Karlin Theory On Growth and Mixing Extended to Linear Differential Equations

Lee Altenberg

`altenber@hawaii.edu`

June 17, 2010

Abstract

Karlin's (1982) Theorem 5.2 shows that linear systems alternating between growth and mixing phases have lower asymptotic growth with greater mixing. Here this result is extended to linear differential equations that combine site-specific growth or decay rates, and mixing between sites, showing that the spectral abscissa of a matrix $\mathbf{D} + m\mathbf{A}$ decreases with m, where $\mathbf{D} \neq c\mathbf{I}$ is a real diagonal matrix, $\mathbf{A}$ is an irreducible matrix with non-negative off-diagonal elements (an ML- or essentially non-negative matrix), and $m \geq 0$. The result is based on the inequality: $\mathbf{u}^{\top}\mathbf{A}\mathbf{v} < r(\mathbf{A})$, where $\mathbf{u}$ and $\mathbf{v}$ are the left and right Perron vectors of the matrix $\mathbf{D} + \mathbf{A}$, and $r(\mathbf{A})$ is the spectral abscissa and Perron root of $\mathbf{A}$. The result gives an analytic solution to prior work that relied on two-site or numerical simulation of models of growth and mixing, such as source and sink ecological models, or multiple tissue compartment models of microbe growth. The result has applications to the Lyapunov stability of perturbations in nonlinear systems.

1 Introduction

Growth and movement are ubiquitous phenomena in physical, biological, and social systems. In particular, site-specific growth and decay rates, combined with movement between sites, can be found throughout nature. An even wider range of phenomena may be included if we consider the formal equivalent: *state*-specific growth rates, and transformations between states.

Karlin (1982) developed two very general theorems on the asymptotic growth rates of systems combining growth and movement. The context that motivated these developments was rather narrow: analysis of the protection of genetic diversity in a subdivided population undergoing natural selection and migration. But the theorems are fundamental, describing the long-term growth behavior of a wide range of coupled linear systems, and the stability of many nonlinear systems.

Karlin's theorems apply to discrete time and discrete space systems in which a growth phase is followed by a movement phase, and these are iterated. This paper focuses on one of the theorems, Theorem 5.2, and extends it to apply to continuous

time and discrete space systems. Extensions to continuous time and continuous space systems entail greater technicalities and are deferred to elsewhere.

Karlin's Theorem 5.2 states simply that in a system of objects undergoing exponential growth or decay at different rates in different sites, *the greater the level of mixing of objects between sites, the lower the long-term growth rate of the whole aggregation of objects*. Here is the theorem, where $\rho(\mathbf{M})$ is the spectral radius of matrix $\mathbf{M}$, the largest magnitude of any eigenvalue of $\mathbf{M}$:

Theorem 5.2 (Karlin 1982). *Let $\mathbf{M}$ be a general nonnegative irreducible backward migration matrix. Consider the family of migration matrices,*

$$\mathbf{M}(\alpha) = (1 - \alpha)\mathbf{I} + \alpha\mathbf{M}.$$

Then for any set of positive fitness values $\mathbf{D} = \mathbf{diag}\Big[d_1, d_2, \ldots, d_n\Big]$,

$$\rho(\mathbf{M}(\alpha)\mathbf{D})$$

is decreasing as α increases (strictly, provided $\mathbf{D} \neq d\mathbf{I}$).

While the result is cast in terms of the specific context, it should be understood that $\mathbf{M}$ can be any irreducible stochastic matrix, and $\mathbf{D}$ any positive diagonal matrix.

Karlin used the theorem to analyze the stability of coupled nonlinear systems, where each site i has a continuous, differentiable map $f_i(x) : [0, 1] \mapsto [0, 1]$, so the coupled system is:

$$x_i(t + 1) = \sum_j M_{ij} f_j(x_j(t))$$

In the population genetics context, $f_i(0) = 0$, and the linearized stability dynamics for small $\mathbf{x}$ are

$$\mathbf{x}(t) = (\mathbf{M}(\alpha)\,\mathbf{D})^t \mathbf{x}(0), \tag{1}$$

where the diagonal elements of $\mathbf{D}$ are $D_i = \dfrac{\mathrm{d}f_i(0)}{\mathrm{d}x}$. The zero solution $\mathbf{x}(t) = 0$ is unstable to perturbation if $\rho(\mathbf{MD}) > 1$ and stable to perturbation if $\rho(\mathbf{MD}) < 1$.

The generality of this result is already evident in that no assumptions are made on $\mathbf{M}$ beyond that it be stochastic, and irreducible, which means there is a path of non-zero elements $M_{ik_1}, M_{k_1 k_2}, \ldots, M_{k_c j}$ between any i and any $j \neq i$. And no assumptions are placed on the set of nonlinear $f_i(x)$ other than that they be differentiable, and their domain and range be the unit interval, and $f_i(0) = 0$.

The first use of this theorem outside its original context was to analyze the evolution of genetic transmission (Altenberg, 1984; Altenberg and Feldman, 1987; Altenberg, 2009). There, instead of objects moving between sites, the objects are genomes transforming between genotypes. Theorem 5.2 translates to the *reduction principle*: the zero solution for a gene that controls 'mixing' between genotypes is unstable to perturbation (i.e. to introduction of the gene to the population) if the gene reduces mixing.

The use of this theorem in additional contexts, has to my knowledge, not yet occurred. This may be due, perhaps, to the limitation of the assumptions of discrete time and discrete space, or simply due to the small community of theoretical population geneticists familiar with it.

Objects:	Sites:	Site specific growth/decay rates:
genes	habitats	fitnesses
genomes	genotypes	fitnesses
organisms	habitats	species survival and reproduction rates
agricultural pests	fields under treatment	replication and survival
microbes	tissue compartments	survival and replication rates
metabolites	tissue compartments	catabolic rates
reactants	reaction-diffusion medium	reaction rates
wastes	reactors	waste breakdown rates
photons	media	absorption rates
particles	heterogenous matter	interaction and decay rates
capital	investments	rates of return

Table 1: A short list of systems exhibiting site-specific growth or decay and mixing between sites.

One may ask whether there is anything about discrete time and space that is essential to the result, or whether there is a more general phenomenon that may extend to continuous time and space. Here I show that the result can be extended to continuous time and discrete space, namely, to coupled linear differential equations. The result here applies to any combination of constant exponential growth or decay rates at different sites, and any constant pattern of movement between sites. The extension of Karlin's theory to linear first order differential equations brings a much wider domain of systems into its purview. One can contemplate a variety of systems that contain the applicable ingredients, shown in Table 1.

2 The System

The system investigated here is of the form

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = (\mathbf{D} + m\mathbf{A})\,\mathbf{x}(t), \tag{2}$$

which has solution

$$\mathbf{x}(t) = e^{(\mathbf{D}+m\mathbf{A})\,t}\,\mathbf{x}(0), \tag{3}$$

where

$\mathbf{x}(t) \in \mathbb{R}^n$ is an n-long vector of the quantities in each site at time t,

$\mathbf{D}$ is a real $n \times n$ diagonal matrix of the growth rates at each site,

$m \geq 0$ is the global rate of mixing between sites, and

$\mathbf{A}$ is a real $n \times n$ matrix that represents the movement distribution among n sites. $\mathbf{A}$ is an *essentially non-negative* matrix (also called a *Metzler*, *Metzler-Leontief*, or *ML-matrix*), defined by $M_{ij} \geq 0$ for $i \neq j$.

The matrices $\mathbf{D}$ and $\mathbf{A}$ hold the specifics for a particular system. The diagonal matrix $\mathbf{D}$ contains the growth rates, and under the action of $\mathbf{D}$ alone, the time trajectories (3) would be

$$x_i(t) = e^{D_i t} x_i(0).$$

Exponential growth or decay is determined by whether $D_i > 0$ or $D_i < 0$.

The matrix $\mathbf{A}$ represents movement between sites (or transformations between states). The form of variation in movement examined here is of the form:

$$\mathbf{F}(m) = \mathbf{D} + m\,\mathbf{A}, \tag{4}$$

where $\mathbf{A}$ represents the *distribution* of movement, while m represents the *rate* of movement. The question I address here is how the global level of 'mixing', m, affects the asymptotic growth rates of (3), and in particular, the stability of the zero solution $\mathbf{x}(t) = 0$ to perturbations.

The form (4) includes, as a special case, the form considered by Karlin (1982), in which a period of growth is followed by a period of movement:

$$\mathbf{F}(m) = [(1-m)\mathbf{I} + m\mathbf{P}]\,\mathbf{D} = \mathbf{D} + m(\mathbf{PD} - \mathbf{D}), \tag{5}$$

where

$\mathbf{I}$ is the identity matrix, and

$\mathbf{P}$ is a stochastic matrix.

In continuous time, both growth and movement occur simultaneously, so the analog to (5) is:

$$\mathbf{F}(m) = \mathbf{D} + m(\mathbf{P} - \mathbf{I}). \tag{6}$$

A typical assumption about movement is that quantity is redistributed but conserved, in which case summing the effect of movement over all destinations produces zero net change:

$$\mathbf{e}^\top \mathbf{A} = \mathbf{e}^\top (\mathbf{PD} - \mathbf{D}) = \mathbf{0}, \text{ and } \mathbf{e}^\top (\mathbf{P} - \mathbf{I}) = \mathbf{0}, \tag{7}$$

where

$\mathbf{e}$ is the vectors of ones, and

$\top$ is the transpose of the vector or matrix.

This class includes the generator matrices of continuous time Markov chains (also called 'intensity' matrices).

If quantity is lost during movement, (7) is replaced by

$$\mathbf{e}^\top \mathbf{A} = \mathbf{e}^\top (\mathbf{PD} - \mathbf{D}) \leq \neq \mathbf{0}, \text{ and } \mathbf{e}^\top (\mathbf{P} - \mathbf{I}) \leq \neq \mathbf{0}. \tag{8}$$

3 Results

The general phenomenon to be shown is the following:

Main Result. *The asymptotic growth rate of*

$$\mathbf{x}(t) = e^{(\mathbf{D}+m\mathbf{A})\,t}\,\mathbf{x}(0),$$

decreases with increasing values of the mixing parameter, m. If this system exhibits net growth, then greater mixing inhibits the rate of growth. If the system exhibits net decay, then greater mixing enhances the rate of decay.

The asymptotic rate of growth or decay in (3) is given by the *spectral abscissa* of $\mathbf{F}(m)$, which is the largest real part of any eigenvalue of $\mathbf{F}(m)$ (Bernstein 2009, p. 734, Gantmacher 1959*a*, pp. 125-129). For real matrices that have non-negative off-diagonal elements (ML-matrices), the spectral abscissa is always an eigenvalue, referred to at the *Perron root*. Irreducible ML-matrices retain many of the properties of irreducible non-negative matrices, including the existence of positive left and right eigenvectors (the Perron vectors), unique up to scaling, that are associated with the spectral abscissa (the Perron root) (Seneta, 1981, pp. 46–47).

This paper considers how variation in the global mixing rate m varies the Perron root of $\mathbf{F}(m)$.

Definitions. *Let:*

$\mathbf{e}_i$ *be the vector with element i equal to 1, and other elements equal to zero;*

$r(\mathbf{M}) := \max_i \mathrm{Re}\lambda_i(\mathbf{M})$ *refer the spectral abscissa of a square matrix $\mathbf{M}$;*

$\mathbf{v}(\mathbf{A}) > 0$ *refer to the right Perron vector of an irreducible ML-matrix $\mathbf{A}$, normalized so that $\mathbf{e}^\top \mathbf{v}(\mathbf{A}) = 1$;*

$\mathbf{u}(\mathbf{A})^\top > 0$ *refer to the left Perron vector of an irreducible ML-matrix $\mathbf{A}$, normalized so that $\mathbf{u}(\mathbf{A})^\top \mathbf{v}(\mathbf{A}) = 1$ (the different normalization is convenient later); So,*

$$\mathbf{u}(\mathbf{A})^\top \mathbf{A} = r(\mathbf{A})\,\mathbf{u}(\mathbf{A})^\top,$$
$$\mathbf{A}\mathbf{v}(\mathbf{A}) = r(\mathbf{A})\,\mathbf{v}(\mathbf{A}),\ \textit{and}$$
$$\mathbf{u}(\mathbf{A})^\top \mathbf{A}\,\mathbf{v}(\mathbf{A}) = r(\mathbf{A}).$$

3.1 Basic Results

The main result that will be used in the analysis is the following:

Theorem: 1. *Let $\mathbf{A}$ be an $n \times n$ irreducible real matrix with non-negative off-diagonal elements (an ML-matrix), and $\mathbf{D}$ be an $n \times n$ diagonal real matrix. Then:*

$$r(\mathbf{A}+\mathbf{D}) - r(\mathbf{A}) \le \mathbf{u}(\mathbf{A}+\mathbf{D})^\top \mathbf{D}\ \mathbf{v}(\mathbf{A}+\mathbf{D}) \tag{9}$$

with equality if and only if $\mathbf{D} = c\,\mathbf{I}$ for some $c \in \mathbb{R}$.

Proof. The spectral abscissa of an irreducible ML-matrix, $\mathbf{A}$, is its Perron root, which is given by this variational formula (Friedland 1981, Corollary 3.1, related to the variational formula of Donsker and Varadhan (1975), and shown to extend to ML-matrices in Altenberg 2009, Lemma 3):

$$r(\mathbf{A}) = \sup_{\mathbf{p} \in \mathcal{P}_n} \inf_{\mathbf{x} > \mathbf{0}} \sum_{i=1}^{n} p_i \frac{[\mathbf{Ax}]_i}{x_i}, \tag{10}$$

where $\mathcal{P}_n = \{\mathbf{p} \colon p_i \geq 0, \sum_{i=1}^{n} p_i = 1\} \subset \mathbb{R}^n$.

Let $\mathbf{x}(\mathbf{A})$ and $\mathbf{p}(\mathbf{A})$ be the vectors, as functions of $\mathbf{A}$, for which the sup and inf are attained, where $\mathbf{x}(\mathbf{A})$ is also normalized so that $\sum_{i=1}^{n} x_i(\mathbf{A}) = 1$. Then $\mathbf{p}(\mathbf{A})$ and $\mathbf{x}(\mathbf{A})$ are unique critical points for a given $\mathbf{A}$ (Friedland and Karlin 1975, Friedland 1981, and Karlin 1982, p. 195),

$$\mathbf{x}(\mathbf{A}) = \mathbf{v}(\mathbf{A}), \tag{11}$$

and

$$\mathbf{p}(\mathbf{A}) = \mathbf{u}(\mathbf{A}) \circ \mathbf{v}(\mathbf{A}), \tag{12}$$

where $\circ$ is element-wise the Schur-Hadamard product.

As utilized in the proof in Karlin (1982, Theorem 5.2), since $\mathbf{x}(\mathbf{A})$ is a unique critical point in (10), the inf means that any $\mathbf{y} \neq \mathbf{x}(\mathbf{A})$ produces:

$$r(\mathbf{A}) = \sup_{\mathbf{p} \in \mathcal{P}_n} \inf_{\mathbf{x} > \mathbf{0}} \sum_{i=1}^{n} p_i \frac{[\mathbf{Ax}]_i}{x_i}$$
$$< \sum_{i=1}^{n} p_i(\mathbf{A}) \frac{[\mathbf{Ay}]_i}{y_i}. \tag{13}$$

Repeating the analogous step in the proof in Karlin (1982, Theorem 5.2), let $\mathbf{y} = \mathbf{v}(\mathbf{A})$. Then

$$\sum_{i=1}^{n} p_i(\mathbf{A} + \mathbf{D}) \frac{[(\mathbf{A} + \mathbf{D})\mathbf{v}(\mathbf{A})]_i}{v_i(\mathbf{A})}$$
$$= r(\mathbf{A}) \sum_{i=1}^{n} p_i(\mathbf{A} + \mathbf{D}) \frac{v_i(\mathbf{A})}{v_i(\mathbf{A})} + \sum_{i=1}^{n} p_i(\mathbf{A} + \mathbf{D}) \frac{[\mathbf{Dv}(\mathbf{A})]_i}{v_i(\mathbf{A})}$$
$$= r(\mathbf{A}) + \sum_{i=1}^{n} p_i(\mathbf{A} + \mathbf{D}) D_i$$
$$= r(\mathbf{A}) + \mathbf{u}(\mathbf{A} + \mathbf{D})^{\top} \mathbf{D}\, \mathbf{v}(\mathbf{A} + \mathbf{D}).$$

Hence

$$r(\mathbf{A}) + \mathbf{u}(\mathbf{A} + \mathbf{D})^{\top} \mathbf{D}\, \mathbf{v}(\mathbf{A} + \mathbf{D}) \geq r(\mathbf{A} + \mathbf{D}),$$

with equality if and only if $\mathbf{v}(\mathbf{A}) = \mathbf{v}(\mathbf{A} + \mathbf{D})$, which entails

$$(\mathbf{A} + \mathbf{D})\mathbf{v}(\mathbf{A}) = r(\mathbf{A} + \mathbf{D})\mathbf{v}(\mathbf{A}) = \mathbf{A}\mathbf{v}(\mathbf{A}) + \mathbf{D}\mathbf{v}(\mathbf{A})$$
$$= r(\mathbf{A})\mathbf{v}(\mathbf{A}) + \mathbf{D}\mathbf{v}(\mathbf{A}),$$

hence $[\, r(\mathbf{A} + \mathbf{D}) - r(\mathbf{A})]\mathbf{I}\mathbf{v}(\mathbf{A}) = \mathbf{D}\mathbf{v}(\mathbf{A})$; and since $\mathbf{v}(\mathbf{A}) > \mathbf{0}$, this implies $\mathbf{D} = c\,\mathbf{I}$, where $c = r(\mathbf{A} + \mathbf{D}) - r(\mathbf{A})$. $\qquad\square$

From Theorem 1 a number of ancillary results can be obtained:

Corollary: 2 (Convexity Derived). *Let* $\mathbf{A}$ *be an irreducible ML-matrix, and* $\mathbf{D}$ *a real diagonal matrix. Then, for* $\beta > 1$:

$$r(\mathbf{A} + \mathbf{D}) - r(\mathbf{A}) \leq r\!\left(\tfrac{1}{\beta}\mathbf{A} + \mathbf{D}\right) - r\!\left(\tfrac{1}{\beta}\mathbf{A}\right) \leq \mathbf{u}(\mathbf{A} + \beta\mathbf{D})^{\top}\,\mathbf{D}\;\mathbf{v}(\mathbf{A} + \beta\mathbf{D}),$$
(14)

with equality if and only if $\mathbf{D} = c\,\mathbf{I}$ *for some* $c \in \mathbb{R}$.

Proof. Cohen (1979) established the convexity relation:

$$(1 - \alpha)\,r(\mathbf{A}) + \alpha\,r(\mathbf{A} + \mathbf{D}) \geq r((1 - \alpha)\mathbf{A} + \alpha(\mathbf{A} + \mathbf{D})) \qquad (15)$$

for non-negative $\mathbf{A}$ and $\mathbf{D}$, and $0 < \alpha < 1$. Friedland (1981) showed that when $\mathbf{A}$ is irreducible, equality holds if and only if $\mathbf{D} = c\,\mathbf{I}$ for some $c \in \mathbb{R}$. This holds when $\mathbf{A}$ and $\mathbf{D}$ are ML-matrices by the relation $r(\mathbf{A} + c\mathbf{I}) - c = r(\mathbf{A})$, since large enough c will guarantee $\mathbf{A} + c\mathbf{I} \geq \mathbf{0}$.

Rearrangement of (15) gives:

$$r(\mathbf{A} + \mathbf{D}) - r(\mathbf{A}) \geq$$
$$\frac{1}{\alpha}\,[\,r((1 - \alpha)\mathbf{A} + \alpha(\mathbf{A} + \mathbf{D})) - r(\mathbf{A})]$$
$$= \left[\, r(\tfrac{1}{\alpha}\mathbf{A} + \mathbf{D}) - r(\tfrac{1}{\alpha}\mathbf{A})\right].$$

Application of (9) gives

$$\mathbf{u}(\mathbf{A} + \mathbf{D})^{\top}\mathbf{D}\mathbf{v}(\mathbf{A} + \mathbf{D}) \geq r(\mathbf{A} + \mathbf{D}) - r(\mathbf{A}) \geq r\!\left(\tfrac{1}{\alpha}\mathbf{A} + \mathbf{D}\right) - r\!\left(\tfrac{1}{\alpha}\mathbf{A}\right).$$

The condition for equality in both places is that $\mathbf{D} = c\,\mathbf{I}$ for some $c \in \mathbb{R}$.

Multiplying by α, letting $\mathbf{D}' := \alpha\mathbf{D}$, and $\beta := 1/\alpha$:

$$\mathbf{u}(\mathbf{A} + \mathbf{D})^{\top}\,\alpha\mathbf{D}\;\mathbf{v}(\mathbf{A} + \mathbf{D}) \geq \alpha\,r(\mathbf{A} + \mathbf{D}) - \alpha\,r(\mathbf{A}) \geq r(\mathbf{A} + \alpha\mathbf{D}) - r(\mathbf{A})$$

$$\Longleftrightarrow$$

$$\mathbf{u}(\mathbf{A} + \beta\mathbf{D}')^{\top}\,\mathbf{D}'\;\mathbf{v}(\mathbf{A} + \beta\mathbf{D}') \geq r\!\left(\tfrac{1}{\beta}\mathbf{A} + \mathbf{D}'\right) - r\!\left(\tfrac{1}{\beta}\mathbf{A}\right) \geq r(\mathbf{A} + \mathbf{D}') - r(\mathbf{A}).$$

Replacement of $\mathbf{D}'$ by $\mathbf{D}$ gives (14). $\qquad\square$

Corollary: 3. *For irreducible ML-matrix* $\mathbf{A}$ *and real diagonal matrix* $\mathbf{D}$:

$$\mathbf{u}(\mathbf{A} + \mathbf{D})^\top \mathbf{A}\, \mathbf{v}(\mathbf{A} + \mathbf{D}) \leq r(\mathbf{A}), \tag{16}$$

with equality if and only if $\mathbf{D} = c\,\mathbf{I}$ *for some* $c \in \mathbb{R}$.

Proof. From Theorem 1:

$$\begin{aligned}
\mathbf{u}(\mathbf{A} + \mathbf{D})^\top &\mathbf{D}\, \mathbf{v}(\mathbf{A} + \mathbf{D}) \\
&\geq r(\mathbf{A} + \mathbf{D}) - r(\mathbf{A}) \\
&= \mathbf{u}(\mathbf{A} + \mathbf{D})^\top\, (\mathbf{A} + \mathbf{D})\, \mathbf{v}(\mathbf{A} + \mathbf{D}) - r(\mathbf{A}) \\
&= \mathbf{u}(\mathbf{A} + \mathbf{D})^\top \mathbf{A}\, \mathbf{v}(\mathbf{A} + \mathbf{D}) \\
&\quad + \mathbf{u}(\mathbf{A} + \mathbf{D})^\top \mathbf{D}\, \mathbf{v}(\mathbf{A} + \mathbf{D}) - r(\mathbf{A}),
\end{aligned}$$

and rearranging,

$$r(\mathbf{A}) \geq \mathbf{u}(\mathbf{A} + \mathbf{D})^\top \mathbf{A}\, \mathbf{v}(\mathbf{A} + \mathbf{D}).$$

The equality condition is unchanged from Theorem 1. $\qquad\square$

Corollary: 4 (Sums). *Let* $\mathbf{D}$ *be a real diagonal matrix, and let* $\mathbf{A} = \sum_{k=1}^{N} \mathbf{A}_k$, *where* $\{\mathbf{A}_k\}$ *are ML-matrices that share a common right [left] Perron vector. Then*

$$\mathbf{u}(\mathbf{A} + \mathbf{D})^\top \mathbf{A}\, \mathbf{v}(\mathbf{A} + \mathbf{D}) \leq \sum_{k=1}^{N} r(\mathbf{A}_k), \tag{17}$$

with equality if and only if $\mathbf{D} = c\,\mathbf{I}$ *for some* $c \in \mathbb{R}$.

Proof. Letting $\mathbf{y}$ be the right common Perron vector, then

$$\mathbf{A}\mathbf{y} = r(\mathbf{A})\mathbf{y} = \sum_{k=1}^{N} \mathbf{A}_i \mathbf{y} = \mathbf{y} \sum_{k=1}^{N} r(\mathbf{A}_i),$$

thus $r(\mathbf{A}) = \sum_{k=1}^{N} r(\mathbf{A}_i)$, and substitution in (16) yields (17). For $\mathbf{y}$ equal to a common left Perron vector, $\mathbf{A}^\top$ is used. $\qquad\square$

Corollary: 5 (The 'Flip' Theorem (Bapat and Raghavan, 1997, Theorem 3.2.5)). *Let* $\mathbf{A}$ *be an irreducible ML-matrix. Let* $\mathbf{y} \circ \mathbf{z} = \mathbf{u}(\mathbf{A}) \circ \mathbf{v}(\mathbf{A})$, *with* $\mathbf{y} > \mathbf{0}$, $\mathbf{z} > \mathbf{0}$. *Then*

$$\mathbf{z}^\top \mathbf{A}\, \mathbf{y} \geq \mathbf{u}(\mathbf{A})^\top \mathbf{A}\, \mathbf{v}(\mathbf{A}) = r(\mathbf{A}).$$

with equality if and only if $\mathbf{y} = \mathbf{v}(\mathbf{A})$ *or* $\mathbf{u}(\mathbf{A}) = \mathbf{v}(\mathbf{A})$.

In particular, when the left and right Perron vectors are 'flipped',

$$\mathbf{v}(\mathbf{A})^\top \mathbf{A}\, \mathbf{u}(\mathbf{A}) > \mathbf{u}(\mathbf{A})^\top \mathbf{A}\, \mathbf{v}(\mathbf{A}) = r(\mathbf{A}),$$

if $\mathbf{u}(\mathbf{A}) \neq \mathbf{v}(\mathbf{A})$.

Proof. This is an alternative proof to that given in Bapat and Raghavan (1997), and this extends Theorem 3.2.5 to ML-matrices.

Substituting $y_i = u_i(\mathbf{A})v_i(\mathbf{A})/z_i$ and (12) into (13), one gets:

$$r(\mathbf{A}) \leq \sum_{i=1}^{n} p_i(\mathbf{A})\frac{[\mathbf{Ay}]_i}{y_i} = \sum_{i=1}^{n} u_i(\mathbf{A})v_i(\mathbf{A})\frac{[\mathbf{Ay}]_i}{u_i(\mathbf{A})v_i(\mathbf{A})/z_i}$$

$$= \sum_{i=1}^{n} z_i[\mathbf{A}\ y]_i = \mathbf{z}^{\top}\mathbf{A}\ \mathbf{y},$$

with equality if and only if $\mathbf{y} = \mathbf{v}(\mathbf{A})$ or $\mathbf{u}(\mathbf{A}) = \mathbf{v}(\mathbf{A})$. $\qquad\square$

3.2   Main Result

These results are now applied to extend Karlin's Theorem 5.2 (Karlin, 1982). Theorem 5.2 applies to matrices $\mathbf{F}(m) = \mathbf{D} + m\,\mathbf{A}$, where $\mathbf{D}$ is positive diagonal matrix, $\mathbf{A} = (\mathbf{P} - \mathbf{I})\mathbf{D}$, $\mathbf{P}$ is an irreducible stochastic matrix, and $0 \leq m \leq 1$. Here, results are extended to $\mathbf{D}$ that may have negative diagonal elements, to arbitrary irreducible and reducible ML-matrices, $\mathbf{A}$, and to any $m \geq 0$.

Theorem: 6 (Growth and Mixing). *Let $\mathbf{D}$ be a real $n \times n$ diagonal matrix, and $\mathbf{A}$ be a real $n \times n$ matrix with non-negative off-diagonal elements (an ML-matrix). Then, for $m \geq 0$:*

$$\frac{\mathrm{d}\,r(\mathbf{D} + m\mathbf{A})}{\mathrm{d}m} \leq r(\mathbf{A}), \tag{18}$$

with equality holding if and only if either:

1. $\mathbf{D} = c\,\mathbf{I}$ *for some $c \in \mathbb{R}$; or,*

2. *(a) $\mathbf{A}$ is reducible; and*

   (b) $\mathbf{D}_\kappa = c\,\mathbf{I}_\kappa$, for some $c \in \mathbb{R}$, for every κ such that:

      i. *κ is an index on the diagonal blocks in the Frobenius normal form of $\mathbf{A}$, each block being defined by a subset of the indices of $\mathbf{A}$,*

      ii. *$\mathbf{D}_\kappa$ and $\mathbf{I}_\kappa$ are the restrictions of $\mathbf{D}$ and $\mathbf{I}$ under the block of indices derived from $\mathbf{A}$, and*

      iii. *on an open neighborhood of m:*

$$r(\mathbf{D}_\kappa + m\mathbf{A}_\kappa) = max_h\ r(\mathbf{D}_h + m\mathbf{A}_h) = r(\mathbf{D} + m\mathbf{A}).$$

Moreover, for reducible $\mathbf{A}$, with κ defined as in 2(b)iii, a sharper inequality obtains:

$$\frac{\mathrm{d}\,r(\mathbf{D} + m\mathbf{A})}{\mathrm{d}m} \leq r(\mathbf{A}_\kappa) \leq r(\mathbf{A}),$$

with equality on the left side only under condition 2b.

Proof. **Case: Irreducible A.** When $\mathbf{F}(m)$ is a C^2 function of m, and is an irreducible ML-matrix on some open set around m, then the derivative of its spectral abscissa is (Altenberg, 2009, Lemma 4):

$$\frac{\mathrm{d}\, r(\mathbf{F}(m))}{\mathrm{d}m} = \sum_{i=1}^{n} p_i(m) \frac{[\frac{\mathrm{d}\mathbf{F}(m)}{\mathrm{d}m}\, \mathbf{x}(m)]_i}{x_i(m)},$$

where for clarity, $\mathbf{p}(m) := \mathbf{p}(\mathbf{F}(m))$ and $\mathbf{x}(m) := \mathbf{x}(\mathbf{F}(m))$. Using (11) and (12) one obtains the classical form (Caswell, 2000, Sec. 9.1.1):

$$\frac{\mathrm{d}\, r(\mathbf{F}(m))}{\mathrm{d}m} = \mathbf{u}(m)^{\top} \frac{\mathrm{d}\mathbf{F}(m)}{\mathrm{d}m} \mathbf{v}(m).$$

Hence,

$$\frac{\mathrm{d}\, r(\mathbf{D} + m\,\mathbf{A}))}{\mathrm{d}m} = \mathbf{u}(m)^{\top} \mathbf{A}\, \mathbf{v}(m).$$

Applying Corollary 3:

$$\mathbf{u}(m)^{\top}(m\,\mathbf{A})\, \mathbf{v}(m) \leq r(m\,\mathbf{A})$$
$$\iff \mathbf{u}(m)^{\top} \mathbf{A}\, \mathbf{v}(m) \leq r(\mathbf{A}),$$

and thus

$$\frac{\mathrm{d}\, r(\mathbf{D} + m\mathbf{A})}{\mathrm{d}m} = \mathbf{u}(m)^{\top} \mathbf{A}\, \mathbf{v}(m) \leq r(\mathbf{A}),$$

with equality if and only if $\mathbf{D} = c\,\mathbf{I}$ for some $c \in \mathbb{R}$.

    Case: Reducible A. If $\mathbf{A}$ is reducible, its Frobenius normal form is utilized. The Frobenius normal form, $\bar{\mathbf{A}}$, of a reducible matrix $\mathbf{A}$ permutes the indices until it has the structure (Gantmacher, 1959*b*, p. 75):

$$\bar{\mathbf{A}} = \left[\begin{array}{cccc|ccc} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} & & & \\ \mathbf{0} & \mathbf{A}_2 & \ddots & \vdots & & \mathbf{0} & \\ \vdots & & \ddots & \mathbf{0} & & & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_t & & & \\ \hline \mathbf{A}_{t+1,1} & \mathbf{A}_{t+1,2} & \cdots & \mathbf{A}_{t+1,t} & \mathbf{A}_{t+1} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \cdots & \cdots & \cdots & \ddots & \mathbf{0} \\ \mathbf{A}_{t+s,1} & \mathbf{A}_{t+s,2} & \cdots & \mathbf{A}_{t+s,t} & \mathbf{A}_{t+s,t+1} & \cdots & \mathbf{A}_{t+s} \end{array}\right] \quad (19)$$

where the diagonal blocks $\mathbf{A}_h$ are irreducible square matrices. The eigenvalues of $\mathbf{A}$ are the eigenvalues of the irreducible diagonal block matrices $\mathbf{A}_h$ (Altenberg, 2009,

Lemma 5). Therefore, the spectral abscissa for $\mathbf{F}(m) = \mathbf{D} + m\,\mathbf{A}$ is the maximum of the spectral abscissae:

$$r(\mathbf{F}(m)) = \max_h r(\mathbf{F}_h(m)).$$

Hence, for every κ that satisfies $r(\mathbf{F}_\kappa(m)) = \max_h r(\mathbf{F}_h(m))$ on some neighborhood of m, the result on irreducible matrices yields:

$$\frac{\mathrm{d}}{\mathrm{d}m}\,r(\mathbf{F}(m)) = \frac{\mathrm{d}}{\mathrm{d}m}\,r(\mathbf{F}_\kappa(m)) \le r(\mathbf{A}_\kappa), \tag{20}$$

with equality holding if and only if $\mathbf{D}_\kappa = c\,\mathbf{I}_\kappa$ for some $c \in \mathbb{R}$ for every κ.

Since $r(\mathbf{A}) = \max_h r(\mathbf{A}_h) \ge r(\mathbf{A}_\kappa)$, (20) is seen to be a sharper inequality than (18):

$$\frac{\mathrm{d}}{\mathrm{d}m}\,r(\mathbf{F}(m)) = \frac{\mathrm{d}}{\mathrm{d}m}\,r(\mathbf{F}_\kappa(m)) \le r(\mathbf{A}_\kappa) \le r(\mathbf{A}). \qquad \square$$

Remark: Blocks $\mathbf{A}_1, \dots, \mathbf{A}_t$ in (19) are called *isolated blocks*. If $\mathbf{x}(0)$ is zero on any isolated block, it remains zero on that block for all $t \ge 0$. Should $r(\mathbf{A})$ derive from isolated blocks, then the system will asymptotically grow at rate $r(\mathbf{A})$ only if $\mathbf{x}(0)$ has non-zero elements for one of those blocks. Therefore, the asymptotic growth rate for reducible $\mathbf{A}$ may depend on the initial value $\mathbf{x}(0)$, whereas for irreducible $\mathbf{A}$ it is independent of any initial value $\mathbf{x}(0) \ge \ne \mathbf{0}$.

3.3 Conservative and Lossy Mixing

Inherent to the concepts of movement and mixing is the idea that the redistributed quantity is conserved or perhaps lost by the movement process, but never created. Theorem 6 is more general and does not assume this. But when the assumption that movement is conservative or lossy is made, it yields the following results as special cases of Theorem 6:

Corollary: 7 (Measure-Preserving Mixing)**.** *If ML-matrix $\mathbf{A}$ is such that mixing preserves the total measure of quantity, then*

$$\frac{\mathrm{d}\,r(\mathbf{D} + m\mathbf{A})}{\mathrm{d}m} \le 0,$$

with equality under the same conditions as Theorem 6.

Proof. Measure-preserving mixing means $\mathbf{e}^\top \mathbf{A} = \mathbf{0}$, in which case $r(\mathbf{A}) = 0$, and this is applied to (18). $\qquad \square$

Corollary: 8 (Lossy Mixing)**.** *If ML-matrix $\mathbf{A}$ is such that quantity is lost in mixing, then*

$$\frac{\mathrm{d}\,r(\mathbf{D} + m\mathbf{A})}{\mathrm{d}m} < 0,$$

for any real diagonal matrix $\mathbf{D}$.

Proof. Lossy mixing means $\mathbf{e}^\top \mathbf{A} \leq\neq \mathbf{0}$, which implies by subinvariance (Seneta, 1981, Corollary 3, p. 52) that $r(\mathbf{A}) < 0$, and this is applied to (18). $\square$

Lemma: 9 (Bounds). *Let $r(\mathbf{A}) = 0$. Then:*

$$r(\mathbf{D} + m\mathbf{A}) \in [\min_i(D_i), \max_i(D_i)]$$

Proof. We know from Cohen (1978) that $\dfrac{\mathrm{d}\, r(\mathbf{A})}{\mathrm{d}A_{ij}} \geq 0$ for any ML-matrix $\mathbf{A}$. So here, $\dfrac{\mathrm{d}}{\mathrm{d}D_i} r(\mathbf{D} + m\mathbf{A}) \geq 0$. Hence

$$\begin{aligned}
r(\max_i D_i\mathbf{I} + m\mathbf{A}) = \max_i D_i + m\, r(\mathbf{A}) &= \max_i D_i \\
\geq r(\mathbf{D} + m\mathbf{A}) &\geq r(\min_i D_i\mathbf{I} + m\mathbf{A}) \\
&= \min_i D_i. \qquad \square
\end{aligned}$$

Theorem: 10 (Limit). *For any choice of $\lambda \in [\max_i(D_i), \min_i(D_i)]$, there exists a family of ML-matrices, $\mathcal{F} \subset \{\mathbf{A}\colon r(\mathbf{A}) = 0\}$, that yields*

$$\lim_{m\to\infty} r(\mathbf{D} + m\,\mathbf{A}) = \lambda \text{ for } \mathbf{A} \in \mathcal{F}.$$

One such family is:

$$\mathcal{F} = \{\boldsymbol{\alpha}\mathbf{e}^\top - \mathbf{I}\colon \sum_{i=1}^n \alpha_i D_i = \lambda,\ \alpha_i \geq 0,\ and\ \mathbf{e}^\top\boldsymbol{\alpha} = 1\}.$$

Proof. First, let $\boldsymbol{\alpha} > \mathbf{0}$, which makes matrices $\boldsymbol{\alpha}\mathbf{e}^\top$ and $\mathbf{F}(m) := \mathbf{D} + m\,\boldsymbol{\alpha}\mathbf{e}^\top$ irreducible, giving $\mathbf{F}(m)$ a unique positive right Perron vector, $\mathbf{v}(m)$. Hence

$$\begin{aligned}
r(\mathbf{F}(m))\,\mathbf{v}(m) &= \mathbf{D}\mathbf{v}(m) + m\,\boldsymbol{\alpha}\,\mathbf{e}^\top\mathbf{v}(m) - m\,\mathbf{v}(m) \\
&= \mathbf{D}\mathbf{v}(m) + m(\boldsymbol{\alpha} - \mathbf{v}(m)) \qquad (21) \\
&\Longleftrightarrow \\
v_i(m) &= \frac{m\,\alpha_i}{r(\mathbf{F}(m)) - D_i + m} = \frac{\alpha_i}{\dfrac{r(\mathbf{F}(m)) - \mathbf{D}}{m} + 1}
\end{aligned}$$

By Lemma 9, $r(\mathbf{F}(m))$ is bounded, hence

$$\lim_{m\to\infty} v_i(m) = \lim_{m\to\infty} \frac{\alpha_i}{\dfrac{r(\mathbf{F}(m)) - \mathbf{D}}{m} + 1} = \alpha_i.$$

When the elements in (21) are summed:

$$\begin{aligned}
r(\mathbf{F}(m))\,\mathbf{e}^\top\mathbf{v}(m) = r(\mathbf{F}(m)) &= \mathbf{e}^\top\mathbf{D}\mathbf{v}(m) + m\,\mathbf{e}^\top(\boldsymbol{\alpha} - \mathbf{v}(m)) \\
&= \mathbf{e}^\top\mathbf{D}\mathbf{v}(m) + m(1 - 1) \\
&= \mathbf{e}^\top\mathbf{D}\mathbf{v}(m).
\end{aligned}$$

Hence,

$$\lim_{m\to\infty} r(\mathbf{F}(m)) = \lim_{m\to\infty} \mathbf{e}^\top \mathbf{D}\mathbf{v}(m) = \mathbf{e}^\top \mathbf{D}\boldsymbol{\alpha} = \sum_{i=1}^n D_i\alpha_i.$$

The assumption that $\boldsymbol{\alpha} > \mathbf{0}$ allows some solution to $\sum_{i=1}^n D_i\alpha_i = \lambda$ for any $\lambda \in (\max_i(D_i), \min_i(D_i))$. But for $\lambda = D_{i_{\max}}$ or $\lambda = D_{i_{\min}}$, where $i_{\max}$ and $i_{\min}$ are the indices producing $\max_i(D_i)$ and $\min_i(D_i)$, then $\boldsymbol{\alpha} = \mathbf{e}_{i_{\max}}$ or $\boldsymbol{\alpha} = \mathbf{e}_{i_{\min}}$, respectively. In these cases, $\mathbf{D} + m\,\boldsymbol{\alpha}\mathbf{e}^\top$ is no longer irreducible. The Frobenius normal form for $\mathbf{F}(m) = \mathbf{D} + m(\mathbf{e}_i\mathbf{e}^\top - \mathbf{I})$ is

$$\bar{\mathbf{F}}(m) = \begin{bmatrix} D_1 - m & & & & & & \mathbf{0} \\ & D_2 - m & & & & & \\ & & \ddots & & & & \\ \mathbf{0} & & & D_j - m & & & \\ & & & & \ddots & & \\ m & m & \cdots & m & \cdots & D_i \end{bmatrix}.$$

Hence, the eigenvalues of $\mathbf{D} + m(\mathbf{e}_i\mathbf{e}^\top - \mathbf{I})$ are the diagonal elements $\{D_j - m\colon j \neq i\} \cup \{D_i\}$. Thus when $m > \max_j(D_j - D_i)$ then $r(\mathbf{F}(m)) = D_i > D_j - m$ for all $j \neq i$. So,

$$\lim_{m\to\infty} r\big(\mathbf{D} + m(\mathbf{e}_i\mathbf{e}^\top - \mathbf{I})\big) = D_i.$$

Letting $i = i_{\max}$ and $i = i_{\min}$ completes the construction. $\qquad\square$

Corollary: 11 (Stability). *Let $\mathbf{F}(m) = \mathbf{D} + m\,\mathbf{A}$ as in Theorem 6. Suppose that $r(\mathbf{A}) \leq 0$. If the solution $\mathbf{x}(t) = 0$ is unstable under (3) for some m^*, then it is unstable for all $0 \leq m \leq m^*$.*

Proof. Instability means that $r(\mathbf{F}(m)) > 0$. By Theorem 6, when $r(\mathbf{A}) \leq 0$,

$$\frac{\mathrm{d}\,r(\mathbf{F}(m))}{\mathrm{d}m} \leq r(\mathbf{A}) \leq 0.$$

Hence, if $r(\mathbf{F}(m^*)) > 0$, then $r(\mathbf{F}(m)) \geq r(\mathbf{F}(m^*)) > 0$ for all $0 \leq m \leq m^*$. $\quad\square$

3.4 Additional Results

Corollary: 12 (Heterogeneity). *Under low enough mixing, heterogeneity of growth rates always produces greater asymptotic growth than the average of the growth rates:*

$$\exists m^* > 0\colon\ \forall\, 0 \leq m < m^*\colon r(\mathbf{D} + m\mathbf{A}) > \frac{1}{n}\sum_{i=1}^n D_i.$$

Proof. When $\mathbf{D} \neq c\,\mathbf{I}$ for any $c \in \mathbb{R}$, then $r(\mathbf{D}) = \max_i(D_i) > \frac{1}{n}\sum_{i=1}^n D_i$. Since the eigenvalues are continuous functions of the entries of $\mathbf{D} + m\,\mathbf{A}$, there is some neighborhood $(0, m^*)$ where $r(\mathbf{D} + m\mathbf{A}) > \frac{1}{n}\sum_{i=1}^n D_i$ for $m \in (0, m^*)$. $\quad\square$

Theorem: 13 (Convexity). *Let* $\mathbf{D}$ *be a real diagonal matrix, and* $\mathbf{A}$ *be an irreducible ML-matrix. Then* $r(\mathbf{D} + m\,\mathbf{A})$ *is convex in* m. *Specifically:*

For any $m_1, m_2 \geq 0, m_1 \neq m_2$ *and* $0 < \alpha < 1$, *one has:*

$$
\begin{aligned}
&r((1 - \alpha)\,(\mathbf{D} + m_1\mathbf{A}) + \alpha\,(\mathbf{D} + m_2\mathbf{A})) \\
&\leq (1 - \alpha)\,r(\mathbf{D} + m_1\mathbf{A}) + \alpha\,r(\mathbf{D} + m_2\mathbf{A}),
\end{aligned}
\tag{22}
$$

with equality if and only if $\mathbf{D} = c\,\mathbf{I}$, *for some* $c \in \mathbb{R}$.

Proof. Convexity of $r(\mathbf{D} + m\,\mathbf{A})$ with respect to diagonal matrix $\mathbf{D}$ was established by Cohen (1979, Theorem 3). Specifically, for real diagonal matrices $\mathbf{D}_i$ and ML-matrix $\mathbf{A}$, for $0 < \beta < 1$:

$$
\begin{aligned}
&r((1 - \beta)\mathbf{D}_1 + \beta\mathbf{D}_2 + \mathbf{A})) \\
&\leq (1 - \beta)\,r(\mathbf{D}_1 + \mathbf{A}) + \beta\,r(\mathbf{D}_2 + \mathbf{A}).
\end{aligned}
\tag{23}
$$

Friedland (1981, Theorem 4.1) showed further that equality in (23) obtains if and only if $\mathbf{D}_1 - \mathbf{D}_2 = c\,\mathbf{I}$ for some $c \in \mathbb{R}$.

Convexity with respect to the diagonal matrix implies convexity with respect to m, which can be seen by morphing (22) into (23) through the identity

$$
r(\mathbf{D} + m\,\mathbf{A}) = m\,r\!\left(\tfrac{1}{m}\mathbf{D} + \mathbf{A}\right):
$$

$$
\begin{aligned}
&r(\mathbf{D} + [(1 - \alpha)m_1 + \alpha m_2]\mathbf{A})) \\
&= [(1 - \alpha)m_1 + \alpha m_2]\,r\!\left(\tfrac{1}{[(1-\alpha)m_1+\alpha m_2]}\mathbf{D} + \mathbf{A}\right) \\
&\leq (1 - \alpha)\,r(\mathbf{D} + m_1\mathbf{A}) + \alpha\,r(\mathbf{D} + m_2\mathbf{A}) \\
&= (1 - \alpha)m_1\,r\!\left(\tfrac{1}{m_1}\mathbf{D} + \mathbf{A}\right) \\
&\quad + \alpha\,m_2\,r\!\left(\tfrac{1}{m_2}\mathbf{D} + \mathbf{A}\right).
\end{aligned}
$$

Dividing both sides by $(1 - \alpha)m_1 + \alpha m_2$ gives:

$$
\begin{aligned}
&r\!\left(\tfrac{1}{[(1-\alpha)m_1+\alpha m_2]}\mathbf{D} + \mathbf{A}\right) \\
&\leq \frac{(1 - \alpha)m_1}{(1 - \alpha)m_1 + \alpha m_2}\,r\!\left(\tfrac{1}{m_1}\mathbf{D} + \mathbf{A}\right) \\
&\quad + \frac{\alpha m_2}{(1 - \alpha)m_1 + \alpha m_2}\,r\!\left(\tfrac{1}{m_2}\mathbf{D} + \mathbf{A}\right).
\end{aligned}
\tag{24}
$$

Now, define

$$
\mathbf{D}_1 := \frac{1}{m_1}\mathbf{D}, \;\; \mathbf{D}_2 := \frac{1}{m_2}\mathbf{D},
$$

and

$$
\beta := \frac{\alpha m_2}{(1 - \alpha)m_1 + \alpha m_2}, \;\; \text{so} \;\; 1 - \beta = \frac{(1 - \alpha)m_1}{(1 - \alpha)m_1 + \alpha m_2}.
$$

Thus (24) becomes:

$$
\begin{aligned}
r((1 &- \beta)\mathbf{D}_1 + \beta\mathbf{D}_2 + \mathbf{A}) \\
&\leq (1 - \beta)\ r(\mathbf{D}_1 + \mathbf{A}) + \beta\ r(\mathbf{D}_2 + \mathbf{A}),
\end{aligned}
\tag{25}
$$

which is (23). The equality condition, $\mathbf{D}_1 - \mathbf{D}_2 = c\,\mathbf{I}$ for some $c \in \mathbb{R}$, becomes $\mathbf{D}_1 - \mathbf{D}_2 = (1/m_1 - 1/m_2)\mathbf{D} = c\,\mathbf{I}$. Since $m_1 \neq m_2$, this is precluded if $\mathbf{D} \neq c\,\mathbf{I}$ for any $c \in \mathbb{R}$, in which case $r(\mathbf{F}(m))$ is strictly convex in m. $\square$

4 Discussion

The proximate motivation for extending Karlin's Theorem 5.2 from maps to differential equations was a theoretical study by Steinmeyer and Wilke (2009) on the effect of tissue compartments for antiviral, lethal mutagenesis therapy. The background of the problem is that a number of anti-viral agents seem to work by mutating the virus to inviability. But different tissues can concentrate the therapeutic mutagen to different concentrations. The greater the concentration, the lower that the viral replication rate becomes. Thus the virus will have different growth rates in different compartments.

Steinmeyer and Wilke (2009) ask how heterogeneity in mutagen concentrations, and movement of virions between compartments, affects the dosage needed to cure the infection. They obtain analytic results for a two compartment model, but for more than two compartments must resort to high m and low m limits, course grained approximations, and numerical studies, to obtain results.

Their results exhibit a number of salient features. Here it is noted where each of these features is proven analytically:

1. Compartmental heterogeneity of growth rates increases long-term viral growth rates above the average of the compartment growth rates (Corollary 12);

2. Greater viral mixing between compartments decreases the long-term total viral growth rate (Theorem 6);

3. The decrease in viral growth rate is convex in the amount of mixing (Theorem 13);

4. The whole body viral growth rate for all levels of mixing is bounded above by the maximum growth rate in any compartment (Lemma 9);

5. As the amount mixing of virions between compartments increases, it converges to a weighted average of the compartmental growth rates (Theorem 10).

As can be seen, each of these qualitative features observed in the two-compartment and numerical results of Steinmeyer and Wilke (2009) are in fact properties that extend to any number of compartments, arbitrary growth rates among compartments, patterns of mixing between compartments, and magnitudes of mixing.

4.1 Wider Application

An intimated in Table 1, the results here may find application in a diverse array of systems. The model for viral replication in multiple tissue compartments in Steinmeyer and Wilke (2009) has the same form as ecological models of sink and source populations (Holt, 1996; Armsworth and Roughgarden, 2005).

In addition to linear dynamical systems of the form (2), many nonlinear systems may have (2) as the dynamic of small perturbations. It would be informative to collect other examples of models of the form (2) from the literature in the various fields mentioned in Table 1.

4.2 Further Extensions

One of the important sources of ML-matrices is the numerical solution of second-order partial differential equations. The second derivative can be discretely approximated by the centered difference method, which in one dimension generates a tri-diagonal ML-matrix (Beattie, 2007, p. 38-4), with 1s along the super- and sub-diagonals, and -2 along the diagonal. Thus, a centered difference approximation to

$$\frac{\partial x(s,t)}{\partial t} = g(s)\,x(s,t) + m\,\frac{\partial^2 x(s,t)}{\partial s^2}, \tag{26}$$

with the proper conditions should be of the form (2), and the results here would apply. The second derivative in higher dimensions, such as the Laplace operator, also has discrete approximations that are ML-matrices (Ng 2007, pp. 40-1 – 40-2; Greenbaum 2007, pp. 41-1 – 41-2), so one can seek analogous results in higher dimensional diffusions.

Under the analytic assumptions for which the solutions of the centered difference approximation converge to positive eigenfunction solutions of (26), one can expect that the Perron root will be a decreasing function of m.

Here I have only touched upon some applications and extensions one might find for the results presented. The fact that Karlin's result — that mixing reduces growth — extends from discrete-time/discrete space systems to continuous-time/discrete space systems, and promises to extend further to continuous time and space systems, and to diffusion operators, suggests something fundamental in the phenomenon.

Acknowledgements

References

Altenberg, L., 1984. A Generalization of Theory on the Evolution of Modifier Genes. Ph.D. thesis, Stanford University. Searchable online and available from University Microfilms, Ann Arbor, MI.

Altenberg, L. 2009. The evolutionary reduction principle for linear variation in genetic transmission. Bulletin of Mathematical Biology 71:1264–1284.

Altenberg, L. and Feldman, M. W. 1987. Selection, generalized transmission, and the evolution of modifier genes. I. The reduction principle. Genetics 117:559–572.

Armsworth, P. R. and Roughgarden, J. E. 2005. The impact of directed versus random movement on population dynamics and biodiversity patterns. American Naturalist 165.

Bapat, R. B. and Raghavan, T. E. S., 1997. Nonnegative Matrices and Applications. Cambridge University Press, Cambridge, UK.

Beattie, C., 2007. Matrix factorizations and direct solution of linear systems. Pages 38–1–38–17 *in* L. Hogben, ed. Handbook of Linear Algebra. Chapman and Hall.

Bernstein, D. S., 2009. Matrix Mathematics: Theorem, Facts, and Formulas. Princeton University Press, Princeton, 2nd edition edition. ISBN 978-0-691-13287-7.

Caswell, H., 2000. Matrix Population Models. Sinauer Associates, 2nd edition. ISBN 978-0878931217.

Cohen, J. E. 1978. Derivatives of the spectral radius as a function of non-negative matrix elements. Math. Proc. Camb. Phil. Soc. 83:183–190.

Cohen, J. E. 1979. Random evolutions and the spectral radius of a non-negative matrix. Math. Proc. Camb. Phil. Soc. 86:345–350.

Donsker, M. D. and Varadhan, S. R. S. 1975. On a variational formula for the principal eigenvalue for operators with maximum principle. Proceedings of the National Academy of Sciences U.S.A. 72:780–783.

Friedland, S. 1981. Convex spectral functions. Linear and Multilinear Algebra 9:299–316.

Friedland, S. and Karlin, S. 1975. Some inequalities for the spectral radius of non-negative matrices and applications. Duke Mathematical Journal 42:459–490.

Gantmacher, F. R., 1959*a*. The Theory of Matrices, volume 1. Chelsea Publishing Company, New York.

Gantmacher, F. R., 1959*b*. The Theory of Matrices, volume 2. Chelsea Publishing Company, New York.

Greenbaum, A., 2007. Iterative solution methods for linear systems. Pages 41–1 – 41–20 *in* L. Hogben, ed. Handbook of Linear Algebra. Chapman and Hall.

Holt, R. D. 1996. Adaptive evolution in source-sink environments: Direct and indirect effects of density-dependence on niche evolution. Oikos 75:182–192.

Karlin, S., 1982. Classification of selection-migration structures and conditions for a protected polymorphism. Pages 61–204 *in* M. K. Hecht, B. Wallace, and G. T. Prance, eds. Evolutionary Biology, volume 14. Plenum Publishing Corporation.

Ng, E. G., 2007. Sparse matrix methods. Pages 40–1 – 40–20 *in* L. Hogben, ed. Handbook of Linear Algebra. Chapman and Hall.

Seneta, E., 1981. Non-negative Matrices and Markov Chains. Springer-Verlag, New York.

Steinmeyer, S. H. and Wilke, C. O. 2009. Lethal mutagenesis in a structured environment. Journal of Theoretical Biology 261:67–73.