The Annals of Probability
2000, Vol. 28, No. 3, 1230–1258

MARKOV ADDITIVE PROCESSES AND PERRON–FROBENIUS EIGENVALUE INEQUALITIES[1]

BY COLM O'CINNEIDE

Purdue University

We present a method for proving Perron–Frobenius eigenvalue inequalities. The method is to apply Jensen's inequality to the change in a "random evolution" over a regenerative cycle of the underlying finite-state Markov chain. One of the primary benefits of the method is that it readily gives necessary and sufficient conditions for strict inequality. It also gives insights into some of the conjectures of J. E. Cohen. Ney and Nummelin's "Hypothesis 2" arises here as a condition for strict inequality, and we explore its ramifications in detail for a special family of Markov additive processes which we call "fluid models." This leads to a connection between Hypothesis 2 and the condition "$P^T P$ irreducible" which arose in the work of Cohen, Friedland, Kato and Kelly.

1. Introduction. This paper concerns a family of inequalities for Perron–Frobenius eigenvalues, going back to Kingman's 1961 log-convexity result [11], and including some of the inequalities of Friedland and Karlin [8], Cohen [5], Cohen, Friedland, Kato and Kelly [7] and Asmussen and O'Cinneide [1]. One contribution of the present paper is a new method of proof of these inequalities which gives a detailed understanding of conditions for strict inequality. Cohen [6] gives several conjectures related to these inequalities. See [12] for a counterexample to one of these. We resolve three others below. This work was motivated by questions related to the tail of the waiting time distribution in queues [1].

The general approach of this paper was anticipated in the final section of [7], in which the authors suggest that the CFKK inequality (Theorem 1 below) might be amenable to proof by applying a majorization argument [9, 14, 18] to a suitable "random evolution." We use the term "random evolution" informally here to mean the exponential of a Markov additive process; see Section 2 for preliminaries on Markov additive processes. In this paper, we see that by considering the change in a random evolution over a suitable regenerative cycle of the underlying Markov chain, one can readily derive many of the Perron–Frobenius eigenvalue inequalities of the literature cited above, and, more importantly, the conditions for strict inequality also become clear. Proving conditions for strict inequality has always been a troublesome point in earlier work, because the techniques needed have been fundamentally different from those invoked to prove the inequalities themselves. Here are some

Received November 1999.

[1]Supported in part by NSF Grant 97-14630-DMII.

AMS 1991 *subject classifications*. Primary 60I10; secondary 15A42, 16J27.

Key words and phrases. Markov additive processes, finite-state Markov chains, Perron–Frobenius eigenvalue inequalities.

instances of this. In [7], the CFKK inequality is proved, with the condition for strict inequality, but the proof is quite long and is based on Kato's perturbation theory for linear operators [10]. Of two short alternative proofs of the inequality, one by Varadhan (described on page 68 of [7]) and another by C. M. Newman (page 45 of [6]), neither appears to present an easy path to the strict inequality. The situation is similar for the inequality of Proposition 1 below. It is given without the condition for strict inequality as Lemma 6 of [7]. In [1], the condition for strict inequality is proved using methods that are of limited scope. The inequality itself is easier to prove. Kingman [7] does not give a condition for strict log-convexity. We provide a natural one here.

A more detailed description of the new method is as follows. We begin with a finite-state Markov chain X, in discrete or continuous time. We devise two Markov additive processes (X, A) and (X, B), X being the Markov component and A and B being the additive components, for which we can prove $\mathbb{E}(e^{A(t)}) \leq \mathbb{E}(e^{B(t)})$, $t \geq 0$, under a suitable initial condition, typically using Jensen's inequality. In some cases, the needed Markov additive processes may be found in the literature. For example, Cohen gives these for the inequality of Proposition 1 below and for the CFKK inequality. By construction, the long-run growth rates of these expectations will be the logs α and β of the eigenvalues to be compared, and so this inequality immediately gives the Perron–Frobenius eigenvalue inequality $e^\alpha \leq e^\beta$. The drawback of this argument is that it requires taking a limit as $t \to \infty$, and this makes deducing a strict inequality awkward. To get the condition for strict inequality, we call on an idea from Ney and Nummelin [15]: that, under an integrability condition, with T denoting the first positive time at which X visits a state i, we have $\mathbb{E}_i(\exp\{A(T) - \alpha T\}) = \mathbb{E}_i(\exp\{B(T) - \beta T\}) = 1$. This often give us detailed eigenvalue information without the need for taking limits.

The idea of looking at random evolutions over regenerative cycles leads to the second theme of this paper. The necessary and sufficient conditions for strict inequality derived in various contexts below always include a condition of [15] known as "Hypothesis 2." See Definitions 1 and 2 of Section 4, where the term "nondegenerate" is used to mean "satisfies Hypothesis 2." For a given discrete-time, finite-state Markov chain X with irreducible transition matrix P, we consider the space $V(P)$ of all vectors v such that the Markov additive process (X, S) defined by

$$S(t) = \sum_{n=1}^{t} v(X(n)) \tag{1.1}$$

is degenerate. We call a Markov additive process of the form (1.1) a "fluid model" here. We give a detailed description of $V(P)$ in terms of the communication relation of the nonnegative matrix $P^T P$. As an example of what is proved, we show that the dimension of $V(P)$ is the number of communicating classes of the matrix $P^T P$. This brings out a connection between the Ney–Nummelin hypothesis and a condition—that $P^T P$ be irreducible—that arises from Kato's perturbation theory in Lemma 5 of [7].

The remainder of the paper is organized as follows. In Section 2, we give some elementary background on Markov additive processes and fluid models associated with finite-state chains. In Section 3 we illustrate the basic approach by proving several known results related to continuous-time fluid models, including the CFKK inequality. Getting strict inequality for continuous-time fluid models is simplified because Ney and Nummelin's Hypothesis 2 can fail only in a trivial way for these processes. In Section 4, we discuss Hypothesis 2 and relate it to the condition for strict inequality in some basic inequalities. In Section 5 we completely characterize the vector space $V(P)$ of degenerate fluid models associated with a given irreducible stochastic matrix P. Section 6 relates this to various results from the literature cited above, giving conditions for strict inequality in Kingman's result [11] and giving a complete picture of the (strict-) convexity properties of the function $\phi(\Lambda) \equiv \log \rho(Pe^{\Lambda})$, of a real diagonal matrix Λ, where ρ is the spectral radius; see Theorems 5 and 6 below. This leads to a resolution of Cohen's conjecture #2, presented as Theorem 7. In Section 7 we prove Cohen's conjecture #8, under an additional condition, using methods unrelated to the main theme of the paper, and also indicate how to prove his conjecture #9.

Some notes on matrix terminology: a *scalar matrix* is a matrix that is a multiple of the identity. An *essentially positive matrix* is a matrix with nonnegative off-diagonal entries. A *generator* is an essentially positive matrix with zero row sums. An essentially positive matrix A is said to be irreducible if $A + \lambda I$ is an irreducible nonnegative matrix for λ sufficiently large. Perron–Frobenius theory extends naturally to such matrices, and the Perron–Frobenius eigenvalue of an essentially positive matrix A is denoted by $\operatorname{dev}(A)$, "dev" abbreviating "dominant eigenvalue." For A nonnegative, of course $\rho(A) = \operatorname{dev}(A)$. For more on nonnegative matrices, see [2] or [17].

Several of the results of this paper for stochastic matrices and generators may be extended to general nonnegative matrices and essentially positive matrices, as follows. If A is an irreducible nonnegative matrix with Perron–Frobenius eigenvector v and Perron–Frobenius eigenvalue λ, then, with $V \equiv \operatorname{diag}(v)$, $P \equiv \lambda^{-1} V^{-1} A V$ is an irreducible stochastic matrix. Similarly, if A is an irreducible essentially positive matrix with Perron–Frobenius eigenvector v and Perron–Frobenius eigenvalue λ, then $Q \equiv V^{-1} A V - \lambda I$ is an irreducible generator, where again $V = \operatorname{diag}(v)$. The assumption of irreducibility generally plays a role below only in getting conditions for strict inequality. If the inequality alone is to be proved, the assumption of irreducibility may often be relaxed using simple arguments based on continuity of eigenvalues and the fact that any nonnegative matrix (respectively, essentially positive matrix) is the limit of a sequence of irreducible nonnegative matrices (respectively, irreducible essentially positive matrices).

2. Preliminaries. Here we give some notation and basic results for Markov additive processes [4, 15, 16]. The Markov additive processes we consider in this paper almost all have the special feature that their additive components are completely determined by the path of the chain. There is only one exception to this, in the proof of Theorem 1 in the next section. Considering more general Markov additive processes does not lead to generalizations of the other inequalities. For this reason, in the present section we focus on Markov additive processes with this special feature.

2.1. *Discrete case.* Let X be an irreducible, discrete-time Markov chain with a finite number m of states. The state space is taken to be $\{1.2, \ldots, m\}$, and we refer to indices in this range as "states" throughout the paper even when no stochastic process has been invoked. Let $P = (p_{jk})$ be the transition matrix of X. Let $F = (f(i, j))$ be an $m \times m$ matrix, where $f(i, j) \in [-\infty, \infty)$ is thought of as a "reward" received when a transition from i to j is made. Define

$$S(t) = \sum_{n=1}^{t} f(X_{n-1}, X_n), \quad t = 0, 1, 2, \ldots, \tag{2.1}$$

which is the total "reward" received by time t. Then the pair (X, S) is here called the *discrete-time Markov additive process with data* (P, F). Note that if $p_{ij} = 0$ the value of $f(i, j)$ has no impact on the law of the process (X, S). Note also that if S ever reaches the value $-\infty$, it remains there from that time on. The mean drift in S is

$$\mu \equiv \lim_{t\to\infty} \frac{\mathbb{E}(S(t))}{t} = \sum_{i, j=1}^{m} \pi_i p(i, j) f(i, j) = \pi(P \odot F)e, \tag{2.2}$$

where e is a column vector of 1's and $\odot$ denotes the "entrywise" matrix product, defined by $(a_{ij}) \odot (b_{ij}) \equiv (a_{ij}b_{ij})$. The limit here is independent of the initial distribution assumed in taking the expectation. It is elementary that

$$\begin{aligned} &\left(\mathbb{E}_i\left[e^{S(t)}; X(t) = j\right]\right)_{i, j\in\{1, 2, \ldots, m\}} = H^t, \\ &\qquad\qquad t = 0, 1, 2, \ldots, \text{ with } H \equiv P \odot F^{\exp}, \end{aligned} \tag{2.3}$$

where $F^{\exp} \equiv (\exp(f(i, j)))$. We term the matrix H here the *fundamental matrix* of the Markov additive process.

In the special case where F has constant columns, we write $\lambda(j) \equiv f(i, j) \in [-\infty, \infty)$, and (2.1) simplifies to

$$S(t) = \sum_{n=1}^{t} \lambda(X_n). \tag{2.4}$$

In this case, the process (X, S) is here called the *discrete-time fluid model with data* (P, Λ), *where* $\Lambda = \operatorname{diag}(\lambda(1), \lambda(2), \ldots, \lambda(m))$. For the discrete-time fluid model, we have

$$H = Pe^{\Lambda}. \tag{2.5}$$

If F is finite, which is to say, if all the $f(i, j)$'s are $> -\infty$, then the fundamental matrix H, like P, is irreducible. Suppose that H is irreducible with right Perron–Frobenius eigenvector denoted by h (which is positive and unique up to scaling) and with Perron–Frobenius eigenvalue denoted by $\lambda = e^{\alpha}$, so that $Hh = \lambda h$. We term the quantity α the *exponential growth parameter* of the Markov additive process. Then it is elementary that the process

$$h(X(t))\exp\{S(t)-\alpha t\}, \quad t = 0, 1, 2, \ldots$$

is a martingale, which we refer to informally as an "exponential martingale." With this notation we have the following basic lemma.

LEMMA 1. *Let i denote a fixed state and let T_k denote the kth positive time that X enters state i, for $k = 1, 2, 3, \ldots$. Then α is the unique solution in v of $\mathbb{E}_i(\exp\{S(T_k) - \vartheta T_k\}) = 1$.*

The proof follows easily from Lemma 4.1 and equation (4.1) of [15], and the discussion of the discrete-state case in Example 6.2 of that paper.

Here is a quick sketch of a direct proof of Lemma 1, for the sake of clarity of exposition. Using $X(T_k) = i$ in the first equality to follow and the optional sampling theorem in the second, we have

$$\begin{aligned}\mathbb{E}_i\big(h(i)\exp\big\{S(T_k)-\alpha T_k\big\}\big) &= \mathbb{E}_i\big(h\big(X(T_k)\big)\exp\big\{S(T_k)-\alpha T_k\big\}\big)\\ &= \mathbb{E}_i\big(h(X(0))\exp\big\{S(0)\big\}\big) = h(i).\end{aligned}$$

(Uniform integrability may be verified by a matrix calculation.) As $h(i) > 0$, this proves that α satisfies the given equation. Uniqueness follows from the observation that, as $T_k > 0$, the quantity $\mathbb{E}_i\big(\exp\{S(T_k) - \vartheta T_k\}\big)$ is strictly decreasing in ϑ whenever it is finite.

For future reference, we note an elementary formula for $\mathbb{E}\big(S(T_k)\big)$ in the context of this lemma,

$$\mathbb{E}\big(S(T_k)\big) = \mu\mathbb{E}(T_k). \tag{2.6}$$

This is a consequence of the laws of large numbers for $S(t)$, $t \to \infty$, and T_k, $k \to \infty$, for example.

2.2. *Continuous case.* Let X be an irreducible, continuous-time Markov chain, again with a finite number m of states. Let $Q = (q_{jk})$ be its generator and let λ be a real-valued function on the state space, which we identify with the real vector $\lambda = (\lambda(1), \lambda(2), \ldots, \lambda(m))$. We think of $\lambda(i)$ as the rate of arrival of a fluid "reward" while X is in state i. Let $F = (f(i, j))$ be a $m \times m$ matrix having zero diagonal entries ($f(i, i) = 0$), the other entries being in $[-\infty, \infty)$. We think of $f(i, j)$ as a fixed "reward" acquired whenever a transition from i to j occurs. The total amount of reward that arrives in $[0, t]$ is given by

$$\begin{aligned}S(t) = &\int_0^t \lambda\big(X(s)\big)ds\\ &+ \sum_{0<s\le t} f\big(X(s^-), X(s)\big), \qquad t \ge 0.\end{aligned}$$

(X, S) is referred to here as the *continuous-time Markov additive process with data* (Q, Λ, F), where $\Lambda = \mathrm{diag}(\lambda(1), \lambda(2), \dots, \lambda(m))$. If $q_{ij} = 0$ for some $i \neq j$, then the value of $f(i, j)$ does not affect the law of the process. The mean drift in S is

$$\mu \equiv \lim_{t\to\infty} \frac{\mathbb{E}\big(S(t)\big)}{t} = \sum_{i=1}^{m} \pi_i \lambda(i) + \sum_{i,\, j=1}^{m} \pi_i q\big(i, j\big) f\big(i, j\big) = \pi\big(\Lambda + Q \odot F\big)e. \tag{2.7}$$

It may be shown that

$$\Big(\mathbb{E}_i\Big[e^{S(t)}; X(t) = j\Big]\Big)_{i,\, j\in\{1, 2, \dots, m\}} = \exp\Big\{\Big(\Lambda + Q \odot F^{\exp}\Big)t\Big\}, \quad t \geq 0. \tag{2.8}$$

In (2.7) and (2.8) we have used the notation appearing in (2.3). If $F = 0$, then S does not jump at transitions of X, and the process is called the *continuous-time fluid model with data* (Q, Λ). In this case $Q \odot F^{\exp} = Q$ and so by (2.8) we have

$$\Big(\mathbb{E}_i\Big[e^{S(t)}; X(t) = j\Big]\Big) = e^{(\Lambda+Q)t}, \quad t \geq 0. \tag{2.9}$$

To construct the appropriate exponential martingale, define the *fundamental matrix* of the present Markov additive process to be $H \equiv \Lambda + Q \odot F^{\exp}$. Suppose that H is irreducible (which it is as long as F is finite), and let h denote its right Perron–Frobenius eigenvector and α its Perron–Frobenius eigenvalue. Again, we call α the *exponential growth parameter* of the Markov additive process. Then, as before, the process

$$h\big(X(t)\big) \exp\big\{S(t) - \alpha t\big\}, \qquad t \geq 0$$

is a martingale, and we have the following analog of Lemma 1, which may be proved in the same way.

LEMMA 2. *Let i denote a fixed state and let T_k denote the kth positive time at which X enters state i. Then α is the unique solution in ϑ to the equation* $\mathbb{E}_i\big(\exp\big\{S(T_k) - \vartheta T_k\big\}\big) = 1$.

We close by noting the continuous-time analog of (2.6),

$$\mathbb{E}\big(S\big(T_k\big)\big) = \mu \mathbb{E}\big(T_k\big). \tag{2.10}$$

3. Illustrations of technique in continuous time. In this section, we prove three simple results for continuous-time fluid models. Each has appeared in the literature before, but in each case the method of proof is new. Proposition 1 says in essence that the long-run growth rate of the expected exponential of a continuous-time fluid model is at least as large as its mean drift. This result is a variant of Lemma 6 of [7] but adds the condition for strict inequality. The proof of the inequality here is a relative of that outlined in [6]—see the first paragraph on page 42. A more general result, in fact, the discrete-time analog, was given as part of Theorem 3.1 in [8]. We treat this in Corollary 1 in the next section, where again we contribute the condition for strict inequality. The second result of this section, Proposition 2, is from [1], where a somewhat more complex proof was given and where the strict inequality was not treated. The last result addressed is the CFKK inequality, of which the condition for strict inequality is given a rather complex proof in [7]. The proof given here is much shorter. Using the methods described in the remark at the end of the introduction, it is easily shown that the inequalities of Propositions 2 and 3 hold assuming only that Q is an essentially positive matrix.

3.1. *Two inequalities for the continuous-time fluid model.*

PROPOSITION 1. *For an irreducible generator Q with steady-state distribution π and a real diagonal matrix Λ, we have*

$$\operatorname{dev}(Q+\Lambda) \geq \mu \equiv \pi\Lambda e.$$

Moreover, there is equality here if and only if Λ is a scalar matrix.

PROOF. Let (X, S) be the continuous-time fluid model with data (Q, Λ). Then μ is the average drift of S by (2.7). Letting T denote the first positive time at which X enters an arbitrary fixed state i, and letting $\alpha \equiv \operatorname{dev}(Q+\Lambda)$, Lemma 2 gives the first step of the following:

$$\begin{aligned} 1 &= \mathbb{E}_i(\exp\{S(T) - \alpha T\}) \geq \exp\{\mathbb{E}_i(S(T) - \alpha T)\} \\ &= \exp\{(\mu - \alpha)\mathbb{E}_i(T)\}. \end{aligned}$$

The inequality here is Jensen's and the last equality is by (2.10) and the fact that $\mathbb{E}_i(T) < \infty$ by irreducibility. Since $\mathbb{E}_i(T) > 0$ we can conclude that $\alpha \geq \mu$, as was to be proved.

By comparing the first and last expressions in the inequality above, we see that there is equality there if and only if $\alpha = \mu$. But there is equality in Jensen's inequality if and only if $S(T) - \alpha T$ is a.s. constant when $X(0) = i$, but by irreducibility this happens only if $\Lambda = \alpha I$. This is because the joint distribution of the occupation times of the m states over the cycle $[0, T]$ has an absolutely continuous component on $\mathbb{R}^m$. This completes the proof. □

The example of Section 6 of [1] may be adapted to show that the next inequality does not necessarily hold for k nonintegral.

PROPOSITION 2. *For an irreducible generator Q and a real diagonal matrix Λ, we have*

$$\operatorname{dev}(Q+\Lambda) \geq \operatorname{dev}(kQ+\Lambda) \quad \textit{for } k = 2, 3, \ldots .$$

Moreover, there is equality here if and only if Λ is a scalar matrix.

PROOF. Let (X, S) be as in the proof of Proposition 1. Let $T_1, T_2, \ldots$ denote the successive positive return times to some particular state i, and set $T_0 \equiv 0$. Set $\tau_n \equiv T_n - T_{n-1}$ and $\Delta_n \equiv S(T_n) - S(T_{n-1})$, $n = 1, 2, \ldots$. Let α denote the eigenvalue on the left of the inequality to be proved. Then Lemma 2 gives the first step here.

$$\begin{aligned}
1 &= \mathbb{E}_i(\exp(S(T_k) - \alpha T_k) \\
&= \mathbb{E}_i(\exp(\Delta_1 + \Delta_2 + \cdots + \Delta_k - \alpha(\tau_1 + \tau_2 + \cdots + \tau_k))) \\
&\leq \mathbb{E}_i\left(\frac{1}{k}\sum_{n=1}^{k} \exp(k\Delta_n - k\alpha\tau_n)\right) \\
&= \mathbb{E}_i(\exp(k\Delta_1 - k\alpha\tau_1)) = \mathbb{E}_i(\exp(kS(T_1) - k\alpha T_1))
\end{aligned} \tag{3.1}$$

The inequality again is Jensen's, and the next-to-last equality is because the cycles delimited by the successive return times to state i are i.i.d. It is easy to verify that (X, S'), where $S'(t) \equiv kS(t)$, is a fluid model with data $(Q, k\Lambda)$. Let $\beta \equiv \operatorname{dev}(Q + k\Lambda)$. Then, again by Lemma 2, we have $\mathbb{E}_i \exp\{S'(T_1) - \beta T_1\} = 1$. By comparing this with the right side of (3.1), using the fact that $0 < T_1 < \infty$ a.s., we conclude that

$$\beta \equiv \operatorname{dev}(Q + k\Lambda) \geq k\alpha = k \operatorname{dev}(Q + \Lambda). \tag{3.2}$$

Equality here is equivalent to equality in (3.1) by Lemma 2. But we have equality in Jensen's inequality if and only if $k\Delta_n - k\alpha\tau_n$, $n = 1, 2, \ldots, k$, are all equal a.s. This in turn is equivalent to the statement that these random variables are constant, being independent, and this, finally, is equivalent to the condition that $\Lambda = \alpha I$, as in the proof of Proposition 1. A simple scaling (replacing Q by kQ) shows that (3.2) is equivalent to the inequality to be proved. □

3.2. *The CFKK inequality.* Consider again the continuous-time fluid model (X, S) with parameters (Q, Λ), where Q is irreducible. The process $X^0 = (X(t)$: $t = 0, 1, 2, \ldots)$ is a discrete-time Markov chain with transition matrix e^Q. We associate with this a natural discrete-time approximation to S, defined as

$$S^0(t) = \sum_{n=1}^{t} \lambda(X(n)) = \sum_{n=1}^{t} \lambda(X^0(n)), \quad t = 0, 1, 2, \ldots .$$

This is again a Markov additive process: it is the discrete-time fluid model with parameters (e^Q, Λ). In [6], Cohen considers the two processes

$$\exp(S(t)) \text{ and } \exp(S^0(t)), \quad t = 0, 1, 2, \ldots .$$

The first may be viewed as the size of a population at integer time t whose growth rate at any time $s \geq 0$ is $\lambda(X(s))$. The second is what you would believe the population size to be at time t if you observe the rate of growth at integer times and assumed that this rate remained constant between observations. Cohen [6] points out that the CFKK inequality may be interpreted as stating that the long-run growth rate of the expectation of the discrete-time process is at least as large as that of the continuous-time version, in the sense that

$$\lim_{t\to\infty} \frac{\log \mathbb{E}(\exp(S(t)))}{t} \leq \lim_{t\to\infty} \frac{\log \mathbb{E}(\exp(S^0(t)))}{t}, \tag{3.3}$$

with equality if and only if the $\lambda(i)$'s are all equal. [*Note added in proof*: This interpretation was first given in Cohen, J. E. (1979) Random evolutions in discrete and continuous time. *Stochastic Process. Appl.* **9** 245–251.] Using (2.3) and (2.8), (3.3) becomes the eigenvalue inequality of the following.

THEOREM 1 (The CFKK inequality [7]). *For Q an irreducible generator and Λ a real diagonal matrix, we have $\rho(e^{Q+\Lambda}) \leq \rho(e^Q e^\Lambda)$, with equality if and only if Λ is a scalar matrix.*

PROOF. For each $\sigma \in [0, 1)$, define the process

$$S^\sigma(t) = \sum_{0<n-\sigma\leq t} \lambda\big(X\big(n-\sigma\big)\big), \quad t = 0, 1, 2, \ldots,$$

which, like the special case S^0 above, is a discretization of the process S. The pair (X^0, S^σ) is a (discrete-time) Markov additive process. S is related to the processes S^σ at integer times t by

$$S(t) = \int_0^1 S^\sigma(t)\, d\sigma, \quad t = 1, 2, \ldots. \tag{3.4}$$

Now let T denote the first positive (integer) time that X^0 enters state i. Let $\alpha \equiv \rho(e^{Q+\Lambda})$ and $\beta = \rho(e^Q e^\Lambda)$. Using Lemma 1 twice (more precisely, an elementary extensions of Lemma 1: see the remark following the proof), we have

$$1 = \mathbb{E}_i(\exp(S(T) - \alpha T)) = \mathbb{E}_i(\exp(S^\sigma(T) - \beta T)) \qquad \text{for all } \sigma \in [0, 1). \tag{3.5}$$

It is intuitive that the exponential growth parameter β of the S^σ's does not depend on σ, since these are all similar discretizations of the same continuous-time process, and it is a simple calculation to verify this. Using (3.4) in the first step to follow and Jensen's inequality in the second, we have

$$\begin{aligned} \exp\big(S(T) - \beta T\big) &= \exp\left\{\int_0^1 \big(S^\sigma(T) - \beta T\big) d\sigma\right\} \\ &\leq \int_0^1 \exp\big\{S^\sigma(T) - \beta T\big\} d\sigma \end{aligned} \tag{3.6}$$

and so

$$\mathbb{E}_i\big(\exp(S(T) - \beta T)\big) \leq \int_0^1 \mathbb{E}_i\big(\exp\big\{S^\sigma(T) - \beta T\big\}\big) d\sigma = 1, \tag{3.7}$$

using (3.5) at the last step. For a given point ω in our heretofore anonymous probability space Ω, there is equality in Jensen's inequality (3.6) if and only if $S^\sigma(T) = S(T)$ for almost all $\sigma \in [0, 1)$. Thus there is equality in (3.7) if and only if $S^\sigma(T) = S(T)$ for almost all (σ, ω) in the product space $[0, 1) \times \Omega$. This is clearly true if Λ is scalar, but otherwise it cannot be true because, for each $\sigma \in [0, 1)$, the distribution of $S^\sigma(T)$ is concentrated on the countable set of finite sums of integer multiples of the $\lambda(i)$'s, whereas that of $S(T)$ has an absolutely continuous component on $\mathbb{R}$. $\square$

We remark that the discrete-time Markov additive processes (X^0, S) (observed at integer times) and (X^0, S^σ), which appeared in (3.5), are exceptional in this paper in one respect: the additive components are not completely determined by the path of the Markov component X^0, as they depend on the evolution of X between the integer times. However, Lemma 1 extends readily to this situation [15].

4. Ney and Nummelin's Hypothesis 2. Ney and Nummelin's Hypothesis 2, given on page 570 of [15], is needed in that paper to get a certain strict convexity property (see their Corollary 3.3). This condition, which is called "nondegeneracy" here, has apparently not previously been connected with the inequalities of [5–8]. Proposition 3 and Corollary 1 below are the most basic examples of the role of nondegeneracy in proving strict inequalities. Corollary 1 is in essence Theorem 3.1 of [8] but adds the necessary and sufficient condition for strict inequality. Lemmas 3 and 4 are graph-theoretical preliminaries needed in the next section.

Let A be a nonnegative $m \times m$ matrix and let $i, j \in \{1, 2, \ldots, m\}$ be a pair of states. A *path of A from i to j*, is a sequence of states $\Pi = \langle i = i(0), i(1), \ldots, i(k) = j\rangle$, where $k \geq 0$, such that $a_{i(l-1), i(l)} > 0$ for $l = 1, 2, \ldots, k$. The *length of the path is k*. A path from i to i is called an *i-cycle*. We allow paths and cycles to revisit the same state. Given an $m \times m$ matrix $F = (f(i, j))$, $f(i, j) \in [-\infty, \infty)$, the ($F$-) *reward accumulated on the path* Π is defined as

$$\text{Reward}(\Pi) \equiv \sum_{l=1}^{k} f(i(l-1), i(l)).$$

DEFINITION 1 (Degeneracy). Let A be a nonnegative $m \times m$ matrix. A real $m \times m$ matrix $F = (f(i, j))$ is said to be *degenerate with respect to* A if there exists a real θ such that, for every cycle $\langle i_0, i_1, \ldots, i_k = i_0\rangle$ of A of length $k \geq 0$, we have

$$\text{Reward}(\Pi) = k\theta. \tag{4.1}$$

Similarly, we say that the discrete-time Markov additive process with data (P, F) is *degenerate* if F is degenerate with respect to P. If (4.1) holds with $\theta = 0$, then we say that (A, F) is *null*.

The simplest example of degeneracy is to take $f(i, j) = c$ for all i and j, in which case (4.1) holds with $\theta = c$. The property of degeneracy depends on the matrix A only through its graph, that is, through the set of pairs (i, j) for which a_{ij} is positive; the actual values of the positive a_{ij}'s are immaterial.

Note that if A is irreducible, then it is enough to suppose that (4.1) holds for all i_0-cycles, for some fixed i_0. This is because, by irreducibility, for any given state i there is an i_0-cycle which visits i exactly once. Such an i_0-cycle may be extended by inserting an arbitrary i-cycle in place of i. The result is a new i_0-cycle. Let k_0 denote the length of the original i_0-cycle and k that of the inserted i-cycle. Then, by (4.1), the reward accumulated over the original i_0-cycle is $k_0\theta$, while the reward accumulated over the elongated i_0-cycle is $(k_0 + k)\theta$. Now their difference, which is the reward accumulated over the inserted i-cycle, is $k\theta$, as required.

To relate this concept to Markov additive processes, let P be an irreducible stochastic matrix and F a real matrix. Consider the discrete-time Markov additive process (X, S) with data (P, F). Let i be an arbitrary fixed state and let $T > 0$ denote the time of first entry of X into i. Then it is immediate from the definition and the discussion of the previous paragraph that F is degenerate with respect to P if and only if, for some constant θ, we have

$$S(T) = \theta T \text{ a.s. under the initial condition } X(0) = i. \tag{4.2}$$

If this holds, then it is clear from Lemma 1 and (2.6) that θ is both the exponential growth parameter and the drift of the Markov additive process (X, S):

$$\alpha = \mu = \theta \tag{4.3}$$

The next result strengthens the connection between degeneracy and Perron–Frobenius eigenvalues by giving a "converse" to (4.3).

PROPOSITION 3. *Consider the discrete-time Markov additive process (X, S) with data (P, F), where P is irreducible and F is real, and let μ denote its drift and α its exponential growth parameter. Then we have $\alpha \geq \mu$, with equality if and only if F is degenerate with respect to P.*

PROOF. Letting T denote the first positive time at which the chain enters a state i, Lemma 2 gives the first step of the following:

$$1 = \mathbb{E}_i(\exp\{S(T) = \alpha T\}) \geq \exp\{\mathbb{E}_i(S(T) - \alpha T)\} = \exp\{(\mu - \alpha)\mathbb{E}_i(T)\}.$$

The inequality here is Jensen's, and the last equality is because of (2.6) and the fact that $\mathbb{E}_i(T)$ is finite by irreducibility of P. Since $\mathbb{E}_i(T) > 0$, it follows that $\alpha \geq \mu$, as was to be proved. We have $\alpha = \mu$ if and only if there is equality in Jensen's inequality here, which in turn holds if and only if $S(T) = \alpha T$ a.s. when $X(0) = i$. Note now that this condition, that "*$S(T) = \alpha T$ a.s. under $X(0) = i$*," is equivalent to the condition that "*$S(T) = \theta T$ a.s. under $X(0) = i$ for some constant θ*," because the latter implies that $\theta = \alpha$ as explained at (4.3). But the second condition is precisely the condition (4.2) for (P, F) to be degenerate. This completes the proof. $\square$

The next result is a direct corollary, stated for the special case of discrete fluid models. It follows at once from Proposition 3 and (2.5). As we mentioned earlier, it is Theorem 3.1 of [8] plus a necessary and sufficient condition for strict inequality. A slightly weaker assumption than irreducibility was adopted in [8], but the extension to the more general case is straightforward. Theorem 2.1 of [8] gives the condition for strict inequality only in the case of strictly positive, doubly-stochastic matrices P.

COROLLARY 1. *For an irreducible stochastic matrix P with steady-state distribution π, and any real diagonal matrix Λ, we have*

$$\rho\big(Pe^{\Lambda}\big) \ge e^{\mu} \quad \textit{where } \mu \equiv \pi\Lambda e.$$

Moreover, the inequality here is strict if and only if (P, Λ) is nondegenerate.

In the remainder of this section we develop various basic facts and ideas related to degeneracy. We begin with a sample characterization of degeneracy with respect to an irreducible matrix.

LEMMA 3. *Let A be an irreducible nonnegative matrix and F a real matrix.*

(a) *F is null with respect to A if and only if F-reward is path-independent with respect to A: for every pair of states i and j, the F-reward accumulated on any A-path from i to j is the same.*

(b) *F is degenerate with respect to A if and only if there exists a real vector g and a scalar θ such that $f(i, j) = g(j) - g(i) + \theta$ for all i, j with $a_{ij} > 0$.*

PROOF. (a) Suppose (A, F) is null. Let i and j be arbitrary states. By irreducibility, there is a path Π from j to i. By concatenating Π with an arbitrary path Π' from i to j, we form an i-cycle, on which the accumulated reward is 0. Thus Reward $(\Pi') = -\text{Reward}(\Pi)$, a quantity which does not depend on the choice of the path Π'.

Conversely, if the reward is path-independent, then the reward accumulated on an i-cycle is the same as the reward on the zero-length i-cycle $\langle i\rangle$, which is zero.

(b) If f is of this form, then (4.1) is seen to hold as the series on the left "telescopes." Conversely, suppose that (4.1) holds. Consider the matrix $F' \equiv (f(i, j)-\theta)$ of modified reward rates. Then (4.1) holds for F' and A with $\theta = 0$, so that (A, F') is null. This implies that the F'-reward is path-independent, by part (a) of this lemma. Since by irreducibility of A there is a path leading from any state to any other, this allows us to uniquely define the F'-reward $r(i, j)$ accumulated over a path from i to j for all i, j. As the reward accumulated on a cycle is zero, it follows by concatenating paths in a natural way that $r(i, j) + r(j, i) = 0$ and $r(i, j) + r(j, k) + r(k, i) = 0$ for all i, j, k. These imply that $r(i, j) = g(j) - g(i)$ where g is defined by $g(i) \equiv r(i_0, i)$, i_0 being any fixed state. Now, if $a_{ij} > 0$, then $\langle i, j\rangle$ is a path, and the F'-reward accumulated on this path is $f'(i, j) = f(i, j) - \theta$ by definition, but it is also $g(j) - g(i)$ by what we have just shown. This completes the proof. □

The next concept, which we call "fluid-degeneracy" of a pair (A, Λ) with Λ diagonal, could in fact be expressed directly in terms of Definition 1. However, the special parametrization of fluid models warrants a new statement. Moreover, this leads naturally to the idea of *degeneracy of a matrix* A, which is the central concept of the next two sections.

DEFINITION 2 (Fluid-degeneracy). For a nonnegative $m \times m$ matrix A, a real $m \times m$ diagonal matrix $\Lambda = \mathrm{diag}(\lambda(1), \lambda(2), \dots, \lambda(m))$ is said to be *(fluid-) degenerate with respect to* A if there exists a real θ such that, for every cycle $\Pi = \langle i_0, i_1, \dots, i_k = i_0 \rangle$ of A, we have

$$\mathrm{Reward}(\Pi) \equiv \sum_{l=1}^{k} \lambda(i_l) = \theta k. \tag{4.4}$$

We also express this condition by saying that "the pair (A, Λ) is *(fluid-)degenerate*." If (A, Λ) is *(fluid-)*degenerate for some *nonscalar* Λ, then A is itself said to be *(fluid-)degenerate*. If (4.4) holds with $\theta = 0$, we say that (A, Λ) is *(fluid-)null*.

The scalar matrix $\Lambda = cI$ is fluid-degenerate with respect to any nonnegative A, and in this case (4.4) holds with $\theta = c$. To relate Definition 2 to Definition 1, note that (A, Λ) is fluid-degenerate if and only if (A, F) is degenerate with $F \equiv (f(i, j))$ where $f(i, j) \equiv \lambda(j)$. Naturally, we say that A is *fluid-nondegenerate* if, for all nonscalar real diagonal matrices Λ, (A, Λ) is fluid-nondegenerate.

If F is a diagonal matrix, there is a possibility of ambiguity in the statement "(A, F) is degenerate," because the parameters could be interpreted as defining a general Markov additive process or a fluid model. We use the prefix "fluid-" if it is unclear from the context whether Definition 1 or Definition 2 applies.

EXAMPLE 1. This example is similar to the one on page 67 of [7]. Consider the fluid model with parameters (P, Λ) given by

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Every cycle beginning and ending in any particular state i makes the same number of visits to state 1 as it does to 2, and so, because the rewards in these states are 1 and -1, respectively, we see that the sum of the rewards over any cycle is zero. Thus Λ is degenerate with respect to P [in fact, null, as $\theta = 0$ in (4.4)]. As Λ is nonscalar, this implies that P is degenerate by Definition 2.

LEMMA 4. *Let A be an irreducible nonnegative matrix and F a real matrix.*

(a) *If for two distinct states j and k both $p(j, j)$ and $p(k, k)$ are positive and $f(j, j) \neq f(k, k)$, then F is nondegenerate with respect to A.*

(b) *If the diagonal entries of A are all positive, then A is nondegenerate.*

(c) *A is degenerate if and only if (A, Λ) is fluid-null for some nonscalar diagonal matrix Λ.*

PROOF. (a) Consider a cycle of A that visits each of j and k exactly, once; we represent it as $\Pi = \langle i_0, i_1, \ldots, j, \ldots, k, \ldots, i_l = i_0\rangle$. Such a cycle exists as A is irreducible. Since $p(j, j)$ and $p(k, k)$ are both positive, a new cycle of A may be constructed which is identical to the original except that n_1 additional visits to j are inserted after the original visit to j and n_2 additional visits to k are inserted after the original visit to k, giving a cycle which may be represented as $\Pi' = \langle i_0, i_1, \ldots, j, j, \ldots, j, \ldots, k, k, \ldots, k, \ldots, i_l = i_0\rangle$. Suppose now that F is degenerate with respect to A, and consider (4.1) for the two paths Π and Π'. If we subtract these two cases of (4.1), we get

$$\begin{aligned}\mathrm{Reward}(\Pi') - \mathrm{Reward}(\Pi) &= n_1 f(j, j) + n_2 f(k, k)\\ &= \theta(n_1 + n_2) \quad \text{for all } n_1, n_2 = 1, 2, \ldots .\end{aligned}$$

This implies that $f(j, j) = f(k, k) = \theta$. It follows that if $f(j, j) \neq f(k, k)$ then F cannot be degenerate with respect to A.

(b) Suppose A satisfies the given condition. Then by part (a) of this lemma, if for any $j \neq k$ we have $\lambda(j) \neq \lambda(k)$, then Λ is not degenerate with respect to A. Thus Λ is not degenerate with respect to A unless its diagonal entries are all equal, which is to say, unless it is scalar matrix. So A itself is nondegenerate by Definition 2.

(c) If A is degenerate, then for some nonscalar $\Lambda', (A, \Lambda')$ is degenerate. Define $\Lambda \equiv \Lambda' - \theta I$, where θ is the value for which (4.4) holds for this model. Then the fluid model (A, Λ) is null, as it satisfies (4.4) with $\theta = 0$. The converse is obvious. $\square$

5. Fluid-degeneracy and the graph of $A^T A$. In this section we completely describe the space of diagonal matrices Λ which are degenerate with respect to a given irreducible nonnegative matrix A. In particular, we give a formula for its dimension. These results explain and further develop an idea appearing in Lemma 5 of [7], in which the condition that $A^T A$ be irreducible arises as a sufficient condition for a certain strict inequality. Here the full significance of this condition becomes clear, and in the next section it is seen to be *necessary and sufficient* for strict inequality or strict convexity in some standard results.

We remark that the condition that $A^T A$ be irreducible is equivalent to the condition that AA^T be irreducible. This is explained just after the proof of Lemma 7. We now introduce some simple and descriptive terminology. A state i is said to be a *parent* of a state j (with respect to A) if $a_{ij} > 0$; in this case we also say that j is the *daughter* of i. States i and j are said to be *sisters*, written $i \approx j$, if, for some state l, $a_{li} a_{lj} > 0$ (they have the same parent l). It is easy to verify that $i \approx j$ if and only if the (i, j)-entry of $A^T A$ is positive. States i and j are said to be *cousins* if they are "related through sisterhood," in the sense that there exists a sequence of states $i_1, i_2, \ldots, i_n$ such that $i \approx i_1 \approx i_2 \approx \cdots \approx i_n \approx j$. The next lemma is immediate from these observations.

LEMMA 5. *Let A be a nonnegative matrix. Then the cousin relation determined by A is the communication relation determined by $A^T A$. In particular, every state is a cousin of every other state with respect to A if and only if $A^T A$ is irreducible.*

The following specializes Lemma 3(b) to the case of fluid models, in which case the function g of that lemma may be identified explicitly up to an additive constant.

LEMMA 6. *Let A be irreducible and suppose that (A, Λ) is null. Then, for any path Π from a state i to one of its cousins j, we have*

$$\textit{Reward } (\Pi) = \lambda(j) - \lambda(i).$$

PROOF. Consider first two sister states $i \approx j$ and an arbitrary A-path $\Pi \equiv \langle i = i_0, i_1, i_2, \ldots, i_k = j\rangle$ from i to j. Such a path exists as A is irreducible. As i and j are sisters, there is a state l such that $a_{li}a_{lj} > 0$. Irreducibility of A implies that there is a path $\langle j, j_1, j_2, \ldots, j_p = l\rangle$ from j to l, and so $\Pi' = \langle i = i_0, i_1, i_2, \ldots, i_k = j, j_1, j_2, \ldots, j_p = l, i\rangle$ is an i-cycle. The fact that $a_{li} > 0$ is used here in "closing the loop" with the last visit to i. The fact that $a_{lj} > 0$ implies that $\Pi'' = \langle j, j_1, j_2, \ldots, j_p = l, j\rangle$ is also a cycle. The reward accumulated around each of these cycles is zero by assumption, and so the difference between these two rewards must also be zero. This gives

$$\text{Reward}(\Pi') - \text{Reward } (\Pi'') = \sum_{q=0}^{k-1} \lambda(i_q) = 0 \text{ or } \sum_{q=1}^{k} \lambda(i_q) = \lambda(j) - \lambda(i),$$

since $i_0 = i$ and $i_k = j$. The second sum here is precisely the reward accumulated on the original path Π from i to j. We have established this for every path from a state i to one of its sisters j. But if i and j are merely *cousins*, they are connected by a series of sisters $i = \iota 0 \approx \iota_1 \approx \cdots \approx \iota_n = j$, and, by concatenating paths from one sister to the next, we can construct a path connecting these cousins i and j on which the reward accumulated is again

$$\begin{aligned}&\big(\lambda(\iota_1) - \lambda(\iota_0)\big) + \big(\lambda(\iota_2) - \lambda(\iota_1)\big) + \cdots + \big(\lambda(\iota_n) - \lambda(\iota_{n-1})\big)\\ &\quad = \lambda(\iota_n) - \lambda(\iota_0) = \lambda(j) - \lambda(i).\end{aligned}$$

This completes the proof. □

LEMMA 7. (i) *Let A be an irreducible nonnegative matrix and let $D_1, D_2, \ldots, D_r$ be the equivalence classes of the states $\{1, 2, \ldots, m\}$ under the cousin relation. Let C_k denote the set of all states that are parents of some state in D_k, for $k = 1, 2, \ldots, r$. Then the C_k's, like the D_k's, form a partition of $\{1, 2, \ldots, m\}$.*

(ii) *Suppose now that* $r > 1$. *Define* $D_k^{\mathrm{exit}} \equiv D_k \cap C_k^c$ *and* $D_k^{\mathrm{enter}} \equiv D_k^c \cap C_k$, *for* $k = 1, 2, \ldots, r$. *These sets are nonempty. Define a diagonal matrix* Λ_k *whose ith diagonal entry is*

$$\lambda_k(i) = \begin{cases} 1, & \text{if } \in D_k^{\mathrm{exit}}, \\ -1, & \text{if } \in D_k^{\mathrm{enter}}, \\ 0, & \text{otherwise.} \end{cases}$$

Then Λ_k *is nonscalar and null with respect to* A *for* $k = 1, 2, \ldots, r$.

(iii) *For each state* $i \in \{1, 2, \ldots, m\}$, *either:*

(a) $\lambda_k(i) = 0$ *for* $k = 1, 2, \ldots, r$, *or*

(b) *there is a unique pair of indices* $h \neq k$ *such that* $\lambda_h(i) = +1$, $\lambda_k(i) = -1$, *and* $\lambda_j(i) = 0$ *for all* $j \neq h, k$.

Finally, we have

$$\sum_{k=1}^{r} \Lambda_k = 0. \tag{5.1}$$

It is a consequence of (i) here that D_k^{exit} is the set of states in D_k all of whose daughters are outside of D_k, and D_k^{enter} is the set of states outside of D_k all of whose daughters are in D_k.

PROOF. (i) That the cousin relation is an equivalence relation follows from Lemma 5. Thus the D_k's form a partition of the state space. To deal with the C_k's, first note that every state is in some C_k because every state is a parent of some other state, by irreducibility. Now if a state were in C_k and C_j for $k \neq j$, then this state would be a parent of at least one state in each of D_k and D_j, implying that these states were sisters, and contradicting the fact that they are in different equivalence classes with respect to the cousin relation. Thus the C_k's also form a partition.

(ii) Suppose now that $r > 1$. The set $D_k^{\mathrm{exit}} = D_k \cap C_k^c$ is nonempty for each $k = 1, 2, \ldots, r$, for otherwise no state of D_k would be a parent of a state outside D_k, contradicting irreducibility of A. Similarly, $D_k^{\mathrm{enter}} = D_k^c \cap C_k$ is nonempty, for otherwise no state of D_k^c would be a parent of a state outside of D_k^c (i.e., in D_k), again contradicting irreducibility of A. These facts imply that each Λ_k is nonscalar, having at least one $+1$ and at least one -1 on its diagonal.

To show that (A, Λ_k) is null, we argue as follows. Example 1 of the previous section illustrates the idea underlying the proof. Without loss of generality, and to simplify indices, take $k = 1$. Consider a cycle starting in a state $i_0 \in D_1^{\mathrm{enter}}$. Thus i_0 is not in D_1, but the second state of the cycle is in D_1. This is the first of possibly several successive visits to states in D_1, among which there is exactly one visit to a state in $D_1^{\mathrm{exit}} \subset C_1^c$, after which the next state is in D_1^c, in which set the cycle remains until it makes precisely one visit to D_1^{enter}. At that point, when the path first returns to the set D_1^{enter}, in which it started, we have accumulated a reward of $+1$ for the single visit to D_1^{exit} and -1 for the single visit to D_1^{enter} [not counting the initial visit to i_0; see (4.4)], giving a total of $+1 - 1 = 0$. However, the chain is not necessarily back in the original state i_0, and, even if it is, the cycle may not yet be complete. We repeat such "tours," paths beginning and ending in D_1^{enter}, each time collecting a total reward of 0, until we eventually visit to i_0 and complete the cycle. The reward accumulated upon completing the cycle is 0. Thus (A, Λ_1) is null.

(iii) Since the D_h, $h = 1, 2, \ldots, r$, and the C_k, $k = 1, 2, \ldots, r$, form partitions of the state space, the r^2 sets $D_h \cap C_k$, $h, k = 1, 2, \ldots, r$, also form a partition. Consider a fixed state i. Then $i \in D_h \cap C_k$ for a unique pair h, k. It turns out that cases (a) and (b) correspond to the conditions $h = k$ and $h \neq k$, respectively.

First, suppose that $h = k$. Then $i \in D_h \cap C_h$ and so i is not in any of the sets $D_j^{\text{exit}} = D_j \cap C_j^c$ or $D_j^{\text{enter}} = D_j^c \cap C_j$, $j = 1, 2, \ldots, r$. Thus $\lambda_k(i) = 0$ for $k = 1, 2, \ldots, r$. This is condition (a) above.

Otherwise, we must have $h \neq k$. Then

$$D_h^{\text{exit}} \cap D_k^{\text{enter}} = \left(D_h \cap C_h^c\right) \cap \left(D_k^c \cap C_k\right) = D_h \cap C_k, \tag{5.2}$$

and so, as $i \in D_h \cap C_k$, it follows that $i \in D_h^{\text{exit}} \cap D_k^{\text{enter}}$. Now (b) is seen to hold by definition of Λ_j; $j = 1, 2, \ldots, r$. In particular, $\lambda_j(i) = 0$ for $i \neq h, k$, because i cannot be in any D_j^{exit}, $j \neq h$, or D_j^{enter}, $j \neq k$.

Finally, we now know that for a given state i either (a) all the $\lambda_k(i)$'s are zero or (b) there is exactly one $+1$ and exactly -1 among them. In either case, the sum of the $\lambda_k(i)$, $k = 1, 2, \ldots, r$, is zero. This implies (5.1). $\square$

We remarked earlier that, for A irreducible, irreducibility of $A^T A$ is equivalent to irreducibility of AA^T. A fuller picture of this fact may now be given. The classes C_k, $k = 1, 2, \ldots, r$, defined in Lemma 7 are the communicating classes of the nonnegative matrix AA^T. We do not use this fact later and so do not present the simple proof.

EXAMPLE 2. Consider the irreducible nonnegative matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \text{ for which } A^T A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}.$$

$A^T A$ has three classes:

$$D_1 = \{1\}; D_2 = \{2, 4\}; D_3 = \{3\}.$$

The C_k's of Lemma 7 are

$$C_1 = \{4\}; C_2 = \{1, 3\}; C_3 = \{2\}.$$

The states in D_k can be reached in one transition from states in C_k. The Λ_k's of Lemma 7 are

$$\Lambda_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}; \quad \Lambda_2 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad \Lambda_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

A special feature of this example is that $D_k^{\text{exit}} = D_k$ for all $k = 1, 2, \dots, r$.

THEOREM 2. *Let A be an irreducible nonnegative matrix. Then A is nondegenerate if and only if $A^T A$ is irreducible.*

PROOF. Suppose first that $A^T A$ is reducible. In Lemma 7 we constructed a nonscalar diagonal matrix Λ_1 such that (A, Λ_1) is degenerate. Therefore, A is degenerate, as required.

Conversely, suppose that $A^T A$ is irreducible and that (A, Λ) is null, so that (4.4) holds with $\theta = 0$. We next deduce from this that $\Lambda = 0$. Once this is done, Lemma 4(c) will imply that A is nondegenerate, and this will complete the proof of the theorem.

By Lemma 5 and the assumption that $A^T A$ is irreducible, all pairs of states are cousins. Now choose two state i and j for which $a_{ij} > 0$. Then $\langle i, j \rangle$ itself is a path from i to j, on which the accumulated reward is $\lambda(j) - \lambda(i)$ by Lemma 6, but is also $\lambda(j)$ by definition of the accumulated reward; see the first equality of (4.4). Thus $\lambda(j) - \lambda(i) = \lambda(j)$ and so $\lambda(i) = 0$. As this holds for all i for which there is a j with $a_{ij} > 0$, we conclude by irreducibility of A that $\Lambda = 0$. □

LEMMA 8. *With the same hypotheses and notation as Lemma* 7, *the vector space*

$$V \equiv \operatorname{span}\{\Lambda_1, \Lambda_2, \dots, \Lambda_r\}$$

has dimension $r - 1$.

PROOF. Consider the directed graph

$$G \equiv \{(h, k) | 1 \le h, k \le r \text{ and } D_h \cap C_k \ne \varnothing\}$$

on nodes $1, 2, \dots, r$. Irreducibility of A implies that this graph is "diconnected" [3]: there is a (directed) path from any node to any other. Therefore there exists a (directed) spanning tree (or "spanning arborescence" [13]) G_0 of G with $r - 1$ arcs,

$$G_0 = \{(h_1, k_1), (h_2, k_2), \dots, (h_{r-1}, k_{r-1})\}.$$

We order these so that (a) h_1 is the root of the directed tree; (b) each h_p, $p > 1$, appears at least once among $h_1, h_2, \dots, h_{p-1}, k_1, k_2, \dots, k_{p-1}$ and (c) no k_p appears among $h_1, h_2, \dots, h_p$ or $k_1, k_2, \dots, k_{p-1}$. That this can be done may be seen by thinking about the obvious algorithm for constructing the tree, in which, beginning with an arbitrary node h_1, at the pth step we connect one new node, k_p, to the current tree. By relabeling states, we can suppose, as we do a little later, that $h_1 = 1$ and $k_p = p+1$ for $p = 1, 2, \dots, r-1$, in which case $h_p < k_p = p+1$.

For each $p = 1, 2, \dots, r-1$, choose a single state i_p from $D_{h_p} \cap C_{k_p}$. The latter set is $D_{h_p}^{\text{exit}} \cap D_{k_p}^{\text{enter}}$ by (5.2). Part (iii)(b) of Lemma 7 then implies that for $p = 1, 2, \dots, r-1$ we have

$$\begin{aligned}&\lambda_{h_p}(i_p) = +1, \lambda_{k_p}(i_p) = -1 \quad \text{and} \\ &\lambda_j(i_p) = 0 \quad \text{for all } j \neq h_p, k_p, 1 \leq j \leq r.\end{aligned} \tag{5.3}$$

Define an $(r-1) \times r$ matrix Z as follows: the entry z_{pk} of Z is $\lambda_k(i_p)$, which is the i_pth entry on the diagonal of Λ_k. In other words, the kth column of Z consists of the $r-1$ diagonal entries of Λ_k with indices i_p, $p = 1, 2, \dots, r-1$. Now invoking (5.3) and the labeling described in the preceding paragraph, in which $h_p < k_p = p+1$, we see that the pth row of Z contains a 1 in the h_pth column and a -1 in the $p+1$st column:

$$Z = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots \\ z_{21} & z_{22} & -1 & 0 & \cdots \\ z_{31} & z_{32} & z_{33} & -1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

(Each z_{pk} here is either 0 or 1, and each row has precisely one $+1$ and precisely one -1.) Therefore, Z is of full rank, that is, rank $r-1$. Since the columns of Z are subarrays of the Λ_k's, the span of the Λ_k's, which is V, must have dimension at least as large as the rank of Z, namely $r-1$. But the dimension of V cannot exceed $r-1$ because of the linear constraint (5.1) satisfied by the Λ_k's. These two facts imply that the dimension of V is $r-1$, as required. □

THEOREM 3. *Let A be an irreducible nonnegative matrix and let r denote the number of classes of $A^T A$. Let* $V_0(A) \equiv \{\Lambda | \Lambda = \operatorname{diag}(v)$ *for some* $v \in \mathbb{R}^m$ and (A, Λ) *is null* $\}$. *Then* $\dim(V_0(A)) = r-1$. *Moreover*

$$V_0(A) = \operatorname{span}\{\Lambda_1, \Lambda_2, \dots, \Lambda_r\},$$

where the Λ_i's are as defined in Lemma 7.

PROOF. Part 2 of Lemma 7 tells us that each Λ_k, $k = 1, 2, \dots, r$, is null with respect to A, and so

$$V_0(A) \supset \operatorname{span}\{\Lambda_1, \Lambda_2, \dots, \Lambda_r\}. \tag{5.4}$$

The dimension of the space on the right is $r-1$, by Lemma 8. We prove that the dimension of $V_0(A)$ is no greater than $r-1$. This forces equality in (5.4) and proves the theorem.

Let Π be a cycle of A, and define the *occupation time of* Π *in state* i, denoted by $\operatorname{Occ}_i(\Pi)$, to be the number of visits the cycle makes to state i (not counting the first visit to the initial state):

$$\text{for } \Pi = \langle i_0, i_1, i_2, \dots, i_k = i_0 \rangle,\ \ \mathrm{Occ}(\Pi) \equiv \#\{l | i_l = i, 1 \le l \le k\}.$$

Let $\mathrm{Occ}(\Pi)$ denote the vector $(\mathrm{Occ}_1(\Pi), \mathrm{Occ}_2(\Pi), \dots, \mathrm{Occ}_m(\Pi))$, which we refer to as the *occupation time vector* of Π. Let $\mathscr{O}(A)$ denote the set of occupation time vectors of all A-cycles:

$$\mathscr{O}(A) = \{\mathrm{Occ}(\Pi) | \Pi \text{ a cycle of } A\}.$$

Then by definition $V_0(A)$ is the set of diagonal matrices that "annihilate" $\mathscr{O}(A)$:

$$V_0(A) = \{\mathrm{diag}\ (v) | v \in \mathbb{R}^m \text{ and } v \cdot w = 0 \text{ for all } w \in \mathscr{O}(A)\}.$$

This implies that

$$\dim\big(V_0(A)\big) + \dim(\mathrm{span}(\mathscr{O}(A))) = m. \tag{5.5}$$

We now modify A in a simple way to produce a matrix A' which is *nondegenerate* and whose space $\mathscr{O}(A')$ of occupation vectors is simply related to $\mathscr{O}(A)$. This will give us the information we need on the dimension of $V_0(A)$.

Consider again the directed graph G constructed in the proof of Lemma 8, its directed spanning tree G_0, and the states i_p chosen from $D_{h_p} \cap C_{k_p}$, for $p = 1, 2, \dots, r-1$. Now we define A' as follows. Set

$$a'(i_p, i_p) = 1,\ \ p = 1, 2, \dots, r-1 \text{ but set } a'(i, j) = a(i, j) \text{ for all other pairs of indices.} \tag{5.6}$$

Note that $a'(i_p, i_p) = 0$, because $i_p \in D_{h_p}$ but all of its daughters are in D_{k_p}. Thus A' differs from A only in having $r-1$ more positive diagonal entries, in positions i_p, $p = 1, 2, \dots, r-1$. We now prove that A' is nondegenerate. First, any states that are cousins with respect to A are also cousins with respect to A', as $A' \ge A$. But under A' we have some new cousin relations: for $p = 1, 2, \dots, r-1$, the states in D_{h_p} are cousins of the states in D_{k_p} with respect to A', because i_p is an A'-parent of a state in the former, namely i_p itself, by (5.6), and also in the latter as $i_p \in C_{k_p}$. However, since G_0 is a directed spanning tree it follows from this that, for all $h, k = 1, 2, \dots, r$, the states in D_h are A'-cousins of those in D_k. Thus all pairs of states are cousins with respect to A', and so, by Lemma 5 and Theorem 2, A' is nondegenerate, as required.

By (5.6), the only transitions allowed by A' but not A are transitions from a state i_p to itself, for $p = 1, 2, \dots, r-1$. Therefore, each trajectory Π' in $\mathscr{O}(A')$ is got from some Π in $\mathscr{O}(A)$ by a process of inserting visits to the states $i_1, i_2, \dots, i_{r-1}$ at certain points. Thus the occupation time vector of an A'-path differs from that of some A-path only in the time spent in the states $i_1, i_2, \dots, i_{r-1}$, and so

$$\mathrm{span}(\mathscr{O}(A')) \subset \mathrm{span}(\mathscr{O}(A)) \oplus \mathrm{span}\{e_{i_1}, e_{i_2}, \dots, e_{i_{r-1}}\},$$

where e_i is the ith unit vector for $i = 1, 2, \ldots, m$ and $\oplus$ denotes the direct sum of vector spaces. It follows that

$$\dim(\operatorname{span}(\mathscr{O}(A'))) \le \dim(\operatorname{span}(\mathscr{O}(A))) + r - 1. \tag{5.7}$$

Since A' is nondegenerate, $\dim\bigl(V_0(A')\bigr) = 0$ by Theorems 2 and 3. Applying (5.5) in the case of A', it follows that $\dim(\operatorname{span}(\mathscr{O}(A'))) = m$. Thus, upon rearranging (5.7), we get $\dim(\operatorname{span}(\mathscr{O}(A))) \ge m - r + 1$. This implies that $\dim\bigl(V_0(A)\bigr) \le r - 1$ by (5.5) and brings us to the point described at the start of the proof: Lemma 8 now forces equality in (5.4), and we are done. $\square$

Theorem 2 may be viewed as the $r = 1$ case of Theorem 3. But of course it is not a corollary as it played a key role in the proof of Theorem 3. For the next result we introduce the space

$$V(A) \equiv \{\Lambda | \Lambda = \operatorname{diag}(v) \text{ for some } v \in \mathbb{R}^m \text{ and } (A, \Lambda) \text{ is degenerate}\}.$$

COROLLARY 2. *Let A and r be as in Theorem 3. Then $V(A) = \operatorname{span}\{I, \Lambda_1, \Lambda_2, \ldots, \Lambda_r\}$, and the dimension of this space is r.*

PROOF. This is a consequence of the following two observations: (a) I is degenerate but not null, because (4.4) holds for $\Lambda = I$ with $\theta = 1$. Thus $I \in V(A)$ but $I \notin V_0(A)$. (b) For $\Lambda \in (A)$, (4.4) holds for some value θ, and it follows that $\Lambda - \theta I \in V_0(A)$. $\square$

6. Inequalities in discrete time.

Here we address various discrete-time inequalities from the work cited in the Introduction. In particular, we give a condition for strict inequality in Kingman's log-convexity result and we resolve Cohen's conjecture #2.

6.1. *Kingman's log-convexity result.*

The formulation of Kingman's result below is slightly different from the original. The hypothesis (6.1) here corresponds to Kingman's log-convexity condition.

THEOREM 4 (Kingman [11]). *Let $A = (a_{ij})$, $B = (b_{ij})$ and $C = (c_{ij})$ be nonnegative $m \times m$ matrices, with Perron–Frobenius eigenvalues a, b and c, respectively. Let p and q satisfy $0 < p = 1 - q < 1$ and suppose that*

$$b_{ij} \le a_{ij}^p c_{ij}^q, \quad i, j = 1, 2, \ldots, m. \tag{6.1}$$

Then

$$b \le a^p c^q. \tag{6.2}$$

Suppose further that A and C are irreducible. Then there is equality in (6.2) if and only if:

(a) A, B, and C all have the same graph;
(b) There is equality in (6.1) for all i, j; and

(c) the matrix F with entries

$$f_{ij} \equiv \begin{cases} \log(a_{ij}) - \log(c_{ij}), & \text{if } b_{ij} > 0, \\ 0, & \text{if } b_{ij} = 0 \end{cases}$$

is degenerate with respect to B.

PROOF. We suppose first that B itself is irreducible. By (6.1), this implies that A and C are also irreducible, so that a, b, and c are all positive. Let $P = (p_{ij})$ be a stochastic matrix satisfying $p_{ij} = 0$ if and only if $a_{ij} + c_{ij} = 0$. Then P is irreducible by (6.1) and irreducibility of B. Let X be the Markov chain on $\{1, 2, \ldots, m\}$ with transition matrix P. Associate with the matrix A a Markov additive process (X, S_A) with parameters $(P, \bar{A})$, where

$$\bar{A} = (\bar{a}_{ij}), \text{ in which } \bar{a}_{ij} = \begin{cases} \log(a_{ij}) - \log(p_{ij}), & \text{if } p_{ij} \neq 0, \\ 0, & \text{if } p_{ij} = 0. \end{cases}$$

Note that some entries here may be $-\infty$. Analogously, define matrices $\bar{B}$ and $\bar{C}$ and corresponding Markov additive processes (X, S_B) and (X, S_C). By calculation, the fundamental matrices H of (2.3) associated with these three Markov additive processes are precisely the matrices A, B, and C. Condition (6.1) is equivalent to

$$\bar{b}_{ij} \leq p\bar{a}_{ij} + q\bar{c}_{ij}. \tag{6.3}$$

Let T denote the first positive time at which X enters a specified state i. Define $\alpha = \log a$, $\beta = \log b$, and $\gamma = \log c$, which are finite as the corresponding matrices are irreducible. We have

$$\begin{aligned} &\mathbb{E}_i \exp\{S_B(T) - (p\alpha + q\gamma)T\} \\ &\quad \leq \mathbb{E}_i \exp\{p(S_A(T) - \alpha T) + q(S_C(T) - \gamma T)\} \end{aligned} \tag{6.4}$$

$$\quad \leq \mathbb{E}_i (\exp\{S_A(T) - \alpha T\})\mathbb{E}_i (\exp\{S_C(T) - \gamma T\}) = 1, \tag{6.5}$$

using (6.3) in the first inequality, then Holder's inequality, and Lemma 1 at the end. Since, again by Lemma 1, $\mathbb{E}_i (\exp\{S_B(T) - \beta T\}) = 1$, we conclude that $\beta \leq p\alpha + q\gamma$, proving the inequality (6.2) when B is irreducible. The general case of (6.2) follows from this by the limit argument explained in the remark in the last paragraph of the Introduction.

Now consider the case in which both A and C are irreducible. We suppose that we have equality in (6.2), and from this we deduce the conditions (a)–(c) of the theorem.

To prove (a), suppose, seeking a contradiction, that $a_{ij} > 0$ and $b_{ij} = 0$ for some pair of states i, j. Let A' denote the matrix formed by replacing the (i, j)-entry of A by 0, and let $a' \equiv \rho(A')$. Then, by Corollary 2.1.5(b) of [2] and irreducibility of A, we have $a' < a$. Now (6.1) still holds when A is replaced by A', and so $b \leq a'^p c^q$ by the first part of the theorem. Irreducibility of C implies that $c > 0$, and so we have $a^p c^q > a'^p c^q \geq b$, implying strict inequality in (6.2), a contradiction of the assumption of equality. This proves that if we have equality in (6.2) then $a_{ij} = 0$ whenever $b_{ij} = 0$. The same argument shows that, under the same condition, $c_{ij} = 0$ whenever $b_{ij} = 0$. This and (6.1) establish (a).

We continue to assume that there is equality in (6.2). Since we now know that A, B and C have the same graph, the assumption of irreducibility of A implies irreducibility of B, and the argument leading to (6.4) and (6.5) is valid. For the same reason, the matrices $\bar{A}$, $\bar{B}$ and $\bar{C}$ are all finite. To have equality at (6.2), we must have equality at (6.4) and (6.5) by Lemma 1. Equality at (6.4) implies equality at (6.1) for all i and j by irreducibility of X. This establishes condition (b) above. Equality in Hölder's inequality at (6.5) implies, for some constant κ, that

$$S_A(T) - \alpha T = \kappa + S_C(T) - \gamma T \qquad \text{a.s. under the initial condition } X(0) = i. \tag{6.6}$$

We can use Lemma 1 twice again to argue from (6.6) that $1 = \mathbb{E}_i(\exp\{S_A(T) - \alpha T\}) = \mathbb{E}_i(\exp\{\kappa + S_C(T) - \gamma T\}) = e^{\kappa}$, showing that $\kappa = 0$. Using (4.2) and (4.3), (6.6) with $\kappa = 0$ implies that the finite matrix $F \equiv A - C$ is degenerate with respect to P, or, equivalently, with respect to B [as B and P have the same graph by part (a)]. This establishes condition (c). We have shown that (a)–(c) are necessary conditions for equality when A and C are irreducible.

By retracing steps it is easy to verify that, if A and C are irreducible, then (a)–(c) imply equality in (6.4) and (6.5), which implies equality in (6.2). □

Condition (c) of this theorem may be replaced by the following: (c′) *There exists a positive vector h and a positive constant ν such that* $h_i a_{ij} = \nu c_{ij} h_j$, $i, j = 1, 2, \dots, m$. This is a consequence of Lemma 3(b).

6.2. *Cohen's conjecture #2 and some related results.* This subsection concerns convexity properties of the function $\phi(\Lambda) \equiv \log \rho(De^{\Lambda})$ of a diagonal matrix Λ, for D nonnegative and irreducible. These convexity properties are related to the inequalities of Lemmas 5 and 7C of [7]. Theorem 5 below explains the convexity and strict-convexity properties of ϕ in terms of degeneracy. This is a direct consequence of our version of Kingman's result. It is a simple confirmation of the importance of the concept of degeneracy. Theorem 6 gives a little more information, expressing ϕ as a "direct sum" of a linear function on $V(D)$ and a strictly convex function on a subspace W complementary to $V(D)$. In Theorem 7 we resolve Cohen's conjecture #2 by giving necessary and sufficient conditions for the strict convexity property in question. Cohen's conjecture is false as it stands, as it fails to mention the degeneracy condition.

THEOREM 5. *For an irreducible nonnegative matrix D, consider the function*

$$\phi(\Lambda) \equiv \log(\rho(De^{\Lambda})),$$

of a real diagonal matrix argument Λ. Let $0 < p = 1 - q < 1$ and let Λ_1 and Λ_2 denote two diagonal matrices. Then

$$\phi(p\Lambda_1 + q\Lambda_2) \le p\phi(\Lambda_1) + q\phi(\Lambda_2),$$

with equality if and only if $\Lambda_1 - \Lambda_2$ is degenarate with respect to D.

PROOF. Let Λ_1 and Λ_2 be given real diagonal matrices, and set Λ to be their average $p\Lambda_1 + q\Lambda_2$. Define $A \equiv De^{\Lambda_1}$, $B \equiv De^{\Lambda}$ and $C \equiv De^{\Lambda_2}$, and let a, b and c denote their Perron–Frobenius eigenvalues. Then (6.1) holds. and (6.2) immediately gives

$$b \le a^p c^q,$$

from which the inequality of the theorem follows.

Clearly A and C are irreducible. Thus, we have strict inequality if and only if one of the conditions (a), (b) or (c) of Theorem 4 fails to hold. In the present setting, A, B and C have the same graph, and so (a) holds. It is easily verified also that we have equality in (6.1), and so (b) holds also. In the present situation, condition (c) of Theorem 4 is equivalent to the statement that $\Lambda_1 - \Lambda_2$ is degenerate with respect to D. This is because the entries f_{ij} of the matrix F of (c) are $\lambda_1(j) - \lambda_2(j)$ (except those for which $d_{ij} = 0$, but these values are irrelevant to the condition for degeneracy). This completes the proof. □

THEOREM 6. *In the setting of Theorem* 5, ϕ *is linear on* $V(D)$. *Let* W *denote a vector space complementary to* $V(D)$ *in* $\{\operatorname{diag}(v) \mid v \in \mathbb{R}^m\}$. *Then* ϕ *is strictly convex on* W. *Let* π *denote the projection onto* W *parallel to* $V(D)$ *and for* $\Lambda \in \{\operatorname{diag}(v) \mid v \in \mathbb{R}^m\}$ *write* $\Lambda = \Lambda' + \Lambda''$ *where* $\Lambda' \equiv \pi(\Lambda)$. *Then* $\phi(\Lambda) = \phi(\Lambda') + \phi(\Lambda'')$.

PROOF. To avoid some routine extensions of material in Section 4 to general nonnegative matrices, we now restrict ourselves to the case D stochastic. The general case follows as described in the last paragraph of the Introduction.

That ϕ is strictly convex on W is direct consequence of Theorem 5. Consider a fluid model (X, S) with data (D, Λ). Parallel to the decomposition of Λ given in the statement, S may be decomposed as $S = S' + S''$, where (X, S') is the fluid model with data (D, Λ') and (X, S'') is the fluid model with data (D, Λ''). The exponential growth parameters of these fluid models are given by $\alpha = \phi(\Lambda)$, $\alpha' = \phi(\Lambda')$, and $\alpha'' = \phi(\Lambda'')$, by (2.5). Now (D, Λ'') is degenerate as $\Lambda'' \in V(D)$. Corollary 2 gives a natural basis for $V(D)$, namely, $\{I, \Lambda_1, \Lambda_2, \dots, \Lambda_r\}$, where $r = \dim(V(D))$. Noting that each Λ_k is null, the exponential growth parameter $\alpha'' = \phi(\Lambda'')$ may be evaluated with respect to this basis through (4.4) and (4.3) as

$$\alpha'' \equiv \phi(\Lambda'') = x_0 \quad \text{where } \Lambda'' = x_0 I + x_1\Lambda_1 + \cdots + x_r\Lambda_r \in V(D).$$

Thus ϕ is linear on $V(D)$.

Let i be an arbitrary state and let T denote the first positive time at which X enters state i. Using (4.2) and (4.3) (specialized to the case of a fluid model) for the second equality to follow and Lemma 1 for the third, we have

$$\begin{aligned}\mathbb{E}_i\big(\exp\{S(T) - (\alpha' + \alpha'')T\}\big) &= \mathbb{E}_i\big(\exp\{(S'(T) - \alpha' T) + (S''(T) - \alpha'' T)\}\big)\\ &= \mathbb{E}_i\big(\exp\{(S'(T) - \alpha' T)\}\big) = 1.\end{aligned}$$

Bringing Lemma 1 to bear on the first expression here, we conclude that $\alpha = \alpha' + \alpha''$. That is, $\phi(\Lambda) = \phi(\Lambda') + \phi(\Lambda'')$ as required. □

COROLLARY 3. *For an irreducible nonnegative matrix D, the function* $\phi(\Lambda) \equiv \log(\rho(De^{\Lambda}))$ *is strictly convex on the subspace* $W \equiv \{\Lambda \mid \Lambda$ *diagonal and* $\mathrm{Trace}(\Lambda) = 0\}$ *if and only if D is nondegenerate.*

PROOF. If D is nondegenerate then $V(D)$ is the space of scalar matrices by Corollary 2, and so in this case W is complementary to $V(D)$. Theorem 6 gives the required strict convexity. Conversely, if D is degenerate then $V(D)$ is of dimension greater than 1, and so must have a nontrivial intersection with the $m-1$-dimensional space W. As ϕ is linear on $V(D)$ by Theorem 6, this implies that it is not strictly convex on W. $\square$

EXAMPLE 3. This is a continuation of Example 2 in Section 5. For the nonnegative matrix A described there, the Perron–Frobenius eigenvalue of Ae^{Λ}, for $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, is, by direct calculation,

$$\rho(Ae^{\Lambda}) = \sqrt{\frac{\exp(\lambda_1+\lambda_4) + \sqrt{\exp(2(\lambda_1+\lambda_4)) + 4\exp(\lambda_1+\lambda_2+\lambda_3+\lambda_4)}}{2}},$$

so that

$$\phi(\Lambda) \equiv \log\ \rho(Ae^{\Lambda}) = \frac{1}{2}\log\left(\frac{1+\sqrt{1+4\exp(-\lambda_1+\lambda_2+\lambda_3-\lambda_4)}}{2}\right) + \frac{\lambda_1+\lambda_4}{2}.$$

This exhibits the decomposition of ϕ into the sum of a strictly convex function on a subspace of dimension $m - r = 1$ and a linear function, as described in Theorem 6.

THEOREM 7 (Cohen's conjecture #2 [6]). *Suppose that* $A_1, A_2, \dots, A_K$ *are nonnegative matrices, and suppose also that the cyclic products* $A_1A_2\cdots A_K$, $A_2A_3\cdots A_1, \dots, A_KA_1\cdots A_{K-1}$ *are irreducible. Then the function*

$$\phi(\mathrm{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_K)) \equiv \log(\rho(A_1e^{\Lambda_1}A_2e^{\Lambda_2}\cdots A_Ke^{\Lambda_K})),$$

of a real diagonal $Km \times Km$ *matrix, is strictly convex on the domain*

$$\mathscr{D} \equiv \{\mathrm{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_k) \mid \Lambda_k \text{ diagonal and } \mathrm{Trace}(\Lambda_k) = 0 \quad \text{for each } k = 1, 2, \dots, K\}$$

if and only if each $A_k^TA_k$, $k = 1, 2, \dots, K$, *is irreducible.*

PROOF. Set

$$A \equiv \begin{pmatrix} 0 & A_1 & 0 & \cdots & 0 \\ 0 & 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_K & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Then $\phi(\Lambda) = \log(\rho(Ae^{\Lambda}))$, where $\Lambda = \mathrm{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_K)$, and the irreducibility assumptions of the theorem are equivalent to irreducibility of A. We write the indices of the matrix A as (k, j), $k = 1, 2, \dots, K$ and

$j = 1, 2, \ldots, m$. The set of indices $L_k \equiv \{(k, j) \mid j = 1, 2, \ldots, m\}$ is referred to as *level* k.

By Theorem 6, ϕ is strictly convex on the domain $\mathscr{D}$ if and only if the intersection of that domain with the space $V(A)$ is $\{0\}$. We compute the dimension of $V(A)$ using Theorem 3. We have

$$A^T A = \operatorname{diag}(A_K^T A_K, A_1^T A_1, A_2^T A_2, \ldots, A_{K-1}^T A_{K-1}). \tag{6.7}$$

The block-diagonal form here implies that the classes of $A^T A$ are subsets of the levels $L_1, L_2, \ldots, L_K$. Let $r \equiv \dim(V(A))$ and $r_k \equiv \dim(V(A_k))$ for $k = 1, 2, \ldots, K$. Then (6.7) and Corollary 2 imply that

$$r = \sum_{k=1}^{K} r_k \geq K. \tag{6.8}$$

The inequality here is because the number of classes under the cousin relation is at least one for any nonnegative matrix. We have

$$V(A) \supset \{\operatorname{diag}(x_1 I, x_2 I, \ldots, x_K I) \mid x_1, x_2, \ldots, x_K \in \mathbb{R}\}, \tag{6.9}$$

as each element of the space on the right is easily seen to be degenerate with respect to A. Now if each $A_k^T A_k$, $k = 1, 2, \ldots, K$, is irreducible, then each r_k is 1 and we have equality in (6.8), but this forces equality in (6.9), as then the dimensions of the two spaces are equal. In this case the space $\mathscr{D}$ is complementary to $V(A)$, and so it follows from Theorem 6 that ϕ is strictly convex on $\mathscr{D}$, as required.

Conversely, if the $A_k^T A_k$'s were not all irreducible, then the dimension r of $V(A)$ would exceed K by (6.8), and so it would have a nontrivial intersection with the $K(m-1)$-dimensional space $\mathscr{D}$. Theorem 6 would then imply that ϕ is not strictly convex on $\mathscr{D}$. $\square$

7. Cohen's conjectures #8 and #9.

Cohen's eighth and ninth conjectures are addressed in the two propositions below. In the first of these, we break our habit of considering only the irreducible case. For the proofs we need the following properties of collections of commuting essentially positive $k \times k$ matrices. (For general results on commuting matrices, see [19].) Here, I_k is the order-k identity matrix.

LEMMA 9. *Let $F(i)$, $i = 1, 2, \ldots, m$, be commuting essentially positive matrices. Let $\mathscr{F} \equiv F(1) + F(2) + \cdots + F(m)$ and let C_l, $l = 1, 2, \ldots, L$, denote the classes of $\mathscr{F}$. Denote by $\mathscr{F}_l$ [respectively, $F_l(i)$] the square submatrix of $\mathscr{F}$ [respectively, $F(i)$] corresponding to class C_l. Let π_l denote the Perron–Frobenius eigenvector of the (irreducible essentially positive) matrix $\mathscr{F}_l$, scaled so that the entries sum to 1. Then:*

(a) *The classes of $F(i)$ are subsets of the classes of $\mathscr{F}$ for $i = 1, 2, \ldots, m$.*

(b) *The vector π_l is an eigenvector of each $F_l(i)$ $i = 1, 2, \ldots, m$.*

(c) *The eigenvalue ρ_{il} of $F_l(i)$ corresponding to the eigenvector π_l is the Perron–Frobenius eigenvalue of $F_l(i)$.*

PROOF. (a) This is because if two states communicate in $F(i)$ for any i, then they also communicate in $\mathscr{F}$.

(b) If $F_l(i) = 0$ the statement is obvious. Suppose $F_l(i) \neq 0$. Commutativity of $F(i)$ and $\mathscr{F}$ implies that each $F_l(i)$ commutes with $\mathscr{F}_l$. Thus

$$(\pi_l F_l(i))\mathscr{F}_l = (\pi_l \mathscr{F}_l) F_l(i) = \rho_l(\pi_l F_l(i)),$$

where ρ_l is the Perron–Frobenius eigenvalue of $\mathscr{F}_l$. Therefore $\pi_l F_l(i)$ is a nonnegative eigenvector of $\mathscr{F}_l$, and is nonzero as $F_l(i) \neq 0$ and $\pi_l > 0$. But, since $\mathscr{F}_l$ is irreducible, $\pi_l F_l(i)$ must then be a multiple of π_l and we get

$$\pi_l F_l(i) = \rho_{il}\pi_l$$

for a scalar ρ_{il}.

(c) The fact that ρ_{il} is the Perron–Frobenius eigenvalue of $F_l(i)$ is because all the components of π_l are positive, being the Perron–Frobenius eigenvector of the irreducible matrix $\mathscr{F}_l$. $\square$

PROPOSITION 4 (Cohen's conjecture #8). *With the notation and hypotheses of Lemma 9, let* $\Lambda_l = \operatorname{diag}(\rho_{1l}, \rho_{2l}, \dots, \rho_{ml})$ *for* $l = 1, 2, \dots, L$. *Let* Q *be an irreducible* $m \times m$ *generator. Then*

$$\rho(\exp(Q \otimes I_k + \Lambda)) \leq \rho(\exp(Q \otimes I_k)e^{\Lambda}) \qquad \textit{where } \Lambda \equiv \operatorname{diag}(F(1), F(2), \dots, F(m)),$$

with equality if and only if, for at least one class l *for which* $\rho(e^Q e^{\Lambda_l})$ *attains its maximum over* $l = 1, 2, \dots, k$, Λ_l *is scalar.*

PROOF. We calculate the Perron–Frobenius eigenvalues in question. Throughout, we use the notation of Lemma 9. Let π_l^* denote the row k-vector which has zeros everywhere except for indices $i \in C_l$, where it has the appropriate entries of π_l. Let e_l^* be the column k-vector which has zeros everywhere except for indices $i \in C_l$, where it has 1's. The following may be proved by straightforward manipulations:

$$(I_m \otimes \pi_l^*)\exp(t(Q \otimes I_k + \Lambda))(I_m \otimes e_l^*) = \exp(t(Q + \Lambda_l)), \tag{7.1}$$

$$(I_m \otimes \pi_l^*)(\exp(Q \otimes I_k)e^{\Lambda})^n(I_m \otimes e_l^*) = (e^Q e^{\Lambda_l})^n. \tag{7.2}$$

From (7.1) it follows upon letting $t \to \infty$ that the spectral radius of $\exp(Q \otimes I_k + \Lambda)$ is the maximum over l of $\rho(\exp(Q + \Lambda_l))$. This is because the components of the vectors $\sum_{l=1}^{L} \pi_l^*$ and $\sum_{l=1}^{L} e_l^*$ are all positive. Similarly, from (7.2) it follows that the spectral radius of $\exp(Q \otimes I_k)e^{\Lambda}$ is the maximum over l of $\rho(e^Q e^{\Lambda_l})$. Relating these using Theorem 1 for inequality we have

$$\begin{aligned}\rho(\exp(Q \otimes I_k + \Lambda)) &= \max_l \rho(\exp(Q + \Lambda_l)) \leq \max_l \rho(e^Q e^{\Lambda_l})\\ &= \rho(\exp(Q \otimes I_k)e^{\Lambda}).\end{aligned}$$

The condition for strict inequality also follows directly from Theorem 1. $\square$

Cohen conjectured strict inequality under the sole condition of irreducibility of the $F(i)$'s, but this is not enough; even when irreducibility holds, we must further insist that the Perron–Frobenius eigenvalues of the $F(i)$'s not be all equal. An entirely parallel argument leads to the following version of Cohen's conjecture #9. We do not give the proof. We also assume irreducibility as Cohen does, avoiding the somewhat cumbersome generality of Proposition 4.

PROPOSITION 5 (Cohen's conjecture #9). *Let $A(i), i = 1, 2, \ldots, m$, be $k \times k$ commuting irreducible nonnegative matrices. Let P be an irreducible stochastic matrix. Then*

$$\rho((P \otimes I_k)A)^2 \leq \rho((P \otimes I_k)^2 A^2) \quad \textit{where } A \equiv \operatorname{diag}(A(1), A(2), \ldots, A(m)),$$

with equality if and only if $(P, \Lambda_\bullet)$ is degenerate, where $\Lambda_\bullet = \operatorname{diag}(\alpha_1, \alpha_2, \ldots, \alpha_m)$ and $\alpha_i = \log \rho(A(i)), i = 1, 2, \ldots, m$.

Irreducibility of a nonnegative matrix implies that its Perron–Frobenius eigenvalue is positive, and so $\Lambda_\bullet$ is a finite matrix here.

8. Concluding remarks. In this paper, we have limited our scope in several ways for brevity. We have not tried to treat the reducible case systematically, and in results involving diagonal matrices we have considered only invertibles ones (of the form e^Λ with Λ finite). There are many results in the literature cited, beyond those treated above, which can be more fully understood in the light of the present work.

Much of the paper concerns fluid models, reflecting the emphasis in the existing literature, but there are other inequalities associated with fundamental matrices of more general Markov additive processes that may be interest. An example is Proposition 3, which places a simple lower bound on the Perron–Frobenius eigenvalue of a matrix of the form $P \odot A$ for P irreducible and A nonnegative. The key technical fact contained in Lemmas 1 and 2 extends readily to more general Markov additive processes having a suitable regenerative structure [15], and the techniques of this paper give a natural approach to inequalities in this larger context.

REFERENCES

[1] ASMUSSEN, S. and O'CINNEIDE, C. A. (1998). On the tail of the waiting time in the Markov-modulated $M/G/1$ queue. Unpublished mansucript.

[2] BERMAN, A. and PLEMMONS, R. J. (1994). *Nonnegative Matrices in the Mathematical Sciences*. SIAM, Philadelphia.

[3] BONDY, J. A. and MURTY, U. S. R. (1976). *Graph Theory with Applications*. North-Holland, Amsterdam.

[4] CINLAR, E. (1972). Markov Additive Processes. I, II. *Z. Wahrsch. Verw. Gebiete.* **24** 85–121.

[5] COHEN, J. E. (1979). Random evolutions and the radius of a nonnegative matrix. *Proc. Cambridge Philos. Soc.* **86** 345–350.

[6] COHEN, J. E. (1984). Eigenvalue inequalities for random evolutions: orgins and open problems. In *Inequalities in Statistices and Probability.* 41–53. IMS, Hayward, CA.

[7] COHEN, J. E. FRIEDLAND, S. KATO, T. and KELLY, F. P. (1982). Eigenvalue inequalities for products of matrix exponentials. *Linear Algebra Appl.* **45** 55–95.

[8] FRIELDLAND, S. and KARLIN, S. (1975). Some inequalities for the spectral radius of nonnegative matrices and applications. *Duke Math. J.* **43** 459–490.

[9] HARDY, G. H. LITTLEWOOD, J. E and PÒLYA, G. (1952). *Inequalities*, 2nd ed. Cambridge Univ. Press.

[10] KATO, T. (1976). *Perturbation Theory for Linerar Operators*, 2nd ed. Springer, New York.

[11] KINGMAN, J. F. C. (1961). A convexity property of positive matrices. *Oxford Quaterly J. Math.* **12** 283–284.

[12] LIEB, E. H. (1990). On the spectral radius of the product of matrix exponentials. *Linear Algebra Appl.* **141** 271–273.

[13] LOVSÁZ, L. (1993). *Combinatorial Problems and Exercises*. North-Holland, Amsterdam.

[14] MARSHAL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

[15] NEY, P. and NUMMELIN, E. (1987). Markov additive processes I. Eigenvalue properties and limit theorems. *Ann. Probab.* **15** 561–592.

[16] NEY, P. and NUMMELIN, E. (1987). Markov additive processes II. Large deviations. *Ann. Probab.* **15** 593–609.

[17] SENETA, E. (1981). *Nonnegative Matrices and Markov Chains.* Springer, New York.

[18] SHAKED, M. and SHANTHIKUMAR, J. G. (1994). *Stochastic Orders and Their Applications*. Academic Press, San Diego.

[19] SUPRUNENKO, D. A. and TYSHKEVICH, R. I. (1968). *Commutative Matrices*. Academic Press, New York.

SCHOOL OF INDUSTRIAL ENGINEERING
PURDUE UNIVERSITY
1287 GRISSOM HALL
WEST LAFAYETTE, INDIANA 47907-1287
E-MAIL: Colm@ecn.purdue.edu