Necessary and sufficient conditions for a nonnegative matrix to be strongly R-positive

Jan M. Swart*

September 27, 2017

Abstract

Using the Perron-Frobenius eigenfunction and eigenvalue, each finite irreducible nonnegative matrix A can be transformed into a probability kernel P. This was generalized by David Vere-Jones who gave necessary and sufficient conditions for a countably infinite irreducible nonnegative matrix A to be transformable into a recurrent probability kernel P, and showed uniqueness of P. Such A are called R-recurrent. Let us say that A is strongly R-positive if the return times of the Markov chain with kernel P have exponential moments of some positive order. Then we prove that strong R-positivity is equivalent to the property that lowering the value of finitely many entries of A lowers the spectral radius. This condition is more robust than the condition of Vere-Jones and can often be checked even in cases where the spectral radius is not known explicitly. We also prove a complementary characterization of R-transience.

MSC 2010. Primary: 60J45. Secondary: 60J10.
Keywords. Strong R-positivity, R-recurrence, geometric ergodicity, Perron-Frobenius theorem, nonnegative matrix, spectral radius, pinning model.
Acknowledgement. Work sponsored by grant 16-15238S of the Czech Science Foundation (GA CR).

1 Introduction and main results

1.1 R-recurrence

A nonnegative matrix $A = (A(x,y))_{x,y\in S}$ indexed by a countable set S is called *irreducible* if for each $x, y \in S$ there exists an $n \geq 1$ such that $A^n(x,y) > 0$; it is moreover *aperiodic* if the greatest common divisor of $\{n \geq 1 : A^n(x,x) > 0\}$ is one for some, and hence for all $x \in S$. The classical Perron-Frobenius theorem [Per07, Fro12] says that if A is an irreducible nonnegative matrix indexed by a finite set S, then it has a unique positive eigenfunction. More precisely, there exists a function $h : S \to (0,\infty)$, which is unique up to scalar multiples, and a unique constant $c > 0$, such that $Ah = ch$. The function h is called the *Perron-Frobenius eigenfunction* and c the *Perron-Frobenius eigenvalue*. We will be interested in generalizations of this theorem to countably infinite matrices.

Let A be an aperiodic irreducible nonnegative matrix indexed by a countable set S. A simple argument based on superadditivity [Kin63], shows that the limit

$$\rho(A) := \lim_{n\to\infty} \big(A^n(x,x)\big)^{1/n} \tag{1.1}$$

*Institute of Information Theory and Automation of the CAS (ÚTIA), Pod vodárenskou věží 4, 18208 Praha 8, Czech Republic; swart@utia.cas.cz

exists in $(0,\infty]$ and does not depend on $x \in S$. If A is periodic, then $\rho(A)$ is defined in the same way except that in (1.1) n ranges only through those integers for which $A^n(x,x) > 0$. Because of its interpretation in the finite case, the quantity $\rho(A)$ is called the *spectral radius* of A. By definition, A is called *R-recurrent*[1] if $\rho(A) < \infty$ and

$$\sum_{n=1}^{\infty} \rho(A)^{-n} A^n(x,x) = \infty \tag{1.2}$$

for some, and hence for all $x \in S$. We observe that a function $h : S \to (0,\infty)$ is an eigenfunction of A with eigenvalue $c > 0$ if and only if

$$P(x,y) := c^{-1} h(x)^{-1} A(x,y) h(y) \qquad (x,y \in S) \tag{1.3}$$

defines a probability kernel on S. As will be shown in Appendix A.1 below, the following theorem follows easily from the work of Vere-Jones [Ver62, Ver67].

Theorem 1 (R-recurrent matrices) *Let A be an R-recurrent irreducible nonnegative matrix indexed by a countable set S. Then there exists a function $h : S \to (0,\infty)$, which is unique up to scalar multiples, and a unique constant $c > 0$, such that (1.3) defines a recurrent probability kernel P. Moreover, $c = \rho(A)$.*

Since finite matrices are R-recurrent (this is proved in [Ver67, Sect. 7] and will also follow from our Theorem 4 below), the classical Perron-Frobenius theorem is implied by Theorem 1. If S is finite, then $\rho(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\}$, where $\sigma(A)$ denotes the set of all complex eigenvalues of A; by contrast, if S is infinite, then it often happens that A has positive eigenfunctions with eigenvalues $c > \rho(A)$ [Ver63]. For such eigenfunctions, the probability kernel in (1.3) is transient. Note that in view of this, the term "spectral radius" for $\rho(A)$ is somewhat of a misnomer if A is infinite, but we retain it for historical reasons. The following theorem, first proved in [Ver67, Thm 4.1], shows that for R-recurrent matrices, there is only one positive eigenfunction associated with the eigenvalue $\rho(A)$, and all other positive eigenfunctions (if there are any) have eigenvalues $c > \rho(A)$.

Theorem 2 (Positive eigenfunction) *Let A be an R-recurrent irreducible nonnegative matrix indexed by a countable set S. Then there exists a function $h : S \to (0,\infty)$, which is unique up to scalar multiples, such that such that $Ah = \rho(A)h$. Moreover, if some function $f : S \to [0,\infty)$ satisfies $Af \le \rho(A)f$, then $f = \lambda h$ for some $\lambda \ge 0$.*

It should be noted that the approach based on R-recurrence is just one of many different ways to generalize the Perron-Frobenius theorem to infinite dimensions. For a more functional analytic approach, see, e.g., [KR48, KR50, Kar59, Sch74, Zer87]. The theory of R-recurrence is treated in the books [Sen81, Woe00] and generalized to uncountable spaces in [Num84].

In view of Theorems 1 and 2, it is clearly very useful to know of a given nonnegative matrix that it is R-recurrent. Unfortunately, it is often not feasible to check this directly from the definition (1.2), since this requires rather subtle knowledge about the asymptotics of the powers of A, while often it is not even possible to obtain the spectral radius $\rho(A)$ in closed form. In the next section, we will show that for a subclass of the R-recurrent matrices, more robust methods are available.

1.2 Strong R-positivity

Let $X = (X_k)_{k\ge 0}$ be a Markov chain with countable state space S and transition kernel P, and let $\sigma_x := \inf\{k > 0 : X_k = x\}$ denote its first return time to a point $x \in S$. Let $\mathbb{P}^x$

[1] Originally, the letter R was mathematical notation for $1/\rho(A)$. For us the 'R' in the words R-recurrence, R-positivity etc. will just be part of the name and not refer to any mathematical constant.

denote the law of X started in $X_0 = x$ and let $\mathbb{E}^x$ denote expectation with respect to $\mathbb{P}^x$. Recall that by definition, x is recurrent if $P^x[\sigma_x < \infty] = 1$ and x is positive recurrent if $\mathbb{E}^x[\sigma_x] < \infty$. We will say that x is *strongly positive recurrent*[2] if $\mathbb{E}^x[e^{\varepsilon\sigma_x}] < \infty$ for some $\varepsilon > 0$. It is well-known that recurrence and positive recurrence are class properties. Kendall [Ken59] proved that the same is true for strong positive recurrence, i.e., if P is irreducible, then $\{x \in S : x \text{ is strongly positive recurrent}\}$ is either S or $\emptyset$. For aperiodic chains, he moreover proved that strong positive recurrence is equivalent to *geometric ergodicity*, in the following sense. In Kendall's formulation, the constant ε in point (iii) was allowed to depend on x, y. Vere-Jones [Ver62] showed that it can be chosen uniformly.

Proposition 3 (Geometric ergodicity) *Let P be the transition kernel of an irreducible, aperiodic, positive recurrent Markov chain with countable state space S, and let π denote its invariant law. Then the following statements are equivalent.*

(i) *P is strongly positive recurrent.*

(ii) *There exist $x \in S$, $\varepsilon > 0$, and $M < \infty$ such that $|P^n(x,x) - \pi(x)| \leq Me^{-\varepsilon n}$ for all $n \geq 0$.*

(iii) *There exist $\varepsilon > 0$ and $M_{x,y} < \infty$ such that $|P^n(x,y) - \pi(y)| \leq M_{x,y}e^{-\varepsilon n}$ for all $n \geq 0$ and $x, y \in S$.*

An R-recurrent irreducible nonnegative matrix A is called *R-positive* if the unique recurrent probability kernel P from Theorem 1 is positive recurrent. We will say that A is *strongly R-positive* if P is strongly positive recurrent. (This is called *geometrically R-recurrent* in [Num84].) An irreducible nonnegative matrix A is called *R-transient* if it is not R-recurrent and *R-null recurrent* if it is R-recurrent but not R-positive. We will say that A is *weakly R-positive* if it is R-positive but not strongly so. The main result of the present paper is the following theorem, that gives necessary and sufficient conditions for strong R-positivity.

Theorem 4 (Strong R-positivity) *Let A be an irreducible nonnegative matrix indexed by a countable set S and assume that $\rho(A) < \infty$. Let $B \leq A$ be another nonnegative matrix such that $B(x,y) > 0$ if and only if $A(x,y) > 0$ $(x, y \in S)$, and $B \neq A$. Then:*

(a) *If A is strongly R-positive, then $\rho(B) < \rho(A)$.*

(b) *If $\rho(B) < \rho(A)$ and the set $\{(x,y) \in S^2 : A(x,y) \neq B(x,y)\}$ is finite, then A is strongly R-positive.*

Theorem 4 says that a nonnegative matrix is strongly R-positive if and only if lowering the value of finitely many entries lowers the spectral radius. In view of this, to prove strong R-positivity, it suffices to prove sufficiently sharp upper and lower bounds on the spectral radii of two nonnegative matrices, which is in general much easier than determining the exact asymptotics as in (1.2).

For R-transience, a complementary statement holds. The following theorem says that a nonnegative matrix is R-transitive if and only if it is possible to increase the value of finitely many entries without increasing the spectral radius.

Theorem 5 (R-transience) *Let $A \leq B$ be irreducible nonnegative matrices indexed by a countable set S and assume that $B \neq A$ and $\rho(B) < \infty$. Then:*

(a) *If A is R-transient and $\{(x,y) \in S^2 : A(x,y) \neq B(x,y)\}$ is finite, then $\rho\big(A + \varepsilon(B - A)\big) = \rho(A)$ for some $\varepsilon > 0$.*

(b) *If $\rho(B) = \rho(A)$, then A is R-transient.*

[2] This should be distinguished from the closely related, but different concepts of *strong ergodicity* and *strong recurrence*, the latter having been defined in [Spi90].

1.3 Discussion

One possible application of Theorem 4 in the study of quasi-stationary laws. The usefulness of R-positivity in the study of quasi-stationary laws has been noticed long ago [SV66, Thm 3.2]; see also [And91, Prop 5.2.10 en 5.2.11] for the continuous-time case.

Another application of Theorems 4 and 5 is in the study of pinning models. In fact, using these theorems, it is easy to prove that for pinning models in the localized regime, return times have exponential moments of some positive order. Moreover, at the critical point separating the localized and delocalized regimes, the model is either null recurrent or weakly positive recurrent. These facts have been noticed before, see [CGZ06, Thm 4.1 and Prop. 4.2] and [Gia07, Thm 2.3]. Note that these references do not explicitly discuss exponential moments of return times; nevertheless, the claims follow from their formulas.

As demonstrated by the following corollaries, Theorem 4 can also be used to prove strong positive recurrence (which, by Proposition 3, in the aperiodic case is equivalent to geometric ergodicity).

Corollary 6 (Conditions for strong positive recurrence) *Let P be an irreducible probability kernel P on a countable set S and let $Q \leq P$ be a subprobability kernel such that $Q(x,y) > 0$ if and only if $P(x,y) > 0$, and $\{(x,y) \in S^2 : Q(x,y) < P(x,y)\}$ is finite. Then P is strongly positive recurrent if and only if $\rho(Q) < \rho(P) = 1$.*

Proof It is easy to see that a recurrent probability kernel P satisfies $\rho(P) = 1$. Moreover, a probability kernel that is strongly positive recurrent is clearly strongly R-positive. In view of this, the necessity of the conditions $\rho(Q) < \rho(P) = 1$ follows from Theorem4 (a).

Conversely, if $\rho(Q) < \rho(P)$, then Theorem4 (b) shows that P is strongly R-positive, i.e., there exists a strongly positive recurrent probability kernel P' and a function $h : S \to (0, \infty)$ such that $P'(x,y) = \rho(P)^{-1}h(x)^{-1}P(x,y)h(y)$. It follows that $Ph = \rho(P)h$. Since $\rho(P) = 1$, the constant function $h \equiv 1$ also solves this equation. Since by Theorem 2, solutions of $Ph = \rho(P)h$ are up to a multiplicative constant unique, we conclude that $P = P'$. ∎

Corollary 7 (Conditions for strong positive recurrence) *Let P be an irreducible probability kernel P on a countable set S and let $S' \subset S$ be finite. Then P is strongly positive recurrent if and only if the following conditions is satisfied:*

(i) $\rho(P) = 1$

(ii) *There exists a function $f : S \to (0, \infty)$ and $0 < \varepsilon < 1$ such that $Pf(x) < \infty$ for all $x \in S'$ and $Pf(x) \leq (1 - \varepsilon)f(x)$ for all $x \in S \backslash S'$.*

Proof If (ii) holds, then we can construct a subprobability kernel Q with $Q(x,y) > 0$ if and only if $P(x,y) > 0$ and $Q(x,y) = P(x,y)$ for all $x \in S\backslash S'$, such that $\{(x,y) \in S^2 : Q(x,y) < P(x,y)\}$ is finite and

$$\sum_y Q(x,y)f(y) \leq (1 - \varepsilon)f(x) \qquad (x \in S'). \tag{1.4}$$

Then $Qf \leq (1 - \varepsilon)f$, which is easily seen to imply $\rho(Q) \leq 1 - \varepsilon$. Together with condition (i), by Corollary 6, this implies the strong positive recurrence of P.

Conversely, if P is strongly positive recurrent, pick some $x_0 \in S'$ and $y_0 \in S$ with $P(x_0, y_0) > 0$ and define $Q(x_0, y_0) := \frac{1}{2}P(x_0, y_0)$ and $Q(x,y) := P(x,y)$ for all $(x,y) \neq (x_0, y_0)$. By Corollary 6, $\rho(Q) < 1$, so by Lemma 13 below there exists a function $f : S \to (0, \infty)$ such that $Qf \leq \rho(Q)f$ and hence $Pf(x) \leq \rho(Q)f(x)$ for all $x \neq x_0$, while $Pf(x_0) \leq 2Qf(x_0) < \infty$. ∎

Corollary 7 is similar to a result of Popov [Pop77], who proved that the function f in condition (ii) can be chosen such that $f \geq 1$, and with this extra condition on f, condition (i) can be dropped.

For a given nonnegative matrix A, let $\rho_\infty(A)$ denote the infimum of all possible values of $\rho(B)$ where $B \leq A$ satisfies $B(x,y) > 0$ if and only if $A(x,y) > 0$ and $\{(x,y) : A(x,y) \neq B(x,y)\}$ is finite. Then Theorem 4 implies that A is stongly R-positive if and only if $\rho_\infty(A) < \rho(A)$. We can describe this condition in words by saying that under a Gibbs measure with transfer matrix A, paths far from the origin carry less mass, on an exponential scale, than paths near the origin. The quantity $\rho_\infty(A)$ has been studied in [MS95, Ign06]. In the latter paper, it is called the *essential spectral radius*.

The rest of the paper is dedicated to proofs. The methods are fairly elementary. One can easily imagine Theorems 4 and 5 having been proved 40 or more years ago, using basically the same proofs. Nevertheless, as far as I have been able to find out, this is not the case. The language of the proofs is somewhat more modern. A result like Lemma 12 below, for example, would in the old days have been proved by manipulation with functions that live on the space S, while here we take the "pathwise" approach, using the function $\mathcal{A}$ that is defined on walks.

The organition of the proofs is as follows. Section 2.1 contains some preliminary definitions and lemmas. Section 2.2 gives a characterization of forms of R-recurrence in terms of a logarithmic moment generating function. Using this, Theorems 4 and 5 are then proved in Sections 2.3 and 2.4, respectively. In Appendix A.1, it is explained how Theorem 1 follows from the work of Vere-Jones. Appendix A.2 contains some general facts about logarithmic moment generating functions.

2 Proofs

2.1 Excursions away from subgraphs

Given a nonnegative matrix A indexed by a countable set S, we define a directed graph $G = (S,E)$ with vertex set S and set of directed edges E given by $E := \{(x,y) \in S^2 : A(x,y) > 0\}$. Alternatively, we denote an edge by $e = (x,y)$ and call $e^- := x$ and $e^+ := y$ its *starting vertex* and *endvertex*, respectively. A *walk* in G is a function $\omega : \{0,\dots,n\} \to S$ with $n \geq 0$ such that

$$\vec{\omega}_k := (\omega_{k-1},\omega_k) \in E \qquad (1 \leq k \leq n). \tag{2.1}$$

We call $\ell_\omega := n \geq 0$ the *length* of ω and we call $\omega^- := \omega_0$ and $\omega^+ := \omega_n$ its *starting vertex* and *endvertex*. We can, and sometimes will, naturally identify walks of length zero and one with vertices and edges, respectively. We let $\Omega = \Omega(G)$ denote the space of all walks in G and write

$$\Omega^n := \{\omega \in \Omega : \ell_\omega = n\} \quad \text{and} \quad \Omega_{x,y} := \{\omega \in \Omega : \omega^- = x,\ \omega^+ = y\} \tag{2.2}$$

and $\Omega^n_{x,y} := \Omega^n \cap \Omega_{x,y}$. We observe that

$$A^n(x,y) = \sum_{\omega \in \Omega^n_{x,y}} \mathcal{A}(\omega) \quad \text{with} \quad \mathcal{A}(\omega) := \prod_{k=1}^{\ell_\omega} A(\omega_{k-1},\omega_k). \tag{2.3}$$

This formula also holds for $n = 0$ provided we define the empty product as $:= 1$.

If $S' \subset S$ is a subset of vertices, then an *excursion away from* S' is a walk $\omega \in \Omega$ of length $\ell_\omega \geq 1$ such that $\omega^\pm \in S'$ and $\omega_k \notin S'$ for all $0 < k < \ell_\omega$. We denote the set of all excursions away from S' by $\widehat{\Omega}(S')$. We sometimes view a graph as the disjoint union of its vertex and edge sets, $G = S \cup E$. A *subgraph* of G is then a set $F \subset G$ such that $e^\pm \in S \cap F$ for all $e \in E \cap F$. Extending our earlier definition, an *excursion away from* F is an element $\omega \in \widehat{\Omega}(F \cap S)$ such that moreover $\omega \notin F \cap E$, where we naturally identify edges with walks of length one. We denote the set of all excursions away from F by $\widehat{\Omega}(F)$ and write $\widehat{\Omega}_{x,y}(F) := \widehat{\Omega}(F) \cap \Omega_{x,y}$, $\widehat{\Omega}^n(F) := \widehat{\Omega}(F) \cap \Omega^n$, etc.

For each subgraph F of G and $x, y \in S \cap F$, we define a *moment generating function* $\phi^F_{x,y}$ and *logarithmic moment generating function* $\psi^F_{x,y}$ by

$$\phi^F_{x,y}(\lambda) := \sum_{\omega \in \widehat{\Omega}_{x,y}(F)} e^{\lambda \ell_\omega} \mathcal{A}(\omega) \quad \text{and} \quad \psi^F_{x,y}(\lambda) := \log \phi^F_{x,y}(\lambda) \qquad (\lambda \in \mathbb{R}). \tag{2.4}$$

Here $\phi^F_{x,y}$ and $\psi^F_{x,y}$ may be ∞ for some values of λ; in addition, $\psi^F_{x,y}(\lambda) := -\infty$ if $\phi^F_{x,y}(\lambda) = 0$. The following lemma lists some elementary properties of $\psi^F_{x,y}$.

Lemma 8 (Logarithmic moment generating functions) *Assume that A is irreducible and $\rho(A) < \infty$. Let F be a subgraph of G, let $x, y \in S \cap F$, and set*

$$\begin{aligned} \lambda_+ = \lambda^F_{x,y,+} &:= \sup\{\lambda \in \mathbb{R} : \psi^F_{x,y}(\lambda) < \infty\}, \\ \lambda_* = \lambda^F_{x,y,*} &:= \sup\{\lambda \in \mathbb{R} : \psi^F_{x,y}(\lambda) < 0\}. \end{aligned} \tag{2.5}$$

Then either $\psi^F_{x,y} \equiv -\infty$ or:

(i) $\psi^F_{x,y}$ *is convex.*

(ii) $\psi^F_{x,y}$ *is lower semi-continuous.*

(iii) $-\infty < \lambda_* < \infty$ *and* $\lambda_* \leq \lambda_+ \leq \infty$.

(iv) $\psi^F_{x,y}$ *is infinitely differentiable on* $(-\infty, \lambda_+)$.

(v) $\psi^F_{x,y}$ *is strictly increasing on* $(-\infty, \lambda_+)$.

(vi) $\lim_{\lambda \to \pm\infty} \psi^F_{x,y}(\lambda) = \pm\infty$.

Proof If $\mathcal{A}(\omega) = 0$ for all $\omega \in \widehat{\Omega}_{x,y}(F)$, then $\psi^F_{x,y} \equiv -\infty$, while otherwise $\psi^F_{x,y}(\lambda) > -\infty$ for all $\lambda \in \mathbb{R}$. Clearly, $\psi^F_{x,y}(\lambda)$ is nondecreasing as a function of λ. Since

$$\phi^F_{x,y}(\lambda) = \sum_{k=1}^{\infty} \sum_{\omega \in \widehat{\Omega}^k_{x,y}(F)} e^{\lambda k} \mathcal{A}(\omega) \leq \sum_{k=0}^{\infty} \sum_{\omega \in \Omega^k_{x,y}} e^{\lambda k} \mathcal{A}(\omega) = \sum_{k=0}^{\infty} e^{\lambda k} A^k(x,y), \tag{2.6}$$

which by (1.1) is finite for $\lambda < -\log \rho(A)$, we see that $-\infty < -\log \rho(A) \leq \lambda_+$. Properties (i)–(iv), except for the fact that $-\infty < \lambda_* < \infty$, now follow from general properties of logarithmic moment generating functions, see Lemma 16 in the appendix. Property (vi) follows by monotone convergence and this implies $-\infty < \lambda_* < \infty$. Since excursions have length ≥ 1, formula (A.6) from Lemma 16 moreover implies property (v). ∎

The following two lemmas allow us to prove properties of $\phi^F_{x,y}$ for finite subgraphs F by induction on the number of vertices and edges.

Lemma 9 (Removal of an edge) *Let A be a nonnegative matrix, let $G = (S, E)$ be its associated graph, and let F be a subgraph of G. Let $e \in F \cap E$ and let $F' := F \backslash \{e\}$. Then*

$$\phi^{F'}_{x,y}(\lambda) = \begin{cases} \phi^F_{x,y}(\lambda) + e^\lambda A(x,y) & \textit{if } e = (x,y), \\ \phi^F_{x,y}(\lambda) & \textit{otherwise} \end{cases} \qquad (\lambda \in \mathbb{R}). \tag{2.7}$$

Proof This is immediate from the definition of the moment generating function in (2.4) and the fact that

$$\widehat{\Omega}_{x,y}(F') = \begin{cases} \widehat{\Omega}_{x,y}(F) \cup \{e\} & \text{if } e = (x,y), \\ \widehat{\Omega}_{x,y}(F) & \text{otherwise,} \end{cases} \tag{2.8}$$

where we identify e with the walk of length 1 that jumps through e. ∎

Lemma 10 (Removal of an isolated vertex) *Let A be a nonnegative matrix, let $G = (S, E)$ be its associated graph, and let F be a subgraph of G. Let $z \in F \cap S$ be a vertex of F. Assume that no edges in $F \cap E$ start or end at z and hence $F' := F \backslash \{z\}$ is a subgraph of G. Then*

$$\phi^{F'}_{x,y}(\lambda) = \phi^{F}_{x,y}(\lambda) + \sum_{k=0}^{\infty} \phi^{F}_{x,z}(\lambda)\phi^{F}_{z,z}(\lambda)^k \phi^{F}_{z,y}(\lambda) \qquad (x, y \in F' \cap S,\ \lambda \in \mathbb{R}), \tag{2.9}$$

Proof Distinguishing excursions away from F' according to how often they visit the vertex z, we have

$$\begin{aligned}\phi^{F'}_{x,y}(\lambda) &= \sum_{\omega_{x,y}} e^{\lambda \ell_{\omega_{x,y}}} \mathcal{A}(\omega_{x,y}) \\ &+ \sum_{k=0}^{\infty} \sum_{\omega_{x,z}} \sum_{\omega_{z,y}} \sum_{\omega^1_{z,z}} \cdots \sum_{\omega^k_{z,z}} e^{\lambda(\ell_{\omega_{x,z}} + \ell_{\omega_{z,y}} + \ell_{\omega^1_{z,z}} + \cdots + \ell_{\omega^k_{z,z}})} \\ &\qquad\qquad \times \mathcal{A}(\omega_{x,z})\mathcal{A}(\omega_{z,y})\mathcal{A}(\omega^1_{z,z}) \cdots \mathcal{A}(\omega^k_{z,z}),\end{aligned} \tag{2.10}$$

where we sum over $\omega_{x,y} \in \widehat{\Omega}_{x,y}(F)$ etc. Rewriting gives

$$\begin{aligned}\phi^{F'}_{x,y}(\lambda) &= \sum_{\omega_{x,y}} e^{\lambda \ell_{\omega_{x,y}}} \mathcal{A}(\omega_{x,y}) \\ &+ \Big(\sum_{\omega_{x,z}} e^{\lambda \ell_{\omega_{x,z}}} \mathcal{A}(\omega_{x,z})\Big)\Big(\sum_{\omega_{z,y}} e^{\lambda \ell_{\omega_{z,y}}} \mathcal{A}(\omega_{z,y})\Big) \sum_{k=0}^{\infty} \Big(\sum_{\omega_{z,z}} e^{\lambda \ell_{\omega_{z,z}}} \mathcal{A}(\omega_{z,z})\Big)^k,\end{aligned} \tag{2.11}$$

which is the formula in the lemma. ∎

2.2 Excursions away from a single point

Let A be a nonnegative matrix with index set S and let $G = (S, E)$ be its associated directed graph. For $z \in S$, we let

$$\psi_z := \psi^{\{z\}}_{z,z}, \quad \lambda_{z,+} := \lambda^{\{z\}}_{z,z,+}, \quad \text{and} \quad \lambda_{z,*} := \lambda^{\{z\}}_{z,z,*}, \tag{2.12}$$

denote the logarithmic moment generating function defined in (2.4) and the constants from (2.5) for the subgraph $F = \{z\}$ which consists of the vertex z and no edges. We also write $\widehat{\Omega}_z := \widehat{\Omega}_{z,z}(\{z\})$ for the space of all excursions away from z. The following proposition links forms of R-recurrence to the shape of ψ_z.

Proposition 11 (Forms of R-recurrence) *Assume that A is irreducible with $\rho(A) < \infty$, and let $z \in S$ be any reference point. Then*

(a) $\lambda_{z,*} = -\log \rho(A) =: \lambda_*$.

(b) *One has $\psi_z(\lambda_*) < 0$ if A is R-transient and $\psi_z(\lambda_*) = 0$ if A is R-recurrent.*

(c) *A is R-positive if and only if the left derivative of ψ_z at λ_* is finite.*

(d) *A is strongly R-positive if and only if $\lambda_* < \lambda_{z,+}$.*

To prove Proposition 11, we need one preparatory definition and lemma. Given a nonnegative matrix A, for each $\lambda \in \mathbb{R}$, we define a *Green's function* $G_\lambda(x, y)$ by

$$G_\lambda(x, y) := \sum_{k=0}^{\infty} e^{\lambda k} A^k(x, y) \qquad (x, y \in S), \tag{2.13}$$

which may be infinite for some values of λ. If A is irreducible, then it is known that [Ver67, Thm A]

$$G_\lambda(x,y) < \infty \text{ for } \lambda < -\log\rho(A) \quad\text{and}\quad G_\lambda(x,y) = \infty \text{ for } \lambda > -\log\rho(A) \tag{2.14}$$

for all $x, y \in S$. The following lemma makes a link between the Green's function and ψ_z.

Lemma 12 (Value on the diagonal) *One has*

$$G_\lambda(z,z) = \begin{cases} (1 - e^{\psi_z(\lambda)})^{-1} & \text{if } \psi_z(\lambda) < 0, \\ \infty & \text{otherwise.} \end{cases} \tag{2.15}$$

Proof Since each $\omega \in \Omega_{z,z}$ can be written as the concatenation of $m \geq 0$ excursions $\omega^{(i)} \in \widehat{\Omega}_z$, using the convention that a product of $m = 0$ factors is 1, we see that

$$\begin{aligned} \sum_{n=0}^{\infty} e^{\lambda n} A^n(z,z) &= \sum_{n=0}^{\infty} \sum_{\omega \in \Omega^n_{z,z}} e^{\lambda \ell_\omega} \mathcal{A}(\omega) = \sum_{\omega \in \Omega_{z,z}} e^{\lambda \ell_\omega} \mathcal{A}(\omega) \\ &= \sum_{m=0}^{\infty} \prod_{i=1}^{m} \Big(\sum_{\omega^{(i)} \in \widehat{\Omega}_z} e^{\lambda \ell_{\omega^{(i)}}} \mathcal{A}(\omega^{(i)}) \Big) = \sum_{m=0}^{\infty} e^{m\psi_z(\lambda)}, \end{aligned}$$

which yields (2.15). ∎

Proof of Proposition 11 Part (a) follows from formula (2.14) and Lemma 12. If A is R-transient, then it is immediate from the definition of R-transience (1.2) that $G_{\lambda_*}(z,z) < \infty$ and hence by Lemma 12 $\psi_z(\lambda_*) < 0$.

On the other hand, if A is R-recurrent, then by Theorem 2 there exists a function $h : S \to (0,\infty)$, which is unique up to scalar multiples, such that such that $Ah = \rho(A)h$. Setting

$$P(x,y) := \rho(A)^{-1} h(x)^{-1} A(x,y) h(y) \qquad (x, y \in S) \tag{2.16}$$

now defines a probability kernel. Since $P^n(x,x) = \rho(A)^{-n} A^n(x,x)$, we see from (1.2) that $\sum_n P^n(x,x) = \infty$ $(x \in S)$, which proves that P is recurrent. The Markov chain with transition kernel P makes i.i.d. excursions away from z with common law

$$\mathcal{P}(\omega) = \prod_{k=1}^{\ell_\omega} P(\omega_{k-1}, \omega_k) = \rho(A)^{-\ell_\omega} \mathcal{A}(\omega) = e^{\lambda_* \ell_\omega} \mathcal{A}(\omega) \qquad (\omega \in \widehat{\Omega}_z). \tag{2.17}$$

In particular, since P is recurrent,

$$1 = \sum_{\omega \in \widehat{\Omega}_z} \mathcal{P}(\omega) = \sum_{\omega \in \widehat{\Omega}_z} e^{\lambda_* \ell_\omega} \mathcal{A}(\omega) = e^{\psi_z(\lambda_*)}. \tag{2.18}$$

This shows that $\psi_z(\lambda_*) = 0$, completing the proof of part (b).

It follows from Lemmas 16 and 17 in the appendix that the left derivative of ψ_z at λ_* is the mean length of excursions away from z under the law in (2.17), proving part (c). Moreover,

$$\sum_{\omega \in \widehat{\Omega}_z} \mathcal{P}(\omega) e^{\varepsilon \ell_\omega} = \sum_{\omega \in \widehat{\Omega}_z} e^{(\lambda_* + \varepsilon)\ell_\omega} \mathcal{A}(\omega) = \psi_z(\lambda_* + \varepsilon) \tag{2.19}$$

is finite for $\varepsilon > 0$ suffiently small if and only if $\lambda_* < \lambda_{z,+}$, proving part (d). ∎

2.3 Characterization of strong R-recurrence

In this section, we prove Theorem 4. We start with two preparatory results.

Lemma 13 (Excessive functions) *Let A be an irreducible nonnegative matrix indexed by a countable set S. Then there exists a function $h : S \to (0, \infty)$ such that $Ah \leq \rho(A)h$.*

Proof We can without loss of generality assume that $\rho(A) < \infty$. If A is R-recurrent, then the statement follows from Theorem 2. If A is R-transient, then $G_{\lambda_*}(x,x) < \infty$ for each $x \in S$ by Proposition 11 (b) and Lemma 12. A simple argument based on irreducibility shows that $G_{\lambda_*}(x,y) < \infty$ for each $x, y \in S$. Since

$$AG_{\lambda_*}(x,z) = \sum_{k=0}^{\infty} e^{\lambda_* k} A^{k+1}(x,z) = e^{-\lambda_*} G_{\lambda_*}(x,z) - 1_{\{x=z\}} = \rho(A) G_{\lambda_*}(x,z) - 1_{\{x=z\}}, \quad (2.20)$$

setting $h(x) := G_{\lambda_*}(x,z)$ $(x \in S)$, where $z \in S$ is any reference point, now proves the claim. ∎

Proposition 14 (Exponential moments of excursions) *Let P be an irreducible subprobability kernel on a countable set S. Let $G = (S, E)$ be the graph associated with P and for any subgraph $F \subset G$, let $\lambda^F_{x,y,+}$ be defined in terms of P as in (2.5). Then, if*

$$\lambda^F_{x,y,+} > 0 \text{ for all } x, y \in F \cap S \quad (2.21)$$

holds for some finite nonempty subgraph F of G, it holds for all such subgraphs.

Proof We need to show that if F, F' are finite nonempty subgraphs of G, then (2.21) holds for F if and only if it holds for F'. It suffices to consider only the following two cases: I. $F' = F \backslash \{e\}$ where $e \in F \cap E$ is some edge in F, and II. $F' = F \backslash \{z\}$ where $z \in F \cap S$ is an isolated vertex in F. By Lemma 9, removing an edge does not change the value of $\lambda^F_{x,y,+}$ for any $x, y \in F \cap S$, so case I is easy.

In case II, we first prove that if (2.21) holds for F, then it also holds for F'. We distinguish two subcases: II.a: there exists an $\omega \in \widehat{\Omega}_{x,y}(F')$ that passes through z, and II.b: no such ω exists. In case II.b,

$$\lambda^{F'}_{x,y,+} = \lambda^F_{x,y,+}, \quad (2.22)$$

so this case is trivial. In case II.a, Lemma 10 tells us that

$$\lambda^{F'}_{x,y,+} = \lambda^F_{x,y,+} \wedge \lambda^F_{x,z,+} \wedge \lambda^F_{z,y,+} \wedge \lambda^F_{z,z,*}. \quad (2.23)$$

We claim that

$$\lambda^F_{z,z,*} > 0 \quad \text{if and only if} \quad \lambda^F_{z,z,+} > 0. \quad (2.24)$$

To see this, we observe that $\sum_{\omega \in \widehat{\Omega}^F_{z,z}} \mathcal{P}(\omega)$ is the probability that the Markov chain with transition kernel P started in z returns to z before visiting any point of F' or being killed. Since there exists an $\omega \in \widehat{\Omega}_{x,y}(F')$ that passes through z, this probability is < 1 and hence

$$\psi^F_{z,z}(0) = \log\Big(\sum_{\omega \in \widehat{\Omega}^F_{z,z}} \mathcal{P}(\omega)\Big) < 0.$$

By Lemma 8 (i) and (ii), $\psi^F_{z,z}$ is continuous on $(-\infty, \lambda^F_{z,z,+}]$, so if $\psi^F_{z,z}(\lambda) < \infty$ for some $\lambda > 0$ then also $\psi^F_{z,z}(\lambda) < 0$ for some $\lambda > 0$, proving (2.24). Combining (2.23) and (2.24), we see that if (2.21) holds for F, then it also holds for F'.

We next show that if (2.21) does not hold for F, then neither does it for F'. We consider four cases: (i) $\lambda^F_{x,y,+} \leq 0$ for some $x, y \in F' \cap S$, (ii) $\lambda^F_{x,z,+} \leq 0$ for some $x \in F' \cap S$, (iii)

$\lambda^F_{z,y,+} \leq 0$ for some $y \in F' \cap S$, and (iv) $\lambda^F_{z,z,+} \leq 0$. In case (i), formulas (2.22) and (2.23) immediately show that $\lambda^{F'}_{x,y} \leq 0$. In case (ii), by irreducibility, we can find some $y \in F'$ and $\omega \in \widehat{\Omega}_{z,y}(F)$; then we are in case II.a and (2.23) implies that $\lambda^{F'}_{x,y} \leq 0$. Case (iii) is similar to case (ii). In case (iv), finally, by irreducibility we can find $x, y \in F'$ and an $\omega \in \widehat{\Omega}_{x,y}(F')$ that passes through z, so (2.23) and (2.24) imply that $\lambda^{F'}_{x,y} \leq 0$. ∎

Proof of Theorem 4 Pick any reference vertex $z \in S$ and let ψ_z be the logarithmic moment generating function of A, as defined in (2.4) and (2.12). Let $\lambda_{z,+}$ and $\lambda_{z,*}$ be defined as in (2.5) and (2.12) and recall from Proposition 11 (a) that $\lambda_{z,*} = \lambda_* := -\log \rho(A)$. Define ψ'_z, $\lambda'_{z,+}$, and λ'_* in the same way for B.

It is immediately clear from the definition of ψ_z and irreducibility that $B \neq A$ implies that $\psi'_z(\lambda) < \psi_z(\lambda)$ for all λ such that $\psi'_z(\lambda) < \infty$. Since A is strongly R-positive, Proposition 11 (d) implies that $\lambda_* < \lambda_{z,+}$. It follows that $\psi'_z(\lambda_*) < \psi_z(\lambda_*) = 0$ while $\lambda'_{z,+} \geq \lambda_{z,+} > \lambda_*$, so by the continuity of ψ'_z on $(-\infty, \lambda'_{z,+}]$ we have

$$-\log \rho(B) = \lambda'_* = \sup\{\lambda \in \mathbb{R} : \psi'_z(\lambda) < 0\} > \lambda_* = -\log \rho(A),$$

which shows that $\rho(B) < \rho(A)$.

To prove part (b), we will show that if $\{(x, y) \in S^2 : A(x, y) \neq B(x, y)\}$ is finite and A is not strongly R-positive, then $\rho(B) = \rho(A)$. By Lemma 13, there exists a function $h : S \to (0, \infty)$ such that $Ah \leq \rho(A)h$. We use this function to define subprobability kernels P and P' by

$$\left.\begin{aligned} P(x,y) &:= \rho(A)^{-1} h(x)^{-1} A(x,y) h(y), \\ P'(x,y) &:= \rho(A)^{-1} h(x)^{-1} B(x,y) h(y), \end{aligned}\right\} \quad (x, y \in S). \tag{2.25}$$

Since $P^n(x, x) = \rho(A)^{-n} A^n(x, x)$, we see that $\rho(P) = 1$, and likewise $\rho(P') = \rho(A)^{-1}\rho(B)$. Thus, to prove that $\rho(B) = \rho(A)$, it suffices to prove that $\rho(P') = 1$.

Fix any reference point $z \in S$ and from now on, let ψ_z denote the logarithmic moment generating function of P (and not of A as before), let $\lambda_{z,+}$ and $\lambda_{z,*}$ be as in (2.5), and let ψ'_z, $\lambda'_{z,+}$, and $\lambda'_{z,*}$ be the same objects defined for P'. By Proposition 11 (a), $\lambda_{z,*} = \lambda_* := -\log \rho(P) = 0$ and $\lambda'_{z,*} = \lambda'_* := -\log \rho(P')$, so we need to show that $\lambda'_{z,*} = 0$.

Let us say that two nonnegative matrices are *equivalent* if they are related as in (1.3) for some function $h : S \to (0, \infty)$ and constant $c > 0$. Then, in the light of Theorem 1, a nonnegative matrix A is strongly R-positive if and only if it is equivalent to some (necessarily unique) strongly positive recurrent probability kernel. Since the subprobability kernel P from (2.25) is equivalent to A, and by assumption, A is not strongly R-positive, it follows that also P is not strongly R-positive. Therefore by Proposition 11 (c), $\lambda_{z,+} = \lambda_* = 0$. Since P' is a subprobability kernel, we have $\psi'_z(0) \leq 0$. Thus, to prove that $\lambda'_* = 0$, it suffices to show that $\psi'_z(\lambda) = \infty$ for all $\lambda > 0$ or equivalently $\lambda'_{z,+} \leq 0$.

Let F be any finite subgraph of G that contains all edges (x, y) where $P'(x, y) < P(x, y)$. Let $\phi^F_{x,y}$ and $\psi^F_{x,y}$ be defined as in (2.4) and $\lambda^F_{x,y,+}$ as in (2.5). Since each excursion away from F has the same weight under $\mathcal{P}$ and $\mathcal{P}'$, it does not matter whether we use P or P' to define these quantities. Applying Proposition 14 to the subgraphs F and $F' := \{z\}$, we see that

$$\lambda_{z,+} \leq 0 \quad \Leftrightarrow \quad \lambda^F_{x,y,+} \leq 0 \text{ for some } x, y \in F \cap S \quad \Leftrightarrow \quad \lambda'_{z,+} \leq 0.$$

In particular, the fact that $\lambda_{z,+} = 0$ implies $\lambda'_{z,+} \leq 0$, as required. ∎

2.4 Characterization of R-transience

In this section we prove Theorem 5. We need one preparatory lemma.

Lemma 15 (Strictly excessive functions) *Let A be an R-transient irreducible nonnegative matrix indexed by a countable set S, and let $S' \subset S$ be finite. Then there exists a function $h : S \to (0, \infty)$ such that $Ah \leq \rho(A)h$ and $Ah < \rho(A)h$ on S'.*

Proof In the proof of Lemma 13, we have seen that setting $h_z(x) := G_{\lambda_*}(x, z)$ defines a function such that $Ah_z(x) = \rho(A)h_z(x) - 1_{\{x=z\}}$. In view of this, the function $h := \sum_{z \in S'} h_z$ has all the required properties. ∎

Proof of Theorem 5 We start by proving part (a). Let $E' := \{(x, y) \in S^2 : A(x, y) \neq B(x, y)\}$. We will show that there exists a matrix $C \geq A$ with $C(x, y) > A(x, y)$ for all $(x, y) \in E'$ and $\rho(C) \leq \rho(A)$. Since $A \leq A + \varepsilon(B - A) \leq C$ for $\varepsilon > 0$ small enough, the claim then follows.

Let $S' := \{x \in S : (x, y) \in E' \text{ for some } y \in S\}$. By R-transience and Lemma 15, there exists a function $h : S \to (0, \infty)$ such that $Ah \leq \rho(A)h$ and $Ah < \rho(A)h$ on S'. It follows that

$$P(x, y) := \rho(A)^{-1}h(x)^{-1}A(x, y)h(y) \qquad (x, y \in S) \tag{2.26}$$

defines a subprobability kernel such that $\sum_y P(x, y) < 1$ for all $x \in S'$. Using this, we can construct a probability kernel $Q \geq P$ such that $Q(x, y) > P(x, y)$ for all $(x, y) \in E'$. Setting

$$C(x, y) := \rho(A)h(x)Q(x, y)h(y)^{-1} \qquad (x, y \in S) \tag{2.27}$$

then defines a nonnegative matrix $C \geq A$ with $C(x, y) > A(x, y)$ for all $(x, y) \in E'$. Since $C^n(x, x) := \rho(A)^n Q^n(x, x)$ and Q is a probability kernel, we see from (1.1) that $\rho(C) \leq \rho(A)$.

To prove also part (b), fix a reference point $z \in S$ and let $\psi_z, \lambda_{z,+}$, and $\lambda_{z,*}$ be defined in terms of A as in (2.4), (2.5), and (2.12). Let $\psi'_z, \lambda'_{z,+}$, and $\lambda'_{z,*}$ be the same objects defined in terms of B. By Proposition 11 $\lambda_{z,*} = -\log \rho(A) = -\log \rho(B) = \lambda'_{z,*}$. The definition of ψ_z and irreducibility imply that $\psi_z(\lambda) < \psi'_z(\lambda)$ for all λ such that $\psi'_z(\lambda) < \infty$, i.e., for $\lambda \leq \lambda'_{z,+}$. In particular, this applies at $\lambda_* = \lambda'_*$ so we see that $\psi_z(\lambda_*) < \psi'_z(\lambda'_*) \leq 0$. By Proposition 11 (b), it follows that A is R-transient. ∎

A Appendix

A.1 R-recurrence

Proof of Theorem 1 By Theorem 2, there exists a function $h : S \to (0, \infty)$, which is unique up to scalar multiples, such that $Ah = \rho(A)h$. Setting

$$P(x, y) := \rho(A)^{-1}h(x)^{-1}A(x, y)h(y) \qquad (x, y \in S) \tag{A.1}$$

defines a probability kernel on S. Since $P^n(x, x) := \rho(A)^{-1}A^n(x, x)$, we see from (1.2) that $\sum_{n=1}^{\infty} P^n(x, x) = \infty$, proving that P is recurrent.

Conversely, assume that for some function $h : S \to (0, \infty)$ and constant $c > 0$, formula (1.3) defines a recurrent probability kernel. Since P is a probability kernel $Ah = ch$. Since $P^n(x, x) := c^{-n}A^n(x, x)$, it follows that $\rho(P) = c^{-1}\rho(A)$. Since P is a recurrent probability kernel, for any $x \in S$,

$$\sum_{k=1}^{\infty} e^{\lambda k} P^k(x, x) < \infty \quad \Leftrightarrow \quad \lambda < 0 \tag{A.2}$$

which shows that $\rho(P) = 1$ and hence $c = \rho(A)$. Thus $Ah = \rho(A)h$ and Theorem 2 tells us that h is uniquely determined up to scalar multiples. ∎

A.2 Logarithmic moment generating functions

Let μ be a nonzero measure on $\mathbb{R}$. We define the *logarithmic moment generating function* of μ as

$$\psi(\lambda) := \log \int_R \mu(\mathrm{d}x) e^{\lambda x} \qquad (\lambda \in \mathbb{R}), \tag{A.3}$$

where $\log \infty := \infty$. We write

$$\mathcal{D}_\psi := \{\lambda \in \mathbb{R} : \psi(\lambda) < \infty\} \quad \text{and} \quad \mathcal{U}_\psi := \mathrm{int}(\mathcal{D}_\psi). \tag{A.4}$$

Lemma 16 (Logarithmic moment generating functions) *Let μ be a nonzero measure on $\mathbb{R}$ and let ψ be its logarithmic moment generating function. Then*

(i) *ψ is convex.*

(ii) *ψ is lower semi-continuous.*

(iii) *ψ is infinitely differentiable on $\mathcal{U}_\psi$.*

Moreover, for each $\lambda \in \mathcal{D}_\psi$, setting

$$\mu_\lambda(\mathrm{d}x) := e^{\lambda x - \psi(\lambda)} \mu(\mathrm{d}x) \tag{A.5}$$

defines a probability measure on $\mathbb{R}$ with

$$\left.\begin{array}{ll} \text{(i)} & \frac{\partial}{\partial\lambda}\psi(\lambda) = \langle \mu_\lambda \rangle, \\ \text{(ii)} & \frac{\partial^2}{\partial\lambda^2}\psi(\lambda) = \mathrm{Var}(\mu_\lambda), \end{array}\right\} \quad (\lambda \in \mathcal{U}_\psi), \tag{A.6}$$

where $\langle \mu_\lambda \rangle$ and $\mathrm{Var}(\mu_\lambda)$ denote the mean and variance of μ_λ, respectively.

Proof Set

$$\Phi(\lambda) := \int_{\mathbb{R}} \mu(\mathrm{d}x) e^{\lambda x} \qquad (\lambda \in \mathbb{R}), \tag{A.7}$$

so that $\psi(\lambda) = \log \Phi(\lambda)$. We claim that $\lambda \mapsto \Phi(\lambda)$ is infinitely differentiable on $\mathcal{U}_\psi$ and

$$\big(\tfrac{\partial}{\partial\lambda}\big)^n \Phi(\lambda) = \int x^n e^{\lambda x} \mu(\mathrm{d}x) \qquad (\lambda \in \mathcal{U}_\psi). \tag{A.8}$$

To justify this, we must show that the interchanging of differentiation and integral is allowed. We observe that

$$\frac{\partial}{\partial\lambda} \int x^n e^{\lambda x} \mu(\mathrm{d}x) = \lim_{\varepsilon \to 0} \int x^n \varepsilon^{-1} (e^{(\lambda+\varepsilon)x} - e^{\lambda x}) \mu(\mathrm{d}x), \tag{A.9}$$

By our assumption that $\lambda \in \mathcal{U}_\psi$, we can choose $\delta > 0$ such that

$$\int_{\mathbb{R}} \mu(\mathrm{d}x) \big[e^{(\lambda-2\delta)x} + e^{(\lambda+2\delta)x} \big] < \infty. \tag{A.10}$$

Since for any $-\delta < \varepsilon < \delta$ with $\varepsilon \neq 0$,

$$\begin{aligned} |x|^n \varepsilon^{-1} \big| e^{(\lambda+\varepsilon)x} - e^{\lambda x} \big| &= |x|^n \Big| \varepsilon^{-1} \int_\lambda^{\lambda+\varepsilon} x e^{\kappa x} \, \mathrm{d}\kappa \Big| \\ &\leq |x|^{n+1} \big[e^{(\lambda-\delta)x} + e^{(\lambda+\delta)x} \big] \qquad (x \in \mathbb{R}), \end{aligned} \tag{A.11}$$

which is integrable by (A.10), we may use dominated convergence in (A.9) to interchange the limit and integral.

Since

$$\int_{\mathbb{R}} \mu_\lambda(\mathrm{d}x) = \frac{1}{\Phi(\lambda)} e^{\lambda x} \mu(\mathrm{d}x) = 1 \qquad (\lambda \in \mathcal{D}_\psi), \tag{A.12}$$

we see that μ_λ is a probability measure for each $\lambda \in \mathcal{D}_\psi$. Formula (A.8) implies that for each $\lambda \in \mathcal{U}_\psi$

$$\begin{aligned}
&\text{(i)} \quad \tfrac{\partial}{\partial\lambda} \log \Phi(\lambda) = \tfrac{\partial}{\partial\lambda} \log \int e^{\lambda x} \mu(\mathrm{d}x) = \frac{\int x e^{\lambda x} \mu(\mathrm{d}x)}{\int e^{\lambda x} \mu(\mathrm{d}x)} = \langle \mu_\lambda \rangle, \\
&\text{(ii)} \quad \tfrac{\partial^2}{\partial\lambda^2} \log \Phi(\lambda) = \frac{\int x^2 e^{\lambda x} \mu(\mathrm{d}x) - (\Phi(\lambda) \int x e^{\lambda x} \mu(\mathrm{d}x))^2}{\Phi(\lambda)^2} \\
&\qquad\qquad = \int x^2 \mu_\lambda(\mathrm{d}x) - \Big(\int x \mu_\lambda(\mathrm{d}x) \Big)^2 = \mathrm{Var}(\mu_\lambda),
\end{aligned} \tag{A.13}$$

proving (A.6).

In particular, if μ is a compactly supported finite measure, then $\mathcal{U}_\psi = \mathbb{R}$ so these formulas prove that ψ is convex and continuous. For locally finite μ, we may find compactly supported finite μ_n such that $\mu_n \uparrow \mu$ and hence the associated logarithmic moment generating functions satisfy $\psi_n \uparrow \psi$. Since the ψ_n are convex and continuous, ψ must be convex and l.s.c. If μ is not locally finite, then $\psi(\lambda) = \infty$ for all $\lambda \in \mathbb{R}$ so there is nothing to prove. ∎

The next lemma says that formula (A.6) (i) holds more generally for $\lambda \in \mathcal{D}_\psi$, when the derivative is appropriately interpreted as a one-sided derivative or limit of derivatives for $\lambda \in \mathcal{U}_\psi$.

Lemma 17 (One-sided derivative) *Let μ be a nonzero measure on $\mathbb{R}$ and let ψ be its logarithmic moment generating function. Assume that $\mathcal{U}_\psi$ is nonempty, $\lambda_+ := \sup \mathcal{D}_\psi < \infty$, and $\psi(\lambda_+) < \infty$. Then*

$$\lim_{\lambda \uparrow \lambda_+} \tfrac{\partial}{\partial\lambda} \psi(\lambda) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \big(\psi(\lambda_+) - \psi(\lambda_+ - \varepsilon) \big) = \langle \mu_{\lambda_+} \rangle. \tag{A.14}$$

Proof Since $\gamma \mapsto x e^{x\gamma}$ is nondecreasing for each $x \geq 0$, we see that

$$\varepsilon^{-1}(e^{\lambda x} - e^{(\lambda-\varepsilon)x}) = \int_{\lambda-\varepsilon}^{\lambda} x e^{x\gamma} \mathrm{d}\gamma \uparrow x e^{\lambda x} \quad \text{as } \varepsilon \downarrow 0 \quad (x \geq 0), \tag{A.15}$$

so by monotone convergence, using notation as in (A.7), we see that

$$\varepsilon^{-1} \big(\Phi(\lambda_+) - \Phi(\lambda_+ - \varepsilon) \big) = \int \varepsilon^{-1} (e^{\lambda_+ x} - e^{(\lambda_+ - \varepsilon)x}) \mu(\mathrm{d}x) \uparrow \int x e^{\lambda_+ x} \mu(\mathrm{d}x) \quad \text{as } \varepsilon \downarrow 0. \tag{A.16}$$

By assumption $\psi(\lambda_+) < \infty$ so $\Phi(\lambda_+) < \infty$, which implies that μ_{λ_+} is well-defined, and the second equality in (A.14) now follows as in (A.13). Since ψ is convex, this also implies the first equality in (A.14). ∎

Acknowledgement

Theorem 4 is based on a suggestion of Rongfeng Sun, who conjectured a version of this result in joint work with Gregory Maillard and Nicolas Pétrélis. Francesco Caravenna helped me a lot orienting myself in the literature on pinning models. I thank Jan Seidler for pointing out the reference [Ign06].

References

[And91] W.J. Anderson. *Continuous-time Markov chains.* Springer, New York, 1991.

[CGZ06] F. Caravenna, G. Giacomin, and L. Zambotti. Sharp asymptotic behavior for wetting models in (1 + 1)-dimension. *Electron. J. Probab.* 11 (2006), Paper no. 14, 345–362.

[Fro12] G. Frobenius. Über Matrizen aus nicht negativen Elementen. *Sitzungsber. Königl. Preuss. Akad. Wiss.* (1912), 456–477.

[Gia07] G. Giacomin. Random polymer models. Hackensack, 2007.

[Ign06] I. Ignatiouk-Robert On the spectrum of Markov semigroups via sample path large deviations. *Probab. Theory Relat. Fields* 134 (2006), 44–80.

[Kar59] S. Karlin. Positive operators. *J. Math. Mech.* 8 (1959), 907-937.

[Ken59] D.G. Kendall. Unitary dilations of Markov transition operators, and the corresponding integral representations for transition-probability matrices. Pages 139–161 in: U. Grenander (ed), *Probability and Statistics. The Harald Cramér volume.* Almqvist & Wiksell, Stockholm, 1959.

[Kin63] J.F.C. Kingman. The exponential decay of Markov transition probabilities. *Proc. London Math. Soc. (3)* 13 (1963), 337–358.

[KR48] M.G. Kreĭn and M.A. Rutman. Linear operators leaving invariant a cone in a Banach space [in Russian]. *Uspehi Matem. Nauk* 3 (1948), No. 1(23), 3–95.

[KR50] M.G. Kreĭn and M.A. Rutman. Linear operators leaving invariant a cone in a Banach space. *Amer. Math. Soc. Transl.* 1950(26) (1950), 128 pages.

[MS95] V.A. Malyshev and F.M. Spieksma. Intrinsic convergence rate of countable Markov chains. *Markov Process. Related Fields* 1 (1995), 203–266.

[Num84] E. Nummelin. *General Irreducible Markov Chains and Non-negative Operators.* Cambridge University Press, 1984.

[Per07] O. Perron. Zur Theorie der Matrices. *Mathematische Annalen* 64(2) (1907), 248–263.

[Pop77] N.N. Popov. Conditions for geometric ergodicity of countable Markov chains [in Russian]. *Soviet Math. Dokl.* 18 (1977), 676–679.

[Sch74] H.H. Schaefer. *Banach Lattices and Positive Operators.* Springer, Berlin, 1974.

[Sen81] E. Seneta. *Non-negative Matrices and Markov Chains. 2nd ed.* Springer, New York, 1981.

[Spi90] F.M. Spieksma. *Geometrically Ergodic Markov Chains and the Optimal Control of Queues.* PhD thesis, University of Leiden, 1990.

[SV66] E. Seneta and D. Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* 3(2) (1966), 403–434.

[Ver62] D. Vere-Jones. Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford Ser. 2* 13 (1962), 7–28.

[Ver63] D. Vere-Jones. On the spectra of some linear operators associated with queueing systems. *Z. Wahrscheinichkeitstheorie* 2 (1963), 12–21.

[Ver67] D. Vere-Jones. Ergodic properties of non-negative matrices - I. *Pacific J. Math.* 22(2) (1967), 361–386.

[Woe00] W. Woess. *Random walks on infinite graphs and groups.* Cambridge University Press, 2000.

[Zer87] M. Zerner. Quelques propriétés spectrales des opérateurs positifs. *J. Funct. Anal.* 72 (1987), 381–417.