hexagon:2611.00416v1

Convergence Criteria for Measure-Valued Population Dynamics with Mutation and Horizontal Transfer

Zhenyu Sun

12 November 2026

Abstract

We study a measure-valued population equation on a compact trait space with positive baseline growth and competition, an additional mutation birth gain, and frequency-dependent horizontal transfer with an antisymmetric net kernel. We give sufficient conditions for convergence to a unique nonzero equilibrium. Separating total mass from trait frequencies yields positive mass bounds. Without mutation, maximal time-averaged fitness on the initial support tends to zero, and the endpoint measure in each joint subsequential limit is supported on the zero set of the limiting averaged fitness. Potential transfer and recipient-independent competition reduce the frequency dynamics to a normalized positive semigroup. For general transfer, a common-reference balance gives an entropy criterion allowing arbitrarily strong circulation. An adjoint-weighted entropy identity includes singular population components and the normalization correction in transfer. Mutation dissipates deviations in weighted composition, while a quadratic-form condition controls the remaining mass deviation; in finite dimensions, mutation can also compensate for negative directions of that form. A separate contraction criterion compares arbitrary trajectories using only coefficient bounds. Invariant resident laws yield invasion indices for recurrent environments. Examples exhibit bistability under complete reset mutation and periodic populations whose time averages are not equilibria. For rank-one mutation, we classify stationary frequencies, prove a critical lower bound ruling out exponential convergence from an initial point mass at a maximizer, and determine the condensate when the initial population assigns positive mass to the maximizing set.

Keywords. Measure-valued population dynamics; horizontal gene transfer; mutation; relative entropy; positive semigroups; invasion; condensation.

Contents

1 Introduction **3**
- 1.1 The population equation . . . 3
- 1.2 Relation to existing work . . . 3
- 1.3 Notation and convergence . . . 4
- 1.4 Convergence mechanisms and their scope . . . 5

2 Measure dynamics and mutation accessibility **6**
- 2.1 Averaged balance in the presence of mutation . . . 10

3 Accumulated fitness and averaged environments **11**
- 3.1 A coefficient criterion for selection of one trait . . . 13

4 Invariant resident laws and invasion **14**
4.1 A finite-dimensional persistence theorem . . . 16
4.2 Mutation classes and positive linear growth . . . 17

5 Potential differences and circulation **18**
5.1 Hamiltonian structure in the finite pure-transfer system . . . 19

6 Potential transfer and Perron convergence **19**
6.1 A positive adjoint eigenfunction and a Doob transform . . . 20
6.2 A coefficient condition that constructs the eigenfunction . . . 21

7 Rank-one mutation, spectral thresholds, and condensation **23**
7.1 Initial mass on the maximizing set selects the condensate . . . 24
7.2 Adjoint weights and convergence rates at the spectral edge . . . 25

8 Convergence through a common reference distribution **27**

9 Adjoint-weighted entropy and convergence **28**
9.1 The frozen operator and its conservative transform . . . 28
9.2 The exact nonlinear entropy identity . . . 30
9.3 A continuum convergence theorem . . . 31
9.4 Finite-dimensional tests . . . 34
9.5 Examples and boundaries of the criteria . . . 35

10 Quantitative contraction under mutation mixing **37**
10.1 Removing composition-neutral terms . . . 38
10.2 Dependence of the equilibrium on the coefficients . . . 40

11 Examples separating the structural conditions **41**
11.1 Complete-reset mutation can coexist with bistability . . . 41
11.2 Periodic residents with a nonstationary population average . . . 42
11.3 A comparison of the convergence mechanisms . . . 43

12 Discussion **43**
12.1 How to use the criteria . . . 43
12.2 What changes at the boundaries of the hypotheses . . . 44
12.3 Questions suggested by the theory . . . 44

1 Introduction

1.1 The population equation

Let $\mathcal{X}$ be a nonempty compact metric space. In particular, $\mathcal{X}$ may be a compact subset of $\mathbb{R}^d$. Let $b, d : \mathcal{X} \to [0, \infty)$, $C : \mathcal{X}^2 \to (0, \infty)$, and $\tau : \mathcal{X}^2 \to [0, \infty)$ be continuous. Assume

$$r = b - d > 0, \qquad 0 \le p \le 1, \qquad \alpha(x, y) = \tau(x, y) - \tau(y, x).$$

Let $m(x, dy)$ be a Borel probability kernel. The measure u_t records population mass by trait, and $N_t = u_t(\mathcal{X})$ is the total mass. The coefficient $C(x, y)$ is the competitive effect of trait y on trait x, while $m(y, dx)$ gives the trait distribution of offspring produced through the mutation term by a parent of trait y. For a nonzero finite nonnegative initial measure u_0, consider

$$\begin{aligned} \frac{d}{dt}\int f(x)u_t(dx) = &\int f(x)\left[r(x) - \int C(x,y)u_t(dy)\right]u_t(dx) \\ &+ p\int b(y)\left[\int f(x)m(y,dx)\right]u_t(dy) \\ &+ \frac{1}{N_t}\iint f(x)\alpha(x,y)u_t(dx)u_t(dy), \qquad N_t = u_t(\mathcal{X}), \end{aligned} \tag{1.1}$$

for $f \in C(\mathcal{X})$. Here $b - d$ is the baseline net growth, and the mutation term is an additional birth gain of rate $pb(y)$ per parent. Thus p scales an additional offspring channel; it does not replace a fraction of the baseline births by mutant births. Integrating this gain gives the effective growth $a = r + pb$ in the mass equation.

The transfer convention is that $\tau(x, y)$ describes conversion of a trait-y individual to trait x through interaction with a trait-x individual. Indeed, the last term of (1.1) equals

$$\frac{1}{N_t}\iint [f(x) - f(y)]\tau(x, y)u_t(dx)u_t(dy).$$

It vanishes for $f = 1$. Its per-capita contribution depends on the trait frequencies u_t/N_t, which explains the term frequency-dependent transfer.

We seek conditions ensuring a unique nonzero equilibrium that attracts all nonzero initial populations, and limiting balance relations for trajectories that need not converge. Conservation of mass under transfer leaves the frequency dynamics free to circulate. Mutation can mix traits while competition sustains multiple attracting equilibria. On a continuous trait space, parent-independent mutation can also coexist with stationary measures having singular components. The convergence criteria below address these distinct obstructions through normalized positive-semigroup representations, entropy dissipation, or contraction of pairs of trajectories.

1.2 Relation to existing work

Measure-valued ecological and evolutionary equations with horizontal transfer arise from individual-based population models; see Billiard et al. [3], Champagnat et al. [6]. The latter survey states the weak population equation and finite-time well-posedness in Proposition 3.1, and records the general long-time problem with transfer as open in Section 3.1. Pure selection and competition have a more developed asymptotic theory, including concentration and entropy methods under additional assumptions on the interaction kernel [9, 13].

For particular horizontal-transfer models, numerical and formal Hamilton–Jacobi investigations reveal several evolutionary regimes [5]. Stationary small-mutation analyses characterize

monomorphic and dimorphic states in a one-dimensional model [11]. The constrained Hamilton–Jacobi theory of Gárriz and Mirrahimi [10] proves dynamic monomorphic regimes under specific hypotheses, while identifying further questions for initial data and parameters leading beyond those regimes. These diffusion and small-mutation results address a different limiting problem from (1.1) with a fixed, general mutation kernel.

Several established analytical methods underlie the results below. The general relative entropy method uses direct and adjoint positive eigenelements [15]. Positive non-conservative semigroups can be studied by conservative transforms, generalized Doeblin conditions, and Harris-type theorems [1, 2]. The role of boundary invariant laws in persistence is developed in Hofbauer and Schreiber [12]. Continuous mutation–selection models also exhibit concentration and changes in asymptotic profile [4]; spectral thresholds and convergence rates for the house-of-cards model are studied in Cloez and Gabriel [7]. Coville and Girardin [8] study existence, uniqueness, and stability for structured population equations whose density-dependent feedback is a fixed positive function times a scalar nonlocal functional. In (1.1), the competition kernel need not be separable, and normalized transfer introduces a further nonlinear interaction.

We give direct proofs of the stated results for (1.1). The principal model-dependent calculations are the joint averaged-environment conditions, the nonlinear remainder in adjoint-weighted entropy, and the coupled contraction estimate for mass and composition. The examples distinguish obstructions to uniqueness from obstructions to convergence: nonlinear competition can produce bistability even for two traits, while a spectral edge can prevent total variation convergence in the continuum even when the stationary population is unique.

1.3 Notation and convergence

Write $\mathcal{M}(\mathcal{X})$ for the finite signed Radon measures, $\mathcal{M}_+(\mathcal{X})$ for its positive cone, and $\mathcal{P}(\mathcal{X})$ for probability measures. The total variation norm is

$$\|\mu\|_{\mathrm{TV}} = |\mu|(\mathcal{X});$$

two mutually singular probabilities have distance 2. Weak convergence means convergence against all functions in $C(\mathcal{X})$. A nonzero equilibrium is globally attracting in a stated topology if every nonzero nonnegative initial measure converges to it in that topology. This assertion concerns the full trajectory, without time averaging. We use $\mu \sim \nu$ for mutual absolute continuity, and $\mu \geq \nu$ for the order of measures. For a real continuous function f on a compact space, set $f_- = \min f$, $f_+ = \max f$, and $\operatorname{osc}(f) = f_+ - f_-$; C_- and C_+ refer to extrema on $\mathcal{X}^2$. In particular $b_- \geq r_- > 0$.

The operators and scalar competition average used below are

$$B\mu(dx) = \int b(y)m(y,dx)\mu(dy), \qquad Q\mu = B\mu - b\mu, \qquad a = r + pb,$$
$$C\mu(x) = \int C(x,y)\mu(dy), \qquad A\mu(x) = \int \alpha(x,y)\mu(dy), \qquad \bar{C}(\nu) = \int C\nu\, d\nu.$$

The same symbol M is used for the Markov kernel acting on functions or measures: $Mf(x) = \int f(y)m(x,dy)$ and $(M\mu)(E) = \int m(x,E)\mu(dx)$. When a scalar maximum is denoted by M in the rank-one section, its meaning is local to that section.

For probabilities π, ν, the relative entropy is

$$\mathcal{H}(\pi \mid \nu) = \begin{cases} \displaystyle\int \log \frac{d\pi}{d\nu}\, d\pi, & \pi \ll \nu, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is extended-valued when the displayed integral diverges. The reference distribution is the first argument in the entropy estimates below. Finiteness requires $\pi \ll \nu$; the evolving measure ν may still have a component singular with respect to π.

1.4 Convergence mechanisms and their scope

A common reset, or mutation minorization, means a bound $m(x, E) \geq \eta q(E)$ for every trait x and every Borel set E, with one probability measure q and one constant $\eta > 0$. Complete reset means $m(x, \cdot) = q$ independently of the parent trait. These conditions provide quantitative mixing for the conservative generator pQ when $p > 0$; the nonlinear population dynamics also depends on selection and competition.

Table 1 summarizes the convergence criteria. The hypotheses concern different features of the equation; the examples show that one criterion may apply when another fails. Each row gives the main conditions, with the complete assumptions in the cited result.

Table 1: Sufficient conditions for convergence. Except for the last row, the conclusions hold for every nonzero nonnegative initial measure.

Criterion and main hypotheses	Conclusion and reference
Potential transfer. $C(x, y) = c(y)$ and $\alpha(x, y) = h(x) - h(y)$; $p > 0$; a positive bounded adjoint eigenfunction bounded away from zero and a common mutation reset, or finitely many traits and irreducible mutation.	Unique nonzero equilibrium; exponential convergence in total variation. Theorem 6.2 and Corollary 6.4.
Common reference law. $Q\pi = 0$, $C\pi$ and $a + A\pi$ constant; the competition form is nonnegative on zero-mass measures; $p > 0$ and $m \geq \eta\pi$.	Unique nonzero equilibrium; exponential entropy decay and exponential convergence in total variation. Theorem 8.1.
Adjoint-weighted entropy. A stationary u_*, a positive bounded adjoint weight bounded away from zero, and two positive semidefinite endpoint kernels; $p > 0$ and $m \geq \eta\nu_*$.	Unique nonzero equilibrium; exponential convergence in total variation. Theorem 9.4.
Mutation-augmented entropy. Finitely many traits, a positive stationary state and adjoint weight; $p > 0$ and $m \geq \eta\nu_*$; both mutation-augmented endpoint matrices are positive definite.	Unique nonzero equilibrium; exponential convergence, allowing some negative competition directions. Corollary 9.7.
Mutation contraction. A common mutation reset satisfies the coupled inequalities (10.6), which control selection, transfer, competition, and the mass–frequency coupling.	Unique nonzero equilibrium; exponential convergence in total variation, without first solving the stationary equation. Theorem 10.2.
Uniform trait dominance. $p = 0$; a trait in the initial support has the fitness advantage specified by (3.7).	Weak concentration at that trait; total variation convergence of frequencies exactly when it has initial atomic mass. Theorem 3.7.

Several general statements precede these criteria. The exact mass equation yields a positive absorbing interval (Theorem 2.1). A Feller mutation kernel gives a weakly continuous

semiflow, a compact attracting set, and existence of a stationary measure (Theorem 2.3). The merely measurable-kernel formulation still gives total variation well-posedness, but may have no stationary measure at all (Example 2.7).

Without mutation, the maximum of averaged fitness over the initial support tends to zero along the full time axis. At a joint subsequential limit, the limiting averaged fitness is nonpositive on that support, and the endpoint population is supported on its zero set. This fitness depends on two separate averages: the population measure and its normalized frequency (Theorems 3.1–3.3). Invariant occupation laws then describe stationary or recurrent resident environments, and their invasion rates have an explicit two-average formula (Proposition 4.3). The rank-one mutation analysis gives exact spectral thresholds, a family of condensed stationary states, and precise obstructions to total variation or exponential convergence (Section 7).

Mutation changes the averaged balance: the local growth coefficient excluding mutation influx can be strictly negative at every resident trait because mutation continuously supplies mass. Proposition 2.8 gives the general averaged inequality and the exact finite-dimensional logarithmic balance, including this incoming mutation flux.

2 Measure dynamics and mutation accessibility

Throughout this section, $\mathcal{X}$ is a compact metric space. All measures are finite Radon measures. We use the full total variation norm, so that $\|\mu-\nu\|_{\mathrm{TV}} \le 2$ for probability measures. The mutation kernel is initially assumed only to be a Borel probability kernel. Set

$$(B\mu)(E) = \int_{\mathcal{X}} b(y)m(y,E)\,\mu(dy), \qquad Q\mu = B\mu - b\mu, \qquad a = r + pb,$$

and write

$$(C\mu)(x) = \int_{\mathcal{X}} C(x,y)\,\mu(dy), \qquad (A\mu)(x) = \int_{\mathcal{X}} \alpha(x,y)\,\mu(dy).$$

For $u \ge 0$ with $N = u(\mathcal{X}) > 0$, define

$$\nu = \frac{u}{N}, \qquad h_u = r - Cu + A\nu, \qquad \mathcal{F}(u) = h_u u + pBu.$$

The equation under consideration is $\dot{u} = \mathcal{F}(u)$. We call h_u the local ecological growth coefficient. It is the multiplication coefficient before adding mutation influx, so for $p > 0$ it need not equal the total per-capita growth of a resident trait. The quantity $a = r + pb$ includes the additional mutation birth gain. The mass equation bounds population size independently of its composition; the frequency equation retains the interactions that determine which traits persist.

Theorem 2.1 (Total variation well-posedness and mass regulation)**.** *Let r, b, C, α satisfy the standing assumptions, with $r_- > 0$, $C_- > 0$, and $\alpha(x,y) = -\alpha(y,x)$. For every nonzero $u_0 \ge 0$, the equation has a unique global solution*

$$u \in C^1([0,\infty); \mathcal{M}(\mathcal{X}))$$

in total variation. It is nonnegative and has positive total mass. Every weak solution tested against $C(\mathcal{X})$ agrees with this solution. Its mass and frequency satisfy

$$N' = N\{\langle a, \nu\rangle - N\langle C\nu, \nu\rangle\}, \tag{2.1}$$

$$\nu' = pQ\nu + \{a - \langle a, \nu\rangle - N(C\nu - \langle C\nu, \nu\rangle) + A\nu\}\nu. \tag{2.2}$$

In particular,

$$a_- N - C_+ N^2 \le N' \le a_+ N - C_- N^2. \tag{2.3}$$

Consequently,

$$\min\{N_0, a_-/C_+\} \le N_t \le \max\{N_0, a_+/C_-\} \qquad (t \ge 0) \tag{2.4}$$

and

$$\liminf_{t\to\infty} N_t \ge a_-/C_+, \qquad \limsup_{t\to\infty} N_t \le a_+/C_-. \tag{2.5}$$

Proof. The operator B is positive and bounded on total variation, with norm at most b_+. Multiplication by a bounded function is also bounded. On every set where the total variation norms are bounded and the signed mass is bounded away from zero, the map

$$u \longmapsto (r - Cu)u + \frac{(Au)u}{u(\mathcal{X})} + pBu$$

is locally Lipschitz in total variation. Banach-space ordinary differential equation theory therefore gives local existence and uniqueness. Positivity follows from the variation-of-constants formula: the local multiplication term has a bounded negative part on bounded positive sets, and the gain pBu is nonnegative.

Integrating the equation eliminates the transfer term because $\iint \alpha(x,y)u(dx)u(dy) = 0$. The mutation gain has mass $p\langle b, u\rangle$, which yields (2.1). Differentiating u/N gives (2.2). Comparison with scalar logistic equations proves (2.3)–(2.5). These estimates prevent both finite-time blow-up and approach to the zero-mass singularity. The solution therefore extends globally.

For completeness, suppose a weak solution is given in the formulation using continuous test functions. Its mass satisfies the same comparison bounds. On a bounded time interval, the weak equation gives

$$\left|\frac{d}{dt}\langle f, u_t\rangle\right| \le K\|f\|_\infty \qquad (f \in C(\mathcal{X}))$$

with K independent of f. The dual characterization of total variation for Radon measures implies $\|u_t - u_s\|_{\mathrm{TV}} \le K|t-s|$. Thus the path is total variation continuous. Since $\mathcal{F}$ is locally Lipschitz, $\mathcal{F}(u_t)$ is total variation continuous. The integrated weak equation is consequently the corresponding Banach-space integral equation. Uniqueness identifies the two solutions. □

Remark 2.2. The decomposition $pBu = pQu + pb\,u$ separates conservative redistribution from the additional growth pb. The original gain pBu has total mass $p\langle b, u\rangle$. Equation (2.5) excludes extinction of the entire population under the standing positive-growth assumptions. It gives no positive lower bound for the mass of each trait. In finite dimensions, coordinate persistence follows under the mutation hypotheses of Proposition 2.6 or the mutation-free boundary-law condition of Theorem 4.4.

Set

$$\ell = \frac{a_-}{2C_+}, \qquad U = \frac{2a_+}{C_-}, \qquad \mathcal{K} = \{u \ge 0 : \ell \le u(\mathcal{X}) \le U\}. \tag{2.6}$$

The mutation kernel is called Feller if

$$Mf(x) := \int_{\mathcal{X}} f(y)m(x,dy) \in C(\mathcal{X}) \quad \text{for every } f \in C(\mathcal{X}).$$

Theorem 2.3 (Weak compactness, attraction, and stationary states)**.** *Assume $p = 0$, or assume that m is Feller. The solution maps $S_t u_0 = u_t$ form a weakly continuous semiflow on the nonzero positive measures. The slab $\mathcal{K}$ is weakly compact, convex, and positively invariant. It absorbs every nonzero orbit, and the absorption time is uniform when the initial masses lie in a fixed compact subinterval of $(0, \infty)$.*

The set

$$\mathfrak{A} = \bigcap_{t \geq 0} S_t \mathcal{K} \tag{2.7}$$

is nonempty, weakly compact, and fully invariant: $S_t \mathfrak{A} = \mathfrak{A}$ for $t \geq 0$. It attracts $\mathcal{K}$ uniformly in any metric inducing the weak topology, and attracts every nonzero orbit. Every omega-limit set is nonempty, weakly compact, connected, and fully invariant. There is at least one nonzero equilibrium in $\mathcal{K}$.

Proof. Compactness of $\mathcal{X}$ makes every positive bounded-mass set weakly relatively compact; weak mass continuity makes $\mathcal{K}$ compact. Its convexity is immediate. Logistic comparison proves positive invariance and the stated absorption properties.

To prove weak continuity, consider $u_{0,n} \rightharpoonup u_0 \neq 0$. On each bounded time interval, their solutions lie in a common mass annulus and have uniformly bounded total variation speeds. The Arzela–Ascoli argument in the metrizable weak topology produces subsequential convergence of paths, uniformly on compact time intervals. For a continuous kernel, weak convergence of bounded measures implies uniform convergence of its potentials. Moreover, $B^* f = bMf$ is continuous under the Feller assumption. These properties allow passage to the limit in the integrated weak equation. Uniqueness from Theorem 2.1 identifies the limit and proves convergence of the entire sequence.

The sets $S_t \mathcal{K}$ are nonempty compact sets decreasing with t. Their intersection is therefore nonempty and compact. For the reverse inclusion in full invariance, take $v \in \mathfrak{A}$ and, for integers n, choose $z_n \in \mathcal{K}$ with $S_{n+t} z_n = v$. A convergent subsequence of $S_n z_n$ has limit in $\mathfrak{A}$, and continuity gives a preimage of v under S_t. Uniform attraction follows by contradiction: any sequence $S_{t_n} z_n$, $t_n \to \infty$, $z_n \in \mathcal{K}$, has limit points in every $S_T \mathcal{K}$, hence in $\mathfrak{A}$. The omega-limit assertions follow from the same compactness argument; connectedness follows by intersecting the nested closures of the continuous orbit tails.

For existence of an equilibrium, choose $h > 0$ so small that

$$h(C_+ U + \|\alpha\|_\infty) \leq 1, \qquad 2hC_+ U \leq 1.$$

Then $T_h u = u + h\mathcal{F}(u)$ is nonnegative for $u \in \mathcal{K}$. Its mass is bounded below and above, respectively, by

$$f_-(N) = N + h(a_- N - C_+ N^2), \qquad f_+(N) = N + h(a_+ N - C_- N^2).$$

Both comparison functions are increasing on $[\ell, U]$, and $f_-(\ell) \geq \ell$, $f_+(U) \leq U$. Thus $T_h \mathcal{K} \subseteq \mathcal{K}$. The map T_h is weakly continuous. The Schauder–Tychonoff fixed point theorem yields $u_* = T_h u_*$, hence $\mathcal{F}(u_*) = 0$. □

Remark 2.4. The attraction in Theorem 2.3 is weak attraction. No total variation compactness is asserted. Uniform absorption over all nonzero initial measures with arbitrarily small masses is also excluded: even a scalar logistic equation requires arbitrarily long times to leave a neighbourhood of zero.

Proposition 2.5 (Mutation accessibility)**.** *Suppose $p > 0$, let $\nu_0 = u_0/N_0$, and define*

$$\Lambda = \sum_{k=0}^{\infty} 2^{-k-1} M^k \nu_0, \qquad S = \operatorname{supp} \Lambda. \tag{2.8}$$

For every $t > 0$, u_t and Λ are mutually absolutely continuous. In particular, $\operatorname{supp} u_t = S$*. If m is Feller, then $m(x, S) = 1$ for every $x \in S$, and S is the smallest closed mutation-invariant set containing* $\operatorname{supp} u_0$*. When $p = 0$, one instead has $u_t \sim u_0$.*

Proof. The standing assumptions imply $b_- \geq r_- > 0$. Along a fixed orbit, h_{u_t} is uniformly bounded by some H. Positive semigroup comparison gives

$$e^{-Ht} e^{ptB} u_0 \leq u_t \leq e^{Ht} e^{ptB} u_0.$$

For every $k \geq 0$,

$$b_-^k M^k u_0 \leq B^k u_0 \leq b_+^k M^k u_0.$$

All coefficients of e^{ptB} are positive when $t, p > 0$, so this series and Λ have exactly the same null sets. For $p = 0$, the usual multiplicative solution formula gives $u_t \sim u_0$.

Since $M\Lambda = 2\Lambda - \nu_0 \leq 2\Lambda$, we have

$$\int m(x, \mathcal{X} \setminus S)\, \Lambda(dx) = 0.$$

For a Feller kernel the integrand is lower semicontinuous, because $\mathcal{X} \setminus S$ is open. Positivity at any point of S would therefore imply positivity on a set of positive Λ mass, a contradiction. This proves mutation invariance of S. If another closed mutation-invariant set contains $\operatorname{supp} u_0$, every $M^k \nu_0$ is supported on that set, which proves minimality. □

Proposition 2.6 (Eventual lower bounds supplied by mutation)**.** *If $p > 0$ and $m(x, \cdot) \geq \eta \pi_0(\cdot)$ for some probability measure π_0 and $\eta > 0$, then there is $\delta > 0$, depending only on the coefficients and the absorbing slab, such that every nonzero orbit eventually satisfies $u_t \geq \delta \pi_0$. For a finite trait space, irreducibility of B is enough to obtain an eventual positive lower bound for every coordinate.*

Proof. On $\mathcal{K}$, let $h_u \geq -H$, with $H > 0$. Then

$$\dot{u} \geq -Hu + p\eta b_- \ell\, \pi_0.$$

Integration on each unit interval after entry into $\mathcal{K}$ gives

$$u_{t+1} \geq p\eta b_- \ell \frac{1 - e^{-H}}{H}\, \pi_0.$$

In the finite irreducible case, the comparison

$$u_{t+1} \geq e^{-H} e^{pB} u_t$$

and strict positivity of every entry of e^{pB}, together with $N_t \geq \ell$, give the coordinate bounds. □

Example 2.7 (A weak limit that is not stationary)**.** The Feller assumption cannot be omitted from the weak dynamical conclusions. Let

$$\mathcal{X} = \{0\} \cup \{1/n : n \geq 1\}, \qquad T(1/n) = 1/(n+1), \qquad T(0) = 1,$$

and put $m(x, \cdot) = \delta_{T(x)}$. Take $b = r = C = p = 1$ and $d = \tau = 0$. All coefficient assumptions hold, but m is not Feller. For $u_0 = 2\delta_1$, the solution is

$$u_t = 2e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \delta_{1/(k+1)}.$$

It has constant mass 2 and converges weakly to $2\delta_0$. The vector field at that limit equals

$$\mathcal{F}(2\delta_0) = 2(\delta_1 - \delta_0) \neq 0.$$

In fact, this example has no nonzero equilibrium. An equilibrium would have mass 2 and its normalized measure would satisfy $M\nu = \nu$. Since T has no preimage of 0, invariance forces $\nu(\{0\}) = 0$; then $\nu(\{1\}) = 0$, and recursively every point has zero mass. This is impossible for a probability measure on the countable space $\mathcal{X}$.

2.1 Averaged balance in the presence of mutation

The coefficient $h_u = r - Cu + A(u/N)$ omits the nonnegative mutation influx. At a positive finite-dimensional equilibrium, a negative value of $h_u(i)$ can be balanced by incoming offspring. The next result separates this local coefficient from total logarithmic growth and shows which averaged inequality survives for a general mutation kernel.

Proposition 2.8 (Averaged mutation-input balance)**.** *Assume $p > 0$, let $S = \operatorname{supp}\Lambda$ be the reachable support in Proposition 2.5, and define*

$$\overline{u}_T = \frac{1}{T}\int_0^T u_t\,dt, \qquad \overline{\nu}_T = \frac{1}{T}\int_0^T \nu_t\,dt, \qquad F_T = r - C\overline{u}_T + A\overline{\nu}_T.$$

Then, without any Feller assumption,

$$\limsup_{T\to\infty} \max_{x\in S} F_T(x) \leq 0. \tag{2.9}$$

In particular, every weak cluster pair $(\overline{u}, \overline{\nu})$ satisfies

$$r - C\overline{u} + A\overline{\nu} \leq 0 \quad \text{on } S.$$

Equality of the maximum to zero and concentration on the zero set are not asserted.

If the trait space is finite and B is irreducible, then for every $s > 0$ and every coordinate i,

$$\frac{1}{T}\int_s^T \left[h_{u_t}(i) + p\frac{(Bu_t)_i}{u_i(t)}\right] dt = \frac{\log u_i(T) - \log u_i(s)}{T} = O(T^{-1}). \tag{2.10}$$

There are coefficient-dependent constants $\kappa_i > 0$ such that

$$\limsup_{T\to\infty} F_T(i) \leq -\kappa_i < 0. \tag{2.11}$$

Proof. Fix $s > 0$. Proposition 2.5 gives $\operatorname{supp} u_s = S$. Along the orbit, let

$$R_+ = \max\{N_0, a_+/C_-\}, \qquad H = \sup_{t\geq 0, x\in\mathcal{X}} |h_{u_t}(x)| < \infty.$$

The nonnegative mutation gain and variation of constants imply the measure inequality

$$u_T(dx) \geq \exp\left(\int_s^T h_{u_t}(x)\,dt\right) u_s(dx), \qquad T \geq s.$$

The functions F_T have a common modulus of continuity

$$\omega(\delta) = \omega_r(\delta) + R_+\omega_C(\delta) + \omega_\alpha(\delta).$$

For fixed $\delta > 0$, compactness of the support gives

$$q_s(\delta) = \inf_{x \in S} u_s(B(x, \delta)) > 0$$

by a finite-cover argument. At a maximizer of $F_T|_S$, the preceding measure inequality therefore yields

$$R_+ \geq N_T \geq q_s(\delta) \exp\{T(\max_S F_T - \omega(\delta)) - Hs\}.$$

Consequently

$$\max_S F_T \leq \omega(\delta) + \frac{Hs + \log(R_+/q_s(\delta))}{T}.$$

First let $T \to \infty$, and then let $\delta \downarrow 0$. Weak convergence of the environmental averages gives uniform convergence of their continuous-kernel potentials, which proves the cluster-pair assertion.

For finite irreducible mutation, all coordinates are positive for every positive time. Proposition 2.6 gives an eventual common lower bound $u_i(t) \geq \delta_* > 0$, while eventually $N_t \leq U$. The component equation can thus be divided by $u_i(t)$ and integrated from any $s > 0$. The logarithms in (2.10) are uniformly bounded for $T \geq s$, proving the $O(T^{-1})$ estimate. Finally,

$$p\frac{(Bu_t)_i}{u_i(t)} \geq \frac{p\delta_*}{U}(B\mathbf{1})_i =: \kappa_i > 0$$

for all sufficiently large t. Every row of an irreducible mutation matrix is nonzero. The bounded contribution of the initial interval does not affect the time average, so (2.11) follows. □

Example 2.9 (Negative local growth balanced by mutation)**.** On a one-point trait space, take $b > 0$, $0 < r \leq b$, $C = c > 0$, $p > 0$, and the unique probability mutation kernel. The positive equilibrium has mass $N_* = (r + pb)/c$, and its local coefficient is

$$h_* = r - cN_* = -pb < 0.$$

The mutation gain exactly compensates this loss. Thus, with mutation present, even a stationary population may have a strictly negative maximal local averaged fitness and an empty zero-fitness set.

3 Accumulated fitness and averaged environments

Assume $p = 0$ throughout this section, and let $S = \operatorname{supp} u_0$. Set

$$\overline{u}_T = \frac{1}{T}\int_0^T u_s\, ds, \qquad \overline{\nu}_T = \frac{1}{T}\int_0^T \nu_s\, ds, \qquad F_T = r - C\overline{u}_T + A\overline{\nu}_T.$$

These integrals may be taken in total variation; equivalently they are characterized by integration against continuous test functions. These two averages enter separately because competition acts through $u_t = N_t\nu_t$, whereas transfer acts through ν_t. Their limits need not describe a stationary population. The mutation-free equation is multiplicative at each trait and gives

$$u_T(dx) = e^{TF_T(x)}u_0(dx). \tag{3.1}$$

Put

$$R_- = \min\{N_0, r_-/C_+\}, \qquad R_+ = \max\{N_0, r_+/C_-\}.$$

Then $R_- \leq N_t \leq R_+$.

Theorem 3.1 (Uniform maximal averaged fitness)**.** *Let $M_T = \max_S F_T$. Then*

$$\lim_{T\to\infty} M_T = 0.$$

More precisely, define

$$q(\delta) = \inf_{x\in S} u_0(B(x,\delta)).$$

For every $\delta > 0$, $q(\delta) > 0$. If ω is any common modulus of continuity of the functions $F_T|_S$, then

$$\frac{\log(R_-/N_0)}{T} \le M_T \le \omega(\delta) + \frac{1}{T}\log\frac{R_+}{q(\delta)}. \tag{3.2}$$

One such modulus is

$$\omega(\delta) = \omega_r(\delta) + R_+\omega_C(\delta) + \omega_\alpha(\delta),$$

where the kernel moduli refer to their first variable, uniformly in the second variable.

If, for $0 < \delta \le \delta_0$,

$$\omega(\delta) \le L\delta^\beta, \qquad q(\delta) \ge c\delta^\kappa$$

with $L, c > 0$, $\beta > 0$, and $\kappa \ge 0$, then, whenever $T^{-1/\beta} \le \delta_0$,

$$M_T \le \frac{(\kappa/\beta)\log T + L + \log(R_+/c)}{T}. \tag{3.3}$$

For every $\varepsilon > 0$ and every T with $T^{-1/\beta} \le \delta_0$, the moving fitness-deficit set satisfies

$$\nu_T\{F_T \le M_T - \varepsilon\} \le \frac{N_0}{c} e^L T^{\kappa/\beta} e^{-\varepsilon T}. \tag{3.4}$$

Proof. Compactness of S and the definition of support imply $q(\delta) > 0$: choose a finite cover of S by balls of radius $\delta/2$, each centred in S. If $x \in B(x_j, \delta/2)$, then $B(x_j, \delta/2) \subset B(x, \delta)$. The minimum of the positive masses of these finitely many balls is a lower bound.

The upper estimate $N_T \le N_0 e^{TM_T}$ proves the lower bound in (3.2). Let $x_T \in S$ maximize F_T. On $B(x_T, \delta) \cap S$, one has $F_T \ge M_T - \omega(\delta)$, and hence

$$N_T \ge q(\delta) e^{T(M_T - \omega(\delta))}.$$

This proves the upper bound. First let $T \to \infty$, then $\delta \downarrow 0$, to obtain $M_T \to 0$. Taking $\delta = T^{-1/\beta}$ gives (3.3). Finally,

$$u_T\{F_T \le M_T - \varepsilon\} \le N_0 e^{T(M_T - \varepsilon)}.$$

Dividing by the preceding lower bound for N_T, with the same choice of δ, proves (3.4). □

Remark 3.2. For a finite support with positive atomic weights, $q(\delta)$ is bounded below uniformly as $\delta \downarrow 0$. The maximal averaged fitness is then $O(T^{-1})$. The logarithmic factor in (3.3) records the cost of finding initial mass in a shrinking neighbourhood of a maximizing trait.

Theorem 3.3 (Joint-limit complementarity)**.** *Suppose $T_j \to \infty$ and*

$$u_{T_j} \rightharpoonup v, \qquad \overline{u}_{T_j} \rightharpoonup \overline{u}, \qquad \overline{\nu}_{T_j} \rightharpoonup \overline{\nu}.$$

Set $F = r - C\overline{u} + A\overline{\nu}$. Then

$$F \le 0 \text{ on } S, \qquad \max_S F = 0, \qquad \operatorname{supp} v \subseteq S \cap \{F = 0\}. \tag{3.5}$$

Moreover,

$$\frac{r_-}{C_+}\overline{\nu} \le \overline{u} \le \frac{r_+}{C_-}\overline{\nu} \tag{3.6}$$

in the order of measures.

Proof. Continuous kernels map weakly convergent bounded measures to uniformly convergent potentials, so $F_{T_j} \to F$ uniformly. Theorem 3.1 gives $\max_S F = 0$, which also implies $F \le 0$. Every point where $F < 0$ has a neighbourhood O and an $\varepsilon > 0$ such that $F_{T_j} \le -\varepsilon$ on O for all sufficiently large j. Formula (3.1) gives $u_{T_j}(O) \le N_0 e^{-\varepsilon T_j}$. The Portmanteau theorem yields $v(O) = 0$. Finally, (2.5) and $u_t = N_t \nu_t$, integrated against a nonnegative continuous function and averaged in time, give (3.6). □

Corollary 3.4 (A convergent averaged environment)**.** *If $F_T \to F$ uniformly as $T \to \infty$, every weak limit of u_T, ν_T, $\overline{u}_T$, or $\overline{\nu}_T$ is supported on $S \cap \{F = 0\}$. In particular, if this zero set is the singleton $\{x_*\}$, then*

$$\nu_t \rightharpoonup \delta_{x_*}, \qquad N_t \longrightarrow \frac{r(x_*)}{C(x_*, x_*)}, \qquad u_t \rightharpoonup \frac{r(x_*)}{C(x_*, x_*)} \delta_{x_*}.$$

Proof. On each compact subset of $\{F < 0\}$, formula (3.1) now gives exponential decay for every sufficiently large time. This decay also forces vanishing of all time-averaged masses there; the positive lower bound on N_t gives the same conclusion for frequencies. If the zero set is a singleton, compactness yields convergence of the entire frequency trajectory. Then $\langle r, \nu_t \rangle \to r(x_*)$ and $\langle C\nu_t, \nu_t \rangle \to C(x_*, x_*)$. The variable $z_t = 1/N_t$ satisfies

$$z_t' = \langle C\nu_t, \nu_t \rangle - \langle r, \nu_t \rangle z_t.$$

Comparison with asymptotically constant linear equations proves the claimed mass limit. □

Corollary 3.5 (An equilibrium obtained from convergent averages)**.** *Assume $p = 0$, $N_t \to N_* > 0$, and $\overline{\nu}_T \rightharpoonup \nu_*$. Then $N_*\nu_*$ is an equilibrium, and*

$$r - N_* C\nu_* + A\nu_* \le 0 \quad \text{on } S, \qquad r - N_* C\nu_* + A\nu_* = 0 \quad \text{on } \operatorname{supp} \nu_*.$$

The frequency trajectory itself need not converge.

Proof. Boundedness and convergence of N_t imply $\overline{u}_T \rightharpoonup N_*\nu_*$. Thus F_T converges uniformly to the displayed fitness. Corollary 3.4 puts ν_* on its zero set, and Theorem 3.3 makes it nonpositive on S. Multiplication by $N_*\nu_*$ yields the stationary equation. □

Remark 3.6. The endpoint limit v in Theorem 3.3 generally differs from both time averages. One cannot replace $\overline{u}$ by v, or $\overline{\nu}$ by $v/v(\mathcal{X})$, without an additional convergence argument. If the two time averages converge along the entire time axis, Corollary 3.4 also gives $\operatorname{supp} \overline{u} = \operatorname{supp} \overline{\nu} \subseteq \{F = 0\}$.

3.1 A coefficient criterion for selection of one trait

The preceding estimates also yield a convergence criterion involving only coefficient inequalities and the initial support.

Theorem 3.7 (Uniform trait dominance)**.** *Assume $p = 0$, let $S = \operatorname{supp} u_0$, and let $\ell = r_-/(2C_+)$, $U = 2r_+/C_-$. Suppose $x_* \in S$ and a continuous function $\Delta : S \to [0, \infty)$ satisfy $\Delta^{-1}(0) = \{x_*\}$ and*

$$\begin{aligned} r(x) - r(x_*) - N\{C(x,y) - C(x_*,y)\} \qquad & x, y \in S, \\ + \alpha(x,y) - \alpha(x_*,y) \le -\Delta(x) \qquad & N \in [\ell, U]. \end{aligned} \tag{3.7}$$

Then

$$\nu_t \rightharpoonup \delta_{x_*}, \qquad u_t \rightharpoonup \frac{r(x_*)}{C(x_*, x_*)} \delta_{x_*}.$$

For each subset E closed in S with $x_* \notin E$, the frequency $\nu_t(E)$ tends to zero exponentially. If $u_0(\{x_*\}) > 0$, the frequency convergence holds in total variation. If $u_0(\{x_*\}) = 0$, then $\|\nu_t - \delta_{x_*}\|_{\mathrm{TV}} = 2$ for every finite t. The coefficient condition need only be checked at $N = \ell$ and $N = U$.

Proof. After a finite time t_0, the mass belongs to $[\ell, U]$. Integrating (3.7) against $\nu_t(dy)$ yields

$$h_{u_t}(x) - h_{u_t}(x_*) \le -\Delta(x), \qquad t \ge t_0.$$

The bounded contribution of the initial interval gives a constant K_0 such that

$$T\{F_T(x) - F_T(x_*)\} \le K_0 - T\Delta(x), \qquad T \ge t_0,\ x \in S. \tag{3.8}$$

Let E be closed and avoid x_*, and put $\Delta_E = \min_E \Delta > 0$. Choose a ball about x_* whose radius is small enough that the common modulus of continuity of F_T is at most $\Delta_E/2$. The initial measure of this ball is positive. Comparing its contribution to (3.1) with the contribution of E, using (3.8), proves $\nu_T(E) \le K_E e^{-\Delta_E T/2}$. This proves weak concentration; the mass limit follows as in Corollary 3.4.

If $a_0 = u_0(\{x_*\}) > 0$, the exact atomic mass is $u_T(\{x_*\}) = a_0 e^{TF_T(x_*)}$. Thus

$$\frac{u_T(S \setminus \{x_*\})}{u_T(\{x_*\})} \le \frac{e^{K_0}}{a_0} \int_{S\setminus\{x_*\}} e^{-T\Delta(x)} u_0(dx) \longrightarrow 0$$

by dominated convergence. This gives total variation convergence. When $a_0 = 0$, absolute continuity with respect to u_0 preserves zero mass at x_*, so the total variation distance between the two probabilities is exactly 2. The final assertion follows from the affine dependence on N. □

Remark 3.8 (The role of the initial support). For $p = 0$, every $x \in \mathcal{X}$ yields an equilibrium

$$u^x = \frac{r(x)}{C(x,x)} \delta_x.$$

Therefore, on a trait space containing more than one point, uniqueness and global attraction over all nonzero initial measures are impossible in the mutation-free model. Theorem 3.7 specifies its basin through the presence of x_* in the initial support. This qualification is essential even when the coefficient inequalities are strict.

4 Invariant resident laws and invasion

Assume the weak semiflow hypotheses of Theorem 2.3. Invariant resident laws below are probability measures on the absorbing compact set $\mathcal{K}$. An occupation law records how time is distributed over population states, rather than over traits. It retains the joint variation of total mass and composition that the two environmental averages alone do not determine. For a fixed orbit, define its occupation law

$$\Pi_T = \frac{1}{T} \int_0^T \delta_{u_t}\, dt.$$

The barycentres of probability laws on bounded-mass sets are understood weakly, by duality with $C(\mathcal{X})$.

Proposition 4.1 (Invariant occupation laws)**.** *Every sequence $T_j \to \infty$ has a subsequence along which Π_{T_j} converges weakly to an invariant probability law Π, supported on $\omega(u_0)$. For every $f \in C(\mathcal{X})$,*

$$\int [\langle f h_u, u\rangle + p\langle f, Bu\rangle]\, \Pi(du) = 0, \tag{4.1}$$

and

$$\int [\langle a, \nu\rangle - N\langle C\nu, \nu\rangle]\, \Pi(du) = 0. \tag{4.2}$$

The barycentres

$$\overline{u}_\Pi = \int u\, \Pi(du), \qquad \overline{\nu}_\Pi = \int \frac{u}{u(\mathcal{X})}\, \Pi(du)$$

are the corresponding limits of the two environmental averages. If $p = 0$ and $S = \operatorname{supp} u_0$, then

$$r - C\overline{u}_\Pi + A\overline{\nu}_\Pi \le 0 \quad \text{on } S, \qquad \max_S (r - C\overline{u}_\Pi + A\overline{\nu}_\Pi) = 0. \tag{4.3}$$

Proof. The orbit lies in a weak compact mass annulus. Occupation laws are therefore relatively compact. For a bounded continuous function Φ on that annulus and $s > 0$,

$$\left| \int (\Phi \circ S_s - \Phi)\, d\Pi_T \right| \le \frac{2s\|\Phi\|_\infty}{T}.$$

Weak continuity of S_s gives invariance in the limit. Orbit tails approach the omega-limit set, which proves the support assertion. Integrating the derivatives of $\langle f, u_t\rangle$ and $\log N_t$, dividing by T, and using their boundedness proves (4.1) and (4.2). The integrands are weakly continuous. The maps $u \mapsto u$ and $u \mapsto u/u(\mathcal{X})$, tested against a continuous function, are also weakly continuous on the mass annulus, which identifies the barycentres. Equation (4.3) follows from Theorem 3.1. □

Remark 4.2. The law $\Pi = \delta_{u_*}$ represents a stationary resident. The time distribution on a periodic orbit represents a periodic resident. More general invariant laws describe recurrent environments without assuming a single limiting state. The barycentre $\overline{u}_\Pi$ need not be an equilibrium: the stationary moment identity contains $\mathbb{E}_\Pi[u \otimes u]$ and $\mathbb{E}_\Pi[u \otimes u/N]$. In particular, $\overline{u}_\Pi = \overline{N}_\Pi\, \overline{\nu}_\Pi$ need not hold.

Proposition 4.3 (Scalar invasion in an ergodic resident environment)**.** *Assume $p = 0$ and let Π be an ergodic invariant resident law. A rare trait x, absent from the resident, has linearized equation*

$$w' = h_{u_t}(x) w.$$

For Π-almost every resident initial state,

$$\lim_{t\to\infty} \frac{1}{t} \log \frac{w_t}{w_0} = \Lambda(x, \Pi) := r(x) - (C\overline{u}_\Pi)(x) + (A\overline{\nu}_\Pi)(x). \tag{4.4}$$

For a finite trait space, each resident coordinate is either zero Π-almost everywhere or positive Π-almost everywhere. In the latter case its invasion index equals zero.

Proof. The first statement follows from the exact logarithmic growth formula and the ergodic theorem, since $h_u(x)$ is bounded and continuous on the resident mass annulus. In finite dimensions, the events $\{u_i = 0\}$ and $\{u_i > 0\}$ are invariant for the mutation-free flow, so ergodicity gives the dichotomy. In the positive case choose $\delta > 0$ with $\Pi\{u_i \ge \delta\} > 0$. There are resident trajectories that are typical for the ergodic theorem and return to this set at arbitrarily large times. Along the return times,

$$\frac{\log u_i(t) - \log u_i(0)}{t} \longrightarrow 0,$$

because $\delta \le u_i(t) \le U$. Its full-time limit is the ergodic mean $\Lambda_i(\Pi)$, so that mean is zero. □

For a nonergodic invariant law, the integral on the right-hand side of (4.4) still defines $\Lambda(x, \Pi)$. It is the average of the growth indices of the ergodic components, rather than a constant growth rate along almost every individual resident trajectory.

4.1 A finite-dimensional persistence theorem

A positive scalar invasion index describes growth in the linearized rare-trait equation. Uniform persistence of all traits requires control of the boundary environments that an interior trajectory can approach. The next theorem obtains this control from a single positive weighted sum of invasion indices over every boundary invariant law. Its proof uses an average Lyapunov function; the broader structured-population theory is developed in [12].

Theorem 4.4 (Persistence from boundary invariant laws)**.** *Let $\mathcal{X} = \{1, \dots, n\}$, $p = 0$, and let*

$$\mathcal{K}_\partial = \{u \in \mathcal{K} : \text{some } u_i = 0\}.$$

Suppose there are $c_i > 0$ such that

$$\sum_{i=1}^{n} c_i \Lambda_i(\Pi) > 0 \tag{4.5}$$

for every invariant probability law supported on $\mathcal{K}_\partial$. Then there is $\delta > 0$, depending only on the coefficients and c_i, such that every solution with $u_i(0) > 0$ for all i satisfies

$$\liminf_{t\to\infty} u_i(t) \ge \delta \qquad (i = 1, \dots, n).$$

Proof. If the boundary is empty, compactness of $\mathcal{K}$ already gives the conclusion. Otherwise, write $g(u) = \sum_i c_i h_u(i)$. The set of invariant probability laws on the compact boundary is compact. Hence (4.5) has a uniform positive lower bound, say $\rho > 0$. There is $T > 0$ such that

$$\frac{1}{T}\int_0^T g(S_s u)\, ds \ge \rho/2 \qquad (u \in \mathcal{K}_\partial).$$

Otherwise, boundary occupation laws along time intervals whose lengths tend to infinity would have an invariant limit with integral at most $\rho/2$. Continuity extends this inequality, with $\rho/4$, to a neighbourhood of the boundary.

Let $P(u) = \prod_i u_i^{c_i}$. It vanishes precisely on the boundary. Choose $\varepsilon > 0$ so that $\{P \le \varepsilon\} \cap \mathcal{K}$ lies in that neighbourhood. For an interior point in this set,

$$P(S_T u) \ge e^{\rho T/4} P(u).$$

Let $G = \sup_{\mathcal{K}} |g|$. For points with $P(u) > \varepsilon$, one always has

$$P(S_T u) \ge e^{-GT}\varepsilon.$$

Consequently the discrete sequence $P(S_{kT}u)$, once the orbit enters $\mathcal{K}$, eventually stays above εe^{-GT}. Between sampling times it stays above εe^{-2GT}. Since every coordinate is at most U, this lower bound for the product gives a positive lower bound for every coordinate, uniformly over positive initial states. □

Example 4.5 (Invadable boundary equilibria without persistence)**.** Consider the mutation-free, transfer-free three-trait system with $r_i = 1$ and

$$C = \begin{pmatrix} 1 & a & b \\ b & 1 & a \\ a & b & 1 \end{pmatrix}, \qquad a > 1 > b > 0, \qquad s := a + b - 2 > 0.$$

This is the classical cyclic competition family studied in [14]. The only positive-mass boundary equilibria are the three unit vectors. Each can be invaded by one missing trait, whose growth rate is $1 - b > 0$. Nevertheless the invariant law

$$\Pi_\partial = \tfrac{1}{3}(\delta_{e_1} + \delta_{e_2} + \delta_{e_3})$$

has

$$\Lambda_i(\Pi_\partial) = \frac{2 - a - b}{3} < 0 \qquad (i = 1, 2, 3).$$

Thus (4.5) fails for every choice of positive weights.

The failure of persistence can also be seen directly. Write $u = Nq$, $\sum_i q_i = 1$, and $H = q_1 q_2 q_3$. For positive solutions,

$$\frac{d}{dt} \log H = -sN \left(1 - 3 \sum_{i<j} q_i q_j \right) = -\frac{sN}{2} \sum_{i<j} (q_i - q_j)^2 \leq 0.$$

If $q(0) \neq (1/3, 1/3, 1/3)$, the limit of H must be zero. Indeed, a positive limit would keep the trajectory away from the boundary. Compactness and invariance would then force its omega-limit set into the equality set $q_1 = q_2 = q_3$, where $H = 1/27 > H(0)$, contradicting monotonicity. Hence such solutions approach the extinction boundary in frequency despite invadability of every boundary equilibrium.

4.2 Mutation classes and positive linear growth

For a finite trait space, write $B_{ij} = b_j m(j, \{i\})$, so columns index parent traits and rows index offspring traits. With mutation, a resident face is invariant only when no mutation flows from that face to its complement. Let R denote the resident trait set and let J be its complement, so that $B_{JR} = 0$. Linearization in the rare traits then gives

$$w' = \left[\operatorname{diag}(h_{u_t}|_J) + pB_{JJ} \right] w. \tag{4.6}$$

The dependence of the resident environment on w contributes only to higher-order terms in this transverse equation. For a stationary resident the relevant rate is the spectral bound of the constant Metzler matrix. For a periodic resident it is the logarithm of the spectral radius of the period map, divided by the period. For an ergodic resident law, subadditivity gives the top matrix Lyapunov exponent. If $p > 0$ and B_{JJ} is irreducible, bounded coefficients and positivity of its exponential identify a common principal growth rate for nonzero positive rare populations. The same discussion applies to an autonomous block of the missing-trait system; couplings from other missing blocks must be retained when they are present.

In an infinite-dimensional trait class, one can similarly define the positive evolution family and an operator-norm growth rate. A principal positive direction and identification of that rate with growth from arbitrary positive initial measures require additional positivity or compactness. Feller continuity alone does not provide those conclusions. Moreover, if mutation directly supplies a missing class, its dynamics are inhomogeneous; an invasion equation of the homogeneous form (4.6) is then inappropriate.

5 Potential differences and circulation

Only the net transfer kernel $\alpha = \tau - \tau^{\mathsf{T}}$ enters the deterministic equation. Conversely, every continuous antisymmetric kernel is realized by the nonnegative continuous rate $\tau(x, y) = \max\{\alpha(x, y), 0\}$.

Proposition 5.1 (Potential differences and triangular circulation)**.** *For a continuous antisymmetric kernel α, the following are equivalent:*

1. *there is $h \in C(\mathcal{X})$ such that $\alpha(x, y) = h(x) - h(y)$;*
2. *$\alpha(x, y) + \alpha(y, z) + \alpha(z, x) = 0$ for all $x, y, z \in \mathcal{X}$.*

The potential h is unique up to an additive constant.

Proof. A potential difference has zero sum around a triangle. Conversely, fix $x_0 \in \mathcal{X}$ and put $h(x) = \alpha(x, x_0)$. The triangle with vertices x, y, x_0 gives the representation. Subtracting two representations shows that their potentials differ by a constant. □

Every antisymmetric kernel on a two-point set satisfies this criterion. Here triangular circulation means the sum of the net transfer coefficients around a three-trait cycle. It can first occur with three traits. The criterion concerns the transfer kernel. Under the additional recipient-independent competition assumption in Section 6, it represents the frequency dynamics by a normalized linear semigroup; the total mass still satisfies a scalar nonlinear equation.

Proposition 5.2 (A reference-dependent orthogonal decomposition)**.** *For $\sigma \in \mathcal{P}(\mathcal{X})$, set*

$$h_\sigma(x) = \int \alpha(x, y)\sigma(dy), \qquad \beta_\sigma(x, y) = \alpha(x, y) - h_\sigma(x) + h_\sigma(y).$$

Then $\sigma(h_\sigma) = 0$, β_σ is antisymmetric, and $\beta_\sigma\sigma = 0$ pointwise. Moreover, for every continuous h,

$$\|\alpha - h(x) + h(y)\|^2_{L^2(\sigma\otimes\sigma)} = \|\beta_\sigma\|^2_{L^2(\sigma\otimes\sigma)} + 2\|h - h_\sigma - \sigma(h)\|^2_{L^2(\sigma)}. \tag{5.1}$$

If σ has full support, h_σ is the unique continuous least-squares potential after imposing $\sigma(h) = 0$.

Proof. Antisymmetry gives $\sigma(h_\sigma) = 0$ and hence $\beta_\sigma\sigma = h_\sigma - h_\sigma + \sigma(h_\sigma) = 0$. For any k,

$$\iint \beta_\sigma(x, y)[k(x) - k(y)]\, \sigma(dx)\sigma(dy) = 0.$$

Thus the two components are orthogonal. Finally,

$$\iint [k(x) - k(y)]^2\, \sigma(dx)\sigma(dy) = 2\|k - \sigma(k)\|^2_{L^2(\sigma)}.$$

Take $k = h - h_\sigma$. Full support turns equality almost everywhere between continuous potentials into pointwise equality. □

For every probability ν, this decomposition gives

$$(a - \nu(a) + A\nu)\nu = \big[a + h_\sigma - \nu(a + h_\sigma) + \beta_\sigma\nu\big]\nu. \tag{5.2}$$

The residual kernel β_σ has the same triangular circulation as α. The potential part is incorporated into the trait-dependent growth term, while $\beta_\sigma\nu$ retains the dependence on composition.

5.1 Hamiltonian structure in the finite pure-transfer system

For a finite trait set, let $A^{\mathsf{T}} = -A$ and consider

$$\dot{q}_i = q_i(Aq)_i, \qquad q_i > 0, \qquad \sum_i q_i = 1.$$

In logarithmic coordinates $y_i = \log q_i$,

$$\dot{y} = A\nabla H(y), \qquad H(y) = \sum_i e^{y_i}. \tag{5.3}$$

A constant antisymmetric matrix defines a Poisson bracket, so this is a Hamiltonian system, possibly with a degenerate Poisson structure. For $v \in \ker A$, the quantity $v \cdot y$ is conserved. In particular, if $\pi > 0$, $\sum_i \pi_i = 1$ and $A\pi = 0$, then

$$\mathcal{H}(\pi \mid q) = \sum_i \pi_i \log(\pi_i/q_i)$$

is conserved. These identities explain why antisymmetry alone does not give dissipation.

For the full finite frequency equation without mutation, write $C = C_s + C_a$, with $C_s^{\mathsf{T}} = C_s$ and $C_a^{\mathsf{T}} = -C_a$. At a fixed mass N, the selection and symmetric-competition part has the replicator-gradient form generated by

$$\Phi_N(q) = a \cdot q - \frac{N}{2} q^{\mathsf{T}} C_s q.$$

The corresponding frequency field has components $q_i[\partial_i \Phi_N(q) - \sum_j q_j \partial_j \Phi_N(q)]$. The matrix $A - NC_a$ supplies the remaining antisymmetric interaction. Variation of N couples these parts to the mass equation; the fixed-mass potential Φ_N is not asserted to be a Lyapunov function for that coupled system.

6 Potential transfer and Perron convergence

Throughout this section the competition and transfer kernels satisfy

$$C(x,y) = c(y), \qquad \alpha(x,y) = h(x) - h(y), \qquad c \in C(\mathcal{X}), \quad \min c > 0, \tag{6.1}$$

where $h \in C(\mathcal{X})$. Competition then imposes the same per-capita loss on every recipient trait, so it cancels from the frequency equation. The transfer contribution is $h - \nu(h)$ and can be combined with selection. Set

$$s = r + h, \qquad a = r + pb, \qquad V = s + pb = a + h, \qquad \mathcal{L}\mu = s\mu + pB\mu = pQ\mu + V\mu.$$

The bounded operator $\mathcal{L}$ generates a positive, uniformly continuous semigroup on the Banach space of finite signed measures with its total variation norm. We use the full total variation norm, so that the distance between two probabilities is at most 2.

Proposition 6.1 (Exact projective reduction)**.** *Under* (6.1)*, every nonzero solution has the form*

$$\nu_t = \frac{e^{t\mathcal{L}}u_0}{(e^{t\mathcal{L}}u_0)(\mathcal{X})}, \qquad N_t' = N_t\left[\int a\, d\nu_t - N_t \int c\, d\nu_t\right]. \tag{6.2}$$

If $\nu_t \to \nu_*$ in total variation, then

$$u_t \longrightarrow N_*\nu_*, \qquad N_* = \frac{\int a\, d\nu_*}{\int c\, d\nu_*}. \tag{6.3}$$

If the frequency convergence is exponential with exponent $\kappa > 0$, then the full convergence is exponential with any exponent $0 < \omega < \min\{\kappa, \min a\}$.

Proof. The original equation becomes

$$\dot{u} = \mathcal{L}u - \left[\int c\,du + \int h\,d\nu\right] u.$$

The last term is a scalar multiple of u. Multiplication by the scalar integrating factor therefore gives the first identity in (6.2); the mass equation follows from the general mass–frequency decomposition. For the convergence claim let $I_t = N_t^{-1}$, $A_t = \int a\,d\nu_t$, and $c_t = \int c\,d\nu_t$. Then

$$I_t' = -A_t I_t + c_t.$$

Here $A_t \geq \min a > 0$, while $A_t \to A_* = \int a\,d\nu_*$ and $c_t \to c_* = \int c\,d\nu_*$. Variation of constants applied to $I_t - c_*/A_*$ proves convergence. If the coefficients converge exponentially, the same formula gives the asserted rates, with a strict inequality in ω covering the possible resonant case. □

The limiting mass in (6.3) depends on the original effective growth $a = r + pb$. If λ is the Perron eigenvalue of $\mathcal{L}$, then $\lambda = \int (a + h)\,d\nu_*$, and consequently $N_* = (\lambda - \int h\,d\nu_*)/\int c\,d\nu_*$. This formula is invariant under addition of a constant to h.

6.1 A positive adjoint eigenfunction and a Doob transform

A positive adjoint eigenfunction identifies a weighted mass that grows at a single exponential rate under the linear semigroup. Dividing out that growth and changing the weight produces a mass-preserving process. A common reset for the transformed process then yields contraction. Conservative transforms of this kind are used in [1, 2]; the next result gives explicit constants for the bounded mutation operator considered here.

Theorem 6.2 (Quantitative convergence from an adjoint eigenfunction)**.** *Assume $p > 0$ and*

$$m(x, \cdot) \geq \eta q(\cdot) \quad (x \in \mathcal{X}) \tag{6.4}$$

for a probability q and $\eta > 0$. Suppose that a bounded Borel function ψ and a real number λ satisfy

$$0 < \psi_- := \inf \psi \leq \sup \psi =: \psi_+ < \infty, \qquad \mathcal{L}^*\psi = \lambda\psi \quad \text{pointwise on } \mathcal{X}. \tag{6.5}$$

Put

$$\theta = p\eta \min b, \qquad \kappa_\psi = \theta \frac{\int \psi\,dq}{\psi_+}.$$

Then there is a unique probability eigenmeasure ν_* of $\mathcal{L}$, and its eigenvalue is λ. More precisely, there is a positive measure γ with $\int \psi\,d\gamma = 1$ such that

$$\left\| e^{-\lambda t} e^{t\mathcal{L}}\mu - \left(\int \psi\,d\mu\right)\gamma \right\|_{\mathrm{TV}} \leq \frac{e^{-\kappa_\psi t}}{\psi_-} \left\| \psi\mu - \left(\int \psi\,d\mu\right)\pi \right\|_{\mathrm{TV}}, \tag{6.6}$$

where $\pi = \psi\gamma$ is a probability. For every nonzero $\mu \geq 0$,

$$\left\| \frac{e^{t\mathcal{L}}\mu}{(e^{t\mathcal{L}}\mu)(\mathcal{X})} - \nu_* \right\|_{\mathrm{TV}} \leq 4\frac{\psi_+}{\psi_-} e^{-\kappa_\psi t}, \qquad \nu_* = \frac{\gamma}{\gamma(\mathcal{X})}. \tag{6.7}$$

The full population converges to (6.3).

Proof. Define, on finite measures,

$$P_t\zeta = \psi\, e^{-\lambda t} e^{t\mathcal{L}}(\zeta/\psi).$$

The adjoint eigenfunction identity implies that P_t preserves mass. Its jump-rate kernel is

$$k_\psi(x, dy) = pb(x)\frac{\psi(y)}{\psi(x)} m(x, dy). \tag{6.8}$$

Indeed, its total jump rate is $pb(x)\int \psi(y)m(x, dy)/\psi(x) = \lambda - s(x)$, so the diagonal term is precisely the negative of this rate. Let $q_\psi(dy) = \psi(y)q(dy)/\int \psi\, dq$. The lower bound

$$k_\psi(x, dy) \geq \kappa_\psi q_\psi(dy)$$

allows us to subtract a common reset generator. The remainder is a Markov generator. On the subspace of measures of zero total mass, the reset generator is $-\kappa_\psi I$, and hence

$$\|P_t\zeta - P_t\widetilde{\zeta}\|_{\mathrm{TV}} \leq e^{-\kappa_\psi t}\|\zeta - \widetilde{\zeta}\|_{\mathrm{TV}}$$

for probabilities $\zeta, \widetilde{\zeta}$. The complete metric space of probabilities therefore has a unique fixed point for P_1. Commutation with P_t shows that this fixed point, denoted by π, is invariant for every P_t. Conjugating the contraction gives (6.6). For $\mu \geq 0$, its right-hand side is at most $2\psi_-^{-1}(\int \psi\, d\mu)e^{-\kappa_\psi t}$, while

$$(e^{-\lambda t}e^{t\mathcal{L}}\mu)(\mathcal{X}) \geq \psi_+^{-1}\int \psi\, d\mu.$$

The elementary normalization estimate $\|v/v(\mathcal{X}) - w/w(\mathcal{X})\|_{\mathrm{TV}} \leq 2\|v - w\|_{\mathrm{TV}}/v(\mathcal{X})$ proves (6.7). Finally, if $\widetilde{\gamma} \geq 0$ is any nonzero eigenmeasure with eigenvalue $\widetilde{\lambda}$, pairing with ψ yields $(\widetilde{\lambda} - \lambda)\int \psi\, d\widetilde{\gamma} = 0$. Thus $\widetilde{\lambda} = \lambda$, and the normalized convergence identifies $\widetilde{\gamma}$ up to scale. □

No Feller assumption is needed in this theorem. If the kernel is Feller, one may seek ψ in $C(\mathcal{X})$; for a merely measurable kernel, the appropriate space in (6.5) is $B_b(\mathcal{X})$. The two-sided bound on ψ makes the weighted and ordinary total variation norms equivalent. At the spectral edge in the rank-one examples below, no strictly positive bounded adjoint eigenfunction exists.

6.2 A coefficient condition that constructs the eigenfunction

Theorem 6.3 (Mixing stronger than the oscillation of the linear potential)**.** *Assume* (6.4)*, and put*

$$\theta = p\eta \min b, \qquad V = r + h + pb.$$

If

$$\theta > \mathrm{osc}(V) := \sup V - \inf V, \tag{6.9}$$

then the hypotheses of Theorem 6.2 hold. They can be realized with

$$\inf V \leq \lambda \leq \sup V, \qquad \frac{1}{\lambda - \inf V + \theta} \leq \psi(x) \leq \frac{1}{\lambda - \sup V + \theta}, \qquad \theta\int \psi\, dq = 1.$$

Consequently

$$\kappa_\psi = \frac{1}{\psi_+} \geq \theta + \lambda - \sup V \geq \theta - \mathrm{osc}(V) > 0. \tag{6.10}$$

In particular, the population has a unique nonzero equilibrium, which attracts every nonzero initial measure in total variation.

Proof. Let $\Pi_q \mu = q\mu(\mathcal{X})$ and set

$$R = pQ - \theta(\Pi_q - I), \qquad T = R + (V - \theta)I.$$

The kernel $pb(x)m(x, dy) - \theta q(dy)$ is nonnegative and has total mass $pb(x) - \theta$. Thus R is a bounded Markov generator, and

$$\mathcal{L} = T + \theta\Pi_q.$$

Positivity and comparison give

$$e^{(\inf V - \theta)t} \le e^{tT^*}1(x) \le e^{(\sup V - \theta)t}.$$

For $z > \sup V - \theta$, define

$$\psi_z = (z - T^*)^{-1}1 = \int_0^\infty e^{-zt}e^{tT^*}1\, dt, \qquad F(z) = \theta \int \psi_z\, dq.$$

The function F is continuous and strictly decreasing. Moreover,

$$\frac{\theta}{z - \inf V + \theta} \le F(z) \le \frac{\theta}{z - \sup V + \theta}.$$

Condition (6.9) places the interval $[\inf V, \sup V]$ inside the domain of this resolvent. Its endpoints satisfy $F(\inf V) \ge 1 \ge F(\sup V)$, including the constant-potential case in which the endpoints coincide. There is therefore a unique λ in this interval with $F(\lambda) = 1$. Set $\psi = \psi_\lambda$. Then

$$\mathcal{L}^*\psi = T^*\psi + \theta \int \psi\, dq = \lambda\psi - 1 + 1 = \lambda\psi.$$

The stated bounds follow from the comparison estimate. They imply strict positivity and boundedness. If m is Feller, T^* acts boundedly on $C(\mathcal{X})$, and its resolvent preserves that space; otherwise the same construction takes place in $B_b(\mathcal{X})$. Finally, the rate follows from Theorem 6.2 and the normalization $\theta q(\psi) = 1$. □

This theorem does not require compactness of the mutation operator. The operator-norm estimate (6.6) gives convergence to a rank-one projection, hence a simple isolated eigenvalue and a spectral gap on the measure space. The estimate is sufficient rather than necessary: the exact rank-one threshold in the next section can yield a positive gap even when (6.9) fails.

Corollary 6.4 (Finite traits and irreducible mutation)**.** *Suppose $\mathcal{X} = \{1, \dots, n\}$, $C_{ij} = c_j > 0$, $\alpha_{ij} = h_i - h_j$, and $p > 0$. Assume that the mutation matrix is irreducible, equivalently that $B_{ij} = b_j m(j, \{i\})$ is irreducible. Let*

$$L = \operatorname{diag}(r + h) + pB.$$

Then L has a simple Perron eigenvalue λ and a unique positive right eigenvector ν_* normalized by $\sum_i \nu_{*,i} = 1$. There are constants $K, \gamma > 0$, depending only on the coefficients, such that every nonzero nonnegative initial measure satisfies

$$\|\nu_t - \nu_*\|_1 \le Ke^{-\gamma t}.$$

The full population has the unique nonzero equilibrium

$$u_* = N_*\nu_*, \qquad N_* = \frac{\sum_i (r_i + pb_i)\nu_{*,i}}{\sum_i c_i\nu_{*,i}} = \frac{\lambda - \sum_i h_i\nu_{*,i}}{\sum_i c_i\nu_{*,i}}, \tag{6.11}$$

and every nonzero initial measure converges to it exponentially. No one-step minorization or oscillation condition is required.

Proof. Since $b_i > 0$, multiplication of the columns of the mutation matrix by b_i preserves its directed graph. Thus L is an irreducible Metzler matrix. A sufficiently large scalar shift makes L a nonnegative irreducible matrix with positive diagonal, so the Perron–Frobenius theorem gives positive right and left eigenvectors ν_*, ψ, with $\psi^\top \nu_* = 1$, and a strictly dominant real eigenvalue λ. Equivalently, e^{tL} is strictly positive for every $t > 0$. The finite-dimensional spectral decomposition gives

$$e^{-\lambda t} e^{tL} = \nu_* \psi^\top + O(e^{-\gamma t})$$

in operator norm for some $\gamma > 0$. In dimension one the same identity is exact. For a nonnegative vector v, one has $\psi^\top v \geq (\min_i \psi_i)\|v\|_1$; moreover,

$$\|e^{-\lambda t} e^{tL} v\|_1 \geq \frac{\psi^\top v}{\max_i \psi_i}.$$

Normalization therefore gives the frequency estimate uniformly over nonzero nonnegative v. Proposition 6.1 proves the full convergence and the first expression for N_*. Summing $L\nu_* = \lambda \nu_*$ gives $\lambda = \sum_i (r_i + h_i + p b_i)\nu_{*,i}$, proving the second expression. Every nonzero equilibrium has a stationary frequency, so the same frequency convergence establishes uniqueness of the full equilibrium. □

7 Rank-one mutation, spectral thresholds, and condensation

Assume (6.1), let $b \equiv b_0 > 0$, and suppose

$$m(x, \cdot) = q(\cdot), \qquad \operatorname{supp} q = \mathcal{X}, \qquad k = p b_0 > 0. \tag{7.1}$$

Then

$$\mathcal{L}\mu = s\mu + kq\mu(\mathcal{X}).$$

Define

$$M = \max_{\mathcal{X}} s, \qquad g = M - s, \qquad K = \{x : g(x) = 0\}, \qquad \rho = k \int_{\mathcal{X}} \frac{1}{g}\, dq \in (0, \infty], \tag{7.2}$$

where $1/g = +\infty$ on K. Thus K is the maximizing set of the linear potential $s = r + h$, and M is the upper spectral edge of multiplication by s. In particular, $\rho < \infty$ implies $q(K) = 0$. The support assumption on q excludes positive eigenmeasures with an eigenvalue below M.

At eigenvalue M, the stationary equation requires $q(K) = 0$ and prescribes the measure $k g^{-1} q$ off K. When $q(K) = 0$, this part has mass ρ. If $\rho < 1$, the remaining mass $1 - \rho$ must lie on K and is singular with respect to q; this is the condensate. If $\rho > 1$, normalization instead requires an eigenvalue strictly above M. The following theorem makes this distinction precise.

Theorem 7.1 (All stationary frequencies in the rank-one model)**.** *Under (7.1), the stationary probabilities of the normalized semigroup are exactly the following.*

1. *If $\rho > 1$, there is a unique $\lambda > M$ satisfying*

$$k \int \frac{q(dx)}{\lambda - s(x)} = 1. \tag{7.3}$$

The unique stationary probability is

$$\nu_*(dx) = \frac{k}{\lambda - s(x)} q(dx). \tag{7.4}$$

2. *If $\rho = 1$, the unique stationary probability is $\nu_* = kg^{-1}q$, and its eigenvalue is M.*

3. *If $\rho < 1$, the stationary probabilities form the family*

$$\nu_\zeta = kg^{-1}q + (1-\rho)\zeta, \qquad \zeta \in \mathcal{P}(K), \tag{7.5}$$

and every member has eigenvalue M.

Every such frequency gives precisely one nonzero equilibrium of the full population equation, with mass $\int (r+k)\,d\nu / \int c\,d\nu$.

Proof. A stationary probability is equivalent to $\mathcal{L}\nu = \lambda\nu$ for some real λ. Thus

$$(\lambda - s)\nu = kq. \tag{7.6}$$

If $\lambda < M$, the nonempty open set $\{s > \lambda\}$ has positive q-mass, contradicting the signs in (7.6). For $\lambda > M$, normalization is exactly (7.3). Its left side is strictly decreasing, tends to zero at infinity, and tends to ρ as $\lambda \downarrow M$. This proves the first case and excludes such a root when $\rho \le 1$. For $\lambda = M$, equation (7.6) fixes $\nu|_{\mathcal{X}\setminus K} = kg^{-1}q$ and leaves a nonnegative measure on K. Normalization fixes its mass to $1-\rho$, proving the remaining cases. The full equilibrium statement follows from Proposition 6.1. □

For a single maximizer and a mutation density, this threshold is the continuous-time house-of-cards threshold studied in [7]. The proof above applies directly to compact spaces and arbitrary full-support mutation measures, and records the freedom on an entire maximizing set. The following results establish global convergence when $\rho > 1$ and identify the selected condensate when $\rho < 1$ and $u_0(K) > 0$.

Corollary 7.2 (The separated eigenvalue regime)**.** *If $\rho > 1$, every nonzero initial measure satisfies*

$$\|\nu_t - \nu_*\|_{\mathrm{TV}} \le 4\frac{\lambda - \min s}{\lambda - M} e^{-(\lambda - M)t}. \tag{7.7}$$

Consequently the full population has a unique nonzero equilibrium, globally attracting in total variation.

Proof. Take $\psi(x) = (\lambda - s(x))^{-1}$. Equation (7.3) gives $kq(\psi) = 1$, and

$$\mathcal{L}^*\psi = s\psi + kq(\psi) = \lambda\psi.$$

The transformed process jumps with rate $\lambda - s(x)$ and destination law ν_*. Its common reset rate is $\lambda - M$. The result is Theorem 6.2 with $\psi_+/\psi_- = (\lambda - \min s)/(\lambda - M)$. □

7.1 Initial mass on the maximizing set selects the condensate

Theorem 7.3 (Total variation convergence in the condensed regime)**.** *Suppose $\rho < 1$ and let $u_0 \ge 0$ be finite and nonzero. Set $A_0 = u_0(K)$. If $A_0 > 0$, then*

$$\nu_t \longrightarrow \frac{k}{g}q + (1-\rho)\frac{u_0|_K}{A_0} \quad \textit{in total variation.} \tag{7.8}$$

In fact, the unnormalized semigroup satisfies

$$e^{-Mt}e^{t\mathcal{L}}u_0 \longrightarrow u_0|_K + \frac{kA_0}{1-\rho}\frac{q}{g} \quad \textit{in total variation.} \tag{7.9}$$

If $u_0(K) = 0$, the left side of (7.9) converges to zero in total variation, and the finite-time frequency assigns zero mass to K. Hence it cannot converge in total variation to any probability in (7.5).

Proof. Put $v_t = e^{-Mt}e^{t\mathcal{L}}u_0$ and $Z_t = v_t(\mathcal{X})$. The variation-of-constants formula is

$$v_t(dx) = e^{-g(x)t}u_0(dx) + kq(dx)\int_0^t e^{-g(x)(t-s)}Z_s\,ds. \tag{7.10}$$

Consequently

$$Z = a_0 + K_0 * Z, \qquad a_0(t) = \int e^{-gt}\,du_0, \qquad K_0(t) = k\int e^{-gt}\,dq.$$

Here $a_0(t) \to A_0$ and $\|K_0\|_{L^1(0,\infty)} = \rho < 1$. The resolvent $R_0 = \sum_{n\geq 1} K_0^{*n}$ belongs to $L^1(0,\infty)$ and has norm $\rho/(1-\rho)$. Thus

$$Z = a_0 + R_0 * a_0, \qquad \sup_{t\geq 0} Z_t \leq \frac{u_0(\mathcal{X})}{1-\rho}, \qquad Z_t \longrightarrow \frac{A_0}{1-\rho}.$$

The final limit follows by dominated convergence in the convolution. The first term of (7.10) converges in total variation to $u_0|_K$. For $g(x) > 0$, the integral in its second term converges to $A_0/((1-\rho)g(x))$ and is bounded by $u_0(\mathcal{X})/((1-\rho)g(x))$. Since $g^{-1} \in L^1(q)$, dominated convergence gives (7.9). If $A_0 > 0$, normalization proves (7.8). If $A_0 = 0$, the linear limit is zero. Moreover, $q(K) = 0$ and (7.10) give $v_t(K) = u_0(K) = 0$. Every ν_ζ has mass $1-\rho$ on K, so

$$\|\nu_t - \nu_\zeta\|_{\mathrm{TV}} \geq 2(1-\rho) \quad (t \geq 0).$$

□

The condition $u_0(K) > 0$ permits several maximizers and determines the relative condensate weights exactly. It also identifies why normalization requires additional analysis when $u_0(K) = 0$: both the numerator and the normalizing mass in the shifted linear evolution tend to zero. In the single-maximizer density setting, [7] establishes general condensed behavior in Cesàro mean and explicitly distinguishes this from convergence without time averaging. We do not extend that theorem to arbitrary maximizing sets here.

7.2 Adjoint weights and convergence rates at the spectral edge

Proposition 7.4 (Failure of a positive bounded adjoint weight)**.** *Suppose $\rho \leq 1$. Every bounded nonnegative Borel function satisfying $\mathcal{L}^*\psi = M\psi$ vanishes on $\mathcal{X}\setminus K$. Consequently there is no strictly positive bounded adjoint eigenfunction. If ψ is required to be continuous, the only nonnegative solution is $\psi = 0$.*

Proof. The eigenfunction equation is $g(x)\psi(x) = k\int \psi\,dq$. At any point of the nonempty set K, it implies $q(\psi) = 0$. Hence $g\psi = 0$, proving the support assertion. For continuous $\psi \geq 0$, full support of q and $q(\psi) = 0$ imply $\psi = 0$ everywhere. Finally, pairing any strictly positive bounded adjoint eigenfunction with a stationary probability from Theorem 7.1 forces its eigenvalue to equal M, so the preceding argument excludes it. □

Proposition 7.5 (An algebraic lower bound at criticality)**.** *Suppose $\rho = 1$, choose $x_* \in K$, and take $\nu_0 = \delta_{x_*}$. Let $G = \max g > 0$. For the unique critical equilibrium $\nu_* = kg^{-1}q$,*

$$\|\nu_t - \nu_*\|_{\mathrm{TV}} \geq \frac{2}{1+Gt} \qquad (t \geq 0). \tag{7.11}$$

Thus no theorem asserting exponential convergence for every initial measure can hold at this threshold.

Proof. For $v_t = e^{-Mt}e^{t\mathcal{L}}\delta_{x_*}$, formula (7.10) gives $v_t(K) = 1$. The quantity

$$W_t = \int_{\mathcal{X}\setminus K} \frac{1}{g}\, dv_t$$

is finite on bounded time intervals because $\int g^{-1}\, dq = 1/k$. To justify its differentiation without using an unbounded test function, define

$$K_0(t) = k\int e^{-gt}\, dq, \qquad \widetilde{K}_0(t) = k\int g^{-1}e^{-gt}\, dq.$$

The integral formula gives $Z = 1 + K_0 * Z$ and $W = \widetilde{K}_0 * Z$. The function $\widetilde{K}_0$ is continuously differentiable, with $\widetilde{K}_0(0) = 1$ and $\widetilde{K}_0' = -K_0$. Consequently

$$W_t' = Z_t - (K_0 * Z)(t) = 1.$$

Therefore $W_t = t$ and $Z_t = 1 + v_t(\mathcal{X}\setminus K) \le 1 + GW_t = 1 + Gt$. Since $\nu_*(K) = 0$ and $\nu_t(K) = 1/Z_t$, the total variation distance is at least $2/Z_t$, which proves the claim. □

Example 7.6 (An explicit compact-trait threshold)**.** Let $\mathcal{X} = [0,1]$, $q(dx) = dx$, $b \equiv 2$, $d(x) = x^\beta$ with $0 < \beta < 1$, $C \equiv c > 0$, and $\alpha = 0$. Then $r = 2 - x^\beta > 0$, $k = 2p$, and

$$\rho = \frac{2p}{1-\beta}, \qquad p_c = \frac{1-\beta}{2}.$$

For $p > p_c$ there is a unique nonzero equilibrium, and every nonzero initial measure converges to it exponentially in total variation. For $p = p_c$ the stationary frequency is $(1-\beta)x^{-\beta}\, dx$; the initial frequency δ_0 yields the lower bound $2/(1+t)$ in (7.11). For $0 < p < p_c$, the unique stationary frequency is

$$\nu_* = 2px^{-\beta}\, dx + \left(1 - \frac{2p}{1-\beta}\right)\delta_0, \qquad u_* = \frac{2}{c}\nu_*.$$

It attracts in total variation every initial measure having positive mass at 0. For initial measures with zero mass at 0, total variation convergence to this equilibrium is impossible. This example separates equilibrium uniqueness, spectral separation, and convergence topology within one family satisfying all the original coefficient assumptions.

Example 7.7 (A continuum of equilibria under complete reset mutation)**.** Let $\mathcal{X} = [0,1]$, $q(dx) = dx$, $b \equiv 2$, $C \equiv c > 0$, $\alpha = 0$, and $r(x) = 2 - \sqrt{x(1-x)}$. Then $K = \{0,1\}$ and

$$\rho = k\int_0^1 \frac{dx}{\sqrt{x(1-x)}} = \pi k, \qquad k = 2p.$$

For $0 < p < 1/(2\pi)$, all stationary frequencies are

$$\nu_\xi(dx) = \frac{k}{\sqrt{x(1-x)}}\, dx + (1-\pi k)\big[\xi\delta_0 + (1-\xi)\delta_1\big], \qquad 0 \le \xi \le 1,$$

and the corresponding equilibrium mass is $2/c$. If $u_0(\{0,1\}) > 0$, Theorem 7.3 selects $\xi = u_0(\{0\})/u_0(\{0,1\})$. Thus full support of a parent-independent mutation law does not imply uniqueness of the equilibrium or independence from the initial measure.

8 Convergence through a common reference distribution

The competition form acting on signed perturbations is

$$\mathcal{C}(\sigma) = \iint C(x,y)\,\sigma(dx)\sigma(dy).$$

It depends only on the symmetric part of C. Pointwise positivity of C does not imply nonnegativity of this form: the matrix $\left(\begin{smallmatrix}1 & 2\\ 2 & 1\end{smallmatrix}\right)$ has value -2 on $(1,-1)$. Positive competition forms and relative entropy are used in pure-selection convergence theory [13].

Here the reference distribution is stationary for conservative mutation and makes both the competition potential and the combined selection–transfer potential constant across traits. These balances cancel the linear terms in the entropy derivative for every total mass. The remaining competition form is nonnegative, and a mutation reset supplies strict entropy decay.

Theorem 8.1 (A common-reference entropy criterion)**.** *Assume $p > 0$. Suppose that there are $\pi \in \mathcal{P}(\mathcal{X})$ and constants $a_*, c_* > 0$ such that*

$$Q\pi = 0, \qquad C\pi(x) = c_*, \qquad a(x) + A\pi(x) = a_* \quad (x \in \mathcal{X}). \tag{8.1}$$

Assume also that

$$\mathcal{C}(\sigma) \geq 0 \quad \text{if } \sigma(\mathcal{X}) = 0, \qquad m(x,\cdot) \geq \eta\pi(\cdot) \quad (x \in \mathcal{X}) \tag{8.2}$$

for every finite signed measure σ and some $\eta > 0$. Then

$$u_* = \frac{a_*}{c_*}\pi$$

is the unique nonzero stationary measure and attracts every nonzero initial measure in total variation.

More precisely, put $\theta = p\eta b_-$. For every $t \geq s > 0$,

$$\begin{aligned} \mathcal{H}(\pi \mid \nu_t) &\leq e^{-\theta(t-s)}\mathcal{H}(\pi \mid \nu_s), \\ \|\nu_t - \pi\|_{\mathrm{TV}} &\leq \sqrt{2\mathcal{H}(\pi \mid \nu_s)}\, e^{-\theta(t-s)/2}. \end{aligned} \tag{8.3}$$

The entropy on the right is finite. Population convergence is exponential with any rate $0 < \omega < \min\{\theta/2, a_-\}$ after a positive initial time. Neither a bound on the magnitude of α nor a Feller assumption is required.

Proof. Antisymmetry gives $\pi(A\pi) = 0$, so $\pi(a) = a_*$. Substitution in the mass and frequency equations verifies stationarity.

Let $w = \nu - \pi$. The contribution of the multiplicative frequency field to the entropy derivative is

$$\begin{aligned} \frac{d}{dt}\mathcal{H}(\pi \mid \nu)\Big|_{\mathrm{mult}} &= \langle a - NC\nu + A\nu, w\rangle \\ &= \langle a + A\pi - NC\pi, w\rangle - N\mathcal{C}(w) = -N\mathcal{C}(w) \leq 0. \end{aligned} \tag{8.4}$$

The first equality uses the conservative rewriting: the multiplicative frequency field is the potential $a - NC\nu + A\nu$ minus its ν-average. This potential includes the growth correction pb from mutation. In the second equality $\langle Aw, w\rangle = 0$. Conditions (8.1) remove the remaining linear term for every value of N.

By Lemma 10.1, the mutation semigroup contracts $\mathcal{H}(\pi \mid \cdot)$ at rate θ. Its infinitesimal entropy dissipation is therefore at least $\theta\mathcal{H}(\pi \mid \nu)$. Together with (8.4), this yields

$$\frac{d}{dt}\mathcal{H}(\pi \mid \nu_t) \le -\theta\mathcal{H}(\pi \mid \nu_t)$$

and the first inequality in (8.3).

We justify the calculation for arbitrary initial measures. The positive gain in the frequency equation is $pB\nu_t \ge \theta\pi$. The remaining multiplication coefficient is bounded below along each bounded mass trajectory. Variation of constants gives, for each $s > 0$, a constant $c_s > 0$ such that $\nu_t \ge c_s\pi$ for all $t \ge s$. Thus $\mathcal{H}(\pi \mid \nu_t) \le -\log c_s < \infty$. Lebesgue projection relative to π and the integrated logarithmic chain rule justify (8.4) and the mutation derivative. The component singular with respect to π remains part of the probability measure ν_t throughout this calculation. Pinsker's inequality gives the second estimate.

Finally, $I = N^{-1}$ satisfies

$$I' = -\nu(a)I + \bar{C}(\nu).$$

Its coefficients converge exponentially to a_* and c_*, while $\nu(a) \ge a_- > 0$. Variation of constants proves the population estimate, including the strict rate inequality at resonance. Any stationary nonzero measure must have the same limit, which proves uniqueness. □

Example 8.2 (Arbitrarily strong circulation)**.** Take constant $b, r > 0$ with $r \le b$, $p > 0$, $m(x, \cdot) = \pi$, and $f \in C(\mathcal{X})$. Set

$$C(x, y) = c_0 + \varepsilon f(x)f(y), \qquad \pi(f) = 0, \qquad \varepsilon \ge 0, \qquad c_0 > \varepsilon\|f\|_\infty^2.$$

Let v, w be continuous functions of π-mean zero, and set

$$\alpha(x, y) = L[v(x)w(y) - w(x)v(y)].$$

Then $Q\pi = 0$, $C\pi = c_0$, $A\pi = 0$, and $\mathcal{C}(\sigma) = \varepsilon(\sigma(f))^2$ for zero-mass σ. Theorem 8.1 applies for every $L \in \mathbb{R}$. The exponential frequency rate is independent of L. The nonnegative rate $\tau = \max\{\alpha, 0\}$ realizes this flux.

The hypotheses require a mutation-invariant probability measure that also satisfies the two ecological balance identities in (8.1). Irreducibility of mutation alone does not ensure these identities. The next criterion starts from a stationary population and an adjoint eigenfunction of its frozen operator.

9 Adjoint-weighted entropy and convergence

The general relative entropy method uses a positive adjoint eigenelement to weight the entropy relative to a stationary solution; see [15]. Here a stationary population determines a frozen linear operator. When its adjoint has a positive eigenfunction bounded above and away from zero, the method applies to the frozen dynamics. The nonlinear remainder contributes a quadratic form that includes the effect of normalizing horizontal transfer.

9.1 The frozen operator and its conservative transform

Write $b_- = \min_{\mathcal{X}} b > 0$. For a nonzero positive measure u, write

$$N = u(\mathcal{X}), \qquad \nu = u/N, \qquad h_u = r - Cu + A\nu.$$

Let u_* be a nonzero stationary measure, and put

$$N_* = u_*(\mathcal{X}), \qquad \nu_* = u_*/N_*, \qquad h_* = r - Cu_* + A\nu_*.$$

The frozen operator is

$$\mathcal{L}_*\mu = h_*\mu + pB\mu, \qquad \mathcal{L}_*u_* = 0. \tag{9.1}$$

We assume that there is a bounded Borel function ψ such that

$$0 < \psi_- \le \psi(x) \le \psi_+ < \infty, \qquad h_*(x)\psi(x) + pb(x)\int_{\mathcal{X}} \psi(y)m(x,dy) = 0 \quad (x \in \mathcal{X}). \tag{9.2}$$

This is a pointwise adjoint equation on the entire trait space. An adjoint equation only on the support of u_* is insufficient for initial measures charging traits outside that support.

Set

$$Z = \int \psi\, du_*, \qquad \varphi = \frac{\psi u_*}{Z}.$$

Multiplication by ψ conjugates $\mathcal{L}_*$ to the conservative generator

$$G\lambda = \psi\mathcal{L}_*(\psi^{-1}\lambda). \tag{9.3}$$

Its jump-rate kernel is

$$k(y,dx) = p\, b(y)\frac{\psi(x)}{\psi(y)}m(y,dx). \tag{9.4}$$

Indeed, its total outgoing rate at y equals $-h_*(y)$ by (9.2). In particular, G preserves mass and $G\varphi = 0$.

For $p > 0$, condition (9.2) implies

$$h_*(x) \le -pb_-\frac{\psi_-}{\psi_+} < 0 \quad (x \in \mathcal{X}). \tag{9.5}$$

Thus an adjoint function bounded above and away from zero is a substantive spectral condition. It can fail at a spectral edge supporting a condensed equilibrium.

Corollary 9.1 (A sufficient condition for the adjoint function)**.** *Given a stationary u_*, define*

$$V_* = h_* + pb = a - Cu_* + A\nu_*, \qquad \mathcal{L}_* = pQ + V_*.$$

Suppose $m(x,\cdot) \ge \eta q(\cdot)$ for a probability measure q and some $\eta > 0$, and

$$\theta = p\eta b_- > \operatorname{osc}(V_*).$$

Then (9.2) has a bounded Borel solution bounded away from zero. If m is Feller, the solution constructed in this way is continuous.

Proof. The resolvent construction in Theorem 6.3 uses only boundedness of the potential and the mutation minorization. Applied to the frozen operator $pQ + V_*$, it gives a positive bounded adjoint eigenfunction ψ and a real eigenvalue λ. Pairing its pointwise eigenfunction equation with u_* yields

$$\lambda\int \psi\, du_* = \int (\mathcal{L}_*^*\psi)\, du_* = \int \psi\, d(\mathcal{L}_*u_*) = 0.$$

The integral is strictly positive, so $\lambda = 0$. The construction takes place in $B_b(\mathcal{X})$; under the Feller assumption it preserves $C(\mathcal{X})$. □

9.2 The exact nonlinear entropy identity

An entropy for the population measure must account for changes in both composition and total mass. We therefore use relative entropy for finite measures rather than normalizing the population at the outset. For probability measures, $\mathcal{H}(\rho \mid \mu)$ denotes the usual relative entropy $\int \log(d\rho/d\mu)\, d\rho$, with value $+\infty$ when appropriate. For finite positive measures λ_*, λ, define

$$\mathcal{H}_{\mathrm{fin}}(\lambda_* \mid \lambda) = \lambda(\mathcal{X}) - \lambda_*(\mathcal{X}) + \int \log \frac{d\lambda_*}{d\lambda}\, d\lambda_*,$$

with value $+\infty$ when the integral is not well defined as a finite relative entropy. The weighted Volterra entropy is

$$\mathcal{V}_\psi(u) = \mathcal{H}_{\mathrm{fin}}(\psi u_* \mid \psi u). \tag{9.6}$$

It is nonnegative and vanishes exactly at $u = u_*$. If

$$u = z u_* + u^\perp, \qquad u^\perp \perp u_*,$$

then

$$\mathcal{V}_\psi(u) = \int \psi(z - 1 - \log z)\, du_* + \int \psi\, du^\perp. \tag{9.7}$$

The component singular with respect to u_* contributes its full ψ-weighted mass. Thus the entropy does not require $u \ll u_*$.

Define the symmetric kernel

$$\begin{aligned} S_N(x,y) = {} & \frac{\psi(x)C(x,y) + \psi(y)C(y,x)}{2} - \frac{\psi(x) - \psi(y)}{2N}\alpha(x,y) \\ & + \frac{\psi(x)(A\nu_*)(x) + \psi(y)(A\nu_*)(y)}{2N}, \end{aligned} \tag{9.8}$$

and its quadratic form

$$\mathcal{S}_N(v) = \iint S_N(x,y)\, v(dx)v(dy).$$

The last term of (9.8) accounts for the change of total mass in the denominator of the transfer term. Antisymmetry alone does not cancel the weighted transfer terms or their normalization correction.

Proposition 9.2 (Balanced flux and weighted entropy)**.** *Assume* (9.1) *and* (9.2)*. The measure*

$$J(dx,dy) = p\psi(x)b(y)m(y,dx)u_*(dy) \tag{9.9}$$

has equal first and second marginals, both equal to $-\psi h_* u_*$*. Let* $u_t = z_t u_* + u_t^\perp$ *be a solution on a time interval on which* z_t *is bounded below by a positive constant, uniformly on compact time subintervals. Then, almost everywhere in time,*

$$\frac{d}{dt}\mathcal{V}_\psi(u_t) = -\mathcal{D}_J(z_t) - \mathcal{D}_\perp(u_t) - \mathcal{S}_{N_t}(u_t - u_*), \tag{9.10}$$

where

$$\mathcal{D}_J(z) = \iint \left[\frac{z(y)}{z(x)} - 1 - \log\frac{z(y)}{z(x)}\right] J(dx,dy), \tag{9.11}$$

$$\mathcal{D}_\perp(u) = p\int \frac{\psi(x)}{z(x)}\, d(Bu^\perp)_{\mathrm{ac}}(x). \tag{9.12}$$

Here the absolutely continuous part in (9.12) *is taken relative to* u_**. Both dissipations are nonnegative.*

Proof. The first marginal in (9.9) is $p\psi Bu_* = -\psi h_* u_*$. The second marginal has density $pb(y)\int \psi(x)m(y,dx) = -\psi(y)h_*(y)$ relative to u_*. This proves the balance assertion without a detailed-balance assumption.

The stationarity equation shows that pB maps measures absolutely continuous with respect to u_* to measures with the same property. For the frozen equation, the linear part of the entropy derivative vanishes by $\mathcal{L}_*^*\psi = 0$. The remaining part is

$$-\int \frac{\psi}{z}\, d(\mathcal{L}_* u)_{\mathrm{ac}} = -\int \psi h_*\, du_* - \iint \frac{z(y)}{z(x)} J(dx,dy) - \mathcal{D}_\perp(u).$$

The equal marginals give

$$\iint [\log z(y) - \log z(x)]\, J(dx,dy) = 0,$$

and hence the frozen contribution is $-\mathcal{D}_J(z) - \mathcal{D}_\perp(u)$.

Write $v = u - u_*$ and $s = N - N_*$. Since

$$\nu - \nu_* = \frac{v}{N} - \frac{s}{N}\nu_*,$$

the nonlinear difference is the multiplication operator

$$h_u - h_* = -Cv + \frac{1}{N}Av - \frac{s}{N}A\nu_*.$$

Its contribution to the entropy derivative is

$$\int \psi(h_u - h_*)\, dv = -\mathcal{S}_N(v),$$

which proves (9.10).

For completeness, Lebesgue projection relative to u_* is a bounded linear projection in total variation. The absolutely continuous density of a C^1 measure-valued solution is therefore an absolutely continuous curve in $L^1(u_*)$. On intervals where z_t has a positive lower bound, the logarithm is Lipschitz on its range and the integrated chain rule applies. No upper bound on z_t, and no absolute continuity of u_t with respect to u_*, is required. □

9.3 A continuum convergence theorem

Let $I = [\ell, U] \subset (0,\infty)$ be a positively invariant mass interval which contains N_* and eventually contains the mass of every nonzero solution. The interval supplied by the logistic comparison is one such choice. Assume

$$\mathcal{S}_N(v) \geq 0 \quad \text{for every finite signed measure } v \text{ and every } N \in I. \tag{9.13}$$

Since $S_N = S_0 + N^{-1}S_1$, condition (9.13) is equivalent to its two endpoint versions at $N = \ell, U$.

Proposition 9.3 (Gram-matrix and finite-feature tests)**.** *Suppose S_N is continuous on $\mathcal{X}^2$, as is the case when ψ is continuous. For fixed N, its quadratic form is nonnegative on all finite signed measures if and only if*

$$[S_N(x_i,x_j)]_{i,j=1}^k \succeq 0$$

for every $k \geq 1$ and every finite collection $x_1,\dots,x_k \in \mathcal{X}$. For the family (9.8), it suffices to make these tests at $N = \ell, U$.

Proof. Testing atomic measures proves necessity. Conversely, approximate the positive and negative parts of a finite signed measure weakly by finite atomic measures, with uniformly bounded total variation. Continuity of S_N on the compact product space passes the quadratic forms to the limit. The endpoint assertion follows because S_N is an affine function of $1/N$. □

A useful sufficient certificate is a finite-feature representation

$$S_N(x,y) = f(x)^\top H_N f(y), \qquad f : \mathcal{X} \longrightarrow \mathbb{R}^k,$$

with f continuous and H_N symmetric. Positive semidefiniteness of H_ℓ, H_U, when H_N is affine in $1/N$, proves (9.13). More precisely, positive semidefiniteness on span $f(\mathcal{X})$ is needed, and it is also necessary for this representation.

Theorem 9.4 (Adjoint-weighted convergence on compact trait spaces)**.** *Assume $p > 0$, (9.1), (9.2), (9.13), and*

$$m(y,\cdot) \geq \eta\, \nu_*(\cdot) \quad (y \in \mathcal{X}) \tag{9.14}$$

for some $\eta > 0$. Then u_* is the unique nonzero stationary measure. Every nonzero solution converges to u_* exponentially in total variation after a finite entrance time. The exponential rate can be chosen to depend only on the coefficients, u_*, ψ, η and I. A Feller assumption on m is not needed.

Proof. The frozen Markov dynamics dissipates changes in weighted composition while preserving weighted mass. We first obtain this composition estimate, then use the weighted quadratic form to control the remaining mass deviation. For a state u with mass in I, put

$$\lambda = \psi u, \qquad \lambda_* = \psi u_*, \qquad M = \lambda(\mathcal{X}), \qquad q = \lambda/M, \qquad \zeta = M/Z.$$

Then

$$\mathcal{V}_\psi(u) = Z\big[\zeta - 1 - \log\zeta + \mathcal{H}(\varphi \mid q)\big]. \tag{9.15}$$

Here $M = \int \psi\, du$ is weighted mass, $q = \psi u/M$ is weighted composition, and ζ is the ratio of weighted masses. Thus the two terms in (9.15) measure deviations in mass and composition, respectively. Write $H = \mathcal{H}(\varphi \mid q)$. The rate kernel (9.4) satisfies

$$k(y,\cdot) \geq \kappa_0 \varphi(\cdot), \qquad \kappa_0 = \frac{p\eta b_- Z}{N_* \psi_+} > 0.$$

Consequently

$$G = \kappa_0(\varphi\langle 1, \cdot\rangle - I) + R,$$

where R is another conservative Markov generator and $R\varphi = 0$. The data-processing inequality for e^{tR}, together with convexity in the second argument of relative entropy, gives

$$\mathcal{H}(\varphi \mid e^{tG} q) \leq e^{-\kappa_0 t}\mathcal{H}(\varphi \mid q).$$

The frozen entropy dissipation at u is therefore at least $\kappa_0 Z H$. In the notation of Proposition 9.2,

$$\mathcal{D}_J(z) + \mathcal{D}_\perp(u) \geq \kappa_0 Z H. \tag{9.16}$$

Set

$$c_0 = \int \psi(x)(Cu_*)(x)\, u_*(dx) > 0, \qquad P = \sup_{N \in I} \|S_N\|_\infty.$$

The two transfer terms in (9.8) cancel on the ray spanned by u_*, giving the identity

$$\mathcal{S}_N(u_*) = c_0 \quad \text{for every } N > 0. \tag{9.17}$$

Let $e = u - \zeta u_*$. By (9.13), the square root of $\mathcal{S}_N$ is a seminorm, so

$$c_0(\zeta - 1)^2 \le 2\mathcal{S}_N(u - u_*) + 2\mathcal{S}_N(e).$$

Moreover,

$$e = M\psi^{-1}(q - \varphi), \qquad \|e\|_{\mathrm{TV}} \le \frac{M}{\psi_-}\|q - \varphi\|_{\mathrm{TV}}.$$

Pinsker's inequality therefore gives

$$c_0(\zeta - 1)^2 \le 2\mathcal{S}_N(u - u_*) + 4P\left(\frac{M_+}{\psi_-}\right)^2 H, \qquad M_+ = \psi_+ U. \tag{9.18}$$

Put $M_- = \psi_- \ell$ and

$$K_R = \frac{1}{2\min\{1, M_-/Z\}^2}.$$

On the allowed mass interval, $\zeta - 1 - \log\zeta \le K_R(\zeta - 1)^2$. Combining (9.15) and (9.18), we obtain

$$\mathcal{V}_\psi(u) \le Z\left[1 + \frac{4K_R P M_+^2}{c_0\psi_-^2}\right] H + \frac{2ZK_R}{c_0}\mathcal{S}_N(u - u_*).$$

Thus (9.10) and (9.16) imply

$$\frac{d}{dt}\mathcal{V}_\psi(u_t) \le -\kappa\mathcal{V}_\psi(u_t), \qquad \kappa = \min\left\{\frac{\kappa_0}{1 + 4K_R P M_+^2/(c_0\psi_-^2)}, \frac{c_0}{2ZK_R}\right\} > 0. \tag{9.19}$$

We justify the use of the entropy for arbitrary initial measures. Once the mass lies in I, $h_{u_t} \ge -K$ for a constant K depending only on the data and I, and

$$pBu_t \ge p\eta b_- \ell\, \nu_*.$$

Variation of constants implies, for every fixed $s > 0$,

$$u_{t+s} \ge \frac{p\eta b_- \ell}{N_*}\frac{1 - e^{-Ks}}{K}\, u_*,$$

with the usual interpretation when $K = 0$. Thus the density relative to u_* has a positive lower bound at positive elapsed times, although a singular component can remain. Proposition 9.2 applies and $\mathcal{V}_\psi(u_{t+s}) < \infty$.

Finally, the elementary finite-measure entropy inequality

$$\|\lambda - \lambda_*\|_{\mathrm{TV}}^2 \le 2\big[\lambda(\mathcal{X}) + \lambda_*(\mathcal{X})\big]\mathcal{H}_{\mathrm{fin}}(\lambda_* \mid \lambda)$$

follows from $(\sqrt{z} - 1)^2 \le z - 1 - \log z$ and Cauchy–Schwarz. It gives

$$\|u_t - u_*\|_{\mathrm{TV}} \le \frac{\sqrt{2(M_+ + Z)}}{\psi_-}\sqrt{\mathcal{V}_\psi(u_t)}. \tag{9.20}$$

Equations (9.19) and (9.20) prove the claimed exponential convergence. Applying this conclusion to a stationary solution proves uniqueness. □

Remark 9.5 (Weak convergence under strict positivity of the weighted form). There is also a useful version without (9.14). Assume additionally that ψ is continuous, and suppose $\mathcal{V}_\psi(u_0) < \infty$, the initial mass lies in I, and $\mathcal{S}_N(v) > 0$ for every nonzero finite signed v and every $N \in I$. Then $u_t \rightharpoonup u_*$. The strict positivity here separates distinct measures; no

uniform coercive lower bound in the total variation norm is assumed. Indeed, approximation of u_0 by $(1-\varepsilon)u_0 + \varepsilon[N_0/N_*]u_*$, which preserves its mass $N_0 = u_0(\mathcal{X})$, yields

$$\mathcal{V}_\psi(u_t) + \int_0^t \mathcal{S}_{N_s}(u_s - u_*)\, ds \le \mathcal{V}_\psi(u_0).$$

Convexity and lower semicontinuity give $\mathcal{V}_\psi(u_0^\varepsilon) \to \mathcal{V}_\psi(u_0)$. The density in each approximating solution remains bounded below on finite time intervals by the multiplicative part of the equation; the entropy inequality then passes to the limit by lower semicontinuity. The integrand is uniformly continuous in time, because the orbit is uniformly Lipschitz in total variation on the mass slab and S_N is continuous and uniformly bounded there. It tends to zero. Weak compactness and continuity of $(N, u) \mapsto \mathcal{S}_N(u - u_*)$ identify every weak cluster point as u_*. This argument gives weak convergence and does not imply total variation convergence.

9.4 Finite-dimensional tests

On a finite trait space, write

$$B_{ij} = b_j m(j, \{i\}), \quad A^\top = -A, \quad \Psi = \operatorname{diag}(\psi_1, \dots, \psi_n).$$

We write $\operatorname{Sym}(M) = (M + M^\top)/2$. For a positive equilibrium u_*, let

$$L_* = \operatorname{diag}(r - Cu_* + A\nu_*) + pB.$$

If $p > 0$ and B is irreducible, the equation $L_* u_* = 0$ implies that zero is the Perron eigenvalue of L_*, and there is a positive left eigenvector ψ. The kernel (9.8) becomes the symmetric matrix

$$S_N = \operatorname{Sym}(\Psi C) - \frac{1}{N}\operatorname{Sym}(\Psi A) + \frac{1}{2N}\left[\mathbf{1}(\Psi A\nu_*)^\top + (\Psi A\nu_*)\mathbf{1}^\top\right]. \tag{9.21}$$

Corollary 9.6 (Two semidefinite tests)**.** *Assume $p > 0$, B is irreducible, and $u_* > 0$ is stationary. If*

$$S_\ell \succeq 0, \qquad S_U \succeq 0,$$

then u_* is the unique nonzero stationary state and every nonzero trajectory converges to it. The convergence is exponential after an initial transient. No one-step minorization is required.

Proof. The affine dependence on $1/N$ gives $S_N \succeq 0$ for all $N \in I$. Every nonzero trajectory enters the strictly positive orthant at every positive time, by irreducibility and comparison with a positive multiple of $e^{tpB}u_0$. For positive u, let $z_i = u_i/u_{*,i}$ and $F_{ij} = p\psi_i B_{ij} u_{*,j}$. The entropy identity reads

$$\dot{\mathcal{V}}_\psi = -\sum_{i,j} F_{ij}\left[\frac{z_j}{z_i} - 1 - \log\frac{z_j}{z_i}\right] - (u - u_*)^\top S_N (u - u_*).$$

The first term vanishes only if $z_i = z_j$ on every mutation edge. Irreducibility therefore forces $u = cu_*$. On this ray the quadratic term is $(c-1)^2 c_0$, which vanishes only at $c = 1$. Entropy sublevel sets inside the mass slab are compact subsets of the positive orthant. LaSalle's principle proves convergence.

For the rate, the quadratic part of mutation dissipation at u_* is

$$\frac{1}{2}\sum_{i,j} F_{ij}\left(\frac{v_j}{u_{*,j}} - \frac{v_i}{u_{*,i}}\right)^2.$$

Its kernel is the one-dimensional space $\operatorname{span}\{u_*\}$. Adding $v^\top S_{N_*} v$ makes the quadratic form positive definite, by (9.17). The entropy itself has a positive definite Hessian at u_*. A local entropy–dissipation estimate gives exponential convergence once the trajectory enters a sufficiently small neighborhood. □

Mutation can also compensate for negative directions of S_N. The following criterion quantifies this possibility.

Corollary 9.7 (Mutation-augmented matrix criterion)**.** *Assume a finite trait space, $p > 0$, a strictly positive stationary state u_*, and a strictly positive adjoint vector ψ satisfying (9.2). Let $I = [\ell, U]$ be a positively invariant absorbing mass interval containing N_*. Assume (9.14), and define*

$$Z = \psi^\top u_*, \qquad \varphi = \frac{\Psi u_*}{Z}, \qquad R = \Psi - \varphi\psi^\top,$$
$$\kappa_0 = \frac{p\eta b_- Z}{N_* \psi_+}, \qquad \beta = \frac{\kappa_0 Z}{2(\psi_+ U)^2}.$$

No positive semidefiniteness assumption on the unaugmented matrices S_N is imposed. If both matrices

$$S_\ell + \beta R^\top R, \qquad S_U + \beta R^\top R \tag{9.22}$$

are positive definite, then u_* is the unique nonzero stationary state and every nonzero trajectory converges to it exponentially after an initial transient.

Proof. For $v = u - u_*$, $M = \psi^\top u$, and $q = \Psi u / M$,

$$Rv = \Psi u - M\varphi = M(q - \varphi).$$

Since $Ru_* = 0$ and $\ker R = \operatorname{span}\{u_*\}$, this term measures departure from the equilibrium composition and vanishes along the equilibrium ray. Pinsker's inequality gives

$$\mathcal{H}(\varphi \mid q) \geq \frac{\|Rv\|_2^2}{2(\psi_+ U)^2}.$$

The frozen entropy dissipation is at least $\kappa_0 Z \mathcal{H}(\varphi \mid q)$, hence

$$\dot{\mathcal{V}}_\psi \leq -v^\top [S_N + \beta R^\top R] v \leq -\lambda_0 \|v\|_2^2$$

for a positive λ_0, by the two endpoint assumptions. Minorization and the mass lower bound place each trajectory in a fixed compact subset of the positive orthant after one additional unit of time in the mass slab. On that compact set, Taylor's formula gives $\mathcal{V}_\psi(u) \leq K\|u - u_*\|_2^2$. The resulting differential inequality proves exponential convergence and uniqueness. □

9.5 Examples and boundaries of the criteria

Example 9.8 (Nonconstant adjoint weight and nonzero transfer)**.** Take two traits and

$$b = (1, 1), \quad r = \left(\frac{3}{10}, \frac{7}{20}\right), \quad p = \frac{1}{5}, \quad m(j, \{i\}) = \frac{1}{2},$$
$$C = \begin{pmatrix} \frac{1}{2} & \frac{1}{10} \\ \frac{1}{20} & \frac{1}{5} \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & \frac{3}{20} \\ -\frac{3}{20} & 0 \end{pmatrix}.$$

Set $d = b - r$ and realize A by the nonnegative transfer rates $\tau_{ij} = \max\{A_{ij}, 0\}$. Then

$$u_* = (1, 2)^\top, \quad \psi = (1, 2)^\top, \quad h_* = \left(-\frac{3}{10}, -\frac{3}{20}\right)^\top$$

satisfy both stationary equations. The logistic mass interval can be chosen as

$$[\ell, U] = \left[\frac{1}{2}, 22\right].$$

The weighted matrices are

$$S_N = \begin{pmatrix} \frac{1}{2} + \frac{1}{10N} & \frac{1}{10} + \frac{3}{40N} \\ \frac{1}{10} + \frac{3}{40N} & \frac{2}{5} - \frac{1}{10N} \end{pmatrix}.$$

At the lower endpoint,

$$S_{1/2} = \begin{pmatrix} \frac{7}{10} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{5} \end{pmatrix}, \qquad \det S_{1/2} = \frac{31}{400} > 0.$$

At the upper endpoint,

$$S_{22} = \begin{pmatrix} \frac{111}{220} & \frac{91}{880} \\ \frac{91}{880} & \frac{87}{220} \end{pmatrix}, \qquad \det S_{22} = \frac{146231}{774400} > 0.$$

Corollary 9.6 proves global convergence. The mutation generator has the unique invariant probability $(1/2, 1/2)$, for which $C\pi = (3/10, 1/8)$ is not constant. Thus the common-reference balance hypotheses do not hold. The coefficient contraction criterion also fails, even after recipient-centering:

$$\delta_{\mathrm{rec}} = \frac{1}{2} \max_j (\max_i C_{ij} - \min_i C_{ij}) = \frac{9}{40}.$$

Any mass interval containing u_* has $U \geq 3$, so its frequency Lipschitz bound contains $5U\delta_{\mathrm{rec}} \geq 27/8$, whereas every common-reset rate satisfies $\theta = p\eta b_- \leq 1/5$.

Example 9.9 (Mutation controls a negative competition direction)**.** Let $n = 3$, $P = I - \mathbf{1}\mathbf{1}^\top/3$, and take

$$b_i = 1, \quad r_i = \frac{1}{2}, \quad p = 1, \quad m(j, \{i\}) = \frac{1}{3}, \qquad C = \mathbf{1}\mathbf{1}^\top - \frac{1}{20}P.$$

Thus $C_{ii} = 29/30$ and $C_{ij} = 61/60$ for $i \neq j$. For an arbitrary $L \geq 0$, set

$$A = L \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Then

$$u_* = \frac{1}{2}\mathbf{1}, \quad \psi = \mathbf{1}, \quad Z = \frac{3}{2}, \quad \varphi = \frac{1}{3}\mathbf{1}, \quad R = P, \quad \kappa_0 = 1.$$

Since $A\varphi = 0$, the transfer terms in S_N vanish for every N, and $S_N = C$. This matrix has eigenvalue $-1/20$ on the zero-sum subspace, so both the semidefinite condition (9.13) and the competition condition on zero-mass measures in (8.2) fail. On the other hand $U = 90/29$ and

$$\beta = \frac{841}{10800}, \qquad C + \beta P = \mathbf{1}\mathbf{1}^\top + \frac{301}{10800}P$$

is positive definite. Corollary 9.7 proves global convergence for every transfer amplitude L. For $L \geq 1/2$, the coefficient contraction condition fails for every potential subtraction: the three-edge circulation has magnitude $3L$, so $\|\beta_h\|_\infty \geq L$ for every h, whereas the common-reset rate is at most 1.

Example 9.10 (A continuum family with nonconstant adjoint weight)**.** Let π_0 be any probability measure on a compact trait space, let $\phi \in C(\mathcal{X})$ be strictly positive and nonconstant, let $0 < p \le 1$, and put $m(y,\cdot) = \pi_0$, $b \equiv b_0 > 0$, $\alpha = 0$, and

$$C(x,y) = \frac{c}{\phi(x)}, \qquad r(x) = \frac{r_0}{\phi(x)},$$

where $c, r_0 > 0$ and $r_0 \le b_0 \min \phi$, so $d = b_0 - r \ge 0$. Writing $m_\phi = \int \phi \, d\pi_0$, define

$$N_* = \frac{r_0 + p b_0 m_\phi}{c}, \qquad \nu_* = \frac{\phi}{m_\phi}\pi_0, \qquad u_* = N_* \nu_*.$$

The adjoint function $\psi = \phi$ satisfies

$$h_* = -\frac{p b_0 m_\phi}{\phi}, \qquad \mathcal{L}_*^* \phi = 0.$$

The weighted kernel is the constant kernel $S_N(x,y) = c$, whose quadratic form is $c[v(\mathcal{X})]^2$. It is nonnegative and vanishes on every zero-mass perturbation. Moreover

$$\pi_0 \ge \frac{m_\phi}{\max \phi}\nu_*.$$

Theorem 9.4 therefore gives global exponential convergence in total variation for every nonzero initial measure. The competition potential $C\pi$ is nonconstant for every probability π, so the common-reference balance assumption cannot apply when ϕ is nonconstant.

Remark 9.11 (The spectral and entropy criteria are complementary). The positive-semigroup reduction for constant competition can imply global convergence even when the weighted quadratic form is indefinite. For example, take two traits, $A = 0$, $C_{ij} = 1$, equal positive birth rates, uniform irreducible mutation, and unequal growth rates. The frozen adjoint vector ψ is nonconstant, whereas

$$S_N = \begin{pmatrix} \psi_1 & (\psi_1+\psi_2)/2 \\ (\psi_1+\psi_2)/2 & \psi_2 \end{pmatrix}, \qquad \det S_N = -\frac{(\psi_1-\psi_2)^2}{4} < 0.$$

Nevertheless normalized Perron–Frobenius dynamics yields global convergence. Failure of one structural criterion therefore identifies a limitation of that argument, rather than instability of the equilibrium.

10 Quantitative contraction under mutation mixing

This criterion compares two evolving populations without specifying a stationary state. Mutation mixing contracts the difference of their frequencies. Selection and transfer can increase this difference, and competition couples it to the difference of their total masses. Two differential inequalities control these effects together. The resulting test uses only coefficient bounds and a minorization of the mutation kernel.

Throughout this section, the total variation norm is the full norm $\|\sigma\|_{\mathrm{TV}} = |\sigma|(\mathcal{X})$. Assume that $p > 0$ and

$$m(x,\cdot) \ge \eta \pi_0(\cdot) \quad (x \in \mathcal{X}), \qquad \theta = p\eta b_- > 0, \tag{10.1}$$

where π_0 is a probability measure and $b_- = \min b$. No invariance assumption is imposed on π_0.

Lemma 10.1 (The common-reset estimate)**.** *For every signed measure σ of mass zero,*

$$\|e^{tpQ}\sigma\|_{\mathrm{TV}} \le e^{-\theta t}\|\sigma\|_{\mathrm{TV}}.$$

If $Q\pi_0 = 0$, then, for probability measures ν,

$$\mathcal{H}(\pi_0 \mid e^{tpQ}\nu) \le e^{-\theta t}\mathcal{H}(\pi_0 \mid \nu),$$

with the usual extended-valued interpretation.

Proof. Write $\mathsf{P}_{\pi_0}\mu = \mu(\mathcal{X})\pi_0$ and

$$R = pQ - \theta(\mathsf{P}_{\pi_0} - I).$$

The jump measure of R from a point y is $pb(y)m(y, dx) - \theta\pi_0(dx)$, which is nonnegative by (10.1); its mass is $pb(y) - \theta$. Thus R is a bounded conservative Markov generator. Its semigroup is a total variation contraction and preserves the zero-mass subspace. On that subspace $pQ = R - \theta I$, which proves the first assertion.

If $Q\pi_0 = 0$, then $R\pi_0 = 0$ and

$$e^{tpQ}\nu = e^{-\theta t}e^{tR}\nu + (1 - e^{-\theta t})\pi_0.$$

Convexity of relative entropy in its second argument, followed by the data-processing inequality for e^{tR}, proves the second assertion. The data-processing inequality follows directly from conditional Jensen's inequality applied to a joint measure with transition kernel e^{tR}. □

10.1 Removing composition-neutral terms

Choose arbitrary continuous functions h, c on $\mathcal{X}$ and set

$$\beta_h(x, y) = \alpha(x, y) - h(x) + h(y), \qquad E_c(x, y) = C(x, y) - c(y), \qquad \delta = \|E_c\|_\infty. \tag{10.2}$$

The frequency equation is unchanged if its selection and transfer terms are written using $a + h$ and β_h. Its competition term depends only on E_c, because the contribution of $c(y)$ cancels upon normalization.

The subtraction changes only the representation of the frequency equation; the mass equation retains the full competition kernel. The smallest possible value of δ is explicit:

$$\delta_{\mathrm{rec}} = \frac{1}{2}\max_{y\in\mathcal{X}}\left(\max_x C(x, y) - \min_x C(x, y)\right). \tag{10.3}$$

It is attained by the continuous function

$$c_{\mathrm{mid}}(y) = \frac{1}{2}\left(\max_x C(x, y) + \min_x C(x, y)\right).$$

Indeed, for each fixed y, the best uniform approximation to $C(\cdot, y)$ by a constant is its midrange. Compactness and continuity imply continuity of both extremal functions of y. In particular, $\delta_{\mathrm{rec}} = 0$ exactly when $C(x, y) = c(y)$.

Fix the positively invariant absorbing mass interval

$$\ell = \frac{a_-}{2C_+}, \qquad U = \frac{2a_+}{C_-}. \tag{10.4}$$

More generally, the proof below applies to any positively invariant absorbing interval $[\ell, U] \subset (0, \infty)$. Define

$$L = \frac{3}{2}\operatorname{osc}(a+h) + 2\|\beta_h\|_\infty + 5U\delta, \qquad J = 2U\delta,$$
$$K = \frac{1}{2}\operatorname{osc}(a) + U\left(\frac{1}{2}\operatorname{osc}(c) + 2\delta\right), \qquad \gamma = \ell C_-. \tag{10.5}$$

The constants J and K bound the influence of a mass difference on frequency and of a frequency difference on mass, respectively. The condition below requires the net mutation contraction and the mass damping to dominate this coupling.

Theorem 10.2 (A coefficient-based global contraction criterion)**.** *Assume* (10.1)*. If h, c can be chosen so that*

$$\theta > L, \qquad (\theta - L)\gamma > JK, \tag{10.6}$$

then the equation has exactly one nonzero stationary measure u_*. Every nonzero initial measure converges to u_* exponentially in total variation after entering the mass interval (10.4)*.*

More precisely, choose $q > 0$ with

$$q\gamma > J, \qquad qK < \theta - L,$$

and put

$$\lambda = \min\{\theta - L - qK,\ \gamma - J/q\} > 0.$$

For two solutions whose masses lie in $[\ell, U]$,

$$\|\nu_t - \widetilde{\nu}_t\|_{\mathrm{TV}} + q|\log N_t - \log \widetilde{N}_t| \le e^{-\lambda t}\left(\|\nu_0 - \widetilde{\nu}_0\|_{\mathrm{TV}} + q|\log N_0 - \log \widetilde{N}_0|\right). \tag{10.7}$$

The conclusions hold for measurable mutation kernels; the Feller property is not needed.

Proof. Set $D = \|\nu - \widetilde{\nu}\|_{\mathrm{TV}}$ and $E = |\log N - \log \widetilde{N}|$. For any bounded function f and zero-mass signed measure σ,

$$|\sigma(f)| \le \frac{1}{2}\operatorname{osc}(f)\|\sigma\|_{\mathrm{TV}}. \tag{10.8}$$

Consequently the selection map $S_f(\nu) = (f - \nu(f))\nu$ satisfies

$$\|S_f(\nu) - S_f(\widetilde{\nu})\|_{\mathrm{TV}} \le \frac{3}{2}\operatorname{osc}(f)D.$$

For a bounded antisymmetric kernel β, the map $T_\beta(\nu) = (\beta\nu)\nu$ satisfies

$$\|T_\beta(\nu) - T_\beta(\widetilde{\nu})\|_{\mathrm{TV}} \le 2\|\beta\|_\infty D.$$

Write

$$H(\nu) = (E_c\nu - \nu(E_c\nu))\nu.$$

Since $\|E_c\|_\infty = \delta$,

$$\|H(\nu)\|_{\mathrm{TV}} \le 2\delta, \qquad \|H(\nu) - H(\widetilde{\nu})\|_{\mathrm{TV}} \le 5\delta D. \tag{10.9}$$

For the second estimate, expand the difference as

$$(E_c\nu - \nu E_c\nu)(\nu - \widetilde{\nu}) + \big(E_c(\nu - \widetilde{\nu}) - [\nu E_c\nu - \widetilde{\nu}E_c\widetilde{\nu}]\big)\widetilde{\nu}.$$

The three resulting bounds are $2\delta D$, δD, and $2\delta D$. Since $|N - \widetilde{N}| \le UE$, the competition difference is bounded by $5U\delta D + 2U\delta E$. Variation of constants and Lemma 10.1 therefore give the upper Dini inequality

$$D^+ \le -(\theta - L)D + JE. \tag{10.10}$$

For the logarithmic mass,

$$(\log N)' = \nu(a) - N\bar{C}(\nu).$$

The decomposition $C = c(y) + E_c$ gives

$$|\bar{C}(\nu) - \bar{C}(\widetilde{\nu})| \le \left(\frac{1}{2}\operatorname{osc}(c) + 2\delta\right) D.$$

Furthermore

$$\operatorname{sgn}(\log N - \log \widetilde{N})(N - \widetilde{N})\bar{C}(\nu) \ge \ell C_- E.$$

Hence

$$E^+ \le -\gamma E + KD. \tag{10.11}$$

Multiplying (10.11) by q and adding (10.10) proves (10.7).

The space $[\ell, U] \times \mathcal{P}(\mathcal{X})$ is complete for the distance in (10.7). Every fixed positive-time map is a strict contraction and has a unique fixed point. Commutation with all other time maps implies that this fixed point is stationary. Every stationary nonzero measure has mass in $[a_-/C_+, a_+/C_-]$, so there are no stationary measures outside the absorbing interval. Finally,

$$\|N\nu - \widetilde{N}\widetilde{\nu}\|_{\mathrm{TV}} \le U(D + E)$$

transfers the estimate to the original population measures. □

Corollary 10.3 (Recipient-independent competition)**.** *If $C(x, y) = c(y)$, it suffices that, for some continuous h,*

$$p\eta b_- > \frac{3}{2}\operatorname{osc}(a + h) + 2\|\alpha - h(x) + h(y)\|_\infty.$$

In particular, neither the size nor the oscillation of $c(y)$ enters the frequency contraction condition.

Proof. Here $\delta = J = 0$. The frequency estimate is autonomous, and q in Theorem 10.2 can be chosen sufficiently small. □

The constants in Theorem 10.2 are explicit sufficient bounds. For finite trait sets, the search over h and c involves only finitely many oscillation and maximum-norm constraints. The midrange choice (10.3) minimizes the direct contribution of competition to the frequency estimate. Competition independent of the recipient trait still affects the mass equation, but contributes no such frequency term.

10.2 Dependence of the equilibrium on the coefficients

The contraction also supplies a quantitative perturbation statement. Write $\mathcal{G} = pQ$, and use tildes for a second coefficient collection.

Corollary 10.4 (A perturbation bound for equilibria)**.** *Suppose that both systems have the common invariant absorbing interval $[\ell, U]$, and that the first system satisfies (10.7) with $q, \lambda > 0$. Suppose the second system has a stationary measure $\widetilde{u}_*$ in this interval. Set*

$$\varepsilon_\nu = \|\mathcal{G} - \widetilde{\mathcal{G}}\|_{\mathcal{M}\to\mathcal{M}} + \operatorname{osc}(a - \widetilde{a}) + 2U\|C - \widetilde{C}\|_\infty + \|\alpha - \widetilde{\alpha}\|_\infty,$$
$$\varepsilon_N = \|a - \widetilde{a}\|_\infty + U\|C - \widetilde{C}\|_\infty.$$

Then

$$\|\nu_* - \widetilde{\nu}_*\|_{\mathrm{TV}} + q|\log N_* - \log \widetilde{N}_*| \le \frac{\varepsilon_\nu + q\varepsilon_N}{\lambda}.$$

In particular, two coefficient collections satisfying a common strict contraction bound have quantitatively close equilibria.

Proof. At a fixed state (N, ν), the two frequency vector fields differ by at most ε_ν and the two logarithmic mass vector fields by at most ε_N. Repeat (10.10)–(10.11) for one trajectory from each system, adding these forcing bounds. The weighted distance d satisfies $d^+ \le -\lambda d + \varepsilon_\nu + q\varepsilon_N$. Apply this inequality to the two stationary trajectories. □

For example, the operator perturbation is bounded by

$$\|pQ - \widetilde{p}\,\widetilde{Q}\| \le 2\|pb - \widetilde{p}\,\widetilde{b}\|_\infty + \|\widetilde{p}\,\widetilde{b}\|_\infty \sup_x \|m(x,\cdot) - \widetilde{m}(x,\cdot)\|_{\mathrm{TV}}.$$

The one-step minorization itself need not persist under small kernel perturbations. The semigroup contraction does persist: on zero-mass measures, variation of constants gives

$$\|e^{t\widetilde{\mathcal{G}}}\sigma\|_{\mathrm{TV}} \le e^{-(\theta - \|\widetilde{\mathcal{G}} - \mathcal{G}\|)t}\|\sigma\|_{\mathrm{TV}}.$$

Thus the contraction argument of Theorem 10.2 remains valid for sufficiently small coefficient perturbations, using the reduced mixing rate and a common absorbing interval. The contraction regime is therefore open in the displayed operator and uniform norms, even though the original one-step minorization need not persist.

11 Examples separating the structural conditions

The preceding theorems impose different conditions on the coefficients. The examples in this section identify several failures of implications between positivity, mixing, uniqueness, and convergence.

11.1 Complete-reset mutation can coexist with bistability

Proposition 11.1 (Two stable equilibria under strictly positive mutation)**.** *Let $\mathcal{X} = \{1, 2\}$, $b = r = 1$, $d = 0$, $\alpha = 0$, and*

$$m(i, \{j\}) = \frac{1}{2}, \qquad C = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}, \qquad 0 < p < 1.$$

Then the population equation has three nonzero equilibria, all strictly positive. Two are asymptotically stable and the third is a saddle.

Proof. Write $N = u_1 + u_2$ and $q = u_1/N$. Direct calculation gives

$$N' = N\big[1 + p - N(1 + 4q(1-q))\big],$$
$$q' = (q - \tfrac{1}{2})\big[4Nq(1-q) - p\big].$$

A nonzero equilibrium cannot have $q = 0$ or $q = 1$, since mutation points into the interval. The three possibilities are

$$(N, q) = \left(\frac{1+p}{2}, \frac{1}{2}\right), \qquad (N, q) = \left(1, \frac{1 \pm \sqrt{1-p}}{2}\right).$$

The Jacobian at the central point is diagonal with eigenvalues $-(1+p)$ and $(1-p)/2$. At either outer equilibrium put $x = 2q - 1$, so $x^2 = 1 - p$. The Jacobian is

$$\begin{pmatrix} -(1+p) & 4x \\ px/2 & -2x^2 \end{pmatrix}.$$

Its characteristic polynomial is $(\lambda+2)(\lambda+1-p)$. Thus the two outer equilibria are hyperbolic sinks and the central point is a hyperbolic saddle. □

The mutation component alone relaxes to the uniform frequency at rate p, since the kernel satisfies a common-reset condition with $\eta = 1$. In the full equation, competition between different traits is three times self-competition, and two asymmetric attracting equilibria remain for $0 < p < 1$. Thus irreducibility and strict positivity of mutation alone do not imply uniqueness. The negative competition direction $(1, -1)$ prevents the common-reference competition condition from holding.

11.2 Periodic residents with a nonstationary population average

Let

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \qquad \pi = \frac{1}{3}(1,1,1)^{\mathsf{T}}. \tag{11.1}$$

The interior solutions of $\dot{q}_i = q_i(Aq)_i$ preserve $q_1 q_2 q_3$. For each value in $(0, 1/27)$, its level set in the open simplex is a smooth closed curve and the vector field never vanishes on that curve. Hence every such level set is a periodic orbit. The only interior stationary frequency is π.

If $b = r = C = 1$, $d = p = 0$, and $N_0 = 1$, these frequency orbits give periodic population solutions of the original model. The transfer rate $\tau_{ij} = \max\{A_{ij}, 0\}$ realizes the required nonnegative horizontal-transfer rates.

The next variant exhibits the correlation between total mass and composition explicitly.

Proposition 11.2 (The mean of a periodic population need not be stationary)**.** *Let $\mathcal{X} = \{1,2,3\}$, $b = r = 1$, $d = p = 0$, let $\alpha = A$ be (11.1), and set*

$$C_{ij} = c_j, \qquad c = (1,2,3)^{\mathsf{T}}.$$

There is a positive periodic population solution $u_t = N_t q_t$ whose time average $\bar{u}$ is not an equilibrium, although

$$r - C\bar{u} + A\bar{q} = 0$$

at every trait.

Proof. Choose a nonconstant interior periodic frequency orbit q_t, extended periodically to all $t \in \mathbb{R}$, and define

$$I_t = \int_0^\infty e^{-s}\, c^{\mathsf{T}} q_{t-s}\, ds, \qquad N_t = I_t^{-1}.$$

Then $I_t' = c^{\mathsf{T}} q_t - I_t$, so

$$N_t' = N_t[1 - N_t c^{\mathsf{T}} q_t].$$

Competition is independent of the recipient, and r is constant. Thus q_t solves the frequency equation independently of N_t, and $u_t = N_t q_t$ is a positive periodic population solution.

Integrating $(\log q_i)' = (Aq)_i$ over one period gives $A\bar{q} = 0$, hence $\bar{q} = \pi$. Averaging the equation for I yields $\bar{I} = c^{\mathsf{T}}\bar{q} = 2$. The function $c^{\mathsf{T}} q_t$ is nonconstant, since a nonconstant closed frequency orbit cannot lie in the affine line $c^{\mathsf{T}} q =$ constant inside the two-dimensional simplex. Consequently I is nonconstant. Strict Jensen inequality gives

$$\bar{N} = \overline{I^{-1}} > \frac{1}{\bar{I}} = \frac{1}{2}.$$

On the other hand, averaging $(\log N)' = 1 - N c^{\mathsf{T}} q$ gives $c^{\mathsf{T}}\bar{u} = 1$. If the strictly positive vector $\bar{u}$ were stationary, its per-capita growth equations would therefore imply $A(\bar{u}/\bar{N}) = 0$. It would follow that $\bar{u} = \bar{N}\pi$, and then $1 = c^{\mathsf{T}}\bar{u} = 2\bar{N}$, a contradiction. Finally, $C\bar{u} = 1$ and $A\bar{q} = 0$, proving the last assertion. □

This example realizes the two distinct environmental averages $\bar{u} = \overline{Nq}$ and $\bar{q}$ in the complementarity and invasion formulas. Their distinction persists even for an exact periodic solution.

11.3 A comparison of the convergence mechanisms

The examples in Section 9 distinguish the convergence criteria. A normalized positive semigroup can converge when the unaugmented weighted kernel is indefinite. The mutation-augmented weighted criterion can also control arbitrarily strong circulation beyond the coefficient contraction bound. Proposition 11.1 shows that a mutation lower bound by itself does not select a unique equilibrium. The spectral-edge examples of Section 7 show that the obstruction can also be infinite dimensional: complete reset mutation may admit a family of stationary measures with nonzero singular components relative to its reset law.

12 Discussion

12.1 How to use the criteria

The mass equation supplies the first reduction in every case. It places every trajectory in a fixed positive mass interval and excludes total extinction under the assumption $r_- > 0$. If $p = 0$, the initial support remains dynamically relevant, and every monomorphic population has its own stationary mass. Global uniqueness over all nonzero initial measures is therefore impossible unless the trait space consists of one point. The averaged-environment and dominance results provide statements with the appropriate support-dependent scope.

For $p > 0$, the accessible mutation class should be identified first. An invariant closed class can be treated as its own trait space. When competition is independent of the recipient and transfer is a potential difference, the frequency equation is exactly a normalized positive semigroup. In finite dimensions irreducibility of the mutation matrix then suffices. In the continuum, Theorem 6.2 combines a common reset with a positive adjoint eigenfunction

bounded above and away from zero. The oscillation criterion constructs that eigenfunction under a stronger coefficient inequality. Complete reset with a full-support law already mixes the conservative mutation process, yet the rank-one examples reach a spectral edge after selection is included. Mixing of mutation alone therefore does not guarantee that the normalized population semigroup converges in total variation to a common stationary frequency.

For general competition and circulation, a common reference measure gives a direct entropy test. If that balance fails, the weighted criterion uses a stationary population and the left eigenfunction of its frozen operator. The corrected quadratic kernel S_N contains both the weighted antisymmetric interaction and the change in the normalization N^{-1}. Once the stationary state, adjoint weight, and required mutation condition are available, the remaining positivity test reduces in finite dimensions to two endpoint matrices. Mutation dissipation can be added before testing positivity, as in Corollary 9.7. Pointwise positivity of C describes competitive mortality; nonnegativity of its quadratic form controls signed perturbations and is a separate condition.

The contraction criterion uses a different set of data. It can be checked through a mutation minorization, coefficient oscillations, and uniform norms without finding a stationary measure. Subtracting a potential from transfer and a recipient-independent term from competition can substantially reduce these bounds. Strict contraction also gives quantitative stability of the equilibrium under coefficient perturbations.

12.2 What changes at the boundaries of the hypotheses

Several boundaries are identified by explicit formulas. Without Feller continuity, weak limits need not be stationary. In the mutation-free cyclic-transfer example with constant competition, conservation of a frequency entropy sustains periodic trajectories. At the critical rank-one threshold $\rho = 1$, an initial frequency δ_{x_*} at a maximizer retains mass at least $(1 + Gt)^{-1}$ on the maximizing set, with $G = \max g$. This lower bound excludes exponential convergence to the stationary frequency, which gives that set zero mass. In the condensed regime $\rho < 1$, an initial population with $u_0(K) = 0$ continues to give K zero mass at every finite time, whereas each stationary frequency gives it mass $1 - \rho$. Total variation convergence is therefore impossible for these initial data, even when K is a singleton and the stationary frequency is unique.

The distinction between endpoint and time-averaged states is equally substantive. A periodic resident can have identically zero averaged fitness while its mean population is not stationary (Proposition 11.2). For invasion and persistence, the relevant environment is an invariant probability law. Checking isolated resident equilibria alone does not control all boundary invariant laws, as the cyclic competition example demonstrates.

12.3 Questions suggested by the theory

The following extensions require additional analysis.

1. For mutation kernels satisfying only a multistep or state-dependent mixing condition, one may seek nonlinear analogues of generalized Doeblin and Harris estimates. The obstacle is to control the time-dependent selection weights while keeping a uniform estimate for differences of normalized trajectories.

2. The weighted identity reduces convergence to a nonlinear quadratic form, which includes competition and transfer corrections, and a frozen Markov dissipation. A useful next step is to replace the uniform reset assumption by a quantitative functional inequality

for the transformed generator. Any such extension must also control the weighted total mass, since dissipation of the normalized weighted composition vanishes on the entire ray of positive multiples of the stationary population.

3. In the rank-one condensed regime, positive initial mass on the maximizing set determines the relative condensate weights explicitly. When that mass is zero, the shifted linear population tends to zero and normalization becomes singular. Multiple maximizers then require estimates of their relative near-maximal contributions. Weak convergence of the frequency trajectory without time averaging cannot be inferred from the stationary classification or from convergence of its Cesàro averages.

4. For recurrent residents, scalar invasion indices are replaced by positive-operator growth rates when mutation couples the rare traits. In finite dimensions, irreducible positive cocycles give a principal exponent. For continuous trait classes, identification of a principal direction requires additional compactness or positivity beyond Feller continuity.

5. Uniform versions of the convergence criteria could justify interchange of large-time and mutation-limit procedures. The present constants may degenerate when a mixing rate or spectral gap tends to zero. Consequently the fixed-kernel results do not settle the monomorphic continuation and polymorphic questions for the noncompact, small-diffusion model studied in Gárriz and Mirrahimi [10]. That model also has a confining growth function that is negative for large traits, so the positive-growth mass bounds used here are unavailable in that form.

References

[1] Vincent Bansaye, Bertrand Cloez, and Pierre Gabriel. Ergodic behavior of non-conservative semigroups via generalized Doeblin's conditions. *Acta Applicandae Mathematicae*, 166:29–72, 2020. doi: 10.1007/s10440-019-00253-5.

[2] Vincent Bansaye, Bertrand Cloez, Pierre Gabriel, and Aline Marguet. A non-conservative Harris ergodic theorem. *Journal of the London Mathematical Society*, 106(3):2459–2510, 2022. doi: 10.1112/jlms.12639.

[3] Sylvain Billiard, Pierre Collet, Régis Ferrière, Sylvie Méléard, and Viet Chi Tran. Stochastic dynamics for adaptation and evolution of microorganisms. In *European Congress of Mathematics, Berlin 2016*, pages 525–550. EMS Press, 2018. doi: 10.4171/176-1/25.

[4] Àngel Calsina, Sílvia Cuadrado, Laurent Desvillettes, and Gaël Raoul. Asymptotic profile in selection–mutation equations: Gauss versus Cauchy distributions. *Journal of Mathematical Analysis and Applications*, 444(2):1515–1541, 2016. doi: 10.1016/j.jmaa.2016.07.028.

[5] Vincent Calvez, Susely Figueroa Iglesias, Hélène Hivert, Sylvie Méléard, Anna Melnykova, and Samuel Nordmann. Horizontal gene transfer: numerical comparison between stochastic and deterministic approaches. *ESAIM: Proceedings and Surveys*, 67:135–160, 2020. doi: 10.1051/proc/202067009.

[6] Nicolas Champagnat, Sylvie Méléard, and Viet Chi Tran. Multiscale eco-evolutionary models: from individuals to populations. In *Proceedings of the International Congress of Mathematicians 2022*, volume 7, pages 5656–5678. EMS Press, 2023. doi: 10.4171/ICM2022/24.

[7] Bertrand Cloez and Pierre Gabriel. Fast, slow convergence, and concentration in the house of cards replicator-mutator model. *Differential and Integral Equations*, 37(7–8):547–584, 2024. doi: 10.57262/die037-0708-547.

[8] Jérôme Coville and Léo Girardin. An abstract framework for a class of nonlocal structured population models: existence, uniqueness and stability of steady states, 2026. URL https://arxiv.org/abs/2512.13232v2. Version 2, 21 September 2026.

[9] Laurent Desvillettes, Pierre-Emmanuel Jabin, Stéphane Mischler, and Gaël Raoul. On selection dynamics for continuous structured populations. *Communications in Mathematical Sciences*, 6 (3):729–747, 2008. doi: 10.4310/CMS.2008.v6.n3.a10.

[10] Alejandro Gárriz and Sepideh Mirrahimi. A Hamilton–Jacobi approach for the evolutionary dynamics of a model with gene transfer: characterizing monomorphic dynamics for non-concave fitness functions. *Journal of Differential Equations*, 472:114402, 2026. doi: 10.1016/j.jde.2026.114402.

[11] Alejandro Gárriz, Alexis Léculier, and Sepideh Mirrahimi. Impact of a unilateral horizontal gene transfer on the evolutionary equilibria of a population. *Mathematical Models and Methods in Applied Sciences*, 34(5):951–991, 2024. doi: 10.1142/S0218202524500192.

[12] Josef Hofbauer and Sebastian J. Schreiber. Robust permanence for interacting structured populations. *Journal of Differential Equations*, 248(8):1955–1971, 2010. doi: 10.1016/j.jde.2009.11.010.

[13] Pierre-Emmanuel Jabin and Gaël Raoul. On selection dynamics for competitive interactions. *Journal of Mathematical Biology*, 63:493–517, 2011. doi: 10.1007/s00285-010-0370-8.

[14] Robert M. May and Warren J. Leonard. Nonlinear aspects of competition between three species. *SIAM Journal on Applied Mathematics*, 29(2):243–253, 1975. doi: 10.1137/0129022.

[15] Philippe Michel, Stéphane Mischler, and Benoît Perthame. General relative entropy inequality: an illustration on growth models. *Journal de Mathématiques Pures et Appliquées*, 84(9):1235–1260, 2005. doi: 10.1016/j.matpur.2005.04.001.