Theory of Growth and Mixing


http://dynamics.org/Altenberg/PAPERS/

Lee Altenberg
altenber@hawaii.edu
http://dynamics.org/Altenberg/

March 16, 2012
Fundamental Actions in Biology

- There are two fundamental processes in biology, and in many physical systems as well:
  1. **Growth** (or negative growth, i.e. decay):
     i.e. Changes in *number*
  2. **Transformation**:
     i.e. Changes in *state*

- By *mixing* here, I refer to transformation between a set of states that already exist, rather than generation of novel states.
The Essence

- **Definitions:**
  - States are $i, j$,
  - $x_i$ is the quantity of state $i$,
  - time is $t$:

- **Growth:**
  $$x_i(t+1) = D_i \ x_i(t)$$

- **Mixing:**
  $$x_i(t+1) = \sum_j M_{ij} \ x_j(t)$$

- **Combined:**
  $$z(t+1) = \mathbf{M} \mathbf{D} \ z(t)$$
## Where Growth and Mixing Combine

<table>
<thead>
<tr>
<th>Objects:</th>
<th>States/Sites:</th>
<th>Site specific growth/decay rates:</th>
</tr>
</thead>
<tbody>
<tr>
<td>genes</td>
<td>habitats</td>
<td>fitnesses</td>
</tr>
<tr>
<td>genomes</td>
<td>genotypes</td>
<td>fitnesses</td>
</tr>
<tr>
<td>organisms</td>
<td>habitats</td>
<td>species survival and reproduction rates</td>
</tr>
<tr>
<td>agricultural pests</td>
<td>fields under treatment</td>
<td>replication and survival</td>
</tr>
<tr>
<td>microbes</td>
<td>tissue compartments</td>
<td>survival and replication rates</td>
</tr>
<tr>
<td>metabolites</td>
<td>tissue compartments</td>
<td>catabolic rates</td>
</tr>
<tr>
<td>reactants</td>
<td>reaction-diffusion medium</td>
<td>reaction rates</td>
</tr>
<tr>
<td>wastes</td>
<td>reactors</td>
<td>waste breakdown rates</td>
</tr>
<tr>
<td>photons</td>
<td>media</td>
<td>absorption rates</td>
</tr>
<tr>
<td>particles</td>
<td>heterogenous matter</td>
<td>interaction and decay rates</td>
</tr>
<tr>
<td>capital</td>
<td>investments</td>
<td>rates of return</td>
</tr>
</tbody>
</table>
The late Sam Karlin at "Feldmania", November, 2007, Stanford University, a meeting upon the 65th birthday of Sam's advisee and my advisor, Marc Feldman.


Sam Karlin, a mathematician with diverse interests, decided to take on population genetics (1966).

- Karlin and his postdoc Shmuel Friedland developed a first installment of general results.
First, some definitions:

- **Nonnegative matrix** \( A \) is **irreducible** if for each \( i, j \), there exists some \( t \geq 1 \): \([A^t]_{ij} > 0\)

- The spectral radius

\[
\rho(MD) := \max_i |\lambda_i(MD)|,
\]

the largest modulus (magnitude) of any eigenvalue of \( A \).

- **Facts for nonnegative irreducible matrix** \( A \):
  - \( \rho(A) \) is an eigenvalue of \( A \) (Perron-Frobenius Theory)
  - \( \rho(A) > 0 \)
  - \( \rho(A) \) is the asymptotic growth rate of a system

\[
\mathbf{x}(t+1) = A \mathbf{x}(t).
\]
Definitions:

- **Perron vectors**: the left and right eigenvectors, $\mathbf{u}^\top$, and $\mathbf{v}$, associated with $\rho(\mathbf{A})$.

  \[ \mathbf{u}^\top \mathbf{A} = \rho(\mathbf{A}) \mathbf{u}^\top, \quad \mathbf{A} \mathbf{v} = \rho(\mathbf{A}) \mathbf{v} \]

- Facts from Perron-Frobenius Theory:
  \[ \mathbf{u}^\top > 0^\top, \quad \mathbf{v} > 0 \]
Theorem (3.1, p. 462)

Let

- $\mathbf{M}$ be an $n \times n$ irreducible non-negative matrix
- $\mathbf{u}^\top$ and $\mathbf{v}$ be the left and right Perron vectors, where $\mathbf{u}^\top \mathbf{v} = 1$.

Then for any positive diagonal matrix $\mathbf{D} \equiv \text{diag} [D_1, \ldots, D_n]$, we have

$$
\rho(\mathbf{DM}) = \rho(\mathbf{MD}) \geq \rho(\mathbf{M}) \prod_{i=1}^{n} D_i^{u_i v_i}.
$$


- A second installment of general results!
- Motivation: to analyze the effect of population subdivision and dispersal on the maintenance of genetic variation by selection.
Two extremely general theorems on greater or less mixing combined with differential growth:

**Theorem (Karlin’s Theorem 5.1, 1982)**

Consider a family of stochastic matrices that commute and are symmetrizable to positive definite matrices:

\[ \mathcal{F} := \{ M_i = D_1 S_i D_2 : M_i M_j = M_j M_i \} , \]

where \( D_1 \) and \( D_2 \) are positive diagonal matrices, and each \( S_i \) is a positive definite symmetric real matrix. Let \( D \) be a positive diagonal matrix. Then for each \( M_i, M_j \in \mathcal{F} \), the spectral radius, \( \rho \), satisfies:

\[ \rho(M_i M_j D) \leq \rho(M_j D) . \]
Theorem (Karlin’s Theorem 5.2, 1982)

Let $\mathbf{M}$ be a non-negative irreducible stochastic matrix. Consider the family of matrices

$$\mathbf{M}(\alpha) = (1-\alpha)\mathbf{I} + \alpha \mathbf{M}, \quad 0 \leq \alpha \leq 1.$$

Then for any positive diagonal matrix $\mathbf{D}$, the spectral radius

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

is decreasing as $\alpha$ increases (strictly provided $\mathbf{D} \neq d \mathbf{I}$).
Karlin’s Theorem 5.2 applies to strongly connected dispersal networks of arbitrary complexity and asymmetry:

Karlin’s proof of Theorem 5.2 used the Donsker-Varadhan (1975) variational formula for the spectral radius:

\[
\rho(A) = \sup_{p \in \mathcal{P}_n} \inf_{x > 0} \sum_{i=1}^{n} p_i \frac{[Ax]_i}{x_i} \quad \text{or} \quad \\
\log \rho(A) = \sup_{p \in \mathcal{P}_n} \inf_{x > 0} \sum_{i=1}^{n} p_i \log \frac{[Ax]_i}{x_i}
\]

where \( A \) is nonnegative, \( \mathcal{P} \) is the space of probability vectors, \( p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \).

Donsker and Varadhan: $\rho(A) = \text{saddle-point of } \sum_i p_i \frac{[Ax]_i}{x_i}$
The Reduction Principle of Marc Feldman

In early theoretical work to understand the evolution of genetic systems, Feldman, colleagues, and others kept finding a common result from each model they examined — be they models for the evolution of recombination, or of mutation, or of dispersal.

Evolution favored reduced levels of these processes in populations near equilibrium under constant environments, and this result was called the Reduction Principle (Feldman, Christiansen, and Brooks, 1980).
Altenberg (1984) found that the models exhibiting the Reduction Principle were of the form
\[ M(\alpha) = (1 - \alpha)I + \alpha P. \]

Thus Karlin’s Theorem 5.2 explains the repeated appearance of the reduction result in the different contexts, and generalizes the result to a whole class of genetic transmission patterns beyond the special cases that had been analyzed.

- e.g. the cultural transmission of traditionalism (Altenberg, 1984)
The Reduction Principle

In detail:

- Suppose the mixing rate $\alpha$ in $\rho(M(\alpha)D)$ is not an extrinsic parameter, but is a variable which is *itself controlled by a gene*.

- Therefore from Karlin’s Theorem 5.2, a gene which decreases $\alpha$ will have a growth advantage over its competitor alleles.

- The action of such modifier genes produces a process that will reduce the rates of dispersal in a population.

- Hence, Theorem 5.2 also means that *differential growth selects for reduced mixing*. 
These results were found for finite-dimensional models.

But the same reduction result has also been found in models for the evolution of dispersal in **continuous space**, in which matrices are replaced by **linear operators**.

This raises the questions of whether this common result, discovered in such a diversity of models, reflects **a single mathematical phenomenon**.

Here, the question is answered affirmatively.
Do all paths lead to Kingman?!?!

**Theorem (Kingman, 1961)**

Let $\mathbf{A}$ be an $n \times n$ matrix whose elements, $A_{ij}(\theta)$, are nonnegative functions of the real variable $\theta$, such that they are ‘superconvex’:

- i.e. for each $i, j$, either $\log A_{ij}(\theta)$ is convex in $\theta$ [$A_{ij}(\theta)$ is log convex], or $A_{ij}(\theta) = 0$ for all $\theta$.

Then the spectral radius of $\mathbf{A}$ is also superconvex in $\theta$. 
**Mathematical path to the new result**

Next: Cohen (1981). But first, more definitions:

- An *essentially negative* matrix is $A + D$, where
  - $A$ is nonnegative and
  - $D$ is any real diagonal matrix.

- The *spectral bound*, $s(A)$, of a matrix $A$, is
  
  $$ s(A) := \max_i \text{Re}(\lambda_i(A)) $$

- For essentially nonnegative matrix $A$, $s(A)$ is an eigenvalue (by Perron-Frobenius theory).
Joel Cohen used Kingman’s theorem to prove this theorem:

**Theorem (Cohen, 1981)**

- Let $\mathbf{A}$ be an essentially nonnegative $n \times n$ matrix.
- Let $\mathbf{D}$ be diagonal real $n \times n$ matrix.

Then $s(\mathbf{A} + \mathbf{D})$ is a convex function of $\mathbf{D}$.

i.e.

$$s(\mathbf{A} + (1 - \alpha)\mathbf{D}_1 + \alpha\mathbf{D}_2) \leq (1 - \alpha)s(\mathbf{A} + \mathbf{D}_1) + \alpha s(\mathbf{A} + \mathbf{D}_2)$$

Equivalently: $s(\mathbf{A} + \beta\mathbf{D})$ is convex in $\beta$. 
First, a rather long list of definitions from the theory of **operators on Banach spaces**:

**Banach Space:** a complete, normed, linear vector space (possibly infinite dimensional). Define:

- $X$ an ordered Banach space or its complexification.
- $X_+$ the proper, closed, positive cone of $X$
- $B(X)$ the set of all bounded linear operators $A: X \to X$.
- $A$ is a **positive operator** if $AX_+ \subset X_+$.
- $C_0 : t \mapsto e^{tA}$, a strongly continuous semigroup generated by $A$ (i.e. a linear dynamical system you might not be able to run backwards).
$R(\xi, A) := (\xi - A)^{-1}$, the resolvent of $A$, where $\xi \in \mathbb{C}$.

- **Resolvent set** $\rho(A) \subset \mathbb{C}$ is values of $\xi$ for which $\xi - A$ is invertible.

- **Spectrum** of $A$, $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is the complement of the resolvent set.

- **Spectral bound** of closed linear operator $A$ is

$$s(A) := \begin{cases} 
\sup \{ \Re(\lambda) : \lambda \in \sigma(A) \} & \text{if } \sigma(A) \neq \emptyset \\
-\infty & \text{if } \sigma(A) = \emptyset.
\end{cases}$$

- **Spectral radius** $r(A) := \sup \{ |\lambda| : \lambda \in \sigma(A) \}$.

- **Type or growth bound** of $A$, when it's an infinitesimal generator of a $C_0$-semigroup, $\{ e^{tA} : t > 0 \}$:

$$\omega(A) := \lim_{t \to \infty} \frac{1}{t} \log \| e^{tA} \| = \log r(e^A).$$
Definition (Resolvent Positive Operator)

Operator $A$ is *resolvent positive* if there is $\xi_0$ such that for all $\xi > \xi_0$:

- $\xi - A$ is invertible, and
- $R(\xi, A) = (\xi - A)^{-1}$ is positive.
Well-known examples of resolvent positive operators:

- Schrödinger operators

\[- \frac{1}{2} \Delta + V \quad \text{on } L^p(\mathbb{R}^N),\]

where

- \( \Delta = \sum_{i=1}^{N} \frac{\partial^2}{\partial x_i^2} \) is the Laplace operator, and
- \( V \) is an operator of multiplication with constraints depending on \( p \) in \( L^p(\mathbb{R}^N) \).
Common Examples

- Second-order elliptic operators on $L^p(\Omega)$,

$$A = -\sum_{j,k=1}^{N} \frac{\partial}{\partial x_k} \left( a_{jk} \frac{\partial}{\partial x_j} \right) + \sum_{j=1}^{N} b_j \frac{\partial}{\partial x_j} + \frac{\partial}{\partial x_j}(c_j.) + a_0$$

where $\Omega \subset \mathbb{R}^N$ is open, coefficients are measurable and bounded, ellipticity conditions apply to $a_{jk}(x)$, and appropriate additional conditions hold for the coefficients, domain and boundary.
Common Examples

- Linear integral operators $A$ on $X = C(\overline{\Omega})$ defined by

\[(Af)(x) := \int_{\Omega} K(x, y) f(y) \, dy + b(x) f(x),\]

where $K \in C(\overline{\Omega} \times \overline{\Omega}, \mathbb{R}^+)$, $\Omega \subset \mathbb{R}^N$ is bounded, and $K(x, y) > 0$, $b(x)$ are measurable functions for $x, y \in \overline{\Omega}$.
Familiar Operator Properties

Relationship to other familiar operator properties:

- Differential operators higher than second order are never resolvent positive.

- If $A$ generates a $C_0$-semigroup $T_t$, then $T_t$ is positive for all $t \geq 0$ if and only if $A$ is resolvent positive.

- If $A$ is a resolvent positive operator defined densely on $X = C(S)$ (the Banach space of continuous complex-valued functions on compact space $S$), then $A$ generates a positive $C_0$-semigroup.
Familiar Operator Properties

- If $A$ is resolvent positive and its domain, $D(A) \subset X$, is dense in $X$, then for every $f \in D(A^2)$, there exists a unique solution to the Cauchy problem

\[ \frac{\partial u}{\partial t} = Au(t) \quad (t \geq 0), \quad u(0) = f. \]

- If $A$ is resolvent positive then:
  - $s(A) < \infty$;
  - if $\sigma(A)$ is nonempty, i.e. $-\infty < s(A)$, then $s(A) \in \sigma(A)$;
  - if $R(\xi, A) \geq 0$ on $\xi \in \mathbb{R} \cap \rho(A)$, then $\xi > s(A)$. 

**Theorem (Kato’s Generalized Cohen’s Theorem, 1982)**

Consider $X = C(S)$ (continuous functions on a compact Hausdorff space $S$) or $X = L^p(S)$, $1 \leq p < \infty$, on a measure space $S$.

Let $A: X \to X$ be a linear operator which is resolvent positive. Let $V$ be an operator of multiplication on $X$ represented by a real-valued function $v$, where $v \in C(S)$ for $X = C(S)$, or $v \in L^\infty(S)$ for the other cases.

Then $s(A + V)$ is a convex function of $V$. If in particular $A$ is a generator of a $C_0$ semigroup, then both $s(A + V)$ and $\omega(A + V)$ are convex in $V$. 
New Result: The Reduction Phenomenon for Resolvent Positive Operators

Theorem (Generalized Karlin’s theorem)

Let $A$ be a resolvent positive linear operator, and $V$ be an operator of multiplication, under the same assumptions as Kato’s Theorem. Then for $\alpha > 0$,

1. $s(\alpha A + V)$ is convex in $\alpha$;

2. For each $\alpha > 0$,

$$\frac{d}{d\alpha} s(\alpha A + V) \leq s(A).$$

If $A$ is a generator of a $C_0$-semigroup, the above apply to both $s(\alpha A + V)$ and $\omega(\alpha A + V)$. 
In particular:

- when \( s(A) = 0 \) then \( s(\alpha A + V) \) is non-increasing in \( \alpha \), and
- when \( s(A) < 0 \) then \( s(\alpha A + V) \) is strictly decreasing in \( \alpha \).

Diffusions with Dirichlet boundary conditions (\( f : \partial \Omega \mapsto 0 \)) typically have \( s(A) < 0 \).

In other words, *greater mixing reduces growth or hastens decay.*
Method of Proof

Lemma (Dual Convexity)

Let $f : (0, \infty) \times [0, \infty) \rightarrow \mathbb{R}$ have the following properties:

$$f(\alpha x, \alpha y) = \alpha f(x, y), \text{ for } \alpha > 0,$$
and

$$f(x, y) \text{ is convex in } y.$$

Then:

- $f(x, y)$ is convex in $x$;
- For each $x \in (0, \infty)$,

$$\frac{\partial}{\partial x} f(x, y) \leq f(1, 0).$$

*
Dual Convexity: *Detailed Version

- For each \( x \in (0, \infty) \), either
  - \( f(x + d, y) < f(x, y) + d f(1, 0) \quad \forall \ d \in (0, \infty) \); or
  - \( f(x + d, y) = f(x, y) + d f(1, 0) \quad \forall \ d \in (0, \infty) \).

- For \( y \neq 0 \), if \( f(x, y) \) is strictly convex in \( y \), then \( f(x, y) \) is strictly convex in \( x \), and
  \( f(x + d, y) < f(x, y) + d f(1, 0) \).

- For each \( x \in (0, \infty) \),
  \[
  \frac{\partial}{\partial x} f(x, y) \leq f(1, 0),
  \]
  except possibly at a countable number of points \( x \), where the one-sided derivatives exist but differ:
  \[
  \frac{\partial}{\partial x_-} f(x, y) < \frac{\partial}{\partial x_+} f(x, y) \leq f(1, 0).
  \]
Proof.

Homogeneity allows two rescalings:

**Key rescaling # 1:** Convexity in $y \implies$ convexity in $x$

$$(1-m)f(x, y_1) + mf(x, y_2) \geq f(x, [1-m]y_1 + my_2)$$

$$\iff$$

$$(1 - m) \frac{y_1}{y} f\left(\frac{xy}{y_1}, y\right) + m \frac{y_2}{y} f\left(\frac{xy}{y_2}, y\right) \geq \frac{(1-m)y_1 + my_2}{y} f\left(\frac{xy}{(1-m)y_1 + my_2}, y\right).$$
Dual Convexity, cont’d

Proof, continued.

Key rescaling # 2: \[ \frac{f(x+d, y) - f(x, y)}{d} \leq f(1, 0) \quad \iff \]

\[ f(x+d, y) \leq f(x, y) + d f(1, 0) \quad \iff \]

\[ (x+d) f(1, \frac{y}{x+d}) \leq x f(1, \frac{y}{x}) + d f(1, 0) \quad \iff \]

\[ f(1, \frac{y}{x+d}) \leq \frac{x}{x+d} f(1, \frac{y}{x}) + \frac{d}{x+d} f(1, 0). \]

i.e. convexity of \( f(x, y) \) in \( y \), since the 2nd arguments for \( f \) are a convex combination:

\[ \frac{y}{x+d} = \frac{x}{x+d} \frac{y}{x} + \left(1 - \frac{x}{x+d}\right) \ast 0. \]
To Summarize:

- So, Kato showed that $s(\alpha A + \beta V)$ is convex in $\beta$.
- With the Dual Convexity Lemma, this implies that
  1. $s(\alpha A + \beta V)$ is convex in $\alpha$, and
  2. $\frac{\partial}{\partial \alpha} s(\alpha A + \beta V) \leq s(A)$,

  which is Karlin’s theorem generalized to resolvent positive operators.
The expression in Karlin’s Theorem 5.2 can be put into the form:

\[
M(\alpha)D = [(1-\alpha)I + \alpha P]D = \alpha (P-I)D + D = \alpha A + \beta D.
\]

where \( A = (P-I)D \), \( \alpha \in (0, 1) \), and \( \beta = 1 \).
A Key Open Problem

In some physical systems, and in biological applications especially, there may be

- multiple, independently varied operators acting on a quantity (e.g. diffusion with independent advection or

- the variation may not scale the mixing process uniformly (e.g. conditional dispersal ),

so that variation is not of the form $\alpha A + V$ but rather $\alpha A + B$, where $B$ is a linear operator other than an operator of multiplication.
A Key Open Problem

- Examples are known where departures from reduction occur, i.e. $ds(\alpha A + B)/d\alpha > s(A)$.

- Results for the form $\alpha A + B$ have been obtained for symmetrizable finite matrices in models of multilocus mutation, and dispersal in random environments.

A Key Open Problem

- A key open problem, then, is to find necessary or sufficient conditions on Banach space operators, $B$, such that $\frac{\partial s(\alpha A + \beta B)}{\partial \alpha} \leq s(A)$ (which may depend on $A$, $\beta/\alpha$, domain, and boundary conditions).

- A sufficient condition is that $s(\alpha A + \beta B)$ be convex in $\beta$, by Dual Convexity.

- Thus, the dual problem is to ask: for which $B$ is $s(\alpha A + \beta B)$ convex in $\beta$?
A Key Open Problem

- Kato’s Theorem used operators of multiplication, \( V \), because the family of operators \( e^{\beta V} \) is semigroup-superconvex in \( \beta \) (definition in Kato).

- But this approach faces the challenge that, “It is in general difficult to find a nontrivial semigroup-superconvex family \( B(h) \).” —Kato (1982)

- I’ll give Sir John Kingman the last word: “It looks very well, and will I hope attract some interest to these problems. Your open problem looks difficult, but not impossible.”

Thank you for your attention!
References


References, cont’d


