MUTATION-SELECTION BALANCE WITH RECOMBINATION: CONVERGENCE TO EQUILIBRIUM FOR POLYNOMIAL SELECTION COSTS

AUBREY CLAYTON AND STEVEN N. EVANS

Abstract. We study a continuous-time dynamical system that models the evolving distribution of genotypes in an infinite, sexually reproducing, haploid population where genomes may have infinitely many or even a continuum of loci, mutations accumulate along lineages without back-mutation, additional mutations reduce fitness, selection costs may be epistatic, and recombination occurs on a faster time scale than mutation and selection. Some features of the model, such as existence and uniqueness of solutions and convergence to the dynamical system of an approximating sequence of discrete time models, were presented in earlier work by Evans, Steinsaltz, and Wachter for quite general selection costs. Here we investigate a special case where the selection cost of a genotype with a given accumulation of ancestral mutations from a wild type ancestor is a sum of costs attributable to each individual mutation plus successive interaction contributions from each k-tuple of mutations for k up to some finite "degree". Using ideas from complex chemical reaction networks and a novel Lyapunov function, we establish that the phenomenon of mutation-selection balance occurs for such selection costs under mild conditions. That is, we show that the dynamical system has a unique equilibrium and that it converges to this equilibrium from all initial conditions.

1. Introduction

Many phenotypic traits, including some genetic disorders, are thought to be *polygenic* and result from the possibly complex *epistatic* (that is, non-additive) interactions between large numbers of mildly deleterious alleles that are slowly weeded out of the population by natural selection but are constantly reintroduced by recurrent mutation. In particular, the Medawar–Williams–Hamilton [Med52, Wil57, Ham66] explanation of the evolution of aging (see, also, [Cha94, Cha01] and the introductory discussions in [SEW05, ESW06, WES08, WSE08]) invokes this mechanism of *mutation-selection balance*.

It has been a challenge to provide a sufficiently usable quantitative description of the mutation-selection phenomenon when many genes are involved (see, for example, [Hol95] or [FK00]: in particular, Section 5.3.2 of the latter is entitled "Need for a mathematical framework" and specifically lays out the necessity for mathematically tractable approaches to outcomes that are modulated by the interaction of large numbers of genes).

Date: February 2, 2009.

1991 *Mathematics Subject Classification.* Primary: 92D10, 37N25; Secondary: 60G57, 37L15.

Key words and phrases. population genetics, Poisson random measure, infinite dimensional dynamical system, global stability, fixed point, continuous flow stirred tank reactor, Lyapunov function, polygenic character.

Research supported in part by NSF grant DMS-0405778.

Evans, Steinsaltz, and Wachter [SEW05, ESW06] describe two related mathematical settings for the mutation-selection process. Here we consider the model of the latter paper. The key assumptions in [ESW06] are:

- the population is infinite,
- the genome may consist of infinitely many or even a continuum of loci,
- reproduction is sexual, in that each individual has two parents and the mechanism of *genetic recombination* randomly shuffles together the genomes of the parents in order to obtain the genome of the offspring,
- mating is random,
- individuals are *haploid*,
- mutations accumulate along a lineage from an ancestral *wild type* genotype (that is, there is no back-mutation canceling out the effect of earlier mutations),
- fitness is calculated for individuals rather than for mating pairs,
- genotypes that have more accumulated mutations are less fit, but otherwise selection costs are arbitrary,
- the effects of recombination are felt on a faster time scale than those of mutation or selection.

For the purpose of explaining our results, we now describe briefly the general model of [ESW06]. Our results concern a particular family of selection costs that we introduce later.

We write $\mathcal{M}$ for the collection of *loci* in the portion of the genome that is of interest to us. We assume that there is a distinguished reference wild type genotype, and each locus represents a "position" at which the genotype of an individual may differ from that of the wild genotype. We have placed the word "position" in quotes, because we do not necessarily take $\mathcal{M}$ to be something like a finite collection of physical DNA base positions or a finite collection of genes. From a modeling point of view, it is convenient to allow $\mathcal{M}$ to be more general. For example, the proposed explanations in Charlesworth [Cha01] for Gompertz mortality curves and mortality plateaus at extreme ages suggest that one relevant choice of $\mathcal{M}$ may be a class of functions from $\mathbb{R}_+$ to $\mathbb{R}_+$: the value of such a function at time $t \geq 0$ represents an additional increment to the mortality hazard rate at age t conferred by a mutation away from the wild type at this locus. Some minimal amount of structure on $\mathcal{M}$ is necessary in order to accommodate probability theory, and so we take $\mathcal{M}$ to be a complete, separable metric space.

We suppose that the genotype of an individual is described completely by the number of times a mutation event has occurred at each locus along the lineage connecting the individual to a wild type ancestor. Similar assumptions are standard in models of the *infinitely-many-alleles*, *infinitely-many-sites*, and *continuum-of-alleles* type. In particular, all potential mutations at a locus that is currently wild type are indistinguishable, and further mutation events cannot undo the results of earlier ones (that is, there is no *back-mutation*). When the space $\mathcal{M}$ is continuous, it will typically be the case that multiple mutations will not occur along a lineage at any locus. For discrete spaces $\mathcal{M}$ that are large, the assumption of mutation accumulation without back-mutation is reasonable when mutation rates are sufficiently low that it is unlikely that any locus will be affected by a mutation more than once along a lineage.

Consequently, we may think of an individual's genotype as an element of the space $\mathcal{G}$ of integer–valued finite Borel measures on $\mathcal{M}$, where the non-negative integer $g(A)$ for $g \in \mathcal{G}$ and $A \subseteq \mathcal{M}$ represents the total number of mutation events that have occurred at loci belonging to the subset A of the genome along the lineage connecting the individual to a wild type ancestor. An element of $\mathcal{G}$ is a finite sum $\sum_i \delta_{m_i}$, where δ_m is the unit point mass at the locus $m \in \mathcal{M}$, corresponding to a genotype where there have been ancestral mutations at loci $m_1, m_2, \ldots$. In particular, the wild type genotype is the null measure. As we remarked above, we do not require that the loci $m_i \in \mathcal{M}$ are distinct. For example, if the set $\mathcal{M}$ of loci is finite, so we might as well take $\mathcal{M} = \{1, 2, \ldots, L\}$ for some positive integer L, then a genotype is of the form $\sum_{j=1}^{L} n_j \delta_j$, indicating that ancestral mutations have occurred n_j times at locus j. We may identify such a genotype with the non-negative integer vector $\mathbf{n} := (n_1, n_2, \ldots, n_L)$, so that $\mathcal{G}$ is essentially the Cartesian product $\mathbb{N}^L$ of L copies the non-negative integers.

We imagine that the population we are interested in is infinite and that all that matters about an individual is the individual's genotype, so that the dynamics of the population can be described in terms of the proportions of individuals with genotypes belonging to the various subsets of $\mathcal{G}$. We therefore seek to model the evolution of probability measures P_t, $t \geq 0$, on $\mathcal{G}$, where $P_t(G)$ represents the proportion of individuals in the population at time t that have genotypes belonging to the set $G \subseteq \mathcal{G}$. Equivalently, $P_t(G)$ can be interpreted as the probability that an individual chosen uniformly at random from the population will have genotype belonging to G, so that P_t is the distribution of a random finite integer-valued measure on $\mathcal{M}$. For example, if $\mathcal{M} = \{1, 2, \ldots, L\}$ and we identify $\mathcal{G}$ with $\mathbb{N}^L$ as above, then $P_t(\{\mathbf{n}\})$ represents the probability that an individual chosen uniformly at random from the population will have n_j ancestral mutations at locus j for $j = 1, 2, \ldots, L$.

We remark that although the introduction of this measure-theoretic formalism might seem somewhat heavy-handed, the use of measures to represent genotypes and probability measures on spaces of measures to represent populations has become quite standard in mathematical population genetics. It is apparent from [EK86] how useful this perspective has been for understanding stochastic models of the infinitely-many-alleles and infinitely-many-sites type. Similarly, the stochastic model of [SH92] is based on a Poisson random measure on a Cartesian product of the space of sites and the unit interval (with the latter coordinate representing frequency of mutant alleles at the corresponding site). Also, the recent papers [BB03, Baa05, Baa07] adopt a measure-theoretic framework to present a deterministic model of recombination based on [Baa01]. The discrete-time mutation-selection models of [Esh71] and [Kin78] describe the population in terms of a probability measure over the real numbers that records the proportions of individuals with various fitnesses. A far-reaching continuous-time generalization of these models is presented in [BB96] (see also [Bür00]) where individuals have an abstract *type* belonging to a locally compact space and the population is described by a probability measure on this space.

In order to make this paper more self-contained and to give the reader a better understanding of the model of [ESW06], we next indicate how mutation, selection and recombination are modeled to obtain the dynamics by which P_t evolves.

Mutation alone. Suppose that there is only mutation and no selection or recombination. Because no genome has a selective advantage, all individuals are dying and reproducing at the same rate, and since there is no recombination the genotype of an offspring is simply that of the parent with possible alterations due to mutation. Somewhat informally, we imagine that there is finite measure ν on the space of loci $\mathcal{M}$ such that for any individual alive at time $t \geq 0$ the probability a mutation occurs in the region $dm \subset \mathcal{M}$ during the time interval dt is $dt \times \nu(dm)$, and we also assume that the occurrences of mutations are independent between different individuals, different times, and different regions of $\mathcal{M}$. If we write $P_t\Phi = \int_{\mathcal{G}} \Phi(g) P_t(dg)$ for some test function $\Phi : \mathcal{G} \to \mathbb{R}$ (that is, $P_t\Phi$ is the expected value of the real-valued random variable obtained by applying the function Φ to the genotype of an individual chosen uniformly at random from the population), then the content of these assumptions is contained formally in the equation

$$\frac{d}{dt} P_t\Phi = P_t \left(\int_{\mathcal{M}} \left[\Phi(\cdot + \delta_m) - \Phi(\cdot) \right] \nu(dm) \right).$$

For example, when $\mathcal{M} = \{1, 2, \ldots, L\}$ we have

$$\frac{d}{dt} P_t(\{\mathbf{n}\}) = \sum_{j=1}^{L} \nu(\{j\}) \left[P_t(\{\mathbf{n} - \mathbf{e_j}\}) - P_t(\{\mathbf{n}\}) \right],$$

where $\mathbf{e_j}$ is the j^{th} coordinate vector. We stress that this equation is classical: it is a special case of the usual equation describing evolution due to mutation of type frequencies in a population where the set of types is $\mathbb{N}^L$ and mutation from type $\mathbf{n}$ to type $\mathbf{n} + \mathbf{e_j}$ occurs at rate $\nu(\{j\})$ – see, for example, Section III.1.2 of [Bür00].

This evolution equation has a simple explicit solution. Let Π denote a Poisson random measure on $\mathcal{M} \times \mathbb{R}_+$ with intensity measure $\nu \otimes \lambda$, where λ is Lebesgue measure; that is, Π is a random integer-valued Borel measure such that:

(1) The random variable $\Pi(A)$ is Poisson with expectation $\nu \otimes \lambda(A)$ for any Borel subset A of $\mathcal{M} \times \mathbb{R}_+$.
(2) If $A_1, A_2, \ldots, A_n$ are disjoint Borel subsets of $\mathcal{M} \times \mathbb{R}_+$, then the random variables $\Pi(A_k)$ are independent.

Define a $\mathcal{G}$-valued random variable Z_t (that is, Z_t is a random finite integer-valued measure on $\mathcal{M}$) by $Z_t := \int_{\mathcal{M} \times [0,t]} \delta_m \, \Pi(d(m,u))$. Then $P_t\Phi = \mathbb{E}\left[\Phi(W + Z_t)\right]$, where W is a random measure on $\mathcal{M}$ that is independent of Π, and W has distribution P_0. In particular, if P_0 is the distribution of a Poisson random measure, then P_t will also be the distribution of Poisson random measure. If we write ρ_t for the intensity measure associated with P_t (that is, ρ_t is the measure on $\mathcal{M}$ defined by $\rho_t(M) := \int_{\mathcal{G}} g(M) \, P_t(dg)$ for $M \subseteq \mathcal{M}$), then ρ_t evolves according to the simple dynamics $\rho_t(M) = \rho_0(M) + t\,\nu(M)$.

Selection alone. Now consider what happens if there is only selection and no mutation or recombination. We assume that there is a *selection cost function* $S : \mathcal{G} \to \mathbb{R}_+$ with the interpretation that $S(g') - S(g'')$ for $g', g'' \in \mathcal{G}$ represents the difference in the rate of sub-population size increase between those individuals with genotype g'' and those individuals with genotype g'. We make the normalizing assumption $S(0) = 0$ and suppose that

$$S(g + h) \geq S(h), \quad g, h \in \mathcal{G}, \tag{1.1}$$

so that if the ancestral mutations from wild type in one genotype are a subset (taking into account multiplicity) of those in a second genotype, then the sub-population of individuals with the second genotype have a lesser rate of increase than the sub-population with the first genotype. Heuristically, we have

$$\frac{d}{dt}P_t(dg') = \frac{d}{du}\frac{\exp(-S(g')u)P_t(dg')}{\int_{\mathcal{G}}\exp(-S(g'')u)\,P_t(dg'')}\Big|_{u=0} = (P_tS - S(g'))\,P_t(dg');$$

That is, at time $t \geq 0$ the relative rate of increase of the proportion of the population of individuals with genotype g' is $P_tS - S(g')$. More formally, we have

$$\frac{d}{dt}P_t\Phi = -\,P_t(\Phi\cdot[S - P_tS]) = -\int_{\mathcal{G}}\Phi(g')\left[S(g') - \int_{\mathcal{G}}S(g'')\,P_t(dg'')\right]\,P_t(dg').$$

For example, when $\mathcal{M} = \{1, 2, \ldots, L\}$ we have

$$\frac{d}{dt}P_t(\{\mathbf{n}'\}) = -\left[S(\mathbf{n}') - \sum_{\mathbf{n}''}P_t(\{\mathbf{n}''\})S(\mathbf{n}'')\right]P_t(\{\mathbf{n}'\}).$$

We again stress that this equation is classical: it is a special case of the usual equation describing evolution due to selection of type frequencies in a population where the set of types is $\mathbb{N}^L$ – see, for example, Section III.1.2 of [Bür00].

If S is *non-epistatic* (that is, S has the additive property $S(\sum_i \delta_{m_i}) = \sum_i S(\delta_{m_i})$), then there is no interaction between the selective effects of ancestral mutations at different loci or multiple ancestral mutations at a single locus: in particular, if P_0 is the distribution of a Poisson random measure then P_t will also be the distribution of a Poisson random measure and, writing ρ_t for the intensity measure associated with P_t as before, we have

$$\rho_t(dm') = \rho_0(dm') - \int_0^t\left[S(\delta_{m'}) - \int_{\mathcal{M}}S(\delta_{m''})\rho_s(dm'')\right]\,\rho_s(dm')\,ds.$$

However, if S is epistatic, then P_t will, in general, not be the distribution of a Poisson random measure – even when P_0 is.

Recombination alone. Lastly, we discuss the incorporation of recombination. Our description will be brief, as the details will not be important to us, because we are interested in an asymptotic regime where recombination occurs on a faster time-scale than mutation or selection, and in the limit the detailed features of the recombination mechanism disappear. The effect of recombination is to choose an individual uniformly at random from the population at some rate and replace the individual's genotype g' with a genotype of the form $g'(\cdot\cap M) + g''(\cdot\cap M^c)$, where g'' is the genotype of another randomly chosen individual, M is a subset of $\mathcal{M}$ chosen according to a suitable random mechanism, and M^c is the complement of M. Thus, recombination randomly shuffles together two different genotypes drawn from the population. If recombination alone is acting, P_0 has the property that there does not exist an $m \in \mathcal{M}$ with $P_0(\{g \in \mathcal{G} : g(\{m\}) > 0\})$, and the mechanism for choosing the so-called *segregating* set M is such that, loosely speaking, if m' and m'' are two loci then there is positive probability that $m' \in M$ and $m'' \in M^c$, so that no region of $\mathcal{M}$ is immune from shuffling, then P_t will converge to a Poisson random measure with the same intensity measure as P_0 as t increases – see [ESW06] for details.

Combining mutation, selection and recombination. We have seen that if P_0 is the distribution of a Poisson random measure, then mutation preserves this property, while epistatic selection and recombination respectively drive the system away from and toward Poisson. If all three processes are operating and we consider a limiting regime where recombination acts on a faster time scale than mutation and selection, then we expect that the resulting system will preserve the Poisson property. Results to this effect are established in [ESW06]: more precisely, it is shown there that if one considers a discrete generation evolution in which mutation and selection are modeled by discrete-time analogues of the differential equations described above, recombination is modeled as above, and time is scaled so that the time between generations converges to zero, then the discrete-time evolution converges to a continuous-time evolution that preserves Poisson initial conditions, provided that asymptotically recombination acts on a faster time scale than mutation and selection. Moreover, the explicit details of the recombination mechanism in the discrete generation evolutions have no influence on the limit: all that matters is that the relative effect of recombination becomes stronger and stronger in the limit, and, as described above, that no region of $\mathcal{M}$ is immune from the shuffling effect of recombination. We refer the reader to [ESW06] for rigorous statements and proofs; however, we do stress that the analysis of [ESW06] does **not** start from an *a priori* assumption that loci are unlinked, rather it delineates how strong the Poissonizing effect of recombination has to be in order to overcome the tendency of non-epistatic selection to induce linkage between loci.

The limiting evolution is thus a family of probability measures P_t on $\mathcal{G}$ that are Poisson with intensity measure ρ_t. If we write X^π for a Poisson random measure on $\mathcal{M}$ with intensity measure π, then, as we expect from combining the observations above, ρ_t satisfies the evolution equation

$$\rho_t(dm) = \rho_0(dm) + t\,\nu(dm) - \int_0^t \mathbb{E}\left[S(X^{\rho_s} + \delta_m) - S(X^{\rho_s})\right]\,\rho_s(dm)\,ds. \tag{1.2}$$

For example, when $\mathcal{M} = \{1, 2, \ldots, L\}$ we have

$$\frac{d}{dt}\rho_t(\{j\}) = \nu(\{j\}) - \rho_t(\{j\}) \sum_{\mathbf{n}} [S(\mathbf{n} + \mathbf{e_j}) - S(\mathbf{n})] \prod_{k=1}^{L} \exp(-\rho_t(\{k\})) \frac{\rho_t(\{k\})^{n_k}}{n_k!}.$$

Note that although there are infinitely many "types" in this latter finitely-many-loci special case, the evolution of the population is described by a system of L differential equations for the L real quantities $\rho_t(\{j\})$, $1 \le j \le L$.

In this paper we consider selection costs S that are, in a suitable sense, *polynomial*. Roughly speaking, this means that, for some finite positive integer N, the selection cost of a given genotype (that is, a given collection of accumulated mutations from the ancestral wild type) is a sum of non-negative contributions from each individual mutation, plus a sum of non-negative contributions due to interactions between the mutations in every pair of mutations at distinct loci, and so on – all they way up to a sum of non-negative contributions due to interactions between the mutations in every N-tuple of mutations at distinct loci.

We show for such selection costs that, under very mild conditions, the system (1.2) has a unique equilibrium and that the system converges to this equilibrium from any initial condition ρ_0 that is absolutely continuous with respect to ν.

Because defining precisely what we mean by a polynomial selection cost and describing our results for a general locus space $\mathcal{M}$ requires a certain amount of extra notation, for the sake of this introduction we just describe our results in the special case when $\mathcal{M}$ is the finite set $\{1, ..., L\}$, and leave the general case to the body of the paper.

In the finitely-many-loci case, the mutation rate measure ν is defined by its value on singleton sets and we write $\nu_i := \nu(\{i\})$. We assume that $\nu_i > 0$ for all i, and let ν also denote the vector $(\nu_1, \ldots, \nu_L)$. As we remarked above, a genotype $g \in \mathcal{G}$ can be encoded by an ordered $L-$tuple of non-negative integers $g = (g_1, \ldots, g_L)$, where g_k represents the number of ancestral mutations that are present at locus k. A polynomial selection cost is one of the form

$$S(g) = \sum_I \alpha_I g^I, \tag{1.3}$$

where the sum is taken over all nonempty subsets $I \subseteq \{1, \ldots, L\}$ and we adopt the convention that for a vector v the notation v^I denotes the product $\prod_{i \in I} v_i$. The constants $\alpha_{\{i\}}$ for $1 \leq i \leq L$ measure the selection cost of a mutation at locus i alone, whereas the constants α_I for subsets $I \subseteq \{1, \ldots, L\}$ of cardinality greater than one measure the selection cost attributable to interactions between the mutations at loci in I over and above that attributable to interactions between mutations in sets of loci that are proper subsets of I. We assume that $\alpha_I \geq 0$ for all I and $\alpha_I = 0$ for sets I with cardinality greater than the "degree" N. This assumption ensures that the monotonicity condition (1.1) holds. It is, of course, not a necessary condition for (1.1) to hold, but it will be crucial in our analysis. Indeed, it will be shown in a forthcoming paper by Evans, Steinsaltz and Wachter that there are some natural infinite dimensional systems that arise in applications to aging where the analogue of (1.3) fails to hold and the resulting system does not converge to an equilibrium.

Write $(\rho_1(t), ..., \rho_L(t))$ for the vector corresponding to the intensity measure of P_t. As we show in Section 2, the evolution equation (1.2) is equivalent to the system of ordinary differential equations

$$\dot{\rho_k} = \nu_k - \sum_{I \in \mathcal{I}_k} \alpha_I \rho^I, \quad 1 \leq k \leq L, \tag{1.4}$$

where $\mathcal{I}_k$ denotes the collection of subsets of $\{1, \ldots, L\}$ that contain k. The following is an immediate corollary of our main result for general locus spaces $\mathcal{M}$, Theorem 3.1.

Corollary 1.1. *Suppose that $\alpha_{\{i\}} > 0$ for $1 \leq i \leq L$. The system (1.4) has a unique equilibrium point in the positive orthant $\mathbb{R}_+^L$, and this equilibrium is globally stable (that is, the system converges to the equilibrium from any initial conditions in $\mathbb{R}_+^L$).*

Note that, in this finitely-many-loci case, the problem of determining the equilibrium point reduces to finding the unique zero in the positive orthant of a set of L equations in L variables, and there is an extensive body of theory in numerical analysis and computational commutative algebra devoted to this problem.

We finish this introduction with a very brief indication of the substantial literature on multilocus deterministic models in population genetics and the particular import of such models for the phenomenon of mutation-selection balance.

The standard mathematical reference on population genetics models in general (both deterministic and stochastic) is [Ewe04] by Warren Ewens. This is a revision of a book from 1979 and the treatment of deterministic models is not expanded significantly from the original. In [ET77], Ewens and Glenys Thomson also made one of the earliest contributions to understanding equilibria of deterministic multilocus genetic systems in which no special assumptions are made about fitness structure. While mutation only occurs from the wild type to the derived type in our model, the models in [ET77] (in common with many other models in the literature) allow several alleles at a locus and mutation from any allelic type to any other.

A very comprehensive recent reference that concentrates on the infinite-population, deterministic aspects of population genetics is Reinhard Bürger's book [Bür00] (see also the Bürger's review paper [Bür98]). In particular, these works consider at length deterministic haploid *continuum-of-alleles* models in which individuals have a type which is thought of as the contribution of a gene to a given quantitative trait. The type belongs to a general state space that represents something like the trait value (in which case the state space is a subset of $\mathbb{R}$) and types have an associated fitness, which is some fairly arbitrary function from the type space to $(0, 1]$. The type may be regarded as the combined effect of a multilocus genotype, but the models do not explicitly incorporate a family of loci, the configuration of alleles present at those loci, or a function describing the fitness of a configuration: everything is cast in terms of how fit each type is and how likely one type is to mutate into another.

The development in [ESW06] leading to the model we consider here was very much influenced by the discrete time multilocus models of [TB90, BT91, KJB02] that incorporate arbitrary forms of selection, general modes of inheritance, and other evolutionary forces such as migration and mutation. Also, the model studied here and in [ESW06] is essentially the result of adding strong recombination to the model of [SEW05], which may itself be thought of as a generalization of the *infinitesimally-rare-alleles* model of [KM66] (see also, [Kon82]) as it is cast in [Daw99].

Certain classes of mutation-selection models without recombination are solved explicitly in [WBG98, BW01] using ideas from statistical mechanics. Such models may be thought of as either multilocus systems with complete linkage or structured single locus systems. Finally, we have already mentioned the constellation of papers [BB03, Baa05, Baa07, Baa01] presenting a deterministic model of population change due to recombination alone.

2. The Model

In this section, we review the details of the model in ([ESW06]) and derive the relevant equations in the special case of polynomial selection costs, to be defined below.

Let $\mathcal{M}$, $\mathcal{G}$, ν, and S be as above. That is, $\mathcal{M}$ is the complete, separable metric space of loci, $\mathcal{G}$ is the space of integer-valued finite measures on $\mathcal{M}$ representing possible genotypes, ν is the finite measure on $\mathcal{M}$ describing the rates at which mutation events occur in different parts of the genome, and $S : \mathcal{G} \to \mathbb{R}_+$ is the selection cost function. As above, we assume that $S(0) = 0$ and $S(g + h) \geq S(g)$ for $g, h \in \mathcal{G}$.

For any finite measure π on $\mathcal{M}$, let X^π denote a Poisson random measure on $\mathcal{M}$ with intensity π (so that X^π is a random variable with values in $\mathcal{G}$). Define a function $F_\pi : \mathcal{M} \to \mathbb{R}_+$ by

$$F_\pi(m) = \mathbb{E}[S(X^\pi + \delta_m) - S(X^\pi)] \tag{2.1}$$

That is, $F_\pi(m)$ measures the average change in selection cost when mutation at locus $m \in \mathcal{M}$ is added to the random genotype X^π.

Let $D\pi$ be the measure on $\mathcal{M}$, absolutely continuous with respect to π, with density (that is, Radon-Nikodym derivative) F_π:

$$\frac{d(D\pi)}{d\pi}(m) = F_\pi(m). \tag{2.2}$$

In this notation, equation (1.2) becomes

$$\rho_t = \rho_0 + t\nu - \int_0^t D\rho_s \, ds, \tag{2.3}$$

where, as above, ρ_t is the finite measure on $\mathcal{M}$ giving the intensity measure of the Poisson random measure with distribution P_t. We may write (2.3) somewhat informally as

$$\rho_t(dm) = \rho_0(dm) + t\nu(dm) - \int_0^t F_{\rho_s}(m)\rho_s(dm)\, ds.$$

It was shown in [ESW06] that (2.3) has a unique solution for any initial condition ρ_0 provided that the selection cost S satisfies a Lipschitz condition with respect to the Wasserstein metric on $\mathcal{G}$. We take a slightly different perspective here. Let $L^\infty(\mathcal{M}, \nu)$ denote the usual Banach space of (equivalence classes of) ν-essentially bounded functions on $\mathcal{M}$, and write $L^\infty_+(\mathcal{M}, \nu)$ for the subset of $L^\infty(\mathcal{M}, \nu)$ consisting of non-negative functions. There is, of course, a bijection between $L^\infty_+(\mathcal{M}, \nu)$ and the space $\mathcal{K}$ of finite measures on $\mathcal{M}$ that are absolutely continuous with respect to ν with ν-essentially bounded Radon-Nikodym derivative, and so we can metrize $\mathcal{K}$ using the $L^\infty(\mathcal{M}, \nu)$ metric on $L^\infty_+(\mathcal{M}, \nu)$. An argument essentially the same as that in [ESW06] establishes the following.

Lemma 2.1. *Suppose that S is such that $F_\pi \in L^\infty_+(M, \nu)$ for $\pi \in \mathcal{K}$ and there exists a function $H : \mathbb{R}_+ \to \mathbb{R}_+$ for which*

$$\left\| \frac{d(D\pi')}{d\nu} - \frac{d(D\pi'')}{d\nu} \right\|_\infty = \|F_{\pi'}\phi' - F_{\pi''}\phi''\|_\infty \le H(\|\phi'\|_\infty \vee \|\phi''\|_\infty)\|\phi' - \phi''\|_\infty$$

when $\pi', \pi'' \in \mathcal{K}$ with $\pi'(dm) = \phi'(m)\,\nu(dm)$ and $\pi''(dm) = \phi''(m)\,\nu(dm)$. Then (2.3) has a unique $\mathcal{K}$-valued solution for any $\rho_0 \in \mathcal{K}$.

From now on we will assume that the selection cost S is *polynomial*, by which we mean that

$$S(g) = \sum_{n=1}^{N} \int_{\mathcal{M}^n} a_n(\mathbf{m})\, g^{\otimes n}(d\mathbf{m}) \tag{2.4}$$

for some positive integer N, where for each n the Borel function $a_n : \mathcal{M}^n \to \mathbb{R}_+$ is permutation-invariant (that is, $a_n(\pi\mathbf{m}) = a_n(\mathbf{m})$ for all permutations π) and has the property that $a_n(\mathbf{m}) = 0$ if there exist $i \ne j$ with $m_i = m_j$. Furthermore, we assume that each function a_n is bounded. The number $n!\, a_n(m_1, \ldots, m_n)$ represents the interactive effect of possessing mutations at the n loci $m_1, \ldots, m_n$

over and above that of possessing any subset of them, and this additional effect is independent of the order in which the mutations are written.

Remark 2.2. When $\mathcal{M}$ is the finite set $\{1, ..., L\}$ and, as above, we encode a genotype $g \in \mathcal{G}$ as an ordered $L-$tuple of non-negative integers $g = (g_1, \dots, g_L)$, where g_k represents the number of times an ancestral mutation has occurred at locus k, then the expression for $S(g)$ in (2.4) coincides with that in (1.3) if we set $\alpha_I = n!\, a_n(\mathbf{m})$ when I is the subset $\{m_1, \dots, m_n\} \subseteq \{1, \dots, L\} = \mathcal{M}$ and $\mathbf{m}$ is the vector $(m_1, \dots, m_n)$.

We will now derive the particular form that the equation (2.3) takes for polynomial selections costs. Fix $\pi \in \mathcal{K}$ and, as above, let X^π be a Poisson random measure with intensity π. Then, for each $m \in \mathcal{M}$, we have

$$\begin{aligned} S(X^\pi + \delta_m) &= \sum_{n=1}^{N} \int_{\mathcal{M}^n} a_n(\mathbf{m})(X^\pi + \delta_m)^{\otimes n}(d\mathbf{m}) \\ &= \sum_{n=1}^{N} \int_{\mathcal{M}^n} a_n(\mathbf{m}) \left(\sum_{k=0}^{n} \binom{n}{k} (X^\pi)^{\otimes k} \otimes \delta_m^{\otimes(n-k)} \right) (d\mathbf{m}). \end{aligned}$$

Note that the symmetry of a_n allows us to rearrange the order of variables in the integral. Thus,

$$\begin{aligned} S(X^\pi + \delta_m) &= \sum_{n=1}^{N} \sum_{k=0}^{n} \binom{n}{k} \int_{\mathcal{M}^n} a_n(\mathbf{m}) \, (X^\pi)^{\otimes k} \otimes \delta_m^{\otimes(n-k)}(d\mathbf{m}) \\ &= \sum_{n=1}^{N} \left[\int_{\mathcal{M}^n} a_n(\mathbf{m})(X^\pi)^{\otimes n}(d\mathbf{m}) + n \int_{\mathcal{M}^n} a_n(\mathbf{m}) \, (X^\pi)^{\otimes(n-1)} \otimes \delta_m(d\mathbf{m}) \right], \end{aligned}$$

because all other integrals are 0 by the assumptions on a_n. Hence,

$$S(X^\pi + \delta_m) = S(X^\pi) + \sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) \, (X^\pi)^{\otimes(n-1)}(d\mathbf{m}), \tag{2.5}$$

where we write $a_n(\mathbf{m}, m) := a_n(m_1, \dots, m_{n-1}, m)$ for $\mathbf{m} = (m_1, \dots, m_{n-1}) \in \mathcal{M}^{(n-1)}$ and $m \in \mathcal{M}$, and by convention, $\int_{\mathcal{M}^0} a_1(\mathbf{m}, m) \, (X^\pi)^{\otimes(0)}(d\mathbf{m}) = a_1(m)$. Therefore,

$$F_\pi(m) = \mathbb{E}\left[S(X^\pi + \delta_m) - S(X^\pi)\right] = \sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) \, \pi^{\otimes(n-1)}(d\mathbf{m}), \tag{2.6}$$

by the independent increments property of Poisson random measures.

Remark 2.3. Suppose that $\mathcal{M}$ is the finite set $\{1, ..., L\}$ and, as above, we encode a genotype $g \in \mathcal{G}$ as an ordered $L-$tuple of non-negative integers $g = (g_1, \dots, g_L)$ and set $\alpha_I = n!\, a_n(\mathbf{m})$ when I is the subset $\{m_1, \dots, m_n\} \subseteq \{1, \dots, L\} = \mathcal{M}$ and $\mathbf{m}$ is the vector $(m_1, \dots, m_n)$. Then

$$F_\pi(k) = \sum_{I \in \mathcal{I}_k} \alpha_I \pi^I, \quad k \in \mathcal{M},$$

where we identify the measure π on $\mathcal{M}$ with the vector $(\pi_1, \dots, \pi_L)$ and we use the notations of the Introduction that $\pi^I = \prod_{i \in I} \pi_i$ and $\mathcal{I}_k$ is the collection of subsets of $\mathcal{M}$ that contain k. Consequently, equation (2.3) becomes equation (1.4) in this special case.

Returning to the general case, it is straightforward to check that the condition of Lemma 2.1 holds for our choice of the cost function S, and so (2.3) has a unique $\mathcal{K}$-valued solution for any initial condition $\rho_0 \in \mathcal{K}$. We will henceforth suppose that $\rho_0 \in \mathcal{K}$.

Write ϕ_t for a choice of the Radon-Nikodym derivative of ρ_t with respect to ν. It will be convenient to turn our measure-valued integral equation (2.3) for $t \mapsto \rho_t$ into a function-valued differential equation for the corresponding Radon-Nikodym derivatives with respect to ν as follows.

For $x \in L^\infty_+(\mathcal{M}, \nu)$, define $Gx \in L^\infty_+(\mathcal{M}, \nu)$ by

$$Gx(m) := \left[\sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) x(m_1) \cdots x(m_{n-1}) \, \nu^{\otimes(n-1)}(d\mathbf{m}) \right] x(m).$$

Lemma 2.4. *Suppose that ρ_t is the solution of (2.3). Then there is a non-negative Borel function $(t, m) \mapsto x_t(m)$ on $\mathbb{R}_+ \times \mathcal{M}$ such that the function $m \mapsto x_t(m)$ is a Radon-Nikodym derivative of ρ_t with respect to ν for all $t \geq 0$, and for ν-a.e. $m \in \mathcal{M}$ the function $t \mapsto x_t(m)$ is differentiable with*

$$\dot{x}_t(m) = 1 - Gx_t(m), \tag{2.7}$$

Proof. For each $t \geq 0$, let ϕ_t be a Radon-Nikodym derivative of ρ_t with respect to ν. We may write $\mathcal{B}(\mathcal{M})$, the Borel σ-field on $\mathcal{M}$, as $\mathcal{B}(\mathcal{M}) = \sigma(\bigcup_k \mathcal{F}_k)$, where $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots$ is a sequence of finitely generated sub-σ-fields. It is therefore clear from a concrete description of the Radon-Nikodym derivative in terms of a limit of $\mathcal{F}_k$-measurable ratios (see, for example, Section III-1 of [Nev75]) that we may suppose that the map $(t, m) \to \phi_t(m)$ is Borel measurable.

By Fubini's theorem,

$$\int_A \phi_t(m) \, \nu(dm) = \int_A \phi_0(m) \, \nu(dm) + t \int_A \nu(dm) - \int_A \left[\int_0^t F_{\rho_s}(m) \phi_s(m) \, ds \right] \nu(dm)$$

for any Borel set $A \subseteq \mathcal{M}$. Thus, for each $t \geq 0$ we have for ν-a.e. $m \in \mathcal{M}$ that

$$\phi_t(m) = \phi_0(m) + t - \int_0^t F_{\rho_s}(m) \phi_s(m) \, ds =: \psi_t(m), \tag{2.8}$$

by the uniqueness of the Radon-Nikodym derivative. A fortiori, (2.8) holds for $\lambda \otimes \nu$-a.e. pair $(t, m) \in \mathbb{R}_+ \times \mathcal{M}$ (recall that λ is Lebesgue measure on $\mathbb{R}_+$). Thus, by Fubini's theorem, it follows that, for ν-a.e. $m \in \mathcal{M}$, the equation (2.8) holds for λ-a.e. $t \geq 0$. Also, ρ_t is the measure $\breve{\psi}_t(dm) := \psi_t(m) \, \nu(dm)$ for all $t \geq 0$, and hence $F_{\rho_t}(m) = F_{\breve{\psi}_t}(m)$ for all $t \geq 0$. Thus, for ν-a.e. $m \in \mathcal{M}$ and for all $t \geq 0$,

$$\int_0^t F_{\breve{\psi}_s}(m) \psi_s(m) \, ds = \int_0^t F_{\rho_s}(m) \phi_s(m) \, ds.$$

By definition, $\psi_0 = \phi_0$, and so, for ν-a.e. $m \in \mathcal{M}$ and for all $t \geq 0$,

$$\psi_t(m) = \psi_0(m) + t - \int_0^t F_{\breve{\psi}_s}(m) \psi_s(m) \, ds.$$

Therefore, for ν-a.e. $m \in \mathcal{M}$ the function $t \mapsto \psi_t(m)$ is differentiable with

$$\dot{\psi}_t(m) = 1 - F_{\breve{\psi}_t}(m) \psi_t(m) = 1 - G\psi_t(m),$$

and setting $x_t = \psi_t$ gives the result. □

3. Equilibria and Stability

We can now state our main theorem.

Theorem 3.1. *Suppose that $E := \inf\{a_1(m) \ : \ m \in \mathcal{M}\} > 0$. Then the system (2.7) has a unique equilibrium point in the space $L^\infty_+(\mathcal{M}, \nu)$. Furthermore, this equilibrium is globally stable, in the sense that all solutions to (2.7) converge to it in the $L^\infty(\mathcal{M}, \nu)$ norm as t goes to infinity.*

The proof will be accomplished in four parts. First, we derive some upper and lower bounds on solutions to (2.7). Then, we establish existence of at least one equilibrium point by a fixed point argument and show uniqueness using a generalization of a method of Craciun and Feinberg [CF05]. Finally, we obtain global stability of the equilibrium via the construction of a Lyapunov function. Since we are dealing with an infinite dimensional dynamical system, all of these steps rely on some technical lemmas concerning the continuity of certain operators on $L^\infty_+(\mathcal{M}, \nu)$, and we defer these to Section 4.

3.1. **Estimates on the solution.** Let x_t be a solution to the equation (2.7). Since $Gx_t(m) \geq a_1(m)x_t(m)$, we see for ν-a.e. $m \in \mathcal{M}$ and all $t \geq 0$ that $\dot{x}_t(m) \leq 1 - a_1(m)x_t(m)$. Hence,

$$x_t(m) \leq \frac{1 - (1 - a_1(m)x_0(m))e^{-a_1(m)t}}{a_1(m)} \tag{3.1}$$

for ν-a.e. $m \in \mathcal{M}$ and all $t \geq 0$. As $t \to \infty$ the right-hand side of (3.1) either increases to the value $1/a_1(m)$, in the case that $x_0(m) < 1/a_1(m)$, or otherwise decreases to the same value. In either case, we have for $\nu-$a.e. $m \in \mathcal{M}$ that $x_t(m) \leq \max\{x_0(m), 1/a_1(m)\}$ for all $t \geq 0$. Thus,

$$\sup_t \|x_t\|_\infty \leq K, \tag{3.2}$$

where $K = \|x_0\|_\infty \vee A$ with $A := E^{-1}$.

Similarly, since for all $t \geq 0$ we have $x_t(m) \leq K$ for ν-a.e. $m \in \mathcal{M}$, it follows that

$$Gx_t(m) \leq \left[\sum_{n=1}^N n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) K^{(n-1)} \nu^{\otimes(n-1)}(d\mathbf{m})\right] x_t(m) \leq C x_t(m), \tag{3.3}$$

for some constant C. Therefore,

$$\begin{aligned} x_t(m) &\geq \frac{1 - (1 - Cx_0(m))e^{-Ct}}{C} \\ &\geq \frac{1 - e^{-Ct}}{C} + x_0(m)e^{-Ct} \geq \frac{1 - e^{-Ct}}{C}. \end{aligned}$$

Hence,

$$\liminf_{t\to\infty} \operatorname{ess\,inf}_m x_t(m) \geq 1/C > 0. \tag{3.4}$$

3.2. **Existence of equilibria.** Observe that $x \in L^\infty_+(\mathcal{M}, \nu)$ is an equilibrium point of (2.7) if and only if $Gx(m) = 1$ for ν-a.e. $m \in \mathcal{M}$, that is, if and only if x is a fixed point of the map $\Gamma : L^\infty_+(\mathcal{M}, \nu) \to L^\infty_+(\mathcal{M}, \nu)$ given by

$$\Gamma y(m) := \left[\sum_{n=1}^N n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) y(m_1) \cdots y(m_{n-1}) \, \nu^{\otimes(n-1)}(d\mathbf{m})\right]^{-1}. \tag{3.5}$$

Note for any $y \in L^\infty_+(\mathcal{M},\nu)$ that

$$a_1(m) \le \sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m},m) y(m_1)\cdots y(m_{n-1})\, \nu^{\otimes(n-1)}(d\mathbf{m}).$$

Hence,

$$\Gamma y(m) \le A, \quad \text{for } \nu\text{-a.e. } m \in \mathcal{M}. \tag{3.6}$$

Thus Γ does indeed map $L^\infty_+(\mathcal{M},\nu)$ into itself. Moreover, if $y(m) \le A$ for ν-a.e. $m \in \mathcal{M}$, then

$$\begin{aligned}\sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} & a_n(\mathbf{m},m) y(m_1)\cdots y(m_{n-1})\, \nu^{\otimes(n-1)}(d\mathbf{m}) \\ &\le \sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m},m) A^{(n-1)}\, \nu^{\otimes(n-1)}(d\mathbf{m}) =: B^{-1},\end{aligned}$$

so that

$$B \le \Gamma y(m), \quad \text{for } \nu\text{-a.e. } m \in \mathcal{M}. \tag{3.7}$$

Therefore, Γ maps the convex set

$$R := \{ y \in L^\infty_+(\mathcal{M},\nu) \ : \ B \le y(m) \le A \} \tag{3.8}$$

into itself. Recall that the Banach space $L^\infty(\mathcal{M},\nu)$ is the dual of the Banach space $L^1(\mathcal{M},\nu)$ under the pairing $\langle x, f\rangle := \int_{\mathcal{M}} x(m) f(m)\, \nu(dm)$ for $x \in L^\infty(\mathcal{M},\nu)$ and $f \in L^1(\mathcal{M},\nu)$. Recall also that the weak* topology on $L^\infty(\mathcal{M},\nu)$ is the weakest topology that makes each of the maps $x \mapsto \langle x, f\rangle$, $f \in L^1(\mathcal{M},\nu)$, continuous. A consequence of the Banach-Alaoglu Theorem is that that any weak*-closed and norm-bounded subset of $L^\infty(\mathcal{M},\nu)$ is weak*-compact (see Corollary 3 of Section V.4.2 of [DS88]), and so the set R is weak*-compact. The map Γ is weak*-continuous on the convex weak*-compact set R by Lemma 4.2 below. An infinite-dimensional extension of the Brouwer Fixed Point Theorem, the Schauder-Tychonoff Theorem (see Theorem 5 of Section V.10.5 in [DS88]), guarantees that Γ has a fixed point in R. Consequently, the system (2.7) has at least one equilibrium point.

Remark 3.2. Note from the argument above that if $x \in L^\infty_+(\mathcal{M},\nu)$ is any equilibrium point of (2.7), then $x(m) = \Gamma x(m) \le A$ for ν-a.e. $m \in \mathcal{M}$ by (3.6), and so $x \in R$ by (3.7).

3.3. **Uniqueness of the equilibrium.** We claim that there is only one solution in $L^\infty_+(\mathcal{M},\nu)$ to the equilibrium equation $Gx(m) = 1$, ν-a.e. $m \in \mathcal{M}$. Our argument follows that used in Theorem 3.1 of [CF05] to establish a criterion for the uniqueness of equilibria in the finite dimensional *mass-action kinetics* systems of differential equations that arise in the study of *continuous flow stirred tank reactors*. We refer the reader to [CF05] for a discussion of the history of related results. A recent paper on the use of algebraic geometry in analyzing instances of such systems is [CDSS07], which also contains several references to other applications of this general class of dynamical systems.

By Remark 3.2, it suffices to show that there is only one solution in the set R. Assume otherwise, that is, that there are two solutions $x \in R$ and $y \in R$ with

$x \neq y$. In particular, we have that $Gx(m) = Gy(m)$ for ν-a.e. $m \in \mathcal{M}$, and so

$$\sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) x(m_1) \cdots x(m_{n-1})\, x(m)\, \nu^{\otimes(n-1)}(d\mathbf{m})$$
$$= \sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) y(m_1) \cdots y(m_{n-1})\, y(m)\, \nu^{\otimes(n-1)}(d\mathbf{m}).$$

Therefore,

$$\sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m}, m) x(m_1) \cdots x(m_{n-1}) x(m)$$
$$\times \left(\frac{y(m_1) \cdots y(m_{n-1}) y(m)}{x(m_1) \cdots x(m_{n-1}) x(m)} - 1 \right) \nu^{\otimes(n-1)}(d\mathbf{m}) = 0$$

for ν-a.e. $m \in \mathcal{M}$. Setting $K_n(\mathbf{m}, m) := a_n(\mathbf{m}, m) x(m_1) \cdots x(m_{n-1}) x(m)$ and $\delta(m) := \log(y(m)/x(m))$, we obtain

$$\sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} K_n(\mathbf{m}, m) \left(e^{\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m)} - 1 \right) \nu^{\otimes(n-1)}(d\mathbf{m}) = 0.$$

Observe that $\delta(m)$ is bounded since x and y are in R. Thus, putting

$$\eta_n(\mathbf{m}, m) := K_n(\mathbf{m}, m) \frac{e^{\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m)} - 1}{\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m)},$$

we get

$$\sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} \eta_n(\mathbf{m}, m) \left(\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m) \right) \nu^{\otimes(n-1)}(d\mathbf{m}) = 0 \tag{3.9}$$

for ν-a.e. $m \in \mathcal{M}$. Note that the function η_n is non-negative, since $\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m)$ and $e^{\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m)} - 1$ have the same sign. Also, η_n is permutation invariant and takes the value 0 whenever two of the coordinates of its argument are equal.

Integrating the left-hand side of (3.9) against the function δ gives

$$0 = \int_{\mathcal{M}} \sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} \eta_n(\mathbf{m}, m)$$
$$\times \left(\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m) \right) \nu^{\otimes(n-1)}(d\mathbf{m}) \delta(m)\, \nu(dm)$$
$$= \sum_{n=1}^{N} n \int_{\mathcal{M}^n} \eta_n(m_1, \ldots, m_n) (\delta(m_1) + \cdots + \delta(m_{n-1}) + \delta(m_n))$$
$$\times \delta(m_n) \nu^{\otimes n}(dm_1, \ldots, dm_n)$$
$$= \sum_{n=1}^{N} \sum_{k=1}^{n} \int_{\mathcal{M}^n} \eta_n(m_1, \ldots, m_n) (\delta(m_1) + \cdots + \delta(m_n))$$
$$\times \delta(m_k)\, \nu^{\otimes n}(dm_1, \ldots, dm_n),$$

by the symmetry of η_n together with the change of variables

$$(m_1, \ldots, m_k, \ldots, m_n) \leftrightarrow (m_1, \ldots, m_n, \ldots, m_k)$$

once for each k.

Therefore,

$$0 = \sum_{n=1}^{N} \int_{\mathcal{M}^n} \eta_n(m_1, \ldots, m_n) \left[\delta(m_1) + \cdots + \delta(m_n)\right]^2 \nu^{\otimes n}(dm_1, \ldots, dm_n).$$

In particular, $0 = \int_{\mathcal{M}} \eta_1(m)\delta(m)^2 \, \nu(dm)$. Since $a_1(m) \geq E > 0$ for all $m \in \mathcal{M}$, it follows that $\eta_1(m) > 0$ for all $m \in \mathcal{M}$, and hence $\delta(m) = 0$ for ν-a.e. $m \in \mathcal{M}$. This contradicts our assumption that $x \neq y$, and so the equilibrium point is unique.

3.4. **Stability.** Let x_t be a solution of (2.7). Write $x^* \in R \subset L^\infty_+(\mathcal{M}, \nu)$ for the unique equilibrium point guaranteed by Subsection 3.2 and Subsection 3.3. We will show that $\lim_{t\to\infty} \|x_t - x^*\|_\infty = 0$ using a variant of *Lyapunov's second method*.

A recent review of Lyapunov function methods for studying the asymptotic behavior of finite dimensional nonlinear systems is [LR04]. Much of the finite dimensional theory considers Lyapunov functions that are continuously differentiable. General stability results for finite dimensional systems that do not assume this degree of smoothness are discussed in [CLH99], where there is a number of references indicating why weaker assumptions are natural in several applications. Infinite dimensional systems are considered in [Daf78], primarily in the context of partial differential equations. Discontinuous Lyapunov functionals appear there because it is natural to work with a weak* topology in order to make norm-bounded sets compact, but functionals that are continuous in the norm topology may cease to be continuous in the weak* topology. This is exactly the situation that confronts us. Because we haven't found a result in the literature that applies directly to establish global stability in our setting, we provide the details of the relatively routine argument.

Define the function $V : L^\infty_+(\mathcal{M}, \nu) \to \mathbb{R} \cup \{+\infty\}$ by

$$V(x) := -\int_{\mathcal{M}} \log(x(m)) \, \nu(dm) + \sum_{n=1}^{N} \int_{\mathcal{M}^n} a_n(\mathbf{m}) x(m_1) \cdots x(m_n) \, \nu^{\otimes n}(d\mathbf{m}).$$

Observe that V is, indeed, $\mathbb{R} \cup \{+\infty\}$-valued and even bounded below, because

$$\begin{aligned} V(x) &= -\int_{\mathcal{M}} \log(x(m)) \, \nu(dm) + \int_{\mathcal{M}} a_1(m) x(m) \, \nu(dm) \\ &\quad + \sum_{n=2}^{N} \int_{\mathcal{M}^n} a_n(\mathbf{m}) x(m_1) \cdots x(m_n) \, \nu^{\otimes n}(d\mathbf{m}) \\ &\geq \int_{\mathcal{M}} \left[a_1(m)x(m) - \log(x(m))\right] \, \nu(dm) \\ &\geq \int_{\mathcal{M}} \left[1 + \log(a_1(m))\right] \, \nu(dm) \\ &\geq (1 + \log(E)) \times \nu(\mathcal{M}) > -\infty, \end{aligned} \tag{3.10}$$

since $au - \log(u) \geq 1 + \log(a)$ for all $u > 0$.

Set $V_t := V(x(t))$. We have

$$\begin{aligned}\frac{d}{dt}V_t &= \lim_{h\to 0}\frac{1}{h}\left[-\int_{\mathcal{M}}\log(x_{t+h}(m))\,\nu(dm)+\int_{\mathcal{M}}\log(x_t(m))\,\nu(dm)\right]\\ &\quad+\lim_{h\to 0}\frac{1}{h}\left[\int_{\mathcal{M}}a_1(m)x_{t+h}(m)\,\nu(dm)-\int_{\mathcal{M}}a_1(m)x_t(m)\,\nu(dm)\right]\\ &\quad+\lim_{h\to 0}\frac{1}{h}\sum_{n=2}^{N}\int_{\mathcal{M}^n}a_n(\mathbf{m})\\ &\qquad\times\left[x_{t+h}(m_1)\cdots x_{t+h}(m_n)-x_t(m_1)\cdots x_t(m_n)\right]\,\nu^{\otimes n}(d\mathbf{m}).\end{aligned}\tag{3.11}$$

From (3.2) and (3.4) there exist constants $C>0$, $K>0$, and $T\geq 0$ such that

$$\frac{1}{2C}\leq \operatorname*{ess\,inf}_m x_t(m)\leq \operatorname*{ess\,sup}_m x_t(m)\leq K,\quad t\geq T.\tag{3.12}$$

Hence, by the Dominated Convergence Theorem, we may interchange the limit and integrals in (3.11) to get

$$\begin{aligned}\frac{d}{dt}V_t &= -\int_{\mathcal{M}}\frac{\dot{x}_t(m)}{x_t(m)}\,\nu(dm)+\int_{\mathcal{M}}a_1(m)\dot{x}_t(m)\,\nu(dm)\\ &\quad+\sum_{n=2}^{N}n\int_{\mathcal{M}^n}a_n(\mathbf{m})x_t(m_1)\cdots x_t(m_{n-1})\dot{x}_t(m_n)\,\nu^{\otimes n}(d\mathbf{m})\\ &= \int_{\mathcal{M}}(-\dot{x}_t(m))\Bigg[\frac{1}{x_t(m)}-a_1(m)\\ &\quad-\sum_{n=2}^{N}n\int_{\mathcal{M}^{(n-1)}}a_n(\mathbf{m})x_t(m_1)\cdots x_t(m_{n-1})\,\nu^{\otimes(n-1)}(d\mathbf{m})\Bigg]\,\nu(dm)\\ &= -\int_{\mathcal{M}}\dot{x}_t(m)\frac{1}{x_t(m)}\dot{x}_t(m)\,\nu(dm)\\ &\leq 0\end{aligned}\tag{3.13}$$

for all $t\geq T$.

Define $\dot{V}: L^\infty_+(\mathcal{M},\nu)\to\mathbb{R}_-\cup\{-\infty\}$ by

$$\dot{V}(x) := -\int_{\mathcal{M}}(1-Gx(m))^2\frac{1}{x(m)}\,\nu(dm).\tag{3.14}$$

From (3.13) and (2.7), we obtain $\frac{d}{dt}V_t=\dot{V}(x_t)$ for $t\geq T$.

By Lemma 4.3, the function $\dot{V}$ is upper semicontinuous in the weak* topology on the domain $\{x\in L^\infty_+(\mathcal{M},\nu):\frac{1}{2C}\leq \operatorname{ess\,inf}_m x(m)\leq \operatorname{ess\,sup}_m x(m)\leq K\}$. Also, if x is in this domain, then $\dot{V}(x)=0$ if and only if $Gx(m)=1$ for $\nu-$a.e. $m\in\mathcal{M}$, that is, if and only if $x=x^*$. It follows that for any $\epsilon>0$ the supremum of $\dot{V}$ on the weak*-compact set

$$\{x\in L^\infty_+(\mathcal{M},\nu):\frac{1}{2C}\leq \operatorname*{ess\,inf}_m x(m)\leq \operatorname*{ess\,sup}_m x(m)\leq K,\ \|x-x^*\|_\infty\geq\epsilon\}$$

is strictly less than 0.

Assume that $\|x_t-x^*\|_\infty$ does not converge to 0 as $t\to\infty$. Then there exists $\delta>0$ and a sequence $t_k\to\infty$ with $t_k\geq T$ for all k such that $\|x_{t_k}-x^*\|_\infty>\delta$ for all k. Now, since $\dot{x}_t$ is uniformly bounded as a consequence of (3.2), (3.3), and (3.4), there is a number $\gamma>0$ independent of k such that $\|x_t-x_{t_k}\|_\infty<\delta/2$ for

$t_k \le t \le t_k + \gamma$, so that $\|x_t - x^*\|_\infty > \delta/2$ for $t_k \le t \le t_k + \gamma$. It follows that there is a constant $\lambda < 0$ independent of k such that

$$\frac{d}{dt} V_t = \dot{V}(x_t) \le \lambda < 0 \tag{3.15}$$

for $t_k \le t \le t_k + \gamma$. Because V_t is nondecreasing for $t \ge T$ by (3.13), the inequality (3.15) contradicts the conclusion from (3.10) that $\inf_t V_t > -\infty$. Thus $\|x_t - x^*\|_\infty$ must converge to 0 as $t \to \infty$.

Remark 3.3. Put $\bar{\nu} := c^{-1}\nu$, where $c := \nu(\mathcal{M})$, and $\dot{y}_t := \frac{d\rho_{ct}}{d\bar{\nu}} = c x_{c^{-1}t}$. Then, $\dot{y}_t(m) = \dot{x}_{c^{-1}t}(m) = 1 - \bar{G} y_t(m)$, where $\bar{G}$ is defined in the same way as G, except that ν is replaced by $\bar{\nu}$. Therefore, by a change of time units, we may suppose without loss of generality that ν is a probability measure. In that case, the Lyapunov function V may be written as

$$V(x) = D(\nu \,\|\, \bar{x} \cdot \nu) - \log\left(\int x(m)\, \nu(dm) \right) + \mathbb{E}\left[S(X^{x \cdot \nu}) \right],$$

where $D(\cdot \,\|\, \cdot)$ denotes the relative entropy or Kullback-Leibler distance between two probability measures, $\bar{x} := (\int x(m)\, \nu(dm))^{-1} x$ is x renormalized to be a probability density with respect to the measure ν, $\bar{x} \cdot \nu$ is the probability measure with density $\bar{x}$ with respect to ν, $x \cdot \nu$ is the finite measure with density x with respect to ν, and $X^{x \cdot \nu}$ is a Poisson random measure with intensity $x \cdot \nu$. Therefore, our system (2.3) evolves in such a way that it tries to decrease the sum of the expected selection cost of an individual chosen at random from the population, $\mathbb{E}[S(X^{\rho_t})]$, and the Kullback-Leibler distance between the probability measures ν and $\rho_t(\mathcal{M})^{-1}\rho_t$, while not allowing the expected total number of ancestral mutation events present in the of a randomly chosen individual, $\rho_t(\mathcal{M}) = \mathbb{E}[X^{\rho_t}(\mathcal{M})]$, to be too small.

We note that the sum of a Kullback-Leibler distance and the logarithm of expected selection cost appears as a Lyapunov function for the continuous-time, single-locus, n-allele, mutation-selection model of [Hof85] in which the mutation rate from allele i to allele j only depends on the target allele j. This Lyapunov function is used there to prove a global stability result. Also, a Kullback-Leibler distance is a Lyapunov function in a neighborhood of the stable state of the frequency-dependent selection model of [Bom91] that builds on the evolutionary game theory work of [Aki82], where a similar quantity also appears.

4. Some technical lemmas

We collect together in this section some results used in the proof of Theorem 3.1. The following result is elementary, but we include it for completeness.

Lemma 4.1. *Suppose that $f \in L^1(\mathcal{M}^n, \nu^{\otimes n})$ for some positive integer n. The function*

$$x \mapsto \int_{\mathcal{M}^n} x(m_1) \cdots x(m_n) f(m_1, \ldots, m_n)\, \nu^{\otimes n}(d\mathbf{m})$$

is weak-continuous.*

Proof. This is obvious for functions of the form

$$f(m_1, \ldots, m_n) = f_1(m_1) \ldots f_n(m_n),$$

where $f_1, \ldots, f_n \in L^1(\mathcal{M}, \nu)$, and the result for general f follows from the fact that finite linear combinations of such functions are dense in $L^1(\mathcal{M}^n, \nu^{\otimes n})$. □

Recall the definition of the map $\Gamma : L^\infty_+(\mathcal{M},\nu) \to L^\infty_+(\mathcal{M},\nu)$ and the set $R \subset L^\infty_+(\mathcal{M},\nu)$ from (3.5) and (3.8), respectively. Recall also that Γ maps R into itself.

Lemma 4.2. *The map $\Gamma : R \to R$ is weak*-continuous.*

Proof. Suppose that x_k is a sequence in R such that x_k converges to $x \in R$ in the weak* topology as $k \to \infty$. Fix a test function $f \in L^1(\mathcal{M},\nu)$. Then

$$\begin{aligned}
&\left|\int_{\mathcal{M}} \Gamma x_k(m) f(m)\,\nu(dm) - \int_{\mathcal{M}} \Gamma x(m) f(m)\,\nu(dm)\right| \\
&\quad = \Big|\sum_{n=2}^{N} n \int_{\mathcal{M}} f(m) \frac{1}{H(m)} \\
&\qquad \times \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m},m)\left[x_k(m_1)\cdots x_k(m_{n-1}) - x(m_1)\cdots x(m_{n-1})\right] \nu^{\otimes(n-1)}(d\mathbf{m}) \\
&\qquad \times \nu(dm)\Big|,
\end{aligned}$$

where

$$\begin{aligned}
H(m) &:= \left(\sum_{n=1}^{N} n \int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m},m) x_k(m_1)\ldots x_k(m_{n-1})\,\nu^{\otimes(n-1)}(d\mathbf{m})\right) \\
&\qquad \times \left(\sum_{n=1}^{N} n \int_{\mathcal{M}_m^n} a_n(\mathbf{m}) x(m_1)\ldots x(m_{n-1})\,\nu^{\otimes(n-1)}(d\mathbf{m})\right) \geq E^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\left|\int_{\mathcal{M}} \Gamma x_k(m) f(m)\,\nu(dm) - \int_{\mathcal{M}} \Gamma x(m) f(m)\,\nu(dm)\right| \\
&\quad \leq \frac{1}{E^2}\sum_{n=2}^{N} \int_{\mathcal{M}} |f(m)| \\
&\qquad \times \left|\int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m},m)\left[x_k(m_1)\cdots x_k(m_{n-1}) - x(m_1)\cdots x(m_{n-1})\right] \nu^{\otimes(n-1)}(d\mathbf{m})\right| \\
&\qquad \times \nu(dm).
\end{aligned}$$

Each integral

$$\int_{\mathcal{M}^{(n-1)}} a_n(\mathbf{m},m)\left[x_k(m_1)\cdots x_k(m_{n-1}) - x(m_1)\cdots x(m_{n-1})\right] \nu^{\otimes(n-1)}(d\mathbf{m})$$

belongs to $L^\infty(\mathcal{M},\nu)$ with norm that is bounded in k. Moreover, each such integral converges to zero as k goes to infinity by Lemma 4.1. It follows from the Dominated Convergence Theorem that

$$\lim_{k\to\infty} \left|\int_{\mathcal{M}} \Gamma x_k(m) f(m)\,\nu(dm) - \int_{\mathcal{M}} \Gamma x(m) f(m)\,\nu(dm)\right| = 0,$$

as required. □

Recall the definition of the function $\dot{V} : L^\infty_+(\mathcal{M},\nu) \to \mathbb{R}$ from (3.14).

Lemma 4.3. *For any constants $0 < a \leq b < \infty$, the function $\dot{V}$ restricted to the set*

$$\{x \in L^\infty_+(\mathcal{M},\nu) : a \leq \operatorname{ess\,inf}_m x(m) \leq \operatorname{ess\,sup}_m x(m) \leq b\}$$

is upper semicontinuous in the weak topology.*

Proof. Fix a domain $D \subset L^\infty_+(\mathcal{M}, \nu)$ of the type considered in the statement of the lemma.

Observe that

$$\dot{V}(x) = -\int_{\mathcal{M}} \frac{1}{x(m)}\,\nu(dm) + H(x),$$

where $H : D \to \mathbb{R}$ is a sum of functions, each of the form

$$x \mapsto \int g(m_1, \ldots, m_k) x(m_1) \cdots x(m_k)\,\nu(dm_1) \cdots \nu(dm_k)$$

for some function $g \in L^\infty(\mathcal{M}^k, \nu^{\otimes k})$. It therefore suffices by Lemma 4.1 to show that the function

$$x \mapsto -\int_{\mathcal{M}} \frac{1}{x(m)}\,\nu(dm) \tag{4.1}$$

is weak*-upper semicontinuous on D

Suppose without loss of generality that ν is a probability measure. Let $\Pi_n = (A_{n_1}, \ldots, A_{n_{p(n)}})$ be a sequence of partitions of $\mathcal{M}$ such that the σ-fields generated by the successive partitions form a filtration $\mathcal{F}_n$ with the property that $\sigma(\bigcup_n \mathcal{F}_n) = \mathcal{B}(\mathcal{M})$, the Borel σ-field on $\mathcal{M}$. Consider $x \in D$ as a random variable on the probability space $(\mathcal{M}, \mathcal{B}(\mathcal{M}), \nu)$, and let $x^{(n)} := \mathbb{E}[x \,|\, \mathcal{F}_n]$ be the conditional expectation of x given $\mathcal{F}_n$. Note that $x^{(n)}$ is a non-negative martingale with $a \le x^{(n)}(m) \le b$ for ν-a.e. $m \in \mathcal{M}$. For $m \in A_{n_q}$ we have

$$x^{(n)}(m) = \frac{1}{\nu(A_{n_q})} \int_{A_{n_q}} x(m)\,\nu(dm),$$

with the convention that $0/0 = 0$. The function taking x to each such integral is weak*-continuous. As a result, the function

$$x \mapsto -\int \frac{1}{x^{(n)}(m)}\,\nu(dm) = -\sum_q \nu(A_{n_q}) \left[\frac{1}{\nu(A_{n_q})} \int_{A_{n_q}} x(m)\,\nu(dm)\right]^{-1}$$

is also weak*-continuous.

By Jensen's inequality for conditional expectation, the sequence of random variables $-1/x^{(n)}$ is a non-positive supermartingale, and so the sequence of expectations

$$-\int_{\mathcal{M}} [x^{(n)}(m)]^{-1}\,\nu(dm) = \mathbb{E}[-1/x^{(n)}]$$

is non-increasing. By the Martingale Convergence Theorem (see, for example, Theorem IV-1-2 of [Nev75]), the sequence $x^{(n)}(m)$ converges to $x(m)$ for ν-a.e. $m \in \mathcal{M}$. Therefore, $-\int_{\mathcal{M}} [x^{(n)}(m)]^{-1}\,\nu(dm)$ decreases to $-\int_{\mathcal{M}} x(m)^{-1}\,\nu(dm)$ by the Dominated Convergence Theorem.

We can thus write the function $x \mapsto -\int_{\mathcal{M}} x(m)^{-1}\,\nu(dm)$ as the infimum of the family of functions $x \mapsto -\int_{\mathcal{M}} [x^{(n)}(m)]^{-1}\,\nu(dm)$, each of which is weak*-continuous. Consequently, the function $x \mapsto -\int_{\mathcal{M}} x(m)^{-1}\,\nu(dm)$ is weak*-upper semicontinuous. □

Acknowledgment: Parts of this paper were written while the second author was a Visiting Fellow at the Mathematical Sciences Institute of the Australian National University. He thanks the members of the Institute for their hospitality. The

authors also thank Reinhard Bürger, Bernd Sturmfels, David Steinsaltz and Ken Wachter for numerous helpful suggestions.

References

[Aki82] Ethan Akin, *Exponential families and game dynamics*, Canad. J. Math. **34** (1982), no. 2, 374–405. MR MR658973 (83m:92052)

[Baa01] Ellen Baake, *Mutation and recombination with tight linkage*, J. Math. Biol. **42** (2001), no. 5, 455–488. MR MR1842838 (2002d:92008)

[Baa05] Michael Baake, *Recombination semigroups on measure spaces*, Monatsh. Math. **146** (2005), no. 4, 267–278. MR MR2191729 (2006k:92055)

[Baa07] ———, *Addendum to: "Recombination semigroups on measure spaces" [Monatsh. Math.* **146** *(2005), no. 4, 267–278; mr2191729]*, Monatsh. Math. **150** (2007), no. 1, 83–84. MR MR2297255

[BB96] Reinhard Bürger and Immanuel M. Bomze, *Stationary distributions under mutation-selection balance: structure and properties*, Adv. in Appl. Probab. **28** (1996), no. 1, 227–251. MR MR1372337 (97h:92008)

[BB03] Michael Baake and Ellen Baake, *An exactly solved model for mutation, recombination and selection*, Canad. J. Math. **55** (2003), no. 1, 3–41. MR MR1952324 (2004a:92015)

[Bom91] Immanuel M. Bomze, *Cross entropy minimization in uninvadable states of complex populations*, J. Math. Biol. **30** (1991), no. 1, 73–87. MR MR1130789 (92j:92012)

[BT91] N. H. Barton and Michael Turelli, *Natural and sexual selection on many loci*, Genetics **127** (1991), 229–255.

[Bür98] Reinhard Bürger, *Mathematical properties of mutation selection models*, Genetica **102/103** (1998), 279–298.

[Bür00] ———, *The mathematical theory of selection, recombination, and mutation*, Wiley Series in Mathematical and Computational Biology, John Wiley & Sons Ltd., Chichester, 2000. MR MR1885085 (2002m:92002)

[BW01] Ellen Baake and Holger Wagner, *Mutation-selection models solved exactly with methods of statistical mechanics*, Genet. Res., Camb. **78** (2001), 93–117.

[CDSS07] Gheorghe Craciun, Alicia Dickenstein, Anne Shiu, and Bernd Sturmfels, *Toric dynamical systems*, 2007, Preprint, available at `http://arxiv.org/abs/0708.3431`.

[CF05] Gheorghe Craciun and Martin Feinberg, *Multiple equilibria in complex chemical reaction networks. I. The injectivity property*, SIAM J. Appl. Math. **65** (2005), no. 5, 1526–1546 (electronic). MR MR2177713 (2006g:92075)

[Cha94] Brian Charlesworth, *Evolution in age-structured populations*, Cambridge University Press, Cambridge, 1994.

[Cha01] ———, *Patterns of age-specific means and genetic variances of mortality rates predicted by the mutation-accumulation theory of ageing*, J. Theor. Biol. **210** (2001), no. 1, 47–65.

[CLH99] VijaySekhar Chellaboina, Alexander Leonessa, and Wassim M. Haddad, *Generalized Lyapunov and invariant set theorems for nonlinear dynamical systems*, Systems Control Lett. **38** (1999), no. 4-5, 289–295. MR MR1754911 (2001h:34069)

[Daf78] C. M. Dafermos, *Asymptotic behavior of solutions of evolution equations*, Nonlinear evolution equations (Proc. Sympos., Univ. Wisconsin, Madison, Wis., 1977), Publ. Math. Res. Center Univ. Wisconsin, vol. 40, Academic Press, New York, 1978, pp. 103–123. MR MR513814 (80i:35019)

[Daw99] Kevin J. Dawson, *The dynamics of infinitesimally rare alleles, applied to the evolution of mutation rates and the expression of deleterious mutations*, Theor. Popul. Biol. **55** (1999), 1–22.

[DS88] Nelson Dunford and Jacob T. Schwartz, *Linear operators. Part I*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1988, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication. MR MR1009162 (90g:47001a)

[EK86] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986, Characterization and convergence. MR MR838085 (88a:60130)

[Esh71] Ilan Eshel, *On evolution in a population with an infinite number of types*, Theoret. Population Biology **2** (1971), 209–236. MR MR0356905 (50 #9373)

[ESW06] Steven N. Evans, David Steinsaltz, and Kenneth W. Wachter, *A mutation-selection model for general genotypes with recombination*, 2006, Preprint, available at `http://arxiv.org/abs/q-bio/0609046`.

[ET77] Warren J. Ewens and Glenys Thomson, *Properties of equilibria in multi-locus genetic systems*, Genetics **87** (1977), no. 4, 807–819. MR MR0682090 (58 #33112)

[Ewe04] Warren J. Ewens, *Mathematical population genetics. I*, second ed., Interdisciplinary Applied Mathematics, vol. 27, Springer-Verlag, New York, 2004, Theoretical introduction. MR MR2026891 (2004k:92001)

[FK00] Caleb E. Finch and Thomas B. L. Kirkwood, *Chance, development, and aging*, Oxford University Press, 2000.

[Ham66] W. D. Hamilton, *The moulding of senescence by natural selection*, J. Theor. Biol. **12** (1966), 12–45.

[Hof85] Josef Hofbauer, *The selection mutation equation*, J. Math. Biol. **23** (1985), no. 1, 41–53. MR MR821683 (87d:92028)

[Hol95] Robin Holliday, *Understanding ageing*, Cambridge University Press, 1995.

[Kin78] J. F. C. Kingman, *A simple model for the balance between selection and mutation*, J. Appl. Probability **15** (1978), no. 1, 1–12. MR MR0465272 (57 #5177)

[KJB02] Mark Kirkpatrick, Toby Johnson, and Nick Barton, *General models of multilocus evolution*, Genetics **161** (2002), 1727–1750.

[KM66] Motoo Kimura and Takeo Maruyama, *The mutational load with epistatic gene interactions in fitness*, Genetics **54** (1966), 1337–1351.

[Kon82] A.S. Kondrashov, *Selection against harmful mutations in large sexual and asexual populations*, Genet. Res. **40** (1982), 325–332.

[LR04] Hartmut Logemann and Eugene P. Ryan, *Asymptotic behaviour of nonlinear systems*, Amer. Math. Monthly **111** (2004), no. 10, 864–889. MR MR2104692 (2005h:34132)

[Med52] Peter Medawar, *An unsolved problem in biology: An inaugural lecture delivered at University College, London, 6 December, 1951*, H. K. Lewis and Co., London, 1952.

[Nev75] J. Neveu, *Discrete-parameter martingales*, revised ed., North-Holland Publishing Co., Amsterdam, 1975, Translated from the French by T. P. Speed, North-Holland Mathematical Library, Vol. 10. MR MR0402915 (53 #6729)

[SEW05] David Steinsaltz, Steven N. Evans, and Kenneth W. Wachter, *A generalized model of mutation-selection balance with applications to aging*, Adv. Appl. Math. **35** (2005), 16–33.

[SH92] Stanley A. Sawyer and Daniel L. Hartl, *Population genetics of polymorphism and divergence*, Genetics **132** (1992), 1161–1176.

[TB90] Michael Turelli and N.H. Barton, *Dynamics of polygenic characters under selection*, Theor. Popul. Biol. **38** (1990), 1–57.

[WBG98] Holger Wagner, Ellen Baake, and Thomas Gerisch, *Ising quantum chain and sequence evolution*, J. Statist. Phys. **92** (1998), no. 5-6, 1017–1052. MR MR1657856 (2000e:92017)

[WES08] Kenneth W. Wachter, Steven N. Evans, and David R. Steinsaltz, *The age-specific forces of natural selection and walls of death*, 2008, Available at `http://arxiv.org/abs/0807.0483`.

[Wil57] George C. Williams, *Pleiotropy, natural selection, and the evolution of senescence*, Evolution **11** (1957), 398–411.

[WSE08] Kenneth W. Wachter, David R. Steinsaltz, and Steven N. Evans, *Vital rates from the action of mutation accumulation*, 2008, Available at `http://arxiv.org/abs/0808.3622`.

E-mail address: `aclayton@math.berkeley.edu`

Department of Mathematics #3840, University of California, Berkeley, 970 Evans Hall, Berkeley, CA 94720-3840, U.S.A.

E-mail address: `evans@stat.Berkeley.EDU`

Department of Statistics #3860, University of California at Berkeley, 367 Evans Hall, Berkeley, CA 94720-3860, U.S.A