Chapter 12

Mutation-Selection Systems

The basic mechanisms of population biology are mutation, selection, recombination and genetic drift. In the previous chapter we concentrated on mutation and genetic drift. In this chapter we introduce mathematical models of recombination and selection. However it should be emphasized that these are idealization of highly complex biological processes and there is an immense biological literature including empirical investigation, theoretical models of varying degrees of complexity and simulation studies. For example the concept of *fitness* is an abstract notion that in the biological context can involve fitness at the level of a single gene, genome or phenotype. At the level of the genome this can involve the interaction between genes (*epistasis*) and various models of such interactions have been proposed (see e.g. Gavrilets [270]). One of the continuing issues is the question of the *levels of selection* (see e.g. Brandon and Burian (1984) [48], Lloyd (2005) [438], Okasha (2006), [495]) which include notions of group selection, kin selection, inclusive fitness (see Hamilton (1964) [295]) and so on. For example, *inclusive fitness* represents to effective overall contribution of an individual including its own reproductive success as well as its contribution (due to its behavior) to the fitness of its genetic kin.

Our aim in this chapter is to introduce some mathematical aspects of the interplay of mutation, selection and genetic drift.

12.1   The infinite population dynamics of mutation, selection and recombination.

12.1.1   Selection

The investigation of infinite population models with mutation, recombination and selection leads to an interesting class of dynamical systems (see Hofbauer and Sigmund (1988) [312] and Bürger [62], [61]). These are obtained as special cases of the general FV process by setting $\gamma = 0$ and serve as approximations to

systems in which the number of individuals N is very large.

One of the objectives of this chapter is to investigate in one setting the extent to which the behavior of the finite system differs from that of the infinite system.

Consider an infinite diploid population without mutation or recombination (i.e. $\gamma = 0$, $A = 0$, $\rho = 0$) with K types of gametes. The unordered pair $\{i, j\}$ represents the genotype determined by the gametes i and j. Let $x_i(t)$ be the amount of copies of gamete i in the population at time t and p_i denote the frequency $p_i = \frac{x_i}{\sum x_i}$.

Let $V(i, j) = V(j, i) = b_{i,j} - d_{i,j}$ where b_{ij} and d_{ij} are the birth and death rates of the genotype. The fitness, $V(i)$ of the ith gamete is defined by

$$V(i) = \sum_j p_j V(i, j)$$

and the mean fitness is defined by

$$\bar{V}(p) = \bar{V} = \sum_i V(i) p_i = \sum_{ij} p_i p_j V(i, j).$$

Then the population sizes x_i satisfy the equations

$$\dot{x}_i = x_i \sum_j V(i, j) \frac{x_j}{|x|}, \ i = 1, \dots, K$$

Proposition 12.1 *The proportions* $\{p_i\}$ *satisfy the equations:*

$$\dot{p}_i = p_i(V(i) - \bar{V}), \ i = 1, \dots, K$$

Proof. This can be derived from the $\dot{x}$ equations by the substitution $x_i = |x| p_i$ giving

$$\dot{p}_i |x| + p_i(\sum \dot{x}_j) = |x| p_i V(i)$$

which yields

$$\dot{p}_i + p_i(\sum_j p_j V(j)) = p_i V(i)$$

and the result immediately follows. ∎

12.1.2   Riemannian structure on Δ_{K-1}

The deterministic differential equations of selection have played an important role in the development of population genetics. A useful tool in their analysis was a

geometrical approach developed by Shahshahani and Akin. We next give a brief introduction to this idea.

Let M be a smooth manifold. The tangent space at x, $T_x M$ can be identified with the space of tangents at x to all smooth curves through x. The tangent bundle $TM = \{(p, v) : p \in M, v \in T_p M\}$.

Definition 12.2 *A Riemannian metric on M is a smooth tensor field*

$$g : C^\infty(TM) \otimes C^\infty(TM) \to C_0^\infty(M)$$

such that for each $p \in M$,

$$g(p)|_{T_p M \otimes T_p M} : T_p M \otimes T_p M \to \mathbb{R}$$

with

$$g(p) : (X, Y) \to \langle X, Y \rangle_{g(p)}$$

where $\langle X, Y \rangle_{g(p)}$ is an inner product on $T_p M$.

Definition 12.3 *The directional derivative in direction v is defined by*

$$\partial_v f(x) = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}$$
$$= \sum v_i \frac{\partial f(x)}{\partial x_i}$$

The gradient $\nabla_g f(x)$ is defined by

$$\langle \nabla_g f(x), v \rangle_g = \partial_v f(x) \quad \forall v \in T_x M.$$

Example 12.4 *Consider the d-dimensional manifold $M = \mathbb{R}^d$ and $\mathbf{a}(\cdot)$ be a smooth map from M to $\mathbb{R}^d \otimes \mathbb{R}^d$ (($d \times d$)-matrices). We will write*

$$\mathbf{a}(x) = (a_{ij}(x))$$
$$\mathbf{a}^{-1}(x) = (a^{ij}(x))$$

Assume that

$$\sum a^{ij}(x) u_i u_j \geq \gamma \sum u_j^2, \ \gamma > 0.$$

The tangent space $T_{\mathbf{x}} M \cong \mathbb{R}^d$ and we define a Riemannian metric on M by

$$g_{\mathbf{a}(x)}(\mathbf{u}, \mathbf{v}) := \sum_{i,j=1}^{d} a_{ij}(x) u^i v^j.$$

The associated Riemannian gradient and norm are

$$(\nabla_{\mathbf{a}} f)^i = \sum_j a^{ij} \frac{\partial f}{\partial x_j}$$
$$\|u\|_{\mathbf{a}(x)}^2 = \sum_{ij} a_{ij}(x) u^i u^j.$$

The Shahshahani metric and gradient on Δ_{K-1}

Let $M_K = \mathbb{R}_+^K := \{x \in \mathbb{R}^K, x = (x_1, \ldots, x_K),\ x_i > 0 \text{ for all } i\}$ is a smooth K-dimensional manifold.

Shahshahani introduced the following Riemannian metric on M_K

$$\langle u, v \rangle_g = g_x(u, v) := \sum_{i=1}^{K} |x| \frac{u_i v_i}{x_i}$$

$$|x| = \sum x_i$$

$\| \ \|_g$ and $\nabla_g F$ will denote the corresponding norm and gradient. We have

$$(\nabla_g F)^i = \sum_i \frac{x^i}{|x|} \frac{\partial F}{\partial x^i} \frac{\partial}{\partial x^i}$$

Recall that the simplex $\Delta_{K-1} := \{(p_1, \ldots, p_K) : p_i \geq 0,\ \sum_{i=1}^{K} p_i = 1\}$. The interior of the simplex $\Delta_{K-1}^0 = \mathbb{R}_+^K \cap \Delta_{K-1}$ is a $(K-1)$-dimensional submanifold of M_K. We denote by $T_p \Delta_{K-1}^0$ the tangent space to Δ_{K-1}^0 at p. Then g induces a Riemannian metric on $T_p \Delta_{K-1}^0$.

Basic Facts

We have the Shahshahani inner product on $\Delta K - 1$ at a point $p \in \Delta K - 1$:

$$(12.1)\quad \langle u, v \rangle_p = \sum_{i=1}^{K} \frac{u_i v_i}{p_i}.$$

1. $T_p \Delta_{K-1}^0$ can be viewed as the subspace of $T_p M_K$ of vectors, v, satisfying $\langle p, v \rangle_g = 0$ if we identify p with an element of $T_p M_K$.

Proof. Recall that $T_p \Delta_{K-1}^0$ is given by tangents to all smooth curves lying in Δ_{K-1}^0. Therefore if $v \in T_p \Delta_{K-1}^0$, then $v = q - p$ where $p, q \in \Delta_{K-1}^0$ and therefore $\sum_{i=1}^{K} v_i = 0$. Therefore,

$$\sum_i p_i \frac{1}{p_i} v_i = 0.$$

2. If $F : \Delta_{K-1}^0 \to \mathbb{R}$ is smooth, then the Shahshahani gradient is

$$(\nabla_g F)_i = p_i \left(\frac{\partial F}{\partial p_i} - \sum_j p_j \frac{\partial F}{\partial p_j} \right).$$

Proof. From the definition, $\nabla_g F$ is the orthogonal projection on the subspace $T_p \Delta_{K-1}^0$ of

$$(\nabla_g F)_i = p_i \frac{\partial F}{\partial p_i}$$

and therefore we must have $\sum_i (\nabla_g F)_i = 0$. This then gives the result.

Remark 12.5 *This (Shahshahani) gradient coincides with the gradient on Δ_{K-1} associated with the K-alleles Wright-Fisher model and appears in the description of the rate function for large deviations from the infinite population limit (see below).*

Theorem 12.6 *The dynamical system $\{\mathbf{p}(t) : t \geq 0\}$ is given by*

$$\dot{\mathbf{p}}(t) = \frac{1}{2}(\nabla_{g(\mathbf{p}(t))}\bar{V})(\mathbf{p}(t)).$$

Proof. From the above, applying the Shahshahani gradient to $\bar{V}$, we get

$$(\nabla_g \bar{V})_i = 2\left(p_i V(i) - p_i \sum p_j V(j)\right)$$
$$= 2p_i(V(i) - \bar{V}).$$

$\blacksquare$

Theorem 12.7 *(Fisher's Fundamental Theorem)*
(a) Mean fitness increases on the trajectories of $p(t)$.
(b) The rate of change of the mean $\bar{V}(t)$ along orbits is proportional to the variance.
(c) At an equilibrium point the eigenvalues of the Hessian must be real.

Proof. (a) follows immediately from (b).
(b)

$$d\bar{V}(t) = \left\langle \nabla_g \bar{V}(\mathbf{p}(t)), \dot{\mathbf{p}}(t) \right\rangle_{g(\mathbf{p}(t))}$$

$$= 2\left\langle \dot{p}(t), \dot{p}(t) \right\rangle_{g(\mathbf{p}(t))} = 2\left(\sum_i p_i(t)(V(i) - \bar{V}(t))^2\right)$$

$$= 2\left(\sum_i p_i(t)V(i)^2 - \bar{V}(t)^2\right)$$

$$= 2Var_{\mathbf{p}(t)}(\mathbf{V}) \geq 0.$$

(b) It also follows from the gradient form that the Hessian is symmetric (matrix of mixed second partials of $\bar{V}$). $\blacksquare$

Theorem 12.8 *(Kimura's Maximum Principle) "Natural selection acts so as to maximize the rate of increase in the average fitness of the population."*

Proof. This simply follows from the property that the directional derivative $\partial_v \bar{V}$ is maximal in the direction of the gradient. $\blacksquare$

Example 12.9 *Consider a two type ($\{1, 2\}$) population with frequencies $(p_1, p_2) = (p, 1 - p)$.*

$$V(i, j) = av(i) + av(j) + c\delta_{ij}$$

(When $c = 0$ we have the additive (or haploid) model. When $a = 0$ and $c > 0$ we have the heterozygote advantage model.)
In this case

$$\begin{aligned}
\bar{V}(p_1, p_2) &= ap_1 v(1) + ap_2 v(2) + cp_1 p_2 \\
&= V(p, 1 - p) = ap(v(1) - v(2)) + av(2) + cp(1 - p)
\end{aligned}$$

Then depending on the choice of $a, c, v(1), v(2)$, the optimum value of p can range between 0 and 1.

Remark 12.10 *For the multilocus situation there is the Fisher-Price-Ewens version (e.g. Frank (1997) [257], Ewens [244]). This is also related to* the secondary theorem of natural selection *of Robertson (1966) [522] which relates the rate of change of a quantitative character under selection in terms of the covariance of the character and fitness.*

The above equations are special cases of the class of *replicator equations* of the form

$$(12.2) \quad \frac{dp_i(t)}{dt} = p_i(t)(f_i(\mathbf{p}(t)) - \sum p_i f_i(\mathbf{p}(t)), \quad i = 1, \dots, K$$

where $\{f_i(\mathbf{p})\}_{i=1m\dots,K}$ is a vector field on Δ_{K-1}. In the linear case $f_i(\mathbf{p}) = \sum_j a_i j p_j$ these are equivalent to the the Lotka-Volterra equations

$$(12.3) \quad \frac{dx_i(t)}{dt} = x_i(t) \left(r_i + \sum_{j=1}^n K_{ij} x_j(t) \right), \quad i = 1, \dots, K - 1$$

by setting $p_i(t) = \frac{x_i(t)}{\sum_i x_i(t)}$.

12.1.3   Mutation-Selection

The replicator equations that include both mutation and selection are given by

$$(12.4) \quad \frac{dp_i(t)}{dt} = p_i(t)(V(i) - \bar{V}) + m(\sum_{j \neq i} q_{ji} p_j - p_i)$$

where m is the mutation rate and for each j, q_{ji}, $i \neq j$ is the probability that type j mutates to type i and $\sum_{i \neq j} q_{ji} = 1$.

Theorem 12.11 *The mutation-selection dynamical system is a Shahshahani gradient system if and only if*

(12.5) $q_{ji} = q_i \forall j$,

(that is type-independent mutation as in the infinitely many alleles model). In the latter case the potential is

(12.6) $W(p) = \bar{V}(p) - H(q|p), \quad H(q|p) = -\sum_{i=1}^{n} q_i \log p_i.$

Proof. See Hofbauer and Sigmund [312], Chapt. VI, Theorem 1. ■

We will see below that there is a far-reaching analogue of this for the stochastic (finite population) generalizations.

Remark 12.12 *In general the deterministic mutation-selection equations are not a gradient system and can exhibit complex dynamics - for example, a stable limit cycle (Hofbauer and Sigmund [312], 25.4). An interesting special case is the diploid case with three types - two favourable and mutation. Baake [23] showed that these can exhibit stable limit cycles. Hofbauer (1985) [311] also showed this for selection mutation models with cyclic mutation.*

Smale [555] pointed out that for n types, $n \geq 5$, dynamical systems on the simplex can have complex behaviour. He gave an example that "may not be approximated by a structurally stable, dynamical system, or it may have strange attractors with an infinite number of periodic solutions". Some further basic results on competitive systems are covered by Hirsch (1982), (1985), (1988) [310] and Liang and Jiang (2003) [424].

12.1.4 Multiple loci and recombination

Multiloci models give rise to dynamical systems that have been extensively studied. They give rise to a large class of dynamical systems that can have complex behaviour. Akin [3] analyzed the simplest two loci model with selection and recombination and proved that in general this is not a gradient system and that periodic orbits can exist. We briefly sketch the simplest example.

Consider a two-loci model with two alleles at each loci. We denote the types by $1 = AB, 2 = Ab, 3 = aB, 4 = ab$ and with gamete frequencies

(12.7)

$p_{AB}, p_{Ab}, p_{aB}, p_{ab}, \quad p_A = p_{AB} + p_{Ab}, \ p_B = p_{AB} + p_{aB}, \ p_a = p_{aB} + p_{ab}, \ p_b = p_{Ab} + p_{ab}.$

Then the *measure of linkage disequilibrium* is defined as

(12.8) $d := p_{AB} p_{ab} - p_{Ab} p_{bA}$

so that $d = 0$ if $p_{AB} = p_A p_B$, etc. The diploid fitness function is denoted by $V(i, j)$. Some natural assumptions are that

(12.9) $m_{ij} = m_{ji}, \quad m_{14} = m_{23} = 0.$

There are 10 zygotic types $AB/AB, Ab/AB, \ldots, ab/ab$ and the corresponding fitness table

$$
\begin{array}{c|cccc}
 & AB & Ab & aB & ab \\
AB & w_{11} & w_{12} & w_{13} & w_{14} \\
Ab & w_{21} & w_{22} & w_{23} & w_{24} \\
aB & w_{31} & w_{32} & w_{33} & w_{34} \\
ab & w_{41} & w_{42} & w_{43} & w_{44}
\end{array}
$$
(12.10)

The recombination vectorfield

(12.11) $R = rbd\xi_i, \ i = 1, 2, 3, 4$

where r is the recombination rate, b is the birth rate for double heterozygotes, d is the linkage disequilibrium and

(12.12) $\xi = (1, -1, -1, 1)$

so that

(12.13) $d\xi = p - \pi(p)$

where $\pi(p)$ has the same marginals as p but in linkage equilibrium (independent loci).

The system of differential equations for the frequencies of types $1, 2, 3, 4$ with selection and recombination are

(12.14) $\dfrac{dp_i}{dt} = p_i(V(i) - \bar{V}) - rbd\xi_i \quad i = 1, 2, 3, 4$

where

(12.15) $V(i) = \displaystyle\sum_{j=1}^{4} p_j V(i, j), \quad \bar{V} = \sum_{i=1}^{4} p_i V(i), \quad d = p_1 p_4 - p_2 p_3.$

In the case $V \equiv 0$ the system approaches linkage equilibrium. However Akin [3] showed that there exist fitness functions V and parameters b, r such that the system exhibits a Hopf bifurcation leading to cyclic behaviour. More generally, multilocus systems can exhibit many types of complex behaviour (see for example, Kirzhner, Korol and Nevo (1996) [405] and Lyubich and Kirzhner (2003) [439]).

12.2 Infinitely many types Fleming-Viot: Dual representation with mutation, selection and recombination

We now consider the Fleming-Viot process with selection and recombination and establish uniqueness using a dual representation of Ethier and Kurtz.

Let $\rho \geq 0$ and $\eta(x_1, x_2, \Gamma)$ be a transition function from $E \times E \to E$. For $i = 1, \ldots, m$ define $R_{im} : \mathcal{B}(E^m) \to \mathcal{B}(E^{m+1})$ by

$$(12.16) \quad R_{im}f(x_1, \ldots, x_{m+1}) = \int f(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_m)\eta(x_i, x_{m+1}, dz)$$

and assume that $R_{im} : C_b(E^m) \to C_b(E^{m+1})$. The R_{im} are called the recombination operators for the process and ρ is called the recombination rate.

Given $V \in \mathcal{B}_{\mathrm{sym}}(E \times E)$, with $\bar{V} := \sup_{x,y,z} |V(x,y) - V(y,z)| < \infty$, define the selection operators

$$(12.17)$$
$$V_{im}f(x_1, \ldots, x_{m+2}) = \frac{V(x_i, x_{m+1}) - V(x_{m+1}, x_{m+2})}{\bar{V}} f(x_1, \ldots, x_m) \quad \text{for } i = 1, \ldots, m.$$

For $f \in \mathcal{D}(A^{(n)}) \cap \mathcal{B}(E^n)$, define $F(f, \mu) = \int f d\mu^n$ and

$$(12.18)$$
$$GF(f, \mu) = F(A^{(n)}f, \mu) + \gamma \sum_{1 \leq i < j \leq n} \left(F(\widetilde{\Theta}_{ij}f, \mu) - F(f, \mu) \right)$$
$$+ \rho \sum_{i=1}^{n} (F(R_{in}f, \mu) - F(f, \mu)) + \bar{V} \sum_{i=1}^{n} F(V_{in}f, \mu).$$

For $f \in C_{\mathrm{sim}}(E^{\mathbb{N}})$, with $\mathfrak{n}(f) = n$, and $f \in \mathcal{D}(A^n) \cap \mathcal{B}(E^n)$, let

$$(12.19) \quad Hf := \sum_{i=1}^{n} A_i f + \gamma \sum_{j=1}^{n} \sum_{k \neq j} [\widetilde{\Theta}_{jk}f - f] + \rho \sum_{i=1}^{n} [R_{in}f - f] + \bar{V} \sum_{i=1}^{n} [V_{in}f - f].$$

If $\beta(f) := \bar{V}\mathfrak{n}(f)$, then

$$(12.20) \quad GF(f, \mu) = F(Hf, \mu) + \beta(f))F(f, \mu),$$

and $\sup_{\mu \in M_1(E)} |F(Hf, \mu)| \leq \mathrm{const} \cdot \mathfrak{n}(f)$.

Theorem 12.13 *Let G satisfy the above conditions and assume that the mutation process with generator A has a version with sample paths in $D_E[0, \infty)$. Then for each $\mu \in \mathcal{P}(E)$ there exists a unique solution P_μ of the martingale problem for G.*

Proof. (Ethier-Kurtz (1987) [222]) We construct a function-valued dual process. Let N be a jump Markov process taking non-negative integer values with transition intensities

(12.21) $q_{m,m-1} = \gamma m(m-1),\ q_{n,m+2} = \bar{V}m,\ q_{m,m+1} = \rho m,\ q_{i,j} = 0$ otherwise.

For $1 \leq i \leq m$, let $\{\tau_k\}$ be the jump times of N, $\tau_0 = 0$, and let $\{\Gamma_k\}$ be a sequence of random operators which are conditionally independent given M and satisfy

(12.22) $P(\Gamma_k = \Theta_{ij}|N) = \dfrac{2}{N(\tau_k-)N(\tau_k)}\mathbf{1}_{N(\tau_k-)-N(\tau_k)=1},\quad 1 \leq i < j \leq N(\tau_k-)$

(12.23) $P(\Gamma_k = R_{im}|N) = \dfrac{1}{m}\mathbf{1}_{\{N(\tau_k-)=m,N(\tau_k=m+1)\}}$

(12.24) $P(\Gamma_k = V_{im}|N) = \dfrac{1}{m}\mathbf{1}_{\{N(\tau_k-)=m,N(\tau_k)=m+2\}}.$

For $f \in C_{\text{sim}}(E^{\mathbb{N}})$, define the $C_{\text{sim}}(E^{\mathbb{N}})$-valued process Y with $Y(0) = f$ by

(12.25) $Y(t) = S_{t-\tau_k}\Gamma_k S_{\tau_k-\tau_{k-1}}\Gamma_{k-1}\ldots\Gamma_1 S_{\tau_1}f,\quad \tau_k \leq \tau_{k+1}.$

Then for any solution P_μ to the martingale problem for G and $f \in C_{\text{sim}}(E^{\mathbb{N}})$ we get the FK-dual representation

(12.26) $P_\mu[F(f, X(t))] = Q_f\left[F(Y(t),\mu)\exp\left(\bar{V}\int_0^t \mathfrak{n}(Y(u))du\right)\right]$

which establishes that the martingale problem for G is well-posed. Since the function $\beta(f) = \bar{V}\mathfrak{n}(f)$ is not bounded we must verify condition (7.29). This follows from the following lemma due to Ethier and Kurtz (1998) [224], Lemma 2.1. ∎

Lemma 12.14 *Let* $N(t) = \mathfrak{n}(Y(t))$ *be as above,* $\tau_K := \inf\{t : N(t) \geq K\}$ *and* $\theta > 0$. *Then there exists a function* $F(n) \geq const \cdot n^2$ *and a constant* $L > 0$ *such that*

(12.27)
$$E\left[F(N(t \wedge \tau_K))\exp\left(\theta\int_0^{t\wedge\tau_K} N(s)ds\right)|N(0) = n\right] \leq F(n)e^{Lt},\quad \forall\, K \geq 1,$$

and given $N(0) = n$, $\left\{N(t \wedge \tau_K)\exp\left(\bar{V}\int_o^{t\wedge\tau_K} N(s)ds\right) : K \geq 1\right\}$ *are uniformly integrable.*

Proof. Take $F(m) := (m!)^{\beta}$, with $\beta < \frac{1}{2}$. Then

$$
\begin{aligned}
QF(m) &+ \theta m F(m) \\
&= \gamma m(m-1)(F(m-1) - F(m)) + \rho m(F(m+1) - F(m)) \\
&\quad + \theta m(F(m+2) - F(m)) + \theta m F(m) \\
&= -\gamma O(m^{2+\beta})(m!)^{\beta} + \rho O(m^2)(m!)^{\beta} + \theta O(m^{2\beta})(m!)^{\beta}
\end{aligned}
$$

Since the negative term dominates for large m if $0 < \beta < \frac{1}{2}$ and $\gamma > 0$, we can choose $L > 0$ such that

$(12.28)\quad QF(m) + \theta m F(m) \leq L.$

The optional sampling theorem implies that for $\tau_K := \inf\{t : N(t) \geq K\}$ and $N(0) = m$

$$
\begin{aligned}
E&\left[\exp\left(\theta \int_0^{t \wedge \tau_k} N(s)ds\right) |N(0) = m\right] \\
&\leq E\left[F(N(t \wedge \tau_k)) \exp\left(\theta \int_0^{t \wedge \tau_k} N(s)ds\right) |N(0) = m\right] \\
&\leq F(m) + E\left[\int_0^{t \wedge \tau_k} \exp\left(\theta \int_0^u N(s)ds\right)(QF(N(u)) + \theta N(u)F(N(u)))du|N(0) = m\right] \\
&\leq F(m) + LE\left[\int_0^{t \wedge \tau_k} \exp\left(\theta \int_0^u N(s)ds\right)du|N(0) = m\right]
\end{aligned}
$$

and the lemma follows by Gronwall's inequality.

∎

In Chapter 6 we showed that the martingale problem for the Fleming-Viot process with mutation selection and recombination is well-posed and defines a $\mathcal{P}(E)$-valued Markov diffusion process. In this chapter we focus on mutation and selection but also give a brief introduction to some aspects of recombination. In evolutionary theory mutation plays an important role in producing novelty and maintaining diversity while selection eliminates deleterious mutations and makes possible the emergence and fixation of rare advantageous mutations. From a more abstract viewpoint this can be viewed as a search process which generates new information.

As above we consider the mutation generator A and the bounded diploid fitness function For $V \in \mathcal{B}_{\text{sym}}(E \times E)$, set $\bar{V} = \sup_{x,y,z} |V(x,y) - V(y,z)|$. Without loss of generality we can assume that $\bar{V} = 1$ and define the selection coefficient $s > 0$ and selection operators

$(12.29)\quad V_{im} f(x_1, \ldots, x_{m+2}) = (V(x_i, x_{m+1}) - V(x_{m+1}, x_{m+2})) f(x_1, \ldots, x_m).$

For $f \in \mathcal{D}(A^{(n)}) \cap \mathcal{B}(E^n)$, define $F(f, \mu) = \int f d\mu^n$ and

(12.30)

$$GF(f, \mu) \; = F(A^{(n)}f, \mu) + \gamma_{1 \leq i < j \leq n} \left(F(\Theta_{ij}f, \mu) - F(f, \mu) \right) + s \sum_{i=1}^{n} F(V_{in}f, \mu).$$

For $f \in C_{\mathrm{sim}}(E^{\mathbb{N}})$, with $\mathfrak{n}(f) = n$, and $f \in \mathcal{D}(A^n) \cap \mathcal{B}(E^n)$, let

$$(12.31) \;\; Kf := \sum_{i=1}^{n} A_i f + \gamma \sum_{j=1}^{n} \sum_{k \neq j} [\Theta_{jk}f - f] + s \sum_{i=1}^{n} [V_{in}f - f].$$

where $\widetilde{\Theta}_{jk}$, $\mathfrak{n}(f)$ are defined as in section 7.5.
If $\beta(f) := s\mathfrak{n}(f)$, then

$$(12.32) \;\; GF(f, \mu) = F(Kf, \mu) + \bar{V}(\mathfrak{n}(f))F(f, \mu),$$

and $\sup_{\mu \in M_1(E)} |F(Kf, \mu)| \leq \mathrm{const} \cdot \mathfrak{n}(f)$.
   The function-valued dual process is constructed as follows. Let N be a jump Markov process taking non-negative integer values with transition intensities

$$(12.33) \;\; q_{m,m-1} = \gamma m(m - 1), \; q_{m,m+2} = sm, \; q_{i,j} = 0 \text{ otherwise.}$$

For $1 \leq i \leq m$, let $\{\tau_k\}$ be the jump times of N, $\tau_0 = 0$, and let $\{\Gamma_k\}$ be a sequence of random operators which are conditionally independent given N and satisfy

$$(12.34) \;\; P(\Gamma_k = \widetilde{\Theta}_{ij}|N) = \frac{2}{N(\tau_k-)N(\tau_k)} \mathbf{1}_{N(\tau_k-)-N(\tau_k)=1}, \quad 1 \leq i < j \leq N(\tau_k-)$$

$$(12.35) \;\; P(\Gamma_k = V_{im}|N) = \frac{1}{m} \mathbf{1}_{\{N(\tau_k-)=m, N(\tau_k)=m+2\}}.$$

   For $f \in C_{\mathrm{sim}}(E^{\mathbb{N}})$, define the $C_{\mathrm{sim}}(E^{\mathbb{N}})$-valued process Y with $Y(0) = f$ by

$$(12.36) \;\; Y(t) = S_{t-\tau_k}\Gamma_k S_{\tau_k - \tau_{k-1}}\Gamma_{k-1} \ldots \Gamma_1 S_{\tau_1}f, \quad \tau_k \leq \tau_{k+1}.$$

   We then have the dual representation: for $f \in C_{\mathrm{sim}}(E^{\mathbb{N}})$, define the $C_{\mathrm{sim}}(E^{\mathbb{N}})$-valued process Y with $Y(0) = f$ by

$$(12.37) \;\; Y(t) = S_{t-\tau_k}\Gamma_k S_{\tau_k - \tau_{k-1}}\Gamma_{k-1} \ldots \Gamma_1 S_{\tau_1}f, \quad \tau_k \leq \tau_{k+1}.$$

Then for any solution P_μ to the martingale problem for G and $f \in C_{\mathrm{sim}}(E^{\mathbb{N}})$ we get the FK-dual representation

$$(12.38) \;\; P_\mu[F(f, X(t))] = Q_f \left[F(Y(t), \mu) \exp \left(\bar{V} \int_0^t \mathfrak{n}(Y(u))du \right) \right]$$

12.2.1 Girsanov formula for Fleming-Viot with Mutation and Selection

Recall that the Fleming-Viot martingale problem $\mathbb{MP}_{(A,\gamma Q,0)}$ corresponds to the case

$$\langle M(A), M(A)\rangle_t = \gamma \int_0^t Q(X_s; A, A)ds$$

where

$$Q(\mu; dx, dy) = \mu(dx)\delta_x(dy) - \mu(dx)\mu(dy).$$

and that M is a worthy martingale measure.

Now consider a time-dependent diploid fitness function $V : [0, \infty) \times E \times E \to \mathbb{R}$ with $\|V\|_\infty < \infty$. . Then the FV martingale problem $\mathbb{MP}_{(A,Q,V)}$ is

$$M^V(\phi)_t$$

$$:= \langle X_t, \phi\rangle - \int_0^t \langle X_s, A\phi\rangle \, ds$$

$$- \int_0^t \int \left[\int V(s, x, y)X_s(dy) - \int \int V(s, y, z)X_s(dy)X_s(dz)\right] \phi(x)X_s(dx)ds$$

$$= \langle X_t, \phi\rangle - \int_0^t \langle X_s, A\phi\rangle \, ds$$

$$- \int_0^t \int \int \left[\left(\int \frac{V(s, y, z)}{\gamma}X_s(dz)\right) \gamma Q(X_s, dx, dy)\right] \phi(x)ds$$

$$\left\langle M^V(\phi)\right\rangle_t = \gamma \int_0^t \int \int \phi(x)\phi(y)Q(X_s, dx, dy)ds.$$

We then apply Girsanov to conclude that this martingale problem has a unique solution $\mathbb{P}^V$ and that the Radon-Nikodym derivative

$$Z_t^V := \frac{d\mathbb{P}^V}{d\mathbb{P}^0}|_{\mathcal{F}_t}$$

where $\mathbb{P}^0$ is the unique solution to $\mathbb{MP}_{(A,\gamma Q,0)}$ is given by

$$Z_t^V := \exp\left(\frac{1}{\gamma} \int_0^t \int V(s, X_s, y)M^0(ds, dy)\right.$$

$$\left. -\frac{1}{2\gamma^2} \int_0^t \int \int V(s, X_s, x)V(s, X_s, y)\gamma Q(X_s; dx, dy)ds\right).$$

where we write

$$V(s, X_s, x) = \int V(s, z, x)X_s(dz).$$

12.3   Long-time behaviour of systems with finite population resampling, mutation and selection

Systems with finite population resampling can have rather different long-time behaviour than the corresponding infinite population systems. One essential difference is that even high fitness types can be lost due to resampling and in the absence of mutation the system can eventually become unitype. On the other hand if the mutation process can regenerate all types, then the system can reach equilibrium in which all types are present. We now consider these two situations.

12.3.1   Fixation in finite population systems without mutation

In the previous section we have considered the infinite population system with selection but no mutation. In this case Fisher's fundamental theorem states that such a system evolves to one of maximal population fitness. But what happens in the finite population case, $\gamma > 0$? We first observe that if $V \equiv 0$, then $\{X_t(A) : t \geq 0\}$ is a bounded martingale and

$$(12.39)\ \ X_t(A) \underset{t \to \infty}{\longrightarrow} \begin{cases} 1 \text{ with probability } X_0(A) \\ 0 \text{ with probability } (1 - X_0(A)). \end{cases}$$

Therefore

$$X_t \overset{t \to \infty}{\Longrightarrow} \delta_x \text{ with } x \in A \text{ with probability } X_0(A)$$

that is, the system experiences ultimate "fixation". If we add selection to this, ultimate fixation still occurs. However if γ is small then the tendency is for the limiting types to be those of higher fitness.

12.3.2   The Equilibrium Infinitely Many Alleles Model with Selection

In order to have a non-degenerate equilibrium a source of new types through mutation is required. In this section we consider the type independent infinitely many alleles mutation together with selection. If ν_0 is a non-atomic measure, then mutation always leads to a new type and thus provides a mechanism to guarantee sufficient diversity on which selection can act.

Let $\mathbb{P}_\infty^0$ denote the probability measure on $C_{\mathcal{P}([0,1])}(-\infty, \infty)$ corresponding to the reversible stationary measure, with one dimensional marginal distribution $\Pi_\gamma^0(d\mu)$, for the neutral infinitely many alleles model (recall the representation in terms of the Moran subordinator). Assume that V is symmetric and $V(s, x, y) = V(x, y) = V(y, x)$.

The following results is the infinitely many types analogue of a result of Wright [616].

Theorem 12.15 *The infinitely many alleles model with selection has a reversible stationary measure given by*

$$\Pi_\gamma^V(d\mu) = \frac{1}{Z} e^{\frac{V(\mu)}{\gamma}} \Pi_\gamma^0(d\mu)$$

where Z is a normalizing constant.

Proof. Let X_0 have distribution

$$\frac{1}{Z} e^{\frac{V(X_0)}{\gamma}} \Pi_\gamma^0(dX_0)$$

Recall that to verify that this is a reversible equilibrium measure it suffices to show that for any two continuous functions, f and g, on $[0, 1]$

$$\mathbb{P}_\infty(f(X_0)g(X_t)) = \mathbb{P}_\infty(g(X_0)f(X_t)).$$

But

$$\mathbb{P}_\infty(f(X_0)g(X_t))$$
$$= \frac{1}{Z} \int f(X_0) e^{\frac{V(X_0)}{\gamma}} g(X_t) \mathbb{P}_{X_0}^V(d\{X_s : 0 \le s \le t\}) \Pi_\gamma^0(dX_0)$$
$$= \frac{1}{Z} \int f(X_0) e^{\frac{V(X_0)}{\gamma}} Z_t^V \mathbb{P}_\infty^0(dX.) g(X_t).$$

By Girsanov

$$Z_t^V := \exp\left(\frac{1}{\gamma} \int_0^t \int V(X_s, y) M^0(ds, dy) \right.$$
$$\left. - \frac{1}{2\gamma^2} \int_0^t \int \int V(X_s, x) V(X_s, y) \gamma Q(X_s; dx, dy) ds \right)$$

where

$$M_s^0(dy) = X_s - \int_0^s A^* X_u du$$
$$= X_s - \int_0^s c[\nu_0 - X_u] du$$

As a preparation, note that by Ito's lemma,

$$d_t\left(\int \int V(x, y) X_t(dx) X_t(dy) \right)$$
$$= \int \int V(x, y) X_t(dx) d_t X_t(dy) + \int \int V(x, y) X_t(dy) d_t X_t(dx)$$
$$+ \int \int V(X_s, x) V(X_s, y) \gamma Q(X_s; dx, dy)$$

Hence by symmetry in x and y and Ito's Lemma,

$$\frac{1}{\gamma}\int_0^t\int\int V(x,y)X_s(dx)d_sX_s(dy) = \frac{1}{2\gamma}[\int\int V(x,y)X_t(dx)X_t(dy)$$

$$-\int\int V(x,y)X_0(dx)X_0(dy)]$$

$$-\frac{1}{2\gamma^2}\int_0^t\int\int V(X_s,x)V(X_s,y)\gamma Q(X_s;dx,dy)$$

Therefore

$$\log(e^{\frac{1}{\gamma}V(X_0)}Z_t^V)$$

$$= \int_0^t\int\frac{1}{\gamma}V(X_s,y)M^0(ds,dy) + \frac{1}{\gamma}\int\int V(x,y)X_0(dx)X_0(dy)$$

$$-\frac{1}{2\gamma^2}\int_0^t\int\int V(X_s,x)V(X_s,y)\gamma Q(X_s;dx,dy)ds$$

$$= \frac{1}{\gamma}\int_0^t\int\int V(x,y)X_s(dx)d_sX_s(dy) - \frac{c}{\gamma}\int_0^t\int\int V(x,y)X_s(dx)(\nu_0(dy)-X_s(dy))ds$$

$$-\frac{1}{2\gamma^2}\int_0^t\int\int V(X_s,x)V(X_s,y)\gamma Q(X_s;dx,dy)ds\Big) + \frac{1}{\gamma}\int\int V(x,y)X_0(dx)X_0(dy)$$

$$= \frac{1}{2\gamma}[\int\int V(x,y)X_t(dx)X_t(dy) + \int\int V(x,y)X_0(dx)X_0(dy)]$$

$$-\frac{1}{\gamma^2}\int_0^t\int\int V(X_s,x)V(X_s,y)\gamma Q(X_s;dx,dy)ds$$

$$-\frac{c}{\gamma}\int_0^t\int\int V(x,y)X_s(dx)(\nu_0(dy)-X_s(dy))ds$$

This is symmetric with respect to the direction of time. Also under $\mathbb{P}_\infty^0$, $\{X_t : t \in \mathbb{R}\}$ is stationary and reversible. Therefore we conclude that

$$E(f(X_0)g(X_t)) = E(f(X_t)g(X_0))$$

Therefore $\frac{1}{Z}e^{\frac{V(\mu)}{2\gamma}}\Pi_\gamma^0(d\mu)$ is a reversible invariant measure.   ∎

Corollary 12.16 *Consider the K-allele case with $c = \gamma$ and $\nu_0(dx) = dx$. Assume that $V(p)$ is continuous and has a unique global maximum $p_0 \in \Delta_{K-1}$. Then as $\gamma \to 0$, $\Pi_\gamma^V \Longrightarrow \delta_{p_0}$.*

Proof. In this case $\Pi_\gamma^0(dp)$ is the Dirichlet (1) distribution on Δ_{K-1}. Let

$$N_p^\varepsilon := \{p : V(p_0) - V(p) \le \varepsilon\}$$

Then for any $\varepsilon > 0$, $\Pi_\gamma^0(N_{p_0}^\varepsilon) > 0$. It is then easy to check that

$$\Pi_\gamma^V((N_{p_0}^\varepsilon)^c) \to 0 \text{ as } \gamma \to 0.$$

∎

Remark 12.17 *One can ask if there is a reversible equilibrium for other mutation processes. The fact that the only mutation process for which the equilibrium is reversible was proved by Li, Shiga and Ya (1999) [432]. This is the analogue of the result of Hofbauer and Sigmund mentioned above.*