Resolvent Positive Linear Operators Exhibit the Reduction Phenomenon

Lee Altenberg
altenber@hawaii.edu
February 27, 2012

Abstract

The spectral bound, $s(\alpha A + \beta V)$, of a combination of a resolvent positive linear operator A and an operator of multiplication V, was shown by Kato to be convex in $\beta \in \mathbb{R}$. This is shown here to imply, through an elementary 'dual convexity' lemma, that $s(\alpha A + \beta V)$ is also convex in $\alpha > 0$, and notably, $\partial\, s(\alpha A + \beta V)/\partial\alpha \leq s(A)$. Diffusions typically have $s(A) \leq 0$, so that for diffusions with spatially heterogeneous growth or decay rates, *greater mixing reduces growth*. Models of the evolution of dispersal in particular have found this result when A is a Laplacian or second-order elliptic operator, or a nonlocal diffusion operator, implying selection for reduced dispersal. These cases are shown here to be part of a single, broadly general, 'reduction' phenomenon.[1][2]

Keywords: perturbation theory — positive semigroup — reduction principle — non-self-adjoint — Schrödinger operator

The main result to be shown here is that the growth bound, $\omega(\alpha A + V)$, of a positive semigroup generated by $\alpha A + V$ changes with positive scalar α at a rate less than or equal to $\omega(A)$, where A is also a generator, and V is an operator of multiplication. Movement of a reactant in a heterogeneous environment is often of this form, where V represents the local growth or decay rate, and α represents the rate of mixing. Lossless mixing means $\omega(A) = 0$, while lossy mixing means $\omega(A) < 0$, so this result implies that greater mixing reduces the reactant's asymptotic growth rate, or increases its asymptotic decay rate. This is a familiar result when A is a diffusion operator, so what is new here is the generality shown for this phenomenon. At the root of this result is a theorem by Kingman on the 'superconvexity' of the spectral radius of nonnegative matrices [1]. The logical route progresses from Kingman through Cohen [2] to Kato [3]. The historical route begins in population genetics.

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 [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14] — 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* [11].

These results were found for finite-dimensional models. But the same reduction result has also been found in models for the evolution of unconditional 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.

The mathematical underpinnings of the reduction principle for finite-dimensional models were discovered by Sam Karlin [15, 16] (although he did not realize it, and he had earlier proposed an alternate to the reduction principle — the *mean fitness principle* [17], which was found to have counterexamples [18]). Karlin wanted to understand the effect of population subdivision on the maintenance of genetic variation. Genetic variation is preserved if an allele has a positive growth rate when it is rare, protecting it from extinction. The dynamics of a rare allele are approximately linear, and of the form

$$\mathbf{x}(t{+}1) = [(1{-}\alpha)\mathbf{I} + \alpha\mathbf{P}]\mathbf{D}\,\mathbf{x}(t), \tag{1}$$

where $\mathbf{x}(t)$ is a vector of the rare allele's frequency at time t among different population subdivisions, α is the rate of dispersal between subdivisions, $\mathbf{P}$ is the stochastic matrix representing the pattern of dispersal, and $\mathbf{D}$ is a diagonal matrix of the growth rates of the allele in each subdivision. The allele is protected from extinction if its asymptotic growth rate when rare is greater than 1. This asymptotic growth rate is the spectral radius,

$$r(\mathbf{A}) := \sup\{|\lambda| \colon \lambda \in \sigma(\mathbf{A})\}, \tag{2}$$

where $\sigma(\mathbf{A})$ is the set of eigenvalues of matrix $\mathbf{A}$.

[1] Dedicated to Sir John F. C. Kingman on the fiftieth anniversary his theorem on the 'superconvexity' of the spectral radius [1], which is at the root of the results presented here.

[2] *Proceedings of the National Academy of Sciences U.S.A.*, in press.

Karlin discovered that for $\mathbf{M}(\alpha) := [(1-\alpha)\mathbf{I} + \alpha\mathbf{P}]$, the spectral radius, $r(\mathbf{M}(\alpha)\mathbf{D})$, is a decreasing function of the dispersal rate α, for arbitrary strongly-connected dispersal pattern:

Theorem 1 (Karlin's Theorem 5.2) *[16, pp. 194–196] Let* $\mathbf{P}$ *be an arbitrary non-negative irreducible stochastic matrix. Consider the family of matrices*

$$\mathbf{M}(\alpha) = (1-\alpha)\mathbf{I} + \alpha\mathbf{P}, \qquad 0 < \alpha < 1.$$

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

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

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

Karlin's Theorem 5.2 means that greater mixing between subdivisions produces lower $r(\mathbf{M}(\alpha)\mathbf{D})$, and if it goes below 1, the allele will go extinct. While this theorem was motivated by the issue of genetic diversity in a subdivided population, its form applies generally to situations where differential growth is combined with mixing. $\mathbf{D}$ could just as well represent the investment returns on different assets and $\mathbf{P}$ a pattern of portfolio rebalancing. Or $\mathbf{D}$ could represent the decay rates of reactant in different parts of a reactor, and $\mathbf{P}$ a pattern of stirring within the reactor. In a very general interpretation, Theorem 5.2 means that *greater mixing reduces growth and hastens decay*.

If the dispersal rate α is not an extrinsic parameter, but is a variable which is itself controlled by a gene, then a gene which decreases α 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. Therefore, Theorem 5.2 also means that *differential growth selects for reduced mixing*.

In the evolutionary context, the generality of the mixing pattern $\mathbf{P}$ in Karlin's Theorem 5.2 makes it applicable to other kinds of 'mixing' besides dispersal. The pattern matrix $\mathbf{P}$ can just as well refer to the pattern of mutations between genotypes, and then α refers to the mutation rate. Or $\mathbf{P}$ can represent the pattern of transmission when two loci recombine, and then α represents the recombination rate. The early models for the evolution of recombination and mutation that exhibited the reduction principle in fact had the same form as (1) for the dynamics of a rare modifier allele. Once this was recognized [19, 20, 21], it was clear that Karlin's theorem explained the repeated appearance of the reduction result in the different contexts, and generalized the result to a whole class of genetic transmission patterns beyond the special cases that had been analyzed.

The dynamics of movement in space have been long modeled by infinite-dimensional models, where space is continuous and the concentrations of a quantity at each point are represented as a function. The dynamics of change in the concentration are modeled as diffusions, where the Laplacian or elliptic differential operator or nonlocal integral operator takes the place of the matrix $\mathbf{P}$ in the finite-dimensional case. When the substance grows or decays at rates that are a function of its location, the system is often referred to as a reaction-diffusion. In reaction-diffusion models for the evolution of dispersal, the reduction principle again makes its appearance [22][23, Lemma 5.2] [24, Lemma 2.1][25]. In nonlocal diffusion models, again the reduction principle appears [26]. This points to the possibility of an underlying mathematical unity.

Here, a broad characterization of this 'reduction phenomenon' is established by generalizing Karlin's theorem to linear operators. The reduction results previously found for various linear operators are, therefore, seen to be special cases of a general phenomenon.

This result is actually implicit in Kato's generalization [3] of Cohen's theorem [2] on the convexity of the spectral bound of essentially nonnegative matrices with respect to the diagonal elements of the matrix. It is educed from Kato's theorem here by means of an elementary 'dual convexity' lemma.

Kato's goal in [3] was to generalize, from matrices to linear operators, Cohen's convexity result [2]:

Theorem 2 (Cohen) *[2] Let* $\mathbf{D}$ *be diagonal real* $n \times n$ *matrix. Let* $\mathbf{A}$ *be an essentially nonnegative* $n \times n$ *matrix. Then* $s(\mathbf{A}+\mathbf{D})$ *is a convex function of* $\mathbf{D}$*.*

Here, $s(\mathbf{A}+\mathbf{D})$ is the spectral bound — the largest real part of any eigenvalue of $\mathbf{A}+\mathbf{D}$. A synonym for the spectral bound used in the matrix literature is the *spectral abscissa* [27, 28]. When the spectral bound is an eigenvalue, it is also referred to as the *principal eigenvalue* [29], *dominant eigenvalue* [30], *dominant root* [31], *Perron-Frobenius eigenvalue* [32], or *Perron root* [33]. 'Essentially nonnegative' means that the off-diagonal elements are nonnegative. Synonyms include 'quasi-positive' [34], 'Metzler', 'Metzler-Leontief', 'ML' [32], and 'cooperative' [35].

Cohen's proof relied upon the following theorem of Kingman:

Theorem 3 (Kingman) *[1] Let* $\mathbf{A}$ *be an* $n \times n$ *matrix whose elements,* $A_{ij}(\theta)$*, are nonnegative functions of the real variable* θ*, such that they are 'superconvex', i.e. for each* i, j*, either* $\log A_{ij}(\theta)$ *is convex in* θ *[*$A_{ij}(\theta)$ *is* log *convex], or* $A_{ij}(\theta) = 0$ *for all* θ*. Then the spectral radius of* $\mathbf{A}$ *is also superconvex in* θ*.*

Kato generalized Cohen's result to linear operators by first generalizing Kingman's theorem. Before presenting Kato's theorem, some terminology needs to be introduced:

X represents an ordered Banach space or its complexification.

X_+ represents the proper, closed, positive cone of X, assumed to be generating and normal (see [3]).

$B(X)$ represents the set of all bounded linear operators $A\colon X \to X$.

A is a *positive operator* if $AX_+ \subset X_+$.

The resolvent of A is $R(\xi, A) := (\xi - A)^{-1}$, the operator inverse of $\xi - A$, $\xi \in \mathbb{C}$.

The resolvent set $\rho(A) \subset \mathbb{C}$ are those values of ξ for which $\xi - A$ is invertible.

The spectrum of $A \in B(X)$, $\sigma(A)$, is the complement of the resolvent set, $\rho(A)$.

The *spectral bound* of closed linear operator A, not necessarily bounded, is

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

The type (growth bound) of an infinitesimal generator, A, of a strongly continuous (C_0) semigroup, $\{e^{tA} \colon t > 0\}$, is

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

Generally, $-\infty \leq s(A) \leq \omega(A) < \infty$, but conditions for $s(A) = \omega(A)$ or $s(A) < \omega(A)$ are part of a more involved theory for the asymptotic growth of semigroups (see [36, 37, 38]).

Definition 1 *Operator A is resolvent positive if there is ξ_0 such that $(\xi_0, \infty) \subset \rho(A)$ and $R(\xi, A)$ is positive for all $\xi > \xi_0$ [39].*

The relationship of the resolvent positive property to other familiar operator properties includes the following list of key results:

1. 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 [38, p. 188].

2. 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 [38, Theorem 3.11.9].

3. 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, $u(t) \in D(A)$ for all $t \geq 0$, $u \in C^1([0, \infty), X)$, to the Cauchy problem [39, Theorem 7.1]

$$\frac{\partial u}{\partial t} = Au(t) \qquad (t \geq 0), \qquad u(0) = f.$$

4. 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 $\xi \in \mathbb{R} \cap \rho(A)$ yields $R(\xi, A) \geq 0$ then $\xi > s(A)$ [3] [38, Proposition 3.11.2].

5. Differential operators higher than second order are never resolvent positive [40, Corollary 2.3][41].

6. Well-known examples of resolvent positive operators include the following (for details see the sample references).

   (a) Schrödinger operators $-\frac{1}{2}\Delta + V$ on $L^p(\mathbb{R}^N)$, where $\Delta = \sum_{i=1}^N \partial^2/\partial x_i^2$ is the Laplace operator, and V is an operator of multiplication with constraints depending on p (see [42, 43, 44]).

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

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

   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 ([45], also e.g. [46, 3, 47]).

   (c) 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 $x, y \in \overline{\Omega}$ [26, 48, 49]. A resolvent positive combination of integral and differential operator is analyzed in [50].

Kato's generalization of Cohen's theorem is as follows:

Theorem 4 (Generalized Cohen's Theorem) *[3]*
Consider $X = C(S)$ (continuous functions on a compact Hausdorff space S) or $X = L^p(S)$, $l \leq p < \infty$, on a measure space S, or more generally, let X be the intersection of two L^p-spaces with different p's and different weight functions. Let $A\colon X \to X$ 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.

Kato's theorem is further generalized to Banach lattices by Arendt and Batty [51] as follows:

Theorem 5 (Generalized Kato's Theorem) *[51, Theorem 3.5]*

Let A be the generator of a positive semigroup on a Banach lattice X. Let $Z(X) := \{T \in \mathcal{L}(X)\colon \exists c \geq 0, |Tx| \leq c|x| \ (x \in X)\}$ refers to the 'center' *of X. Then the functions $V \mapsto s(A+V)$ and $V \mapsto \omega(A+V)$ from $Z(X)$ into $[-\infty, \infty)$ are convex.*

Results

Theorem 6 (Generalized Karlin's theorem) *Let A be a resolvent positive linear operator, and V be an operator of multiplication, under the same assumptions as Theorem 4 (or let A and V be as in Theorem 5). Then for $\alpha > 0$,*

1. *$s(\alpha\, A+V)$ is convex in α;*
2. *For each $\alpha > 0$, either*
   (a) *$s((\alpha+d)A+V) < s(\alpha\, A+V)+d\; s(A) \quad \forall\; d > 0$, or*
   (b) *$s((\alpha+d)A+V) = s(\alpha\, A+V)+d\; s(A) \quad \forall\; d > 0$;*
3. *In particular, when $s(A)=0$ then $s(\alpha\, A+V)$ is non-increasing in α (the 'reduction phenomenon'), and when $s(A) < 0$ then $s(\alpha\, A+V)$ is strictly decreasing in α;*
4. *For each $\alpha > 0$,*
$$\frac{\mathrm{d}}{\mathrm{d}\alpha} s(\alpha\, A+V) \leq s(A), \tag{3}$$
except possibly at a countable number of points α, where the one-sided derivatives exist but differ:
$$\frac{\mathrm{d}}{\mathrm{d}\alpha_-} s(\alpha\, A+V) < \frac{\mathrm{d}}{\mathrm{d}\alpha_+} s(\alpha\, A+V) \leq s(A). \tag{4}$$

If A is a generator of a C_0-semigroup, then the above relations on $s(\alpha\, A+V)$ also apply to the type $\omega(\alpha\, A+V)$.

Proof: We consider the general form
$$\phi(\alpha, \beta) := s(\alpha A + \beta V) \text{ or } \omega(\alpha A + \beta V) \tag{5}$$
where $\alpha > 0, \beta \in \mathbb{R}$. Kato [3] explicitly shows that $\phi(1, \beta)$ is convex in β (which he points out is equivalent to varying V). This is shown to imply the properties claimed for $s(\alpha\, A+V) = \phi(\alpha, 1)$ with respect to variation in α, by Lemma 1, to follow.

Lemma 1 (Dual Convexity) *Let $x \in \mathcal{D}_1 = (0, \infty)$ and $y \in \mathcal{D}_2 = [0, \infty)$. Let $f\colon \mathcal{D}_1 \times \mathcal{D}_2 \to \mathbb{R}$ have the following properties:*
$$f(\alpha x, \alpha y) = \alpha f(x,y), \text{ for } \alpha > 0, \qquad \text{and} \tag{6}$$
$$f(x,y) \text{ is convex in } y. \tag{7}$$
Then:

1. *$f(x,y)$ is convex in x;*
2. *For each $x \in \mathcal{D}_1$, either*
   (a) *$f(x+d,y) < f(x,y)+d\,f(1,0) \quad \forall\; d \in \mathcal{D}_1$; or*
   (b) *$f(x+d,y) = f(x,y)+d\,f(1,0) \quad \forall\; d \in \mathcal{D}_1$.*

   For $y \neq 0$, if $f(x,y)$ is strictly convex in y, then $f(x,y)$ is strictly convex in x, and $f(x+d,y) < f(x,y)+d\,f(1,0)$.
3. *For each $x \in \mathcal{D}_1$,*
$$\frac{\partial}{\partial x} f(x,y) \leq f(1,0),$$
except possibly at a countable number of points x, where the one-sided derivatives exist but differ:
$$\frac{\partial}{\partial x_-} f(x,y) < \frac{\partial}{\partial x_+} f(x,y) \leq f(1,0).$$

   The lemma holds if we substitute $\mathcal{D}_1 = (-\infty, 0)$ or $\mathcal{D}_2 = (-\infty, 0]$ or both.

Proof:

1. *$f(x,y)$ is convex in x.*

The relation $f(\alpha x, \alpha y) = \alpha f(x,y)$ (f is homogeneous of degree one) allows a set of rescalings that transform convexity in y into convexity in x. It is perhaps worth noting that this relation is actually a homomorphism, which can be put into a more familiar form by defining a product $x \star y := f(x,y)$, and function $\psi(x) := \alpha x$, which gives $\psi(x) \star \psi(y) = \psi(x \star y)$.

For the case $y = 0$, (6) gives $f(\alpha x, 0) = \alpha f(x, 0)$, so f is trivially convex in x.

For $y \neq 0$, the following derivations have the constraints $y, y_1, y_2 \in \mathcal{D}_2$, $y, y_1, y_2 \neq 0$, and $0 < m < 1$, so that $\{y, y_1, y_2, m, 1{-}m, (1{-}m)y_1 + my_2\}$ are nonzero and their ratios and reciprocals are always defined, and ratios of y_i terms always positive. This keeps the arguments of f within their domains throughout the rescalings.

Convexity of $f(x,y)$ in y gives
$$(1{-}m) f(x, y_1) + m f(x, y_2) \geq f(x, (1{-}m)y_1 + my_2), \tag{8}$$
for $m \in (0,1)$, $y_1 \neq y_2$. Applying (6) to (8), with respective substitutions $\alpha = y_1/y$, $\alpha = y_2/y$, and $\alpha = [(1{-}m)y_1 + my_2]/y$ in the three f terms, yields:
$$(1-m)\frac{y_1}{y} f(\frac{xy}{y_1}, y) + m\frac{y_2}{y} f(\frac{xy}{y_2}, y) \geq \frac{(1{-}m)y_1 + my_2}{y} f(\frac{xy}{(1{-}m)y_1 + my_2}, y). \tag{9}$$

Let $x_1 := xy/y_1$ and $x_2 := xy/y_2$ represent the rescaled first arguments for f on the left side of (9) (so $x, x_1, x_2 \in \mathcal{D}_1$). We try the ansatz that x_1 and x_2 can be combined convexly to yield the third rescaled argument on the right side of (9):

$$\frac{xy}{(1-m)y_1+my_2} = (1-h)x_1+hx_2=(1-h)\frac{xy}{y_1}+h\frac{xy}{y_2}.$$

The ansatz has solution

$$h = \frac{my_2}{(1-m)y_1+my_2}, \text{ and } 1-h = \frac{(1-m)y_1}{(1-m)y_1+my_2}.$$

Note that $h \in (0,1)$ is assured because y_1 and y_2 have the same sign, $y_1 \neq y_2$, and $m \in (0,1)$.

Define $\phi := [(1-m)y_1 + my_2]/y$. Then $\phi > 0$ since y, y_1, y_2 all have the same sign. Substitution gives $(1-m)y_1/y = (1-h)\phi$, and $my_2/y = h\phi$, and (9) becomes:

$$(1-h)\phi f(x_1,y) + h\phi f(x_2,y) \geq \phi f((1-h)x_1 + hx_2, y).$$

After dividing both sides by $\phi > 0$,

$$(1-h)\, f(x_1,y) + h\, f(x_2,y) \geq f((1-h)x_1 + hx_2, y), \quad (10)$$

which is convexity in x (for each (x_1,x_2,h), $\exists(y_1,y_2,m)$). The case of strict convexity follows by substituting $>$ for $\geq$ throughout.

2. *Either* $f(x+d,y) < f(x,y) + d\, f(1,0)\ \forall\ d \in \mathcal{D}_1$, *or* $f(x+d,y) = f(x,y) + d\, f(1,0)\ \forall\ d \in \mathcal{D}_1$.

If $y = 0$, then case 2b in Lemma 1 holds by (6):

$$f(x+d,0) = (x+d)f(1,0) = f(x,0) + d\, f(1,0).$$

For $y \neq 0$, the strategy will be to show first that $f(x+d,y) \leq f(x,y) + d\, f(1,0)$. Next, it is shown that if $f(x+d,y) < f(x,y) + d\, f(1,0)$ for any $d \in \mathcal{D}_1$, then it is true for all $d \in \mathcal{D}_1$.

The steps are shown here only for $x, d \in \mathcal{D}_1 = (0,\infty)$, but they are readily applied to $\mathcal{D}_1 = (-\infty, 0)$. By (6), for $x, d > 0$, the following are equivalent:

$$f(x+d,y) \leq f(x,y) + d\, f(1,0); \quad (11)$$

$$(x+d)\, f\big(1, \frac{y}{x+d}\big) \leq x\, f\big(1,\frac{y}{x}\big) + d\, f(1,0); \quad \text{and}$$

$$f\big(1,\frac{y}{x+d}\big) \leq \frac{x}{x+d}\, f\big(1,\frac{y}{x}\big) + \frac{d}{x+d}\, f(1,0). \quad (12)$$

Since $0 < d/(x+d), x/(x+d) < 1$, the second arguments for f in (12) are related by convex combination,

$$\frac{y}{x+d} = \frac{x}{x+d}\,\frac{y}{x} + \left(1 - \frac{x}{x+d}\right) * 0,$$

so (12) is just a statement of the convexity of $f(x,y)$ in y, as hypothesized. Strict convexity of $f(x,y)$ in y replaces $\leq$ with $<$ throughout (11) and (12), yielding case 2a in Lemma 1.

Now, with $x > 0$, suppose that for some $d_1 > 0$,

$$f(x+d_1,y) < f(x,y) + d_1\, f(1,0). \quad (13)$$

We shall see that convexity then prevents $f(x+d,y)$ from ever returning to the line $f(x,y) + d\, f(1,0)$ for $d > 0$.

We consider five points:

$$0 < x < x+d_0 < x+d_1 < x+d_2 < x+d_3. \quad (14)$$

For readability, write $g(x) \equiv f(x,y)$ and $F \equiv f(1,0)$. By convexity (10), and hypothesis (13),

$$\begin{aligned} g(x+d_0) &\leq \left(1-\frac{d_0}{d_1}\right)g(x) + \frac{d_0}{d_1}g(x+d_1) \\ &< \left(1-\frac{d_0}{d_1}\right)g(x) + \frac{d_0}{d_1}(g(x)+d_1F) = g(x) + d_0\, F, \end{aligned}$$

and, by (10), (13), and (11) (line 3 below),

$$\begin{aligned} g(x+d_2) &\leq \frac{d_3-d_2}{d_3-d_1}g(x+d_1) + \frac{d_2-d_1}{d_3-d_1}g(x+d_3) \\ &< \frac{d_3-d_2}{d_3-d_1}(g(x)+d_1F) + \frac{d_2-d_1}{d_3-d_1}g(x+d_3) \\ &\leq \frac{d_3-d_2}{d_3-d_1}(g(x)+d_1F) + \frac{d_2-d_1}{d_3-d_1}(g(x)+d_3F) \\ &= g(x) + d_2\, F. \end{aligned}$$

For the case where $\mathcal{D}_1 = (-\infty, 0)$, the direction of inequalities in (14) needs to be reversed, and all the subsequent relations are preserved.

3. *For each* $x \in \mathcal{D}_1$, $\partial f(x,y)/\partial x \leq f(1,0)$, *except possibly at a countable number of points* x, *where the one-sided derivatives exist but differ:* $\dfrac{\partial f(x,y)}{\partial x_-} < \dfrac{\partial f(x,y)}{\partial x_+} \leq f(1,0)$.

Rearrangement of (11) gives

$$\frac{f(x+d,y) - f(x,y)}{d} \leq f(1,0), \qquad \text{so} \quad (15)$$

$$\lim_{d \downarrow 0} \frac{f(x+d,y)-f(x,y)}{d} =: \frac{\partial f(x,y)}{\partial x_+} \leq f(1,0). \quad (16)$$

For $y = 0$, equality holds in (15) for all $d > 0$. For $y \neq 0$, because $f(x,y)$ is convex in x on the open interval $\mathcal{D}_1$, the left-sided and right-sided derivatives always exist, and differ at most at a countable number of points, at which the right-sided derivative (16) is greater than the left-sided derivative [52, Proposition 17, pp. 113–114]. ☐

A concavity version of the lemma may be trivially produced by reversal of the convexity inequalities.

Remark 1 It would be clearly desirable to characterize the conditions for strict convexity in Kato's theorem, so that by Lemma 1, one would obtain strict convexity in Theorem 6, item 1, and strict monotonicity in items 3 and 4. Item 2 is the best that can be offered in the way of strict inequality without strict convexity. But the problem is more technical and is deferred to elsewhere.

It is reasonable, nevertheless, to conjecture that the properties which produce strict convexity in the matrix case [53, Theorem 4.1] [54, Theorem 1.1] extend to their Banach space versions: i.e. for $a > 0$, when resolvent positive operator A is irreducible [55, p. 250] [51, p. 41], then $s(\alpha A + \beta V)$ is strictly convex in β if and only if V is not a constant scalar.

The conjectured sharpening of Theorem 6 to strict inequality would have application to continuous-space models for the evolution of dispersal, and show populations to be invadable by less-dispersing organisms when they experience spatially heterogeneous growth rates (a 'selection potential' as defined in [20]). This is a key element of the Reduction Principle, and is first stated generally for finite matrix models in [19, pp. 118, 126, 137, 195, 199] and [20, Results 2, 3]. Its primary implication is that for a population to be non-invadable, it must experience no spatial heterogeneity of growth rates where it has positive measure, and this points toward ideal free distributions (defined to be those which spatially equalize the growth rates when this is possible), as the evolutionarily stable states. For reviews and recent developments, see [56, 57, 58, 59, 60].

A Third Proof of Karlin's Theorem 5.2

Karlin's proof is based on the Donsker-Varadhan variational formula for the spectral radius [61]. Kirkland et al. [56] recently discovered another proof using entirely structural methods. A third distinct proof of Karlin's theorem is seen as follows by application of Lemma 1 to Cohen's theorem, combined with Friedland's equality condition [53, Theorem 4.1] (see also [54] and [62] for other proof methods).

The expression in Karlin's Theorem 5.2 can be put into the form used in Theorem 6:

$$\mathbf{M}(\alpha)\mathbf{D} = [(1-\alpha)\mathbf{I} + \alpha\mathbf{P}]\mathbf{D} = \alpha(\mathbf{P}-\mathbf{I})\mathbf{D} + \mathbf{D} = \alpha\,\mathbf{A} + \beta\,\mathbf{D}.$$

where $\mathbf{A} = (\mathbf{P}-\mathbf{I})\mathbf{D}$, $\alpha \in (0, 1)$, and $\beta = 1$. Since $\mathbf{M}(\alpha)\mathbf{D}$ is a nonnegative matrix when $\alpha \in (0, 1)$, by Perron-Frobenius theory its spectral bound $s(\mathbf{M}(\alpha)\mathbf{D})$ equals its spectral radius $r(\mathbf{M}(\alpha)\mathbf{D})$.

Cohen's theorem gives that $s(\alpha\,\mathbf{A} + \beta\,\mathbf{D})$ is convex in β, and thus by Lemma 1, $s(\alpha\,\mathbf{A} + \beta\,\mathbf{D})$ is convex in $\alpha > 0$ and

$$\frac{\partial s(\alpha\,\mathbf{A} + \beta\,\mathbf{D})}{\partial\alpha} \leq s(\mathbf{A}) = s((\mathbf{P}-\mathbf{I})\mathbf{D}) = 0, \tag{17}$$

the right identity seen since $\mathbf{e}^\top(\mathbf{P}-\mathbf{I})\mathbf{D} = (\mathbf{e}^\top - \mathbf{e}^\top)\mathbf{D} = \mathbf{0}$, where $\mathbf{e}$ is the vector of ones, and $\mathbf{e}^\top$ is its transpose.

Strict convexity in β is shown by Friedland [53, Theorem 4.1] to occur when $\mathbf{P}$ is irreducible and $\mathbf{D} \neq c\,\mathbf{I}$, for any $c > 0$. Strict convexity in β implies, by Lemma 1, that $r(\mathbf{M}(\alpha)\mathbf{D}) = s(\mathbf{M}(\alpha)\mathbf{D})$ is strictly convex and decreasing in α. □

Remark 2 The core of Kirkland et al.'s [56] proof of Theorem 1 is their Lemma 4.1, which can be expressed as

$$\mathbf{e}^\top\mathbf{A}\,(\mathbf{u}(\mathbf{A}) \circ \mathbf{v}(\mathbf{A})) \geq \mathbf{u}(\mathbf{A})^\top\mathbf{A}\,\mathbf{v}(\mathbf{A}) = s(\mathbf{A}),$$

with equality only when $\mathbf{e}^\top\mathbf{A} = s(\mathbf{A})\,\mathbf{e}^\top$, where $\mathbf{u}(\mathbf{A})^\top$ and $\mathbf{v}(\mathbf{A})$ are the left and right eigenvectors of $\mathbf{A}$ associated with the Perron root $s(\mathbf{A})$, and $\mathbf{u} \circ \mathbf{v}$ is the Schur-Hadamard (elementwise) product.

Their result, except for the equality condition, is a special case of [63, Theorem 3.2.5] that $\mathbf{x}^\top\mathbf{A}\mathbf{y} \geq s(\mathbf{A})$ for any $\mathbf{x}, \mathbf{y} \geq \mathbf{0}\colon \mathbf{x} \circ \mathbf{y} = \mathbf{u}(\mathbf{A}) \circ \mathbf{v}(\mathbf{A})$. To obtain the equality condition requires an approach their novel proof provides.

Remark 3 Schreiber and Lloyd-Smith [64, Appendix B, Lemma 1] followed the reverse path and extended Kirkland et al's result on $s(\mathbf{M}(\alpha)\mathbf{D})$ to the form $s(\alpha\mathbf{A}+\mathbf{D})$, where $\mathbf{A}$ is essentially nonnegative and $\mathbf{D}$ any diagonal matrix.

Remark 4 Kato [3] notes that the Donsker-Varadhan formula provides another route besides Kingman's theorem to his generalization of Cohen's theorem (but with more restrictive conditions). Indeed, Friedland [53] uses the Donsker-Varadhan formula to prove Cohen's theorem augmented by strict convexity. The 'dual convexity' relationship shown here between Cohen's and Karlin's theorems means that both routes of proof apply as well to Karlin's theorem. Given these parallels, the relationship between the theorem of Kingman and the theorem of Donsker and Varadhan invites deeper study.

Lemma 1 combined with Cohen's theorem can also be used to give a new proof of an inequality of Lindqvist, the special case considered in [65, Theorem 2, pp. 260–261].

Theorem 7 (Lindqvist) *[65, Theorem 2, subcase]*
Let $\mathbf{A}$ be an irreducible $n \times n$ real matrix such that 1) $A_{ij} \geq 0$ for $i \neq j$, and 2) The left and right eigenvectors of $\mathbf{A}$, $\mathbf{u}(\mathbf{A})^\top$ and $\mathbf{v}(\mathbf{A})$, associated with eigenvalue $s(\mathbf{A})$, satisfy $\mathbf{u}(\mathbf{A})^\top\mathbf{v}(\mathbf{A}) = 1$. Let $\mathbf{D}$ be an $n \times n$ real diagonal matrix. Then

$$s(\mathbf{A}+\mathbf{D}) - s(\mathbf{A}) \geq \mathbf{u}(\mathbf{A})^\top\,\mathbf{D}\,\mathbf{v}(\mathbf{A}). \tag{18}$$

Proof: Since $\mathbf{A}$ is an irreducible essentially nonnegative matrix, $s(\mathbf{A})$ is an eigenvalue of multiplicity 1. Consider the representation $\mathbf{A} = \alpha\mathbf{B} - \mathbf{D}$, where $\mathbf{B}$ is essentially nonnegative and $\alpha > 0$. Write $s \equiv s(\mathbf{A})$. Since $\mathbf{A}$ is irreducible, it has unique $\mathbf{u} \equiv \mathbf{u}(\mathbf{A})$ and $\mathbf{v} \equiv \mathbf{v}(\mathbf{A})$ given $\mathbf{u}^\top\mathbf{v} = \mathbf{e}^\top\mathbf{v} = 1$, and all the derivatives exist [28] in the following derivation [66, Sec. 9.1.1]:

$$\begin{aligned}\mathbf{u}^\top \frac{\partial(\mathbf{A}\mathbf{v})}{\partial\alpha} &= \mathbf{u}^\top \left(\frac{\partial \mathbf{A}}{\partial\alpha}\mathbf{v} + \mathbf{A}\frac{\partial \mathbf{v}}{\partial\alpha}\right)\\ &= \mathbf{u}^\top \mathbf{B}\mathbf{v} + s\,\mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial\alpha}\\ = \mathbf{u}^\top \frac{\partial}{\partial\alpha}(s\,\mathbf{v}) &= \mathbf{u}^\top \left(\frac{\partial s}{\partial\alpha}\,\mathbf{v} + s\,\frac{\partial \mathbf{v}}{\partial\alpha}\right)\\ &= \frac{\partial s}{\partial\alpha} + s\,\mathbf{u}^\top \frac{\partial \mathbf{v}}{\partial\alpha}.\end{aligned}$$

Cancellation of terms $s\,\mathbf{u}^\top \partial\mathbf{v}/\partial\alpha$ gives

$$\frac{\partial s(\mathbf{A})}{\partial\alpha} = \mathbf{u}(\mathbf{A})^\top \frac{\partial \mathbf{A}}{\partial\alpha}\mathbf{v}(\mathbf{A}) = \mathbf{u}(\mathbf{A})^\top \mathbf{B}\mathbf{v}(\mathbf{A}) \leq s(\mathbf{B}),$$

the inequality derived as in (17). Scaling by α, subtracting $\mathbf{D}$, and substituting $\alpha\mathbf{B} = \mathbf{A} + \mathbf{D}$, we get

$$\begin{aligned}\mathbf{u}(\mathbf{A})^\top(\alpha\mathbf{B} - \mathbf{D})\mathbf{v}(\mathbf{A}) = s(\mathbf{A}) \leq s(\alpha\mathbf{B}) - \mathbf{u}(\mathbf{A})^\top \mathbf{D}\mathbf{v}(\mathbf{A})\\ \iff \mathbf{u}(\mathbf{A})^\top \mathbf{D}\mathbf{v}(\mathbf{A}) \leq s(\mathbf{A}{+}\mathbf{D}) - s(\mathbf{A}). \qquad \square\end{aligned}$$

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 [67, eq. 2.9]), or the variation may not scale the mixing process uniformly (e.g. conditional dispersal [68, 69]), 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. Examples are known where departures from reduction occur, i.e. $\mathrm{d}s(\alpha A{+}B)/\mathrm{d}\alpha > s(A)$. Results for the form $\alpha A{+}B$ have been obtained for symmetrizable finite matrices in models of multilocus mutation [70], and dispersal in random environments [71]. Some results for Banach space models have also been obtained [67, 72, 68, 73, 69, 74, 75, 76, 77].

A key open problem, then, is to find necessary or sufficient conditions on Banach space operators, B, such that $\partial s(\alpha A + \beta B)/\partial\alpha \leq s(A)$ (which may depend on A, β/α, domain, and boundary conditions). A sufficient condition is that $s(\alpha A{+}\beta B)$ be convex in β, by Lemma 1. Thus, the dual problem is to ask: for which B is $s(\alpha A + \beta B)$ convex in β? Some results towards this are in [78]. Kato obtained Theorem 4 with operators of multiplication, V, because the family of operators $e^{\beta V}$ is semigroup-superconvex in β (definition in [3]), but this approach faces the challenge that, "It is in general difficult to find a nontrivial semigroup-superconvex family B(h)" [3].

Acknowledgements

I thank Prof. Shmuel Friedland for introducing me to the papers of Cohen, and for inviting me to speak on the early state of this work at the 16th International Linear Algebra Society Meeting in Pisa; Prof. Mustapha Mokhtar-Kharroubi and anonymous reviewers for pointing out an error in a definition in the first version; anonymous reviewers for helpful suggestions; Laura Marie Herrmann for assistance with the literature search; and Arendt and Batty [51] for guiding me to Kato [3].

References

[1] Kingman JFC (1961) A convexity property of positive matrices. *The Quarterly Journal of Mathematics* 12:283–284.

[2] Cohen JE (1981) Convexity of the dominant eigenvalue of an essentially nonnegative matrix. *Proceedings of the American Mathematical Society* 81:657–658.

[3] Kato T (1982) Superconvexity of the spectral radius, and convexity of the spectral bound and the type. *Mathematische Zeitschrift* 180:265–273.

[4] Gadgil M (1971) Dispersal: population consequences and evolution. *Ecology* 52:253–261.

[5] Feldman MW (1972) Selection for linkage modification: I. Random mating populations. *Theoretical Population Biology* 3:324–346.

[6] Balkau B, Feldman MW (1973) Selection for migration modification. *Genetics* 74:171–174.

[7] Karlin S, McGregor J (1974) Towards a theory of the evolution of modifier genes. *Theoretical Population Biology* 5:59–103.

[8] Feldman MW, Krakauer J (1976) in *Population Genetics and Ecology*, eds Karlin S, Nevo E (Academic Press, New York), pp 547–583.

[9] Teague R (1976) A result on the selection of recombination altering mechanisms. *Journal of Theoretical Biology* 59:25–32.

[10] Teague R (1977) A model of migration modification. *Theoretical Population Biology* 12:86–94.

[11] Feldman MW, Christiansen FB, Brooks LD (1980) Evolution of recombination in a constant environment. *Proceedings of the National Academy of Sciences U.S.A.* 77:4838–4841.

[12] Holt R (1985) Population dynamics in two-patch environments: Some anomalous consequences of an optimal habitat distribution. *Theoretical Population Biology* 28:181–208.

[13] Feldman MW, Liberman U (1986) An evolutionary reduction principle for genetic modifiers. *Proceedings of the National Academy of Sciences U.S.A.* 83:4824–4827.

[14] McPeek M, Holt R (1992) The evolution of dispersal in spatially and temporally varying environments. *American Naturalist* 140:1010–1027.

[15] Karlin S (1976) in *Population Genetics and Ecology*, eds Karlin S, Nevo E (Academic Press, New York), pp 616–657.

[16] Karlin S (1982) in *Evolutionary Biology*, eds Hecht MK, Wallace B, Prance GT (Plenum Publishing Corporation, New York) Vol. 14, pp 61–204.

[17] Karlin S, McGregor J (1972) The evolutionary development of modifier genes. *Proceedings of the National Academy of Sciences U.S.A.* 69:3611–3614.

[18] Karlin S, Carmelli D (1975) Numerical studies on two-loci selection models with general viabilities. *Theoretical Population Biology* 7:399–421.

[19] Altenberg L (1984) *A Generalization of Theory on the Evolution of Modifier Genes. Ph.D. dissertation* (Stanford University).

[20] Altenberg L, Feldman MW (1987) Selection, generalized transmission, and the evolution of modifier genes. I. The reduction principle. *Genetics* 117:559–572.

[21] Altenberg L (2009) The evolutionary reduction principle for linear variation in genetic transmission. *Bulletin of Mathematical Biology* 71:1264–1284.

[22] Hastings A (1983) Can spatial variation alone lead to selection for dispersal? *Theoretical Population Biology* 24:244–251.

[23] Hutson V, López-Gómez J, Mischaikow K, Vickers G (1995) in *Dynamical systems and applications*, ed Agarwal RP (World Scientific, Singapore), pp 343–358.

[24] Dockery J, Hutson V, Mischaikow K, Pernarowski M (1998) The evolution of slow dispersal rates: A reaction diffusion model. *Journal of Mathematical Biology* 37:61–83.

[25] Cantrell R, Cosner C, Lou Y (2010) in *Spatial Ecology*, Mathematical and Computational Biology, eds Cantrell R, Cosner C, Ruan S (Chapman & Hall/CRC Press, London), pp 213–229.

[26] Hutson V, Martinez S, Mischaikow K, Vickers G (2003) The evolution of dispersal. *Journal of Mathematical Biology* 47:483–517.

[27] Lozinskiy S (1969) On an estimate of the spectral radius and spectral abscissa of a matrix. *Linear Algebra and Its Applications* 2:117–125.

[28] Deutsch E, Neumann M (1985) On the first and second order derivatives of the Perron vector. *Linear Algebra and Its Applications* 71:57–76.

[29] Keilson J (1964) A review of transient behavior in regular diffusion and birth-death processes. *Journal of Applied Probability* 1:247–266.

[30] Horn RA, Johnson CR (1985) *Matrix Analysis* (Cambridge University Press, Cambridge).

[31] Gantmacher F (1959) *Applications of the Theory of Matrices* (Interscience, New York).

[32] Seneta E (1981) *Non-negative Matrices and Markov Chains* (Springer-Verlag, New York).

[33] Bellman R (1955) On an iterative procedure for obtaining the Perron root of a positive matrix. *Proceedings of the American Mathematical Society* 6:719–725.

[34] Hadeler K, Thieme H (2008) Monotone dependence of the spectral bound on the transition rates in linear compartment models. *Journal of Mathematical Biology* 57:697–712.

[35] Birindelli I, Mitidieri E, Sweers G (1999) The existence of the principal eigenvalue for cooperative elliptic systems in a general domain. *Differential Equations* 35:326–334.

[36] Nagel R, ed (1986) *One-parameter Semigroups of Positive Operators*, Lecture Notes in Mathematics (Springer-Verlag, Berlin) Vol. 1184.

[37] Van Neerven J (1996) *The Asymptotic Behaviour of Semigroups of Linear Operators*, Operator Theory Advances and Applications (Birkhäuser, Basel, Switzerland) Vol. 88.

[38] Arendt W, Batty C, Hieber M, Neubrander F (2011) *Vector-valued Laplace Transforms and Cauchy Problems*, Monographs in Mathematics (Birkhäuser, Basel) Vol. 96, Second edition.

[39] Arendt W (1987) Resolvent positive operators. *Proceedings of the London Mathematical Society* 3:321–349.

[40] Arendt W, Batty C, Robinson D (1990) Positive semigroups generated by elliptic operators on Lie groups. *Journal of Operator Theory* 23:369 – 407.

[41] Ulm M (1999) The interval of resolvent-positivity for the biharmonic operator. *Proceedings-American Mathematical Society* 127:481–490.

[42] Simon B (1982) Schrödinger semigroups. *Proceedings of the American Mathematical Society* 7:447–526.

[43] Voigt J (1986) Absorption semigroups, their generators, and Schrödinger semigroups. *Journal of Functional Analysis* 67:167–205.

[44] Mokhtar-Kharroubi M (2009) On Schrödinger semigroups and related topics. *Journal of Functional Analysis* 256:1998–2025.

[45] Ouhabaz EM (2005) *Analysis of Heat Equations on Domains*, London Mathematical Society Monograph Series (Princeton University Press, Princeton, NJ) Vol. 31.

[46] Donsker M, Varadhan S (1976) On the principal eigenvalue of second-order elliptic differential operators. *Communications on Pure and Applied Mathematics* 29:595–621.

[47] Berestycki H, Nirenberg L, Varadhan S (1994) The principal eigenvalue and maximum principle for second-order elliptic operators in general domains. *Communications on Pure and Applied Mathematics* 47:47–92.

[48] Grinfeld M, Hines G, Hutson V, Mischaikow K, Vickers G (2005) Non-local dispersal. *Differential and Integral Equations (Athens)* 18:1299–1320.

[49] Bates P, Zhao G (2007) Existence, uniqueness and stability of the stationary solution to a nonlocal evolution equation arising in population dispersal. *Journal of Mathematical Analysis and Applications* 332:428–440.

[50] Chabi M, Latrach K (2002) On singular monoenergetic transport equations in slab geometry. *Mathematical Methods in the Applied sciences* 25:1121–1147.

[51] Arendt W, Batty CJK (1995) Principal eigenvalues and perturbation. *Operator Theory: Advances and Applications* 75:39–55.

[52] Royden HL (1988) *Real Analysis* (Macmillan, New York), 3rd edition.

[53] Friedland S (1981) Convex spectral functions. *Linear and Multilinear Algebra* 9:299–316.

[54] Nussbaum D (1986) Convexity and log convexity for the spectral radius. *Linear Algebra and Its Applications* 73:59–122.

[55] Greiner G, Voigt J, Wolff M (1981) On the spectral bound of the generator of semigroups of positive operators. *J. Operator Theory* 5:245–256.

[56] Kirkland S, Li CK, Schreiber SJ (2006) On the evolution of dispersal in patchy landscapes. *SIAM Journal on Applied Mathematics* 66:1366–1382.

[57] Cantrell R, Cosner C, Deangelis D, Padron V (2007) The ideal free distribution as an evolutionarily stable strategy. *Journal of Biological Dynamics* 1:249–271.

[58] Cantrell R, Cosner C, Lou Y (2010) Evolution of dispersal and the ideal free distribution. *Mathematical Biosciences and Engineering* 7:17–36.

[59] Averill I, Lou Y, Munther D (2011) On several conjectures from evolution of dispersal. *Journal of Biological Dynamics* 0:1–14.

[60] Cosner C, Dávila J, Martínez S (2011) Evolutionary stability of ideal free nonlocal dispersal. *Journal of Biological Dynamics* in press DOI:10.1080/17513758.2011.588341.

[61] Donsker MD, Varadhan SRS (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.

[62] O'Cinneide C (2000) Markov additive processes and Perron-Frobenius eigenvalue inequalities. *The Annals of Probability* 28:1230–1258.

[63] Bapat RB, Raghavan TES (1997) *Nonnegative Matrices and Applications* (Cambridge University Press, Cambridge, UK).

[64] Schreiber SJ, Lloyd-Smith JO (2009) Invasion dynamics in spatially heterogeneous environments. *American Naturalist* 174:490–505.

[65] Lindqvist BH (2002) On comparison of the Perron-Frobenius eigenvalues of two ML-matrices. *Linear Algebra and Its Applications* 353:257–266.

[66] Caswell H (2000) *Matrix Population Models* (Sinauer Associates, Sunderland, MA), 2nd edition, p 722.

[67] Belgacem F, Cosner C (1995) The effects of dispersal along environmental gradients on the dynamics of populations in heterogeneous environments. *Canadian Applied Mathematics Quarterly* 3:379–397.

[68] Hambrock R (2007) Ph.D. dissertation (The Ohio State University, Columbus, OH).

[69] Hambrock R, Lou Y (2009) The evolution of conditional dispersal strategies in spatially heterogeneous habitats. *Bulletin of mathematical biology* 71:1793–1817.

[70] Altenberg L (2011) An evolutionary reduction principle for mutation rates at multiple loci. *Bulletin of Mathematical Biology* 73:1227–1270.

[71] Altenberg L (2011) The evolution of dispersal in random environments and the principle of partial control. *Submitted*.

[72] Lou Y (2008) Some challenging mathematical problems in evolution of dispersal and population dynamics. *Tutorials in Mathematical Biosciences IV* pp 171–205.

[73] Chen X, Hambrock R, Lou Y (2008) Evolution of conditional dispersal: A reaction–diffusion–advection model. *Journal of Mathematical Biology* 57:361–386.

[74] Bezugly A (2009) Ph.D. dissertation (The Ohio State University, Columbus, OH).

[75] Godoy T, Gossez J, Paczka S (2010) On the asymptotic behavior of the principal eigenvalues of some elliptic problems. *Annali di Matematica Pura ed Applicata* 189:497–521.

[76] Bezuglyy A, Lou Y (2010) Reaction–diffusion models with large advection coefficients. *Applicable Analysis* 89:983–1004.

[77] Kao CY, Lou Y, Shen W (2012) Evolution of mixed dispersal in periodic environments. *Discrete and Continuous Dynamical Systems B* In Press.

[78] Webb G (1993) in *Semigroups of Linear Operators and Applications*, eds Goldstein G, Goldstein J (Kluwer, Dordrecht), pp 259–270.