bounds which connect the conditioning of Kemeny's constant to well-studied condition numbers for the stationary vector of the Markov chain. We also investigate the behaviour of this structured condition number for several infinite families of Markov chains.

Key words. Kemeny's constant, condition number, Markov chains, group inverse

AMS subject classifications. 60J22, 15B51, 15A09, 60J10

1. Introduction.

A Markov chain is a mathematical model which may be used to describe a dynamical system which transitions between a finite number of distinct states, in discrete time-steps. The movement from one state to another is dictated by a prescribed transition probability; in particular, given a finite set of states $\{s_1, \ldots, s_n\}$, the probability of the system moving from s_i to s_j in a single time-step is given by some $t_{i,j} \in [0,1]$. In this way, the evolution of the Markov chain is a stochastic process which is memoryless, in that the behaviour of the system in the next time-step depends only on the current state of the system. More formally, a Markov chain is thought of as a sequence of random variables $\{X_k \mid k = 0, 1, 2, \ldots\}$, where each X_k takes on a value from $\{s_1, \ldots, s_n\}$ with some probability, and this stochastic process satisfies

$$\mathbb{P}[X_{k+1} = x_{k+1} \mid X_k = x_k, \ldots, X_0 = x_0] = \mathbb{P}[X_{k+1} = x_{k+1} \mid X_k = x_k].$$

The transition probability $\mathbb{P}[X_{k+1} = s_j \mid X_k = s_i]$ is written as $t_{i,j}$.

The Markov chain is represented completely by its *probability transition matrix* $T = [t_{i,j}]$, which is a nonnegative, row-stochastic matrix—i.e. all rows sum to one, which we write as $T\mathbb{1} = \mathbb{1}$, where $\mathbb{1}$ represents the vector of all ones. Given an initial probability distribution vector u_0, the probability distribution after k time-steps is given by $u_k^\top = u_0^\top T^k$. Furthermore, if T is irreducible, the long-term probability distribution of the Markov chain is given by the left eigenvector w of T corresponding to the eigenvalue 1, normalised so that $w^\top \mathbb{1} = 1$. This is referred to as the *stationary distribution vector* of the chain, having the property that $w^\top T = w^\top$. The short-term behaviour of the Markov chain is described by the *mean first passage times* $m_{i,j}$, for $i, j = 1, \ldots, n$. The mean first passage time from s_i to s_j, $m_{i,j}$, gives the expected number of time-steps before the system reaches s_j, given that it starts in s_i; that is,

$$m_{i,j} = \mathbb{E}[\, k \mid X_k = s_j, X_0 = s_i],$$

where $\mathbb{E}[\cdot]$ denotes the expected value. Note that $m_{i,i}$ denotes the *mean first return time* of state i, and is given by $\frac{1}{w_i}$.

*The results of this article also appear in the Ph.D. thesis of the first author (see [3]).

†Department of Mathematics, Iowa State University, Ames, IA 50011, USA (breenj@iastate.edu.)

‡Department of Mathematics, University of Manitoba, Winnipeg, MB R3T 2N2 (stephen.kirkland@umanitoba.ca).

An extremely interesting quantifier of the behaviour of an irreducible Markov chain is *Kemeny's constant*. This was first defined in [18], and can be written as

$$K_i(T) := \sum_{j \neq i} w_j m_{i,j},$$

which may be interpreted in terms of the expected number of time-steps required to reach a randomly-chosen destination state from a fixed starting state s_i. Remarkably, this was shown to be independent of the index i, and so is named Kemeny's constant, and denoted by $\mathcal{K}(T)$. It is easily seen that Kemeny's constant may also be written

$$\mathcal{K}(T) = \sum_i \sum_{j \neq i} w_i m_{i,j} w_j,$$

admitting the interpretation of $\mathcal{K}(T)$ in terms of the expected length of a random trip between states in the chain, where the initial and terminal states are chosen randomly according to the stationary distribution of the Markov chain. Levene and Loizou showed in [27] that Kemeny's constant can also be expressed in terms of the eigenvalues $1, \lambda_2, \ldots, \lambda_n$ of T, with

$$\mathcal{K}(T) = \sum_{j=2}^{n} \frac{1}{1-\lambda_j}. \tag{1.1}$$

Remark 1.1. There is some inconsistency in the literature regarding the definition of Kemeny's constant as

$$\mathcal{K}(T) = \sum_{\substack{j=1 \\ j \neq i}}^{n} w_j m_{i,j},$$

for any index i. In particular, the quantity $\sum_{j=1}^{n} w_j m_{i,j}$ is often considered, which is equal to $\mathcal{K}(T)+1$. There is no great difference in the analysis of these two quantities, but it is worth noting. The issue is further confounded by the fact that some consider first *hitting times* as opposed to first passage times, which are equivalent save for the convention that $m_{i,i} = 0$ when considering hitting times. When expressing Kemeny's constant in terms of hitting times, $\mathcal{K}(T) = \sum_j w_j m_{i,j}$ and $\mathcal{K}(T) = \sum_{j \neq i} w_j m_{i,j}$ are interchangeable, although we note that in [2], the former is used since equality still holds in the expression for continuous-time Markov chains.

There are many applications of irreducible Markov chains to real dynamical systems, including urban road network dynamics (see [9]), molecular conformation dynamics (see [10]), and the spread of infectious disease (see [1]). In each of these, Kemeny's constant is a valuable measure: in urban road networks, the value of $\mathcal{K}(T)$ provides insight into how well-connected the urban area is; in molecular conformational dynamics, the value of $\mathcal{K}(T)$ could indicate the presence or absence of metastable sets (which is extremely useful in computational drug design); in an infectious disease setting, $\mathcal{K}(T)$ provides a measure of how quickly epidemic levels are approached.

Given the utility of Kemeny's constant in practical applications, it is worthwhile to consider how sensitive it is to perturbations in the transition probabilities. For those modeling with Markov chains using real data, the transition probabilities derived from these data are usually only sample estimates, and not true values. The transition matrix of this model can then be viewed as a perturbation of the 'true' transition matrix, and answering the question of how sensitive the calculation of $\mathcal{K}(T)$ is to errors in the data gives a measure of confidence in the computed value for Kemeny's constant. Further, numerical techniques for computing Kemeny's constant are subject to round-off errors, and hence an understanding of the conditioning of Kemeny's constant is needed in that setting as well.

In this article, we develop a structured condition number for Kemeny's constant in Section 3. In Section 4, we determine bounds which provide insight into the transition matrices for which Kemeny's constant is poorly-conditioned, and also connect the conditioning of Kemeny's constant to other notions of conditioning in Markov chain theory. We also explore the value of this structured condition number for some infinite families of matrices in Section 5. The results of this article also appear in part in [3].

Central to the results of this article is a certain type of *generalized matrix inverse* of a singular matrix known as the *group inverse*. The singular matrix of concern in Markov chain theory is $I-T$, which has 0 as a simple eigenvalue when T is irreducible. From the eigenvalue expression (1.1) for $\mathcal{K}(T)$, and from spectral properties of the group inverse, it may be shown that $\mathcal{K}(T)$ is equal to the trace of the group inverse of $I - T$. We give a short introduction to the group inverse of a singular matrix here before proceeding with the subject of the article.

1.1. The group generalized inverse of a singular matrix.

DEFINITION 1.2. *Let A be a complex singular matrix for which the algebraic and geometric multiplicities of the eigenvalue 0 of A are equal (that is, 0 is a semisimple eigenvalue). Then the* group inverse *of A, denoted $A^\#$, is the unique matrix satisfying*

$$AA^\#A = A; \qquad A^\#AA^\# = A^\#; \qquad AA^\# = A^\#A.$$

To find the group inverse of a singular matrix A for which the eigenvalue 0 is semisimple, one may consider the Jordan form of A. That is, there exists an invertible matrix P such that

$$A = P \left[\begin{array}{c|c} B & O \\ \hline O & O \end{array} \right] P^{-1},$$

such that B is invertible. Then the matrix

$$X = P \left[\begin{array}{c|c} B^{-1} & O \\ \hline O & O \end{array} \right] P^{-1}$$

can readily be seen to satisfy the three equations of the above definition. To prove uniqueness, one must consider the range and null space of X and of A; see [26, Section 2.1] for more details. We remark that the group inverse is a special case of the *Drazin inverse* of a singular matrix.

We give some spectral properties of $A^\#$ which will be useful in the remainder of the paper. For detailed discussion on the group inverse, the interested reader may refer to [26].

LEMMA 1.3. *Let A be a singular complex $n \times n$ matrix with 0 as a semisimple eigenvalue, and let $A^\#$ be the group inverse of A. Then $A^\#$ has the following spectral properties:*

(a) $A^\#$ has 0 as a semisimple eigenvalue, and its multiplicity is equal to the multiplicity of 0 as an eigenvalue of A.

(b) For a vector v, $Av = 0$ if and only if $A^\#v = 0$. Similarly, $v^\top A = 0$ if and only if $v^\top A^\# = 0$.

(c) $\lambda \neq 0$ is an eigenvalue of A of multiplicity m if and only if $\frac{1}{\lambda}$ is an eigenvalue of $A^\#$ of multiplicity m.

(d) $Av = \lambda v$ *if and only if* $A^{\#}v = \frac{1}{\lambda}v$*. Similarly,* $v^\top A = \lambda v^\top$ *if and only if* $v^\top A^{\#} = \frac{1}{\lambda}v^\top$*.*

(e) The matrix $I - AA^{\#}$ *is the eigenprojection of* A *onto the eigenspace of* A *corresponding to the eigenvalue* 0*. In particular, if* 0 *is a simple eigenvalue of* A *(having multiplicity* 1*), with right and left null vectors* v *and* $u^\top$ *respectively, normalised so that* $u^\top v = 1$*, then* $AA^{\#} = A^{\#}A = I - vu^\top$*.*

From the statement of this lemma, we know the following about the group inverse of $Q = I - T$ where T is the transition matrix of an irreducible Markov chain, with stationary vector w:

(a) $Q^{\#}$ has 0 as an eigenvalue of multiplicity 1.

(b) $Q^{\#}\mathbb{1} = 0$, and $w^\top Q^{\#} = 0$.

(c) If $1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of T, the eigenvalues of $Q^{\#}$ are given by $0, \frac{1}{1-\lambda_2}, \ldots, \frac{1}{1-\lambda_n}$. Hence $\mathcal{K}(T) = \text{trace}(Q^{\#})$.

(e) $QQ^{\#} = Q^{\#}Q = I - \mathbb{1}w^\top$.

Finally, we make a short remark about the computation of the group inverse of $I - T$ where T is the $n \times n$ transition matrix of an irreducible Markov chain. One method involves the inversion of any $(n-1) \times (n-1)$ principal submatrix of $I - T$, which is discussed in [26, Prop. 2.5.1], with a cost of approximately $2n^3$ floating point operations (or *flops*). Another method involves the QR factorisation of the matrix $I - T$, which may be accomplished with approximately $\frac{11}{3}n^3$ flops, and we remark that this method is known to be more computationally stable. The interested reader may find further discussion in [26, Chapter 8].

2. Conditioning problems in Markov chain theory. Suppose T is an irreducible stochastic matrix representing a Markov chain, with stationary vector w. Then suppose that T is perturbed to form some new irreducible stochastic matrix $\tilde{T}$, with stationary vector $\tilde{w}$. How different can w and $\tilde{w}$ be, relative to the magnitude of the perturbation? An answer to this question determines how sensitive the long-term behaviour of a system modelled by a Markov chain can be to small changes in the transition probabilities.

The above problem is referred to as *conditioning* of the stationary vector, and is formalised as follows: Given T, an irreducible stochastic matrix with stationary vector w, we wish to determine some function $f(T)$, such that if $\tilde{T} = T + E$ is also irreducible, nonnegative and stochastic with stationary vector $\tilde{w}$, then for some appropriate p, q,

$$\|w - \tilde{w}\|_p \leq \|E\|_q \cdot f(T). \tag{2.1}$$

This function $f(T)$ is referred to as a *condition number*. The norms we will most frequently discuss are the ∞-norm and the 1-norm. We recall that for any real $m \times n$ matrix A,

$$\|A\|_\infty = \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{i,j}|,$$

and

$$\|A\|_1 = \max_{j=1,\ldots,n} \sum_{i=1}^{m} |a_{i,j}|,$$

so that $\|A^\top\|_1 = \|A\|_\infty$. The matrix norm $\|\cdot\|_\infty$ is sometimes referred to as the *absolute row sum norm*, and $\|\cdot\|_1$ as the *absolute column sum norm*. For more on vector and matrix norms, see [13, Chapter 5].

Solutions to these conditioning problems usually rely upon some generalized inverse of the singular matrix $I - T$; one example of this is the 'fundamental matrix', defined in [18] as $Z = (I - T + \mathbb{1}w^\top)^{-1}$. However, in the landmark article [30] by Meyer, the group inverse is introduced as the generalized inverse of choice to be used in Markov chain theory, with the author stating that "If T is the one-step transition matrix of a finite homogeneous Markov chain and if $A = I - T$, it will be shown that once the group inverse, $A^\#$, of A is known, then the answer to every important question concerning the chain can be obtained from $A^\#$."

The group inverse of $Q = I - T$ is valuable in answering questions regarding the conditioning of the stationary vector due to the following argument: Given T, w, $\tilde{T}$ and $\tilde{w}$ as above (so that $\tilde{T} = T + E$), from the eigenequation $\tilde{w}^\top \tilde{T} = \tilde{w}^\top$, it follows that $\tilde{w}^\top (T + E) = \tilde{w}^\top$, and so $\tilde{w}^\top E = \tilde{w}^\top (I - T)$. Multiplying on the right by $Q^\# = (I - T)^\#$, we have $\tilde{w}^\top E Q^\# = \tilde{w}^\top Q Q^\#$. Since $I - QQ^\#$ is the eigenprojection of Q onto the eigenspace corresponding to the eigenvalue 0, $QQ^\# = I - \mathbb{1}w^\top$ (from Lemma 1.3 (e).) Hence

$$\tilde{w}^\top E Q^\# = \tilde{w}^\top (I - \mathbb{1}w^\top) = \tilde{w}^\top - w^\top.$$

It is from this relationship that many condition numbers of the type in (2.1) are derived; hence this $f(T)$ is frequently some function of the entries of the group inverse $(I - T)^\#$.

We note that Meyer's assertion in [30] that the group inverse is the "correct" choice of generalized inverse in Markov chain theory is not universally accepted; in [14], Hunter presents a more comprehensive study of the uses of generalized inverses in Markov chain theory, giving expressions for the stationary vector, mean first passage times and their moments in terms of multiple classes of generalized inverses. As is elaborated in, for example [15, 16], Hunter shows how many Markov chain theory results can be expressed in more general terms via any choice of a generalized inverse. A discussion of the usefulness of choosing other generalized inverses is given in [14, Section 7]. We exclusively consider the group inverse in this article, but remark that when T is irreducible with stationary vector w, and if G is any generalized inverse of $I - T$, then $Q^\# = (I - \mathbb{1}w^\top) G (I - \mathbb{1}w^\top)$ (as shown in [14, Thm 6.3]), and conceivably the results of the present paper could be generalized using this observation to results concerning other choices of generalized inverses.

Originally, Schweitzer approached conditioning theory regarding the stationary vector of a Markov chain using the fundamental matrix of the chain, $Z = (I - T + \mathbb{1}w^\top)^{-1}$, and showed in [31] that

$$\|\tilde{w} - w\|_1 \leq \|Z\|_\infty \|E\|_\infty.$$

This was followed by Meyer in [28], who instead used the group inverse $Q^\#$ and showed that

$$\|\tilde{w} - w\|_1 \leq \|Q^\#\|_\infty \|E\|_\infty.$$

Succeeding these are a long list of improvements to and variations of these condition numbers, along with new approaches to analysing the sensitivity of the stationary vector by determining bounds on the condition numbers in terms of the eigenvalues of the matrix (see [29]) and the mean first passage times (see [8]), as well as determining the sensitivity of a single entry of the stationary distribution vector (see [23]). A survey is given in [7], and we also refer the reader to further work since then in [22, 23].

We now give two examples of condition numbers of the stationary vector which will also be used throughout this article.

Let T be an irreducible stochastic matrix, and let $Q = I - T$. Define

$$\kappa_3(T) := \frac{1}{2} \max_{1 \le i,j \le n} (q^{\#}_{j,j} - q^{\#}_{i,j}), \tag{2.2}$$

and

$$\kappa_6(T) := \frac{1}{2} \max_{1 \le i,j \le n} \sum_{k=1}^{n} |q^{\#}_{i,k} - q^{\#}_{j,k}|. \tag{2.3}$$

Both $\kappa_3(T)$ and $\kappa_6(T)$ act as condition numbers, proven by Haviv and Van der Heyden (see [11]) and by Seneta (see [34]), respectively. In particular:

THEOREM 2.1 ([11, 34]). *Let T be an irreducible stochastic matrix with stationary vector w, and let $\tilde{T} = T + E$ also be an irreducible stochastic matrix for some matrix E, with stationary vector $\tilde{w}$. Then:*
(a)

$$\|\tilde{w} - w\|_\infty \le \|E\|_\infty \kappa_3(T).$$

(b)

$$\|\tilde{w} - w\|_1 \le \|E\|_\infty \kappa_6(T).$$

As we will show in Theorem 4.2 and Proposition 4.4 below, the condition numbers $\kappa_3(T)$ and $\kappa_6(T)$ are closely connected with the conditioning properties of Kemeny's constant. Note that $\kappa_6(T)$ is also a special case of the *coefficient of ergodicity* of the group inverse $Q^{\#}$ (see [33, 34]), and is sometimes denoted as $\tau(Q^{\#})$. We note that the numbering of these condition numbers originated in the survey paper [7] by Cho and Meyer, in which the most prominent condition numbers in the literature at the time were listed and compared.

The condition numbers $\kappa_3(T)$ and $\kappa_6(T)$ are well-known as the "most optimal" condition numbers in Markov chain theory. In particular, a study of condition numbers for the stationary vector is given in [20], in which the authors show that if $f(T)$ is any condition number with respect to the (p, ∞) norm pair—that is, satisfying $\|w - \tilde{w}\|_p \le \|E\|_\infty f(T)$—then $\tau_p(Q^{\#}) \le f(T)$, where

$$\tau_p(Q^{\#}) := \sup_{\substack{y^\top \mathbb{1} = 0 \\ y \ne 0}} \frac{\|y^\top Q^{\#}\|_p}{\|y^\top\|_1}.$$

Furthermore, $\tau_p(Q^{\#})$ is shown to be a condition number for the stationary vector of a Markov chain, and so together this gives that it is the minimum of all condition numbers with respect to the (p, ∞) norm pair. Since $\tau_1(Q^{\#}) = \kappa_6(T)$ and $\tau_\infty(Q^{\#}) = \kappa_3(T)$, and since the 1- and ∞-norms are the most commonly-used vector norms in considering conditioning problems of the stationary vector, it is arguable that these two condition numbers are the most useful ones to consider. Indeed, much of the literature focuses on these two in particular (see for example [22, 19], and an overview in [26, Section 5.3]).

Remark 2.2. Note that in [34], it is shown that for an irreducible stochastic matrix T of order n, with eigenvalues $1, \lambda_2, \ldots, \lambda_n$,

$$\kappa_6(T) \leq \sum_{j=2}^{n} \frac{1}{1-\lambda_j},$$

i.e. $\kappa_6(T) \leq \mathcal{K}(T)$. Therefore $\|\tilde{w} - w\|_1 \leq \|E\|_\infty \mathcal{K}(T)$, so that Kemeny's constant is itself a condition number for the stationary distribution of the chain. Note that this was shown independently in [15].

The body of work on perturbation analysis and condition numbers for stationary distribution vectors has grown and developed since the 1960s, and the field is well-established. In this article, we begin the development of a body of work on perturbation analysis and condition numbers for Kemeny's constant. That is, we wish to tackle the question of how sensitive Kemeny's constant is to perturbations or errors in the transition probabilities of the Markov chain, for a given transition matrix T. More formally, given an irreducible stochastic matrix T and perturbing matrix E (such that $T+E$ is also stochastic and irreducible), can we determine an upper bound for $|\mathcal{K}(T+E) - \mathcal{K}(T)|$ in terms of $\|E\|$ (for some choice of norm $\|\cdot\|$) and some function of T? In the following sections we will prove some results to this end, and determine a *structured condition number* for $\mathcal{K}(T)$—that is, a condition number under the restriction that the size of the perturbation is small.

Definition 2.3. *Let T be an irreducible stochastic matrix. The* structured condition number for Kemeny's constant *is defined as*

$$\mathcal{C}(T) := \limsup_{\varepsilon \to 0^+} \left\{ \frac{|\mathcal{K}(T+E) - \mathcal{K}(T)|}{\varepsilon} \,\middle|\, T+E \text{ is irreducible, stochastic; } \|E\|_\infty \leq \varepsilon \right\}.$$

The structured condition number for $\mathcal{K}(T)$ may be interpreted as a measure of the maximum change in $\mathcal{K}(T)$ when T undergoes some perturbation, where it is assumed that the norm of the perturbing matrix is vanishingly small. This lends itself more to the application of considering numerical errors in a computational setting, with $\mathcal{C}(T)$ interpreted in terms of how robust the calculation of $\mathcal{K}(T)$ is. A similar concept has been examined in the context of generalized eigenvalue problems in [12].

It is important to note that while the natural inclination is to attempt to derive an expression of the form

$$|\mathcal{K}(T+E) - \mathcal{K}(T)| \leq \|E\|_\infty c(T),$$

where $c(T)$ depends only on T, the following key example shows that there is no possibility of such a general expression.

Remark 2.4. Let T be any irreducible stochastic matrix, and consider a perturbation which results in a convex combination of T and the identity matrix; that is, for some $a \in (0,1]$, let $E_a = (1-a)(I-T)$, so that

$$T + E_a = aT + (1-a)I.$$

Then $\|E_a\|_\infty \leq 2(1-a)$ by the triangle inequality. However, it is clear that as $a \to 0$, $\mathcal{K}(T+E_a) \to \infty$, since the group inverse of $I - (T+E_a)$ is $\frac{1}{a}(I-T)^{\#}$; that is, $\mathcal{K}(T+E_a) = \frac{1}{a}\mathcal{K}(T)$. In particular, since $\|E_a\|_\infty$ is bounded above by 2 there is no general expression of the type

$$|\mathcal{K}(T+E) - \mathcal{K}(T)| \leq \|E\|_\infty c(T)$$

that holds for all admissible perturbing matrices E. That observation further motivates our interest in analysing the situation where the norm of E is small, as anticipated by Definition 2.3.

An analysis of the behaviour of Kemeny's constant of a Markov chain under perturbation has also been considered in [5]. The authors consider two specific types of perturbations: the first, when $E = e_i u^\top$ and only one row of T is changed; the second, when $E = \mathbb{1}u^\top$, so that every row is perturbed in the same way. In both cases the vector u is chosen appropriately so that nonnegativity, irreducibility, and stochasticity of $T + E$ is preserved.

3. A structured condition number for Kemeny's constant.

Throughout this section, T is considered to be a nonnegative stochastic matrix of order n with 1 as an algebraically simple eigenvalue, and w denotes the stationary vector of T. Let E denote some perturbation matrix of T; that is, E is an $n \times n$ matrix whose rows sum to zero, such that $\tilde{T} = T + E$ is also nonnegative and stochastic, with 1 as an algebraically simple eigenvalue. Let $Q = I - T$, and $\tilde{Q} = I - \tilde{T}$.

In [28], the following is proven to give an expression for $\tilde{Q}^\#$ in terms of $Q^\#$ and E.

THEOREM 3.1 ([28]). *Let T, E, $\tilde{T}$, w, Q, and $\tilde{Q}$ be defined as above. Then $I - EQ^\#$ is invertible, and*

$$\tilde{Q}^\# = Q^\#(I - EQ^\#)^{-1} - \mathbb{1}w^\top(I - EQ^\#)^{-1}Q^\#(I - EQ^\#)^{-1}.$$

We now use this perturbation formula to derive an expression for $\mathcal{K}(T + E)$.

LEMMA 3.2. *Let T, E, $\tilde{T}$, Q, $\tilde{Q}$ be defined as above. If the spectral radius $\rho(EQ^\#) < 1$, then*

$$\mathcal{K}(\tilde{T}) = \mathcal{K}(T) + \sum_{j=1}^{\infty} \operatorname{trace}\left(Q^\#(EQ^\#)^j\right).$$

Proof. Recall that $\mathcal{K}(T) = \operatorname{trace}(I - T)^\#$. Hence from Theorem 3.1,

$$\begin{aligned}\mathcal{K}(T + E) &= \operatorname{trace}(\tilde{Q}^\#)\\ &= \operatorname{trace}(Q^\#(I - EQ^\#)^{-1}) - \operatorname{trace}(\mathbb{1}w^\top(I - EQ^\#)^{-1}Q^\#(I - EQ^\#)^{-1}).\end{aligned}$$

Then since the trace of any rank-one matrix $uv^\top$ is $v^\top u$, the trace of the second term above is

$$\operatorname{trace}(\mathbb{1}w^\top(I - EQ^\#)^{-1}Q^\#(I - EQ^\#)^{-1}) = w^\top(I - EQ^\#)^{-1}Q^\#(I - EQ^\#)^{-1}\mathbb{1}.$$

Now, $I - EQ^\#$ is invertible, and if $\rho(EQ^\#) < 1$, then

$$(I - EQ^\#)^{-1} = I + EQ^\# + (EQ^\#)^2 + (EQ^\#)^3 + \cdots$$

and hence, since $Q^\#\mathbb{1} = 0$, $(I - EQ^\#)^{-1}\mathbb{1} = \mathbb{1}$. Moreover,

$$\operatorname{trace}(\mathbb{1}w^\top(I - EQ^\#)^{-1}Q^\#(I - EQ^\#)^{-1}) = 0,$$

again because $Q^{\#}\mathbb{1} = 0$. So

$$\begin{aligned}
\mathcal{K}(\tilde{T}) &= \operatorname{trace}(\tilde{Q}^{\#}) \\
&= \operatorname{trace}(Q^{\#}(I - EQ^{\#})^{-1}) \\
&= \operatorname{trace}(Q^{\#} + Q^{\#}EQ^{\#} + Q^{\#}(EQ^{\#})^2 + \cdots) \\
&= \operatorname{trace}(Q^{\#}) + \sum_{j=1}^{\infty} \operatorname{trace}(Q^{\#}(EQ^{\#})^j) \\
&= \mathcal{K}(T) + \sum_{j=1}^{\infty} \operatorname{trace}(Q^{\#}(EQ^{\#})^j).
\end{aligned}$$

□

Remark 3.3. The result above requires $\rho(EQ^{\#}) < 1$. However, it is well-known (see [13, Theorem 5.6.9]) that for any matrix norm $\|\cdot\|$ and $n \times n$ matrix A,

$$\rho(A) \leq \|A\|.$$

That is, given any matrix norm $\|\cdot\|$, it is a sufficient condition for Lemma 3.2 that $\|EQ^{\#}\| < 1$. Consequently, by the submultiplicativity of matrix norms, it is hence sufficient that $\|E\| < \frac{1}{\|Q^{\#}\|}$. Thus for any matrix norm, if $\|E\|$ is sufficiently small, the expression for $\mathcal{K}(\tilde{T})$ given in the result above will hold. We will typically use the absolute row sum norm $\|\cdot\|_\infty$.

THEOREM 3.4. *Let T be an irreducible stochastic $n \times n$ matrix; let $Q = I - T$; and let $q^{\#(2)}_{i,j}$ denote the (i,j) entry of $(Q^{\#})^2$. Then*

$$\mathcal{C}(T) \;=\; \frac{1}{2}\sum_{j=1}^{n} \max\left\{\max_i\{q^{\#(2)}_{i,j}\} - \alpha(j),\ \beta(j) - \min_i\{q^{\#(2)}_{i,j}\}\right\}, \tag{3.1}$$

where

$$\alpha(j) := \min_i\{q^{\#(2)}_{i,j} | t_{j,i} > 0\};$$

$$\beta(j) := \max_i\{q^{\#(2)}_{i,j} | t_{j,i} > 0\}.$$

Proof. Let $\varepsilon > 0$ be given. Let T be an irreducible stochastic matrix of order n, and let E be a matrix of order n with zero row sums, such that $\tilde{T} = T + E$ is irreducible, nonnegative and stochastic, and $\|E\|_\infty \leq \varepsilon$. From Lemma 3.2, we have

$$\mathcal{K}(\tilde{T}) - \mathcal{K}(T) = \sum_{j=1}^{\infty} \operatorname{trace}(Q^{\#}(EQ^{\#})^j).$$

We first concentrate our attention on $\operatorname{trace}(Q^{\#}EQ^{\#}) = \operatorname{trace}(E(Q^{\#})^2)$. Representing the rows of E by $u_i^\top$, $i = 1, \ldots, n$, and letting e_j denote the j^{th} standard basis vector, we can write

$$\operatorname{trace}(E(Q^{\#})^2) = \sum_{j=1}^{n} u_j^\top (Q^{\#})^2 e_j.$$

For every j, $u_j^\top = e_j^\top E$ can be written as $x^\top - y^\top$, where x and y are nonnegative vectors, and $x^\top\mathbb{1} = y^\top\mathbb{1} \leq \frac{\varepsilon}{2}$. Note that if $y_i > 0$, then $t_{j,i} > 0$, since $T + E$ is nonnegative.

Fixing j, we have

$$\begin{aligned} u_j^\top (Q^\#)^2 e_j &= x^\top (Q^\#)^2 e_j - y^\top (Q^\#)^2 e_j \\ &= \sum_{l=1}^n x_l q_{l,j}^{\#(2)} - \sum_{l=1}^n y_l q_{l,j}^{\#(2)} \\ &\leq \sum_{l=1}^n x_l \cdot \max_l \{q_{l,j}^{\#(2)}\} - \sum_{l=1}^n y_l \cdot \alpha(j) \end{aligned}$$

where $\alpha(j) = \min_l\{q_{l,j}^{\#(2)} \mid t_{j,l} > 0\}$. Therefore

$$u_j^\top (Q^\#)^2 e_j \leq \frac{\varepsilon}{2} \left(\max_l \{q_{l,j}^{\#(2)}\} - \alpha(j) \right). \tag{3.2}$$

Also consider that

$$\begin{aligned} u_j^\top (Q^\#)^2 e_j &= \sum_{l=1}^n x_l q_{l,j}^{\#(2)} - \sum_{l=1}^n y_l q_{l,j}^{\#(2)} \\ &\geq \sum_{l=1}^n x_l \cdot \min_l \{q_{l,j}^{\#(2)}\} - \sum_{l=1}^n y_l \cdot \beta(j) \end{aligned}$$

where $\beta(j) = \max_l\{q_{l,j}^{\#(2)} \mid t_{j,l} > 0\}$. Therefore

$$u_j^\top (Q^\#)^2 e_j \geq \frac{\varepsilon}{2} \left(\min_l \{q_{l,j}^{\#(2)}\} - \beta(j) \right). \tag{3.3}$$

Hence from (3.2) and (3.3),

$$|u_j^\top (Q^\#)^2 e_j| \leq \frac{\varepsilon}{2} \cdot \max \left\{ \max_l \{q_{l,j}^{\#(2)}\} - \alpha(j), \beta(j) - \min_l \{q_{l,j}^{\#(2)}\} \right\},$$

and so

$$\begin{aligned} |\operatorname{trace}(Q^\# E Q^\#)| &\leq \sum_{j=1}^n |u_j^\top (Q^\#)^2 e_j| \\ &\leq \frac{\varepsilon}{2} \sum_{j=1}^n \max \left\{ \max_i \{q_{i,j}^{\#(2)}\} - \alpha(j),\ \beta(j) - \min_i \{q_{i,j}^{\#(2)}\} \right\}. \end{aligned} \tag{3.4}$$

Finally, we conclude

$$\begin{aligned} \frac{|\mathcal{K}(\tilde{T}) - \mathcal{K}(T)|}{\varepsilon} &= \frac{1}{\varepsilon} \left| \operatorname{trace}(Q^\# E Q^\#) + \sum_{j=2}^n \operatorname{trace}(Q^\# (E Q^\#)^j) \right| \\ &\leq \frac{1}{2} \sum_{j=1}^n \max \left\{ \max_i \{q_{i,j}^{\#(2)}\} - \alpha(j),\ \beta(j) - \min_i \{q_{i,j}^{\#(2)}\} \right\} + \mathcal{O}(\varepsilon), \end{aligned}$$

and hence as $\varepsilon \to 0^+$, the supremum is bounded above by

$$\frac{1}{2} \sum_{j=1}^n \max \left\{ \max_i \{q_{i,j}^{\#(2)}\} - \alpha(j),\ \beta(j) - \min_i \{q_{i,j}^{\#(2)}\} \right\}. \tag{3.5}$$

To show that the supremum is in fact equal to (3.5), it suffices to show that for any matrix T, there is some matrix E for which this bound is achieved by $|\operatorname{trace}(Q^{\#}EQ^{\#})|$. We will demonstrate how to choose the matrix E—in particular, the u_j—so that equality holds in the upper bound (3.4) on $|\operatorname{trace}(Q^{\#}EQ^{\#})|$.

Fix j, and for conciseness, let $\mathfrak{a}_j = \max_i\{q_{i,j}^{\#(2)}\} - \alpha(j)$, and $\mathfrak{b}_j = \beta(j) - \min_i\{q_{i,j}^{\#(2)}\}$. Let r_1 be an index such that

$$q_{r_1,j}^{\#(2)} = \alpha(j) = \min_l\{q_{l,j}^{\#(2)} \mid t_{j,l} > 0\},$$

and r_2 be an index such that

$$q_{r_2,j}^{\#(2)} = \beta(j) = \max_l\{q_{l,j}^{\#(2)} \mid t_{j,l} > 0\}.$$

Let s_1 be an index such that

$$q_{s_1,j}^{\#(2)} = \max_l\{q_{l,j}^{\#(2)}\},$$

and s_2 be an index such that

$$q_{s_2,j}^{\#(2)} = \min_l\{q_{l,j}^{\#(2)}\}.$$

Then the j^{th} row of E, $u_j^\top = e_j^\top E$, is chosen as follows:

$$u_j^\top = \begin{cases} \frac{\varepsilon}{2}(e_{s_1}^\top - e_{r_1}^\top) & \text{if } \max\{\mathfrak{a}_j, \mathfrak{b}_j\} = \mathfrak{a}_j; \\ \frac{\varepsilon}{2}(e_{s_2}^\top - e_{r_2}^\top) & \text{if } \max\{\mathfrak{a}_j, \mathfrak{b}_j\} = \mathfrak{b}_j. \end{cases}$$

Then

$$u_j^\top(Q^{\#})^2 e_j = |u_j^\top(Q^{\#})^2 e_j| = \max\{\mathfrak{a}_j, \mathfrak{b}_j\}.$$

Choosing in this way for each j, we have $E = \sum_{j=1}^n u_j e_j^\top$, with $\|E\|_\infty = \varepsilon$, and with

$$|\operatorname{trace}(Q^{\#}EQ^{\#})| = \frac{\varepsilon}{2}\sum_{j=1}^n \max\left\{\max_i\{q_{i,j}^{\#(2)}\} - \alpha(j),\ \beta(j) - \min_i\{q_{i,j}^{\#(2)}\}\right\}.$$

Furthermore the (i,j) entry of E is negative only if $t_{i,j} > 0$; hence $T + E$ is nonnegative (for appropriate ε). ∎

We present the following small example to further reinforce the distinction between a condition number and a structured condition number and why it is important to keep in mind that the structured condition number only provides information of value when it is assumed that the norm of the perturbing matrix is small. This is a subcase of Remark 2.4.

Example 3.5. Consider the 2×2 stochastic matrix

$$T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

which has $\mathcal{K}(T) = \frac{1}{2}$. Furthermore,

$$Q^{\#} = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \quad \text{and} \quad (Q^{\#})^2 = \begin{bmatrix} \frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{1}{8} \end{bmatrix}.$$

Hence $\mathcal{C}(T)$ can be calculated to be $\frac{1}{4}$.

Now consider the perturbing matrix

$$E = \begin{bmatrix} a & -a \\ -a & a \end{bmatrix},$$

for $0 < a < 1$ so that $\|E\|_\infty = 2a$ and

$$\tilde{T} = T + E = \begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$$

which is a stochastic matrix with eigenvalues $1, 2a-1$. Hence $\mathcal{K}(\tilde{T}) = \frac{1}{2-2a}$, and

$$\begin{aligned} |\mathcal{K}(\tilde{T}) - \mathcal{K}(T)| &= \frac{a}{2-2a} \\ &= \frac{\|E\|_\infty}{4-4a} \\ &> \frac{1}{4}\|E\|_\infty = \mathcal{C}(T)\|E\|_\infty. \end{aligned}$$

In fact, as $a \to 1$, the associated Markov chain with transition matrix $\tilde{T}$ approaches a chain which is completely decoupled, and so $\mathcal{K}(\tilde{T}) \to \infty$.

3.1. Interpretations for $\mathcal{C}(T)$ and $(Q^\#)^2$. In this section, we give some exploratory observations which connect the expression for $\mathcal{C}(T)$ in terms of $(Q^\#)^2$ with some other properties of the chain. It is expected that these may lead to other work regarding the nature of Kemeny's constant and how it is intricately interconnected with other key quantifiers of a Markov chain's behaviour, such as first passage times.

We note that the basic building block of the formula for $\mathcal{C}(T)$ in Theorem 3.4 is the term $q_{i,j}^{\#(2)} - q_{k,j}^{\#(2)}$, for some i, j, k. This term and alternate expressions for it are the subject of this section.

First, we remark that if one has already computed the mean first passage matrix M and the stationary vector w for a chain, then the group inverse $Q^\#$ can be readily computed from these. In particular, it is known (see [30]) that

$$M = (I - Q^\# + JQ^\#_{dg})W^{-1},$$

where $W = \text{diag}(w)$, A_{dg} represents the diagonal matrix with entries $a_{i,i}$ on the diagonal, and $J = \mathbb{1}\mathbb{1}^\top$, the $n \times n$ all-ones matrix. Hence

$$MW - I = -Q^\# + JQ^\#_{dg}.$$

Multiplying on the left by $w^\top$, we obtain

$$w^\top(MW - I) = \mathbb{1}^\top Q^\#_{dg};$$

hence

$$\begin{aligned} Q^\# &= I + JQ^\#_{dg} - MW \\ &= I + \mathbb{1}w^\top(MW - I) - MW \\ &= (I - \mathbb{1}w^\top)(I - MW). \end{aligned}$$

Note that this argument is given in [26, Remark 6.1.2].

From this, we can derive an expression for $(Q^{\#})^2$ in terms of M and w.

$$\begin{aligned}(Q^{\#})^2 &= ((I - \mathbb{1}w^\top)(I - MW))^2 \\ &= (I - MW - \mathbb{1}w^\top + \mathbb{1}w^\top MW)^2.\end{aligned}$$

We now consider the difference of two entries in the same column of $(Q^{\#})^2$:

$$\begin{aligned}q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} &= (e_i - e_k)^\top (Q^{\#})^2 e_j \\ &= (e_i - e_k)^\top (I - MW - \mathbb{1}w^\top + \mathbb{1}w^\top MW)^2 e_j \\ &= (e_i - e_k)^\top (I - MW)(I - MW - \mathbb{1}w^\top + \mathbb{1}w^\top MW) e_j,\end{aligned}$$

since $(e_i - e_k)^\top \mathbb{1} = 0$. From here, we have

$$\begin{aligned}q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} &= (e_i - e_k)^\top (I - MW) e_j \\ &\quad - (e_i - e_k)^\top MW(I - MW - \mathbb{1}w^\top + \mathbb{1}w^\top MW) e_j \\ &= (e_i - e_k)^\top (I - MW)(I - MW) e_j \\ &\quad + (e_i - e_k)^\top MW\mathbb{1}w^\top (I - MW) e_j,\end{aligned}$$

and since the matrix $MW\mathbb{1}w^\top = (\mathcal{K}(T) + 1)\mathbb{1}w^\top$, and all rows are equal, it follows that

$$\begin{aligned}q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} &= (e_i - e_k)^\top (I - MW)^2 e_j \\ &= (e_i - e_k)^\top (I - 2MW + (MW)^2) e_j.\end{aligned}$$

This expression shows the dependence of the sensitivity of Kemeny's constant on relationships between mean first passage times and the stationary vector.

Next we relate the term $q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j}$ with variances of first passage times. We find it particularly interesting that the sensitivity of Kemeny's constant should depend on how widely varying first passage times are in the chain. Recall from [4, Theorem 8.4.4] that the matrix V of variances of first passage times is given by $V = B - M_s$, where

$$B = M(2Q^{\#}_{dg}W^{-1} + I) + 2(Q^{\#}M - J(Q^{\#}M)_{dg}),$$

and $M_s = [(m_{i,j})^2]$. That is, B is the matrix of second moments of first passage times. Since M can be expressed in terms of $Q^{\#}$, and the term $Q^{\#}M$ is present in the above, it should be possible to write $(Q^{\#})^2$ in terms of M, W, and B. In particular, we have

$$\begin{aligned}& B = 2MQ^{\#}_{dg}W^{-1} + M + 2Q^{\#}(I - Q^{\#} - JQ^{\#}_{dg})W^{-1} - 2J(Q^{\#}M)_{dg} \\ \implies\ & BW = 2MQ^{\#}_{dg} + MW + 2Q^{\#} - 2(Q^{\#})^2 - 2J(Q^{\#}M)_{dg}W \\ \implies\ & (Q^{\#})^2 = MQ^{\#}_{dg} + \tfrac{1}{2}MW + Q^{\#} - \tfrac{1}{2}BW - J(Q^{\#}M)_{dg}W.\end{aligned}$$

So

$$\begin{aligned}q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} &= (e_i - e_k)^\top (Q^{\#})^2 e_j \\ &= (m_{i,j} - m_{k,j})q^{\#}_{j,j} + \tfrac{1}{2}(m_{i,j} - m_{k,j})w_j + (q^{\#}_{i,j} - q^{\#}_{k,j}) - \tfrac{1}{2}(b_{i,j} - b_{k,j})w_j.\end{aligned}$$

Since $m_{r,s} = \frac{q^{\#}_{s,s} - q^{\#}_{r,s}}{w_s}$ if $r \neq s$, we can rewrite

$$q^{\#}_{i,j} - q^{\#}_{k,j} = -w_j(m_{i,j} - m_{k,j}) + \delta_{i,j} - \delta_{k,j}, \tag{3.6}$$

where $\delta_{r,s}$ is the Kronecker delta function, accounting for the cases where $i = j$ or $k = j$.

To further analyse this expression, recall (see [25]) that the quantity

$$\alpha_j := \sum_{\substack{k=1 \\ k \neq j}}^{n} w_k m_{k,j}$$

is known as the *accessibility index* of the j^{th} state of the Markov chain, with an interpretation in terms of the expected time to reach state j, beginning at a random state (distinct from j) in the chain. Recalling also that the accessibility index may be written $\alpha_j = \frac{q^{\#}_{j,j}}{w_j}$ (see [25, Theorem 1.1(a)]), we use this in the above to write:

$$q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} = (m_{i,j} - m_{k,j})\alpha_j w_j - \tfrac{1}{2} w_j[(m_{i,j} - m_{k,j}) + (b_{i,j} - b_{k,j})] + (\delta_{i,j} - \delta_{k,j}),$$

or

$$q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} = w_j(m_{i,j} - m_{k,j})(\alpha_j - \tfrac{1}{2}) - \tfrac{1}{2} w_j(b_{i,j} - b_{k,j}) + (\delta_{i,j} - \delta_{k,j}).$$

From these expressions, we can see that the sensitivity of Kemeny's constant (i.e. the value of $\mathcal{C}(T)$) depends on the differences between first and second moments of first passage times from distinct pairs of states (i and k) to the same state (j). Both the importance of that state j (as described by the corresponding entry of the stationary distribution) and the accessibility of that state play a role in this expression.

We note that in [17] an expression is given for the entries of the group inverse in terms of the accessibility indices, stationary vector entries, and mean first passage times:

$$q^{\#}_{i,j} = w_j(\alpha_j - 1 - m_{i,j}) + \delta_{i,j}.$$

Given that $q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} = \sum_{l=1}(q^{\#}_{i,l} - q^{\#}_{k,l})q^{\#}_{l,j}$, along with (3.6) and the above, we have

$$\begin{aligned}
q^{\#(2)}_{i,j} - q^{\#(2)}_{k,j} &= \sum_{l=1}^{n} -w_l(m_{i,l} - m_{k,l})q^{\#}_{l,j} + \sum_{l=1}^{n}(\delta_{i,l} - \delta_{k,l})q^{\#}_{l,j} \\
&= -\sum_{l=1}^{n} w_l(m_{i,l} - m_{k,l})\left(w_j(\alpha_j - 1 - m_{l,j}) + \delta_{l,j}\right) + (q^{\#}_{i,j} - q^{\#}_{k,j}) \\
&= -w_j(\alpha_j - 1)\sum_{l=1}^{n}\left(w_l(m_{i,l} - m_{j,l})\right) + \sum_{l=1}^{n} w_l(m_{i,l} - m_{k,l})m_{l,j} \\
&\qquad -w_j(m_{i,j} - m_{k,j}) - w_j(m_{i,j} - m_{k,j}) + \delta_{i,j} - \delta_{k,j} \\
&= -w_j(\alpha_j - 1)\left(\sum_{l=1}^{n} w_l m_{i,l} - \sum_{l=1}^{n} w_l m_{j,l}\right) + \sum_{l=1}^{n} w_l(m_{i,l} - m_{k,l})m_{l,j} \\
&\qquad -2w_j(m_{i,j} - m_{k,j}) + \delta_{i,j} - \delta_{k,j}.
\end{aligned}$$

Note that $\mathcal{K}(T)+1=\sum_{l=1}^{n} w_l m_{r,l}$ for any r, the first two summations cancel, and we have

$$q_{i,j}^{\#(2)} - q_{k,j}^{\#(2)} = \sum_{l=1}^{n} w_l (m_{i,l} - m_{k,l}) m_{l,j} - 2w_j(m_{i,j} - m_{k,j}) + \delta_{i,j} - \delta_{k,j}.$$

This is an expression for a key quantity which is present in our expression for $\mathcal{C}(T)$, $q_{i,j}^{\#(2)} - q_{k,j}^{\#(2)}$, in terms of the stationary vector entries and mean first passage times. It is difficult to interpret this as it involves products of various quantities, and these products do not have a natural or intuitive interpretation as to their role in the Markov chain. In particular, the summation terms $\sum_{l=1}^{n} w_l m_{i,l} m_{l,j}$ and $-\sum_{l=1} w_l m_{k,l} m_{l,j}$ make it difficult to work towards an intuitive understanding of the behaviour of Kemeny's constant. One could argue intuitively that if i indexes a state which has poor access to relatively important states in the chain, which in turn have poor access to state j, then the value of $\sum_{l=1}^{n} w_l m_{i,l} m_{l,j}$ will be large, where we interpret w_l as a measure of the 'importance' of state l, and $m_{l,j}$ as a measure of how 'accessible' state j is from state l. These descriptions would need to be extended to the other terms in the expression, and then finally rephrased in terms of their influence on the value of $\mathcal{C}(T)$ itself, in order to obtain further understanding of the circumstances under which the value of Kemeny's constant is sensitive to perturbations. More success might be achieved if one considers only very simple perturbations of the transition matrix. For now, it is enough for us to say that it is clear that the sensitivity of $\mathcal{K}(T)$ appears to depend on the stationary probabilities, the mean first passage times, and the accessibility of each state, as well as the variances of the first passage times.

4. Bounds on $\mathcal{C}(T)$.

4.1. An upper bound for $\mathcal{C}(T)$. While the expression of $\mathcal{C}(T)$ in Theorem 3.4 is accurate, it is a complex expression and provides little direct insight into the nature of Kemeny's constant and how it acts under perturbation of the transition probabilities. We provide below an upper bound which does supply some insight, after the following technical lemma, originally proven in [32], and of which a proof may be found in [26, Lemma 5.3.4].

LEMMA 4.1. *Let v be a vector in $\mathbb{R}^n$ such that $v^\top \mathbb{1} = 0$.*
(a) Suppose that A is an $n \times n$ matrix with complex entries. Then

$$\|A^\top v\|_1 \leq \|v\|_1 \cdot \tfrac{1}{2} \max_{i,j} \sum_{k=1}^{n} |a_{i,k} - a_{j,k}|.$$

(b) Suppose that $z \in \mathbb{C}^n$. Then

$$|v^\top z| \leq \|v\|_1 \cdot \max_{i,j} \frac{|z_i - z_j|}{2}.$$

THEOREM 4.2. *Let T be an $n \times n$ irreducible stochastic matrix. Then*

$$\mathcal{C}(T) \leq n \cdot \kappa_3(T) \cdot \kappa_6(T). \tag{4.1}$$

Proof. Let T be an irreducible stochastic matrix of order n, and let E be a matrix with zero row sums such that $\tilde{T} = T + E$ is also irreducible and stochastic. We consider

$$\text{trace}(Q^\# E Q^\#) = \text{trace}(E(Q^\#)^2) = \sum_{i=1}^{n} e_i^\top E (Q^\#)^2 e_i.$$

For any i, we have from Lemma 4.1(b):

$$
\begin{aligned}
|e_i^\top E(Q^\#)^2 e_i| &\le \|e_i^\top E Q^\#\|_1 \cdot \max_{j,k}\left(\frac{q^\#_{j,i}-q^\#_{k,i}}{2}\right)\\
&= \tfrac{1}{2}\|e_i^\top E Q^\#\|_1 \cdot \max_k (q^\#_{i,i}-q^\#_{k,i}),
\end{aligned}
$$

since $q^\#_{i,i} > q^\#_{j,i}$, for all i, and $j \neq i$.

Next, consider that

$$
\begin{aligned}
\|e_i^\top E Q^\#\|_1 &= \|(Q^\#)^\top E^\top e_i\|_1\\
&\le \|E^\top e_i\|_1 \cdot \tfrac{1}{2}\max_{i,j}\sum_{k=1}^n |q^\#_{i,k}-q^\#_{j,k}| \quad \text{(by Lemma 4.1(a))}\\
&= \|e_i^\top E\|_1 \kappa_6(T).
\end{aligned}
$$

Hence

$$
|e_i^\top E(Q^\#)^2 e_i| \le \tfrac{1}{2}\|e_i^\top E\|_1 \kappa_6(T)\max_k(q^\#_{i,i}-q^\#_{k,i}).
$$

Finally, we have

$$
\begin{aligned}
|\operatorname{trace}(E(Q^\#)^2)| &= \left|\sum_{i=1}^n e_i^\top E(Q^\#)^2 e_i\right|\\
&\le \sum_{i=1}^n |e_i^\top E(Q^\#)^2 e_i|\\
&\le \tfrac{1}{2}\sum_{i=1}^n \|e_i^\top E\|_1 \kappa_6(T)\max_k(q^\#_{i,i}-q^\#_{k,i})\\
&\le \tfrac{1}{2}\|E\|_\infty \kappa_6(T)\sum_{i=1}^n \max_k(q^\#_{i,i}-q^\#_{k,i})\\
&\le \tfrac{1}{2}\|E\|_\infty \kappa_6(T)\cdot n\max_{i,k}(q^\#_{i,i}-q^\#_{k,i})\\
&= n\|E\|_\infty \kappa_6(T)\kappa_3(T) \qquad \text{(from (2.2))}.
\end{aligned}
$$

It follows that $\mathcal{C}(T) \le n\cdot\kappa_3(T)\kappa_6(T)$. ◻

Remark 4.3. Since we have observed in Remark 2.2 that $\kappa_6(T) \le \mathcal{K}(T)$ this means that

$$
\mathcal{C}(T) \le n\cdot\mathcal{K}(T)\kappa_3(T).
$$

This furnishes another relative bound, where both the original size of Kemeny's constant and the size of the perturbation are taken into account. That is,

$$
\limsup_{\varepsilon\to 0}\left\{\frac{|\mathcal{K}(T+E)-\mathcal{K}(T)|}{\varepsilon\cdot\mathcal{K}(T)}\,\middle|\, T+E \text{ irreducible, stochastic; } \|E\|_\infty \le \varepsilon\right\} \le n\cdot\kappa_3(T).
$$

4.2. Lower bounds for $\mathcal{C}(T)$. In this section we give a lower bound for the structured condition number in terms of Kemeny's constant $\mathcal{K}(T)$, and also in terms of $\kappa_3(T)$, $\kappa_6(T)$, and as a function of n, the number of states.

PROPOSITION 4.4. *Let T be an irreducible stochastic matrix. Then*

$$\mathcal{C}(T) \geq \frac{1}{2(1-\min_i t_{i,i})}\mathcal{K}(T) \geq \frac{1}{2}\mathcal{K}(T).$$

Proof. Since $\mathcal{C}(T)$ is defined as a limit-supremum over all admissible perturbing matrices E with $\|E\|_\infty \leq \varepsilon$ as $\varepsilon \to 0$, a lower bound may be produced by determining this supremum over some subfamily of perturbing matrices E. In particular, for a given $\varepsilon > 0$, by choosing $E := \frac{\varepsilon}{2}(I-T) = \frac{\varepsilon}{2}Q$, we have $E\mathbb{1} = 0$, and $\tilde{T} = T + E$ is nonnegative and irreducible for small enough ε. Furthermore,

$$\|E\|_\infty = \varepsilon \max_i \{1 - t_{i,i}\} \leq \varepsilon.$$

Then, since $EQ^{\#} = \frac{\varepsilon}{2}QQ^{\#} = \frac{\varepsilon}{2}(I - \mathbb{1}w^\top)$, we have

$$\begin{aligned}
\mathcal{K}(T+E) - \mathcal{K}(T) &= \operatorname{trace}(Q^{\#}EQ^{\#}) + \operatorname{trace}(Q^{\#}(EQ^{\#})^2) + \cdots \\
&= \frac{\varepsilon}{2}\operatorname{trace}(Q^{\#}) + \frac{\varepsilon^2}{4}\operatorname{trace}(Q^{\#}) + \cdots \\
&= \frac{\varepsilon}{2-\varepsilon}\operatorname{trace}(Q^{\#}).
\end{aligned}$$

So

$$\frac{|\mathcal{K}(T+E) - \mathcal{K}(T)|}{\varepsilon} = \frac{1}{2-\varepsilon}\operatorname{trace}(Q^{\#})$$

and

$$\mathcal{C}(T) \geq \frac{1}{2}\operatorname{trace}(Q^{\#}).$$

Finally, note that by choosing

$$E := \frac{\varepsilon}{2 \cdot \max_i\{1 - t_{i,i}\}}(I - T)$$

we have $\|E\|_\infty = \varepsilon$, and obtain the improvement

$$\mathcal{C}(T) \geq \frac{1}{2(1-\min_i t_{i,i})}\operatorname{trace}(Q^{\#}).$$

Since $\mathcal{K}(T) = \operatorname{trace}(Q^{\#})$, the result follows. ∎

Remark 4.5. Since $\kappa_6(T) \leq \mathcal{K}(T)$ as shown in [34] (and referenced in Remark 2.2 above) and since it is shown in [7] that $\kappa_3(T) \leq \kappa_6(T)$, we obtain from Proposition 4.4 that

$$\mathcal{C}(T) \geq \frac{1}{2}\kappa_6(T)$$

and

$$\mathcal{C}(T) \geq \frac{1}{2}\kappa_3(T).$$

In conclusion, Theorem 4.2 and Proposition 4.4 indicate that the conditioning of Kemeny's constant is closely tied with the conditioning of the stationary vector.

PROPOSITION 4.6. *Let T be an $n \times n$ irreducible stochastic matrix. Then*

$$\mathcal{C}(T) > \frac{n-1}{4}.$$

Proof. We have that

$$\mathcal{C}(T) \geq \tfrac{1}{2}\mathcal{K}(T), \tag{4.2}$$

and it is well-known (see, for example, [15]) that

$$\mathcal{K}(T) \geq \frac{n-1}{2}. \tag{4.3}$$

The result follows. We remark that the inequality is strict since the characterization of equality in (4.3) is shown in [24] that T must be permutation equivalent to

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

and equality does not hold in (4.2) for this matrix. ∎

5. Examples. In this section, we investigate the structured condition number $\mathcal{C}(T)$ for some infinite families of matrices. We also examine the upper bound of Theorem 4.2 and determine some families for which $\mathcal{C}(T)$ is on the same order of magnitude as this upper bound.

Example 5.1. Let $w = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}^\top$ be any positive vector such that $\sum_i w_i = 1$, and form T as a convex combination of the identity matrix and the rank-one stochastic matrix $\mathbb{1}w^\top$; that is, for some $c \in [0,1]$,

$$T = cI + (1-c)\mathbb{1}w^\top.$$

Then we have $Q = (1-c)(I - \mathbb{1}w^\top)$ and so

$$Q^\# = \frac{1}{1-c}(I - \mathbb{1}w^\top)$$

and

$$(Q^\#)^2 = \frac{1}{(1-c)^2}(I - \mathbb{1}w^\top).$$

It is easily calculated that

$$\begin{aligned} \mathcal{C}(T) &= \frac{1}{2}\sum_{j=1}^{n} \frac{1}{(1-c)^2} \\ &= \frac{n}{2(1-c)^2}. \end{aligned}$$

Meanwhile,

$$\kappa_3(T) = \frac{1}{2(1-c)}, \quad \text{and} \quad \kappa_6(T) = \frac{1}{1-c}.$$

Hence the upper bound (4.1) is $\frac{n}{2(1-c)^2}$, coinciding with the value of $\mathcal{C}(T)$.

Note also that in the special case that $c = 0$ and $w^\top = \frac{1}{n}\mathbb{1}^\top$, we have $T = \frac{1}{n}J$ and equality holds in both the lower bound of Prop. 4.4 and the upper bound (4.1).

Example 5.2. Consider the Markov chain whose transition matrix is the adjacency matrix of the directed cycle

$$T = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

That is, we consider the random walk on the directed cycle on n vertices. To determine $\mathcal{C}(T)$, we require $(Q^{\#})^2$. In fact, we require the maximum and minimum entries of each column of $(Q^{\#})^2$, along with $\alpha(j)$ and $\beta(j)$, which in this example are both equal to $q^{\#(2)}_{j+1,j}$, for each $j = 1, \ldots, n-1$, and $\alpha(n) = \beta(n) = q^{\#(2)}_{1,n}$.

This is an example of a periodic Markov chain, and there is an expression for the group inverse of $I - T$ (see [21]) which we can use, producing

$$Q^{\#} = \frac{1}{2n} \begin{bmatrix} n-1 & n-3 & n-5 & \cdots & -(n-3) & -(n-1) \\ -(n-1) & n-1 & n-3 & n-5 & \cdots & -(n-3) \\ -(n-3) & -(n-1) & n-1 & n-3 & \cdots & -(n-5) \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ n-3 & n-5 & \cdots & -(n-3) & -(n-1) & n-1 \end{bmatrix}.$$

Alternatively,

$$q^{\#}_{i,j} = \begin{cases} \frac{n-1}{2n} - \frac{j-i}{n}, & \text{if } i \leq j; \\ \frac{n-1}{2n} - \frac{n+j-i}{n}, & \text{if } i > j. \end{cases}$$

Since $Q^{\#}$ is a circulant matrix (that is, each row is a shift to the right of the one preceding it), $(Q^{\#})^2$ will also be a circulant matrix. Hence every term in the sum indexed by j in (3.1) is equal, and it suffices to determine only the first term, and then multiply by $\frac{1}{2}n$; that is,

$$\mathcal{C}(T) = \frac{n}{2} \max \left\{ \max_i \{q^{\#(2)}_{i,1}\} - \alpha(1), \beta(1) - \min_i \{q^{\#(2)}_{i,1}\} \right\}.$$

Some tedious computation produces

$$q^{\#(2)}_{k,1} = \frac{1}{4n^2} \left(-\frac{1}{3}n^3 + (2k-4)n^2 - \frac{23}{3}n + (8k - 2k^2)n \right),$$

if $k \neq 1$, and $q^{\#(2)}_{1,1} = \frac{-1}{12n}(n-1)(n-5)$. It is not difficult to show that

$$\min_k q^{\#(2)}_{k,1} = q^{\#(2)}_{2,1} = -\frac{n^2-1}{12n},$$

while

$$\max_k q^{\#(2)}_{k,1} = \begin{cases} q^{\#(2)}_{\frac{n+4}{2},1} = \dfrac{n^2+2}{24n} & \text{if } n \text{ is even;} \\ q^{\#(2)}_{\frac{n+5}{2},1} = \dfrac{n^2-1}{24n} & \text{if } n \text{ is odd.} \end{cases}$$

Hence

$$\mathcal{C}(T) = \begin{cases} \dfrac{n^2}{16} & \text{if } n \text{ is even;} \\ \dfrac{n^2-1}{16} & \text{if } n \text{ is odd.} \end{cases}$$

However, with some computation we find

$$\kappa_6(T) = \begin{cases} \dfrac{n}{4} & \text{if } n \text{ is even;} \\ \dfrac{n^2-1}{4n} & \text{if } n \text{ is odd;} \end{cases}$$

and $\kappa_3(T) = \frac{n-1}{2n}$. The upper bound for $\mathcal{C}(T)$ given in Theorem 4.2 is then equal to

$$\begin{cases} \dfrac{n^2-n}{8} & \text{if } n \text{ is even;} \\ \dfrac{n^3-n^2-n+1}{8n} & \text{if } n \text{ is odd;} \end{cases}$$

Hence for n large enough, $\mathcal{C}(T) \sim \frac{1}{2}n\kappa_3(T)\kappa_6(T)$.

Example 5.3. Consider the random walk on the path on n vertices. The transition matrix of this Markov chain is

$$T = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \cdots & 0 \\ & & \ddots & \ddots & \ddots & \\ 0 & 0 & \cdots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

From [26, Example 5.5.1], we have the following formula for the entries of $(I-T)^{\#}$:

$$q^{\#}_{i,1} = \frac{4n^2-8n+3}{12(n-1)} - \frac{(i-1)(2n-i-1)}{2(n-1)}, \quad \text{for } i = 1, \dots, n$$

$$q^{\#}_{i,n} = \frac{4n^2-8n+3}{12(n-1)} - \frac{(n-i)(n+i-2)}{2(n-1)}, \quad \text{for } i = 1, \dots, n$$

$$(5.1) \quad q^{\#}_{i,j} = \frac{4n^2-8n+3}{6(n-1)} + 2(n - \max\{i,j\}) - \frac{(n-i)(n+i-2)}{(n-1)} - \frac{(n-j)(n+j-2)}{2(n-1)}, \quad \text{for } j = 2, \dots, n-1 \text{ and } i = 1, \dots, n.$$

The group inverse of $I - T$ and in particular its square do not follow as neat a pattern as the previous example, so we do not produce here a closed-form expression for $\mathcal{C}(T)$. However, we can determine a lower bound by choosing, for each index j, indices for the terms in the sum (3.1) which may not necessarily be maximum. In particular, for the path on n vertices, we have

$$\mathcal{C}(T) \geq \frac{1}{2}\sum_{1\leq j\leq \frac{n}{2}} (q^{\#(2)}_{1,j} - q^{\#(2)}_{j+1,j}) + \sum_{\frac{n}{2}<j\leq n} (q^{\#(2)}_{n,j} - q^{\#(2)}_{j-1,j})$$
$$= \sum_{1\leq j\leq \frac{n}{2}} (q^{\#(2)}_{1,j} - q^{\#(2)}_{j+1,j}),$$

where the equality comes from the structure in $Q^{\#}$ and $(Q^{\#})^2$. In particular, by examining the expressions in (5.2), one can see that $q^{\#}_{i,j} = q^{\#}_{n+1-i,n+1-j}$; the same relationships hold for entries of $(Q^{\#})^2$.

Some tedious computation with the aid of symbolic computation software allows us to compute the following lower bounds for $\mathcal{C}(T)$, where T is the transition matrix of the random walk on a path on n vertices:

$$\mathcal{C}(T) \geq \begin{cases} \dfrac{11n^5 + 65n^4 + 100n^3 - 1040n^2 + 1704n - 720}{2880(n-1)} & \text{if } n \text{ is even} \\ \dfrac{11n^5 + 80n^4 - 50n^3 - 1040n^2 + 1959n - 720}{2880(n-1)} & \text{if } n \text{ is odd.} \end{cases}$$

Note that both lower bounds are on the order of n^4.

Next, we examine the upper bound given in Section 2.2. From [26, Example 5.5.1] it is known that

$$\kappa_3(T) = \frac{(n-2)^2}{2(n-1)}.$$

For $c_1(T) = \max_{i,j}\{\sum_{k=1}^{n} |q^{\#}_{i,k} - q^{\#}_{j,k}|\}$, it is not difficult to show that the maximum is attained when $i = 1$ and $j = n$ (or vice versa). A proof of this claim may be found in the Appendix. Hence we have $\kappa_6(T) = \sum_{k=1}^{n} |q^{\#}_{1,k} - q^{\#}_{n,k}|$ and with some more tedious computation we find that

$$\kappa_6(T) = \begin{cases} \dfrac{(n-1)^2 + 1}{4} & \text{if } n \text{ is even} \\ \dfrac{(n-1)^2}{4} & \text{if } n \text{ is odd.} \end{cases}.$$

Hence the upper bound is on the order of $\frac{n^4}{8}$ (ignoring lower order terms), while $\mathcal{C}(T)$ is bounded below by a function which is also on the order of n^4. That is, $\mathcal{C}(T) = \Theta(n^4)$, and it is on the same order as the upper bound.

Such a high order of magnitude indicates that Kemeny's constant is extremely poorly-conditioned for the random walk on a path on n vertices, particularly since Kemeny's constant for this Markov chain with transition matrix T is known to be $\frac{2n^2-4n+3}{6}$.

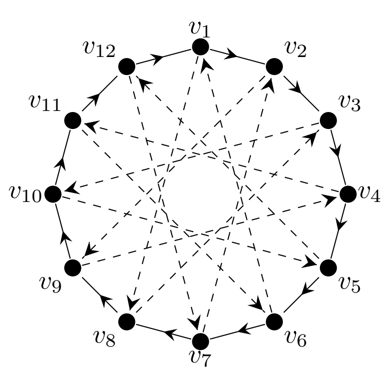

Fig. 1: The directed cycle on 12 vertices, for which a 'bad' perturbation introduces new transitions displayed here as dashed arcs.

In each of these examples where the conditioning of Kemeny's constant is poor, the structure of the transition matrix is very specific. Furthermore, the perturbation which produces such a large difference in Kemeny's constant breaks the structure completely. In the directed cycle example, we observe that this 'worst-case' perturbation introduces many new possible transitions into the chain, taking what is essentially a deterministic process and making it much more stochastic; see Fig. 1 for the new transitions introduced under this perturbation to the directed cycle on twelve vertices.

It is natural, then, to ask about the conditioning of Kemeny's constant where perturbations must respect the given structure of the transition matrix; that is, consider only perturbations where zero entries are preserved. More formally, consider a directed graph D, consisting of a vertex set $V = \{1, \dots, n\}$ and a directed edge set $E \subseteq V \times V$; then define $\mathcal{S}_D$ as the set of all stochastic irreducible matrices T such that $t_{i,j} > 0$ only if $(i, j) \in E$. With this definition, we can re-frame the above as an examination of the conditioning of Kemeny's constant for a matrix $T \in \mathcal{S}_D$ (for a given D) where we consider only the perturbations $T + E$ of T where $T + E \in \mathcal{S}_D$. While this is an interesting and natural question, we remark that there is an entire family of directed graphs given in [5] for which the value of $\mathcal{K}(T)$ depends only on the directed graph, and not on the values of the transition probabilities. Directed graphs with this property are characterised by the following conditions:

1. Every vertex of D has positive outdegree.
2. There exists an integer k such that all cycles of D have length k.
3. There is a vertex in D that lies on every cycle in D.

Then $\mathcal{K}(T) = \frac{2n-k-1}{2}$, for all irreducible $T \in \mathcal{S}_D$, where n is the number of vertices in D. An example of such a directed graph is displayed in Fig. 2.

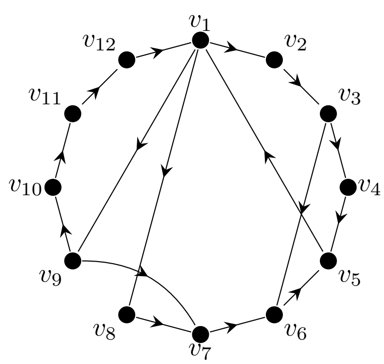

Fig. 2: A directed graph D on twelve vertices for which every $T \in \mathcal{S}_D$ has $\mathcal{K}(T)$ equal to nine.

Acknowledgements:. This research was supported in part by the University of Manitoba Graduate Fellowship (Jane Breen), and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada under grant number RGPIN-2019-05408 (Steve Kirkland).

The authors are grateful to three anonymous referees, whose constructive comments helped to improve this article.

REFERENCES

[1] Linda J.S. Allen and Amy M. Burgin. Comparison of deterministic and stochastic SIS and SIR models in discrete time. *Mathematical Biosciences*, 163(1):1–33, 2000.

[2] Dario Bini, Jeffrey J. Hunter, Guy Latouche, Beatrice Meini, and Peter Taylor. Why is Kemeny's constant a constant? *J. Appl. Probab.*, 55(4):1025–1036, 2018.

[3] Jane Breen. *Markov Chains under Combinatorial Constraints: Analysis and Synthesis.* PhD thesis, University of Manitoba, 2018.

[4] Stephen L. Campbell and Carl D. Meyer. *Generalized inverses of linear transformations*, volume 56 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1979.

[5] Minerva Catral, Steve Kirkland, Michael Neumann, and Nung-Sing Sze. The Kemeny constant for finite homogeneous ergodic Markov chains. *Journal of Scientific Computing*, 45(1-3):151–166, 2010.

[6] Yonghong Chen, Stephen J. Kirkland, and Michael Neumann. Group generalized inverses of *M*-matrices associated with periodic and nonperiodic Jacobi matrices. *Linear and Multilinear Algebra*, 39(4):325–340, 1995.

[7] Grace Cho and Carl Meyer. Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335(1):137–150, 2001.

[8] Grace E. Cho and Carl D. Meyer. Markov chain sensitivity measured by mean first passage times. *Linear Algebra and its Applications*, 316(1-3):21–28, 2000.

[9] Emanuele Crisostomi, Stephen Kirkland, and Robert Shorten. A Google-like model of road network dynamics and its application to regulation and control. *International Journal of Control*, 84(3):633–651, 2011.

[10] Peter Deuflhard and Christof Schütte. Molecular conformation dynamics and computational drug design. *Applied Mathemetics Entering the 21st Century. Proceedings ICIAM*, pages 91–119, 2004.

[11] Moshe Haviv and Ludo Van der Heyden. Perturbation bounds for the stationary probabilities of a finite Markov chain. *Advances in Applied Probability*, 16(4):804–818, 1984.

[12] Desmond J Higham and Nicholas J Higham. Structured backward error and condition of generalized eigenvalue problems. *SIAM Journal on Matrix Analysis and Applications*, 20(2):493–512, 1998.

[13] Roger Horn and Charles Johnson. *Matrix Analysis.* Cambridge University Press, 2nd edition, 2013.

[14] Jeffrey J. Hunter. Generalized inverses and their application to applied probability problems. *Linear Algebra Appl.*, 45:157–198, 1982.

[15] Jeffrey J. Hunter. Mixing times with applications to perturbed Markov chains. *Linear Algebra Appl.*, 417(1):108–123, 2006.

[16] Jeffrey J. Hunter. Variances of first passage times in a Markov chain with applications to mixing times. *Linear Algebra Appl.*, 429(5-6):1135–1162, 2008.

[17] Jeffrey J. Hunter. Generalized inverses of Markovian kernels in terms of properties of the Markov chain. *Linear Algebra Appl.*, 447:38–55, 2014.

[18] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains.* The University Series in Undergraduate Mathematics. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1960.

[19] S. Kirkland. On a question concerning condition numbers for Markov chains. *SIAM J. Matrix Anal. Appl.*, 23(4):1109–1119, 2002.

[20] Stephen J. Kirkland, Michael Neumann, and Nung-Sing Sze. On optimal condition numbers for Markov chains. *Numer. Math.*, 110(4):521–537, 2008.

[21] Steve Kirkland. The group inverse associated with an irreducible periodic nonnegative matrix. *SIAM Journal on Matrix Analysis and Applications*, 16(4):1127–1134, 1995.

[22] Steve Kirkland. Conditioning properties of the stationary distribution for a Markov chain. *Electronic Journal of Linear Algebra*, 10(1):1, 2003.

[23] Steve Kirkland. A combinatorial approach to the conditioning of a single entry in the stationary distribution for a Markov chain. *Electronic Journal of Linear Algebra*, 11(1):14, 2004.

[24] Steve Kirkland. Fastest expected time to mixing for a Markov chain on a directed graph. *Linear Algebra and its Applications*, 433(11):1988–1996, 2010.

[25] Steve Kirkland. Random walk centrality and a partition of Kemeny's constant. *Czechoslovak Math. J.*, 66(141)(3):757–775, 2016.

[26] Steve Kirkland and Michael Neumann. *Group Inverses of M-Matrices and their Applications.* CRC Press, 2013.

[27] Mark Levene and George Loizou. Kemeny's constant and the random surfer. *Amer. Math. Monthly*, 109(8):741–745, 2002.

[28] Carl D. Meyer. The condition of a finite Markov chain and perturbation bounds for the limiting

probabilities. *SIAM Journal on Algebraic Discrete Methods*, 1(3):273–283, 1980.

[29] Carl D. Meyer. Sensitivity of the stationary distribution of a Markov chain. *SIAM Journal on Matrix Analysis and Applications*, 15(3):715–728, 1994.

[30] Carl D. Meyer, Jr. The role of the group generalized inverse in the theory of finite Markov chains. *SIAM Rev.*, 17:443–464, 1975.

[31] Paul J Schweitzer. Perturbation theory and finite Markov chains. *Journal of Applied Probability*, 5(2):401–413, 1968.

[32] E. Seneta. *Non-negative matrices and Markov chains.* Springer Series in Statistics. Springer, New York, 2006. Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544].

[33] Eugene Seneta. Perturbation of the stationary distribution measured by ergodicity coefficients. *Advances in Applied Probability*, 20(1):228–230, 1988.

[34] Eugene Seneta. Sensitivity analysis, ergodicity coefficients, and rank-one updates for finite Markov chains. *Numerical Solutions of Markov Chains*, pages 121–129, 1991.

Appendix. The following is a short proof of a claim used in Example 5.3 in the computation of $\kappa_6(T) := \max_{i,j}\{\sum_{k=1}^{n} |q^{\#}_{i,k} - q^{\#}_{j,k}|\}$, where T is the transition matrix of the random walk on a path on n vertices.

LEMMA 5.4. *Let T be the transition matrix for the random walk on a path. Then*

$$\kappa_6(T) = \sum_{k=1}^{n} |q^{\#}_{1,k} - q^{\#}_{n,k}|.$$

Proof. First, we note that by the symmetry in the entries of $Q^{\#}$, that

$$\sum_{k=1}^{n} |q^{\#}_{i,k} - q^{\#}_{j,k}| = 2\sum_{k=1}^{\frac{n}{2}} |q^{\#}_{i,k} - q^{\#}_{j,k}|,$$

if n is even, and with an extra term corresponding to $k = (n+1)/2$ if n is odd.

Next we show that for $1 \le k \le \lceil \frac{n}{2} \rceil$, $\min_j\{q^{\#}_{j,k}\} = q^{\#}_{n,k}$. This follows from [6], in which it is proven that for a tridiagonal stochastic matrix, the group inverse has the property that

$$q^{\#}_{k,k} > q^{\#}_{k+1,k} > \cdots > q^{\#}_{n,k},$$

and that

$$q^{\#}_{1,k} < q^{\#}_{2,k} < \cdots < q^{\#}_{k,k}.$$

To show that $q^{\#}_{n,k}$ is a minimal entry in the first $\lceil \frac{n}{2} \rceil$ columns of $Q^{\#}$, it suffices to show that $q^{\#}_{n,k} < q^{\#}_{1,k}$. This is easily confirmed from the formulas given for the entries of $Q^{\#}$ in Example 5.3.

Hence

$$\max_{i,j}\left\{\sum_{k=1}^{\lceil \frac{n}{2} \rceil} |q^{\#}_{i,k} - q_{j,k}|\right\} = \max_{1\le i\le n}\left\{\sum_{k=1}^{\lceil \frac{n}{2} \rceil} (q^{\#}_{i,k} - q^{\#}_{n,k})\right\}.$$

It remains to show that this maximum is obtained for $i = 1$. Some computation with the formulas given in (5.2) produces the following:

$$\sum_{k=1}^{\lceil \frac{n}{2} \rceil} (q^{\#}_{i,k} - q^{\#}_{n,k}) = \begin{cases} \dfrac{n^2 - 2n - 2i^2 + 4i}{4} & \text{if } n \text{ is even} \\ \dfrac{n^3 - 3n^2 + n - 2i^2(n-2) + 4i(n-2) + 3}{4(n-1)} & \text{if } n \text{ is odd.} \end{cases}$$

Both expressions are decreasing functions in i for $i > 1$; hence the maximum is attained for $i = 1$. ∎