Rate of Adaptation Under Weak Selection

Etheridge, Alison
Department of Statistics, University of Oxford
1 South Parks Road
Oxford, OX1 3TG, UK
E-mail: etheridg@stats.ox.ac.uk

Yu, Feng
Department of Mathematics, University of Bristol
University Walk
Bristol, BS8 3AN, UK
E-mail: feng.yu@bristol.ac.uk

1 Introduction

Population genetics is the study of genetic composition of populations, which may be affected by selection, mutation, recombination, migration and other genetic, ecological and evolutionary factors. It is a vital ingredient in the modern synthesis of Mendelian genetics and Darwinian evolution. Mathematical thinking and methodology have been entrenched in population genetics since its early days. The past 30 years have seen a flourishing of the application of mathematical, especially probabilistic, techniques to the study of population genetics and indeed to problems from across the biological sciences.

Recently there has been much renewed interest among both biologists (see e.g. Gerrish & Lenski 1998, Wilke 2004) and physicists (see e.g. Higgs & Woodcock 1995, Rouzine *et al* 2003, Desai & Fisher 2007, Brunet *et al* 2008, Park *et al* 2010) in an old question in the mathematical theory of evolution: how quickly can large asexual populations adapt to a novel environment by incorporating beneficial mutations? This elementary question is surprisingly difficult to answer, even for the simple 'toy models' usually discussed. This difficulty is partly due to the involvement of selection – population genetics models are rendered nonlinear by selection – and partly due to its stochastic nature. Taking the infinite population limit and obtaining a deterministic system, if possible at all, reveals very little about the behaviour of the system for a finite population size. An essential assumption (to make the problem mathematically tractable) that is generally made is that the effect of mutations on log-fitness is additive. In this case the type of an individual is characterised by its location within the fitness space (which can be thought of as $\mathbb{Z}$). For small selection coefficients, the situation is then analogous to the noisy travelling front problem which is extensively studied in statistical physics. In that setting, the behaviour at the front of the wave has a subtle yet important stochastic effect on the speed of the wave, which equates to the rate of adaptation in the biological context. Much of the recent work on the rate of adaptation problem was performed by physicists. This body of work provides many valuable insights and useful formulae, but it is non-rigorous and focusses on the asymptotics in the (very) large population limit. There are remarkably few rigorous results available. Our goal here is to outline a mathematically rigorous route to an expansion formula for the rate of adaptation for a Fleming-Viot process with selection.

We consider an asexually reproducing population which is not spatially structured and which is of constant size. We are concerned with the effects of the combined forces of mutation, genetic drift and natural selection. Mutation constantly introduces variation in traits, whereas genetic drift and natural selection causes variants to be more or less common in the population. Genetic drift is directionless and produces entirely random changes in the frequency of a trait. Natural selection, on the other hand, tends to increase the frequency of traits that give its carrier a higher *fitness*, thus

causing a species to adapt to its environment. In this context, fitness is a measure of the reproductive success of an organism. A *beneficial* mutation increases the fitness of its carrier by increasing the number of offspring that survive to adulthood, *deleterious* mutations decrease fitness, while *neutral* mutations have no effect on fitness. Although we have a fairly detailed understanding of the interplay between mutation and genetic drift, models that incorporate selection acting on multiple loci are much less tractable, even in asexually reproducing populations. In this case, different selected alleles compete with each other, resulting in 'clonal interference' (Gerrish & Lenski 1998). A shortage of tools for rigorous analysis has resulted in a poor understanding of such models.

The process whereby a beneficial mutation arises (in what is generally assumed to be a large and otherwise neutral population) and eventually spreads to the entire population is called a selective sweep. If mutations are so rare that a selected mutation is likely to spread to the entire population or go extinct before the next mutation, then each mutation can be considered in isolation. In this case, the proportion of the population possessing this mutation can be studied using the diffusion approximation and applying well-known results from one-dimensional diffusion theory. For isolated selective sweeps, the fixation probability (i.e. the probability of the mutation spreading to the entire population) is not sensitive to population size (provided it is reasonably large). The problem is far more difficult if one considers overlapping selected sweeps at linked loci. If there is no recombination, in order for two favoured mutations to fix, one must arise in an individual carrying the other. This provides an evolutionary advantage to sex and recombination (the Hill-Robertson effect). Even in the presence of recombination, fixation probabilities can be greatly reduced. Previous work (Barton 1995, Yu & Etheridge 2010, Cuthbertson *et al.* 2010) shows that in the presence of recombination, even understanding the case of two overlapping sweeps is very involved.

If the population size is large, then new mutations will continually fall on the population before the fate of mutations already present in the population has been decided. The rate of adaptation is defined to be the expected increase of the mean fitness of the population per generation. It has been studied since the 1930's. In 1930, Fisher stated his famous *Fundamental Theorem of Natural Selection*: 'The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time.' This law, which appears in calculations of almost every subsequent model with selection, uncovers a considerable challenge: as observed in Higgs & Woodcock (1995), the moments of the fitness distribution do not form a closed system of equations (unlike the neutral case, see discussion after (1) below).

2 The Model

We first describe an individual based model. For mathematical convenience we will use a Moran (overlapping generation) model rather than the (discrete generation) Wright-Fisher model that is more commonly employed in the biology literature. Since we will later pass to a diffusion approximation, this does not affect our results. Let $\tilde{\mu}$ and $\tilde{s}$ be the *unscaled* mutation and selection parameters, respectively. When parameters are small, one typically approximates the Wright-Fisher model by separating the mutation and reproduction mechanisms, and letting the waiting time between 'clicks' of each mechanism be exponentially distributed. We assume an 'infinitely-many-sites' model in which each new mutation arises at a different location on the genome and, further, we follow previous work by assuming that each mutation will either increase or decrease an individual's log-fitness by the fixed amount $\tilde{s}$. Our unscaled model involves the following two mechanisms. For each individual i in a population $\{1, \ldots, N\}$ of constant size N, we write X_i for its current 'fitness class' (that is the number of favourable mutations minus the number of unfavourable mutations that it carries). It can experience two types of event:

1' a mutation event at rate $\tilde{\mu}$: with probability q, X_i changes to $X_i + 1$; otherwise X_i changes to $X_i - 1$;

2' a reproduction event at rate 1: an arbitrary individual j is picked, and then individual i replaces individual j with probability $(0 \vee \frac{1}{2}(1+\tilde{s}(X_i - X_j))) \wedge 1$; otherwise, individual j replaces individual i.

The time scaling here is such that one unit of time roughly corresponds to one generation. One may study this unscaled particle model directly, e.g. in Yu *et al* (2010), we established an asymptotic lower bound of $\log^{1-\delta} N$ (δ is any positive number and N is the population size) on the rate of adaptation for large populations. However, a drawback of the particle system approach is that it does not yield explicit expressions for the rate of adaptation.

In this work, as first outlined in Yu & Etheridge (2008), we take the same model and apply a 'diffusion approximation'. First we speed up time by a factor N. Writing $\mu = N\tilde{\mu}$ and $s = N\tilde{s}$ for the scaled mutation and selection parameters, we obtain that each individual i experiences:

1. a mutation event at rate μ: with probability q, X_i changes to $X_i + 1$; otherwise X_i changes to $X_i - 1$;

2. a reproduction event at rate N: an arbitrary individual j is picked, and then individual i replaces individual j with probability $\frac{1}{2}(1 + \frac{s}{N}(X_i - X_j))$; otherwise, individual j replaces individual i.

We then consider *frequencies* of individuals in each fitness class in the population and take the limit as $N \to \infty$. Mechanism 2, like 2', should really include $0\vee$ and $\wedge 1$. However, $|X_i - X_j|$ is overwhelmingly likely to be much smaller than $\mathcal{O}(N)$. To see why, notice that under neutral reproduction, the genealogy of the whole population is given by Kingman's coalescent and so the entire population had its most recent common ancestor at a time of $\mathcal{O}(1)$ before the present. The differences between the numbers of mutations accumulated along different lineages is then of the same order. The presence of selection only serves to shorten the genealogical tree and so reduce variability in fitness. Diffusion approximations are advantageous in two ways. First, they are robust in the sense that many population models with similar population dynamic features are well approximated by the same diffusion approximation when population size is large. Second, in the resulting large population limit where $\mu = N\tilde{\mu}$ and $s = N\tilde{s}$ are both held constant as $N \to \infty$, the frequency of alleles is described by a (possibly multi- or infinite-dimensional) diffusion process with the stochastic effect of genetic drift preserved under passage to the limit.

If we take $\tilde{s} = 0$, then we obtain the neutral model, where all individuals in the population are equally successful at reproducing and quantities such as the long-term rate of accumulation of mutations can be easily calculated. It is simply the rate of mutations falling on a single individual, since most mutations that appear in an individual will have fallen on the common ancestor of all individuals prior to the time when all lineages from the present have merged.

A model with beneficial and deleterious mutations in addition to neutral ones is much more mathematically challenging. The main reason for this is that the probability distribution of the number of offspring of each individual will depend on its fitness at the time of reproduction, which in turn depends on the genealogy of the population prior to that time. In other words, the key property that makes the neutral model amenable to analysis, that is the independence of the genealogy and mutation processes, no longer holds in the selected case. Of all previous work on the rate of adaptation of which we are aware, that most closely related to our approach is Rouzine *et al* (2003). They treat P_k (the proportion of individuals of fitness class k) as a travelling wave in fitness space that has a roughly Gaussian shape. The variance of P_k is determined by stochastic effects at the edge of the wave where the population density is $\mathcal{O}(1/N)$. The travelling wave bears some similarity to the solution

to the stochastic Fisher-KPP equation, in which stochastic effects at the front determine the speed of the wave. Here, however, the analysis is considerably more difficult since, as $N \to \infty$, the speed of the wave is roughly $\log N$ (instead of approaching a limit as in the Fisher-KPP case), so that there is no limiting solution. The critical assumption in Rouzine *et al* (2003) is that the shape of the wave is approximately deterministic (i.e. the variance of P is approximately a constant rather than a random variable) when population size is very large. This assumption has not been verified mathematically and as a result it seems difficult to make this approach rigorous.

If mutations in a finite asexual population are all deleterious, then selection, perhaps counter-intuitively, is unable to counter-balance the effect of constant accumulation of deleterious mutations and the mean fitness of the population deteriorates inexorably. This is commonly known as Muller's ratchet, which was first described by H. J. Muller in 1964. Due to page constraints, we will not go into further details. Muller's ratchet gives a very appealing explanation for the advantage of sexual reproduction and the evolution of recombination. However, just as the case of beneficial mutations, the process is surprisingly difficult to analyse for one so simply defined. There are almost no rigorous results in the literature in the finite population case. For example, the rate of deterioration of the mean fitness is not known.

In this work, we start with a diffusion approximation (obtained by passing to a limit from mechanisms 1 and 2) and follow a rigorous approach. Through some standard calculations involving martingale decomposition (see e.g. Yu & Etheridge 2008 for more details), we obtain that P_k, the proportion of individuals with fitness type k, obeys the following infinite system of interacting SDE's:

$$(1) \quad dP_k = [\mu(qP_{k-1} - P_k + (1-q)P_{k+1}) + s(k - m(P))P_k]\ dt + \sum_{l\in\mathbb{Z}} \sqrt{P_k P_l}\ dW_{kl},\ k \in \mathbb{Z},$$

where $m(p) = \sum_k kp_k$ is the mean of distribution p, $\{W_{kl} : k, l \in \mathbb{Z}, k > l\}$ are independent Brownian motions and $W_{kl} = -W_{lk}$. For convenience we also set $W_{kk} = 0$. The terms on the right hand side of (1) correspond to the effects of, respectively, mutation (individuals with fitness type k change to type $k+1$ at rate μq and to type $k-1$ at rate $\mu(1-q)$), selection (the size of fitness class k increases/decreases exponentially at rate $s(k - m(P))$), and genetic drift (the stochastic term).

From (1), one obtains an evolution equation for the mean fitness $m(P)$ of the population:

$$\begin{aligned} dm(P) &= d\sum_k kP_k \stackrel{m}{=} \sum_k k[\mu(qP_{k-1} - P_k + (1-q)P_{k+1}) + s(k - m(P))P_k]\ dt \\ &= (\mu(2q-1) + sc_2(P))\ dt \end{aligned}$$

where $\stackrel{m}{=}$ means the left and right hand sides differ by a martingale (hence have the same expectation), and $c_2(p)$ is the variance of distribution p. This is consistent with Fisher's Fundamental Theorem of natural selection: the mean fitness $m(P)$ evolves at a speed proportional to the fitness variance $c_2(P)$ of the population. This results from the nonlinearity of the selection term. More generally, the moments of P, unlike in the neutral case, do not form a closed system: the first moment depends on the second, the second moment on the third, and so on.

3 Application of Girsanov Transformation

Recall that the Girsanov Theorem is a powerful tool in probability theory with wide-ranging applications and no counterpart in classical calculus. It tells us how solutions to SDE's change under change of probability measures. More specifically, it allows us to transform the probability measure that corresponds to the solution of one SDE to that corresponding to an SDE with the same noise component, but different drift component (drift here in the mathematical sense). This transformation of probability measures can be extended to multi- or even infinite-dimensional diffusions. We shall use

the infinite-dimensional version of Girsanov Theorem, developed by Dawson (see e.g. Dawson 1993) for superprocesses. In fact, Dawson (1993) remarks that a Girsanov transform can be applied to a neutral model to obtain one with selection, but to the best of our knowledge this has not been fully exploited in the study of models with selection.

Let $\mathbb{P}_s$, $\mathbb{E}_s$, and $\mathcal{A}_s$ denote respectively the probability measure, the expectation under $\mathbb{P}_s$, and the generator corresponding to solutions of (1). Moreover, let $\tilde{P}$ be P centred about its mean, $\tilde{\mathbb{P}}_s$ be the probability measure of $\tilde{P}$, and $\tilde{\mathbb{E}}_s$ be the corresponding expectation. In particular, $\mathbb{P}_0$ is the probability measure corresponding to solutions of the SDE system evolving according to the neutral generator $\mathcal{A}_0$:

$$(2) \quad dP_k = \mu(qP_{k-1} - P_k + (1-q)P_{k+1})\ dt + \sum_{l\in\mathbb{Z}} \sqrt{P_k P_l}\ dW_{kl},\ k \in \mathbb{Z},$$

where the W_{kl}'s are as in (1). We define

$$(3) \quad M(t) = m(P(t)) - \mu(2q-1)t$$

to be the martingale part of the mean fitness (with the convention $M(0) = 0$) and Z to satisfy $dZ = sZ\ dM$, i.e.

$$(4) \quad Z(t) = \exp\left\{ sM(t) - \frac{s^2}{2}\int_0^t c_2(P(u))\ du \right\}.$$

Then provided that $Z(t)$ is a martingale, Dawson's Girsanov Theorem states that

$$\left.\frac{d\mathbb{P}_s}{d\mathbb{P}_0}\right|_{\mathcal{F}_t} = Z(t).$$

In one-dimensional stochastic calculus, in order to verify that $Z(t)$ is a martingale, one usually checks the Novikov condition: $\mathbb{E}_0\left[\exp\left\{\langle M\rangle(t)\right\}\right] < \infty$ for all $t > 0$. In our case, however, it is easier to verify the Kazamaki condition:

$$\mathbb{E}_0\left[e^{\frac{s}{2}M(t)}\right] < \infty \text{ for all } t > 0.$$

The Kazamaki condition above is not immediate in this infinite-dimensional case and we spend a paragraph giving an outline of the proof of its validity. We shall use Kingman's coalescent (Kingman 1982) to calculate $M(t)$. For simplicity, we assume that all mutations are beneficial (i.e. $q = 1$), but this is not necessary. Recall that in the neutral model the genealogy of the population at time t is determined by Kingman's coalescent and mutations can be superposed according to an independent Poisson process of rate μ along each ancestral lineage in the genealogy. For $k \in \mathbb{N}$, let T_k ($T_1 < T_2 < \ldots < t$) be the first time (before t) when there are k lineages left in the coalescent, and let $S_k = 0 \vee T_k$ (with $S_0 = 0$). Let X_{kl} ($l = 1, \ldots, k$) be the number of mutations falling on the lth lineage during $(S_{k-1}, S_k]$. Then

$$m(P(t)) = \sum_{k=1}^{\infty}\sum_{l=1}^{k} a_{kl}X_{kl},$$

where $a_{kl} \geq 0$ denotes the 'weight' of lineage l during $(S_{k-1}, S_k]$, that is the proportion of the population at time t which is descended from the lineage labelled l at time S_k. In particular, $\sum_{l=1}^{k} a_{kl} = 1$. Since conditioning on $\{T_k, k \in \mathbb{N}\}$, all the X_{kl}'s are independent with $X_{kl} \sim Poisson(\mu(S_k - S_{k-1}))$,

we have

$$\begin{aligned}
\mathbb{E}_0\left[e^{\frac{s}{2}M(t)}\middle|\, T_k\text{'s, } a_{kl}\text{'s}\right] &= \mathbb{E}_0\left[e^{\frac{s}{2}(m(P(t))-\mu t)}\middle|\, T_k\text{'s, } a_{kl}\text{'s}\right] = \mathbb{E}_0\left[e^{-\frac{\mu st}{2}}\prod_{k=1}^{\infty}\prod_{l=1}^{k}e^{\frac{s}{2}a_{kl}X_{kl}}\middle|\, T_k\text{'s, } a_{kl}\text{'s}\right] \\
&= e^{-\frac{\mu st}{2}}\prod_{k=1}^{\infty}\prod_{l=1}^{k}\exp\{\mu(S_k-S_{k-1})(e^{sa_{kl}/2}-1)\} = e^{-\frac{\mu st}{2}}\prod_{k=1}^{\infty}\exp\left\{\mu(S_k-S_{k-1})\sum_{l=1}^{k}(e^{sa_{kl}/2}-1)\right\} \\
&\leq e^{-\frac{\mu st}{2}}\prod_{k=1}^{\infty}\exp\left\{\mu(S_k-S_{k-1})(e^{s/2}-1)\right\} = e^{-\frac{\mu st}{2}}\exp\left\{\mu t(e^{s/2}-1)\right\}.
\end{aligned}$$

Hence $\mathbb{E}_0\left[e^{\frac{s}{2}M(t)}\right] < e^{-\frac{\mu st}{2}}\exp\left\{\mu t(e^{s/2}-1)\right\} < \infty$, and Kazamaki's condition is verified.

4 The Expansion Formula

The process Z in (4) can be expanded as a power series in s. We define

$$M^{(n)}(t) = \int_0^t\int_0^{t_1}\cdots\int_0^{t_{n-1}} dM(t_n)\cdots dM(t_2)\, dM(t_1) \tag{5}$$

to be the n^{th} iterated integral of the martingale M, with the convention that $M^{(0)} = 1$ and $M^{(1)} = M$. Then

$$Z(t) = 1 + \sum_{n=1}^{\infty} s^n M^{(n)}.$$

This expansion formula for Z can be verified by simply taking the differential and noting that it satisfies $dZ = sZ\, dM$. The expected mean fitness of the model with selection becomes

$$\mathbb{E}_s[m(P(t))] = \mathbb{E}_0[m(P(t))Z(t)] = \mu(2q-1)t + \sum_{n=1}^{\infty} s^n\mathbb{E}_0[M^{(n)}(t)M^{(1)}(t)]. \tag{6}$$

The use of the iterated integral representation for Z is strongly reminiscent of Wiener chaos expansion (see e.g. Nualart 1995), where noise is taken to be Brownian motion W. This has been central to the theory of Malliavin calculus, thanks in part to the orthogonality of the iterated integrals of W, i.e. $\mathbb{E}[M^{(m)}(t)M^{(n)}(t)] = 0$ for $m \neq n$ if we take replace M by W in (5). The orthogonality is mainly due to the fact that Brownian motion has stationary and independent increments. Instead, in the case of the noise term M defined in (3), we have

$$\begin{aligned}
d\langle M\rangle &= d\left\langle\sum_{k,l} k\sqrt{P_kP_l}\, W_{kl}\right\rangle = \sum_{k,l}\sum_{k',l'} kk'\sqrt{P_kP_l}\sqrt{P_{k'}P_{l'}}\, d\langle W_{kl}, W_{k'l'}\rangle \\
&= \sum_{k,l}(k^2-kl)P_kP_l\, dt = c_2(P)\, dt,
\end{aligned}$$

where the third equality above is because for $k \neq l$, $d\langle W_{kl}, W_{k'l'}\rangle = 1$ if $kl = k'l'$, -1 if $kl = l'k'$, and 0 otherwise. This means that the evolution of M at any time depends on the fitness variance of the population at that time, which in turn depends on the evolution of the population prior to that time. It also means that the $M^{(n)}$'s are not orthogonal. In order to calculate each term in the expansion (6), we observe that

$$d(M^{(n)}M^{(1)}) = M^{(n-1)}d(M^{(1)})^2 \overset{m}{=} M^{(n-1)}c_2(P)\, dt.$$

Therefore the adaptation rate in the selected model (with selection coefficient s) is equal to

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t}\mathbb{E}_s[m(P(t))] &= (2q-1)\mu + \lim_{t\to\infty}\sum_{n=1}^{\infty} s^n \mathbb{E}_0[M^{(n-1)}(t)c_2(P(t))] \\ &= (2q-1)\mu + \lim_{t\to\infty}\sum_{n=1}^{\infty} s^n \tilde{\mathbb{E}}_0[M^{(n-1)}(t)c_2(P(t))]. \qquad (7)\end{aligned}$$

Shiga (1982) showed that there is an ergodic stationary distribution for the neutral process centred about its mean and from now on we abuse notation by writing $\tilde{\mathbb{E}}_0$ for expectations with respect to this stationary distribution. The problem now boils down to calculating $\tilde{\mathbb{E}}_0[M^{(n)}c_2(P)]$ for $n \in \mathbb{Z}^+$. The formula in (7) is rigorous for s in a certain radius of convergence (which may be ∞), although the proof will appear elsewhere.

Rather than working with central moments, it is easier to work with the cumulants κ_n of p. To see how these are defined, suppose for simplicity that p is a distribution on $\mathbb{Z}$. Then the cumulants κ_n are defined through the cumulant generating function:

$$g(x) = \log\sum_k p_k e^{kx} = \sum_{n=1}^{\infty}\kappa_n\frac{x^n}{n!}.$$

In particular, $\kappa_2 = c_2$, $\kappa_3 = c_3$, and for $n \geq 4$, κ_n is an n^{th}-degree polynomial in the central moments $c_2, \ldots, c_n$. Since

$$\mathcal{A}_0 f(p) = \mu\sum_k (qp_{k-1} - p_k + (1-q)p_{k+1})\frac{\partial f(p)}{\partial p_k} + \frac{1}{2}\sum_{k,l} p_k(\delta_{kl} - p_l)\frac{\partial^2 f(p)}{\partial p_k \partial p_l},$$

we can calculate the effect of $\mathcal{A}_0$ on the cumulants, using the cumulant generating function $g(x)$, to obtain:

$$\mathcal{A}_0 g(x) = \mu(qe^x - 1 + (1-q)e^{-x}) + (1 - e^{g(2x)-2g(x)}).$$

In particular, differentiating the above twice with respect to x yields $\mathcal{A}_0 c_2 = \mu - c_2$, which implies that $\tilde{\mathbb{E}}_0[c_2] = \mu$. Similarly, for the next few cumulants, we obtain

$$\begin{aligned}\mathcal{A}_0\kappa_3 &= \mu(2q-1) - 3\kappa_3 \\ \mathcal{A}_0\begin{bmatrix}\kappa_4 \\ \kappa_2^2\end{bmatrix} &= \begin{bmatrix}\mu \\ 2\mu\kappa_2\end{bmatrix} + \begin{bmatrix}-7 & -6 \\ 1 & 0\end{bmatrix}\begin{bmatrix}\kappa_4 \\ \kappa_2^2\end{bmatrix} \\ \mathcal{A}_0\begin{bmatrix}\kappa_5 \\ \kappa_3\kappa_2\end{bmatrix} &= \begin{bmatrix}(2q-1)\mu \\ (2q-1)\mu\kappa_2 + \mu\kappa_3\end{bmatrix} + \begin{bmatrix}-15 & -10 \\ 1 & 0\end{bmatrix}\begin{bmatrix}\kappa_5 \\ \kappa_3\kappa_2\end{bmatrix}.\end{aligned}$$

If we write $i = (i_1, \ldots, i_I)$ and $\kappa_i = \kappa_{i_1}\ldots\kappa_{i_I}$, then theoretically we can such solve linear systems for each $|i|$ and obtain expressions for all $\tilde{\mathbb{E}}_0[\kappa_i]$. However, it remains unclear whether there is a systematic way of solving these systems analytically.

The quantities $\tilde{\mathbb{E}}_0[\kappa_i]$ are explicitly related to the terms $\tilde{\mathbb{E}}_0[M^{(n)}c_2(P)]$ in the expansion (6). For example, for $n = 1$, because

$$d(M\kappa_2) \overset{m}{=} M\ d\kappa_2 + d\langle M, \kappa_2\rangle = (-M\kappa_2 + \kappa_3)\ dt,$$

we have

$$\tilde{\mathbb{E}}_0[M\kappa_2] = \tilde{\mathbb{E}}_0[\kappa_3] = \mu(2q-1)/3.$$

Similar calculations imply that

$$\begin{aligned}\tilde{\mathbb{E}}_0[M^{(2)}\kappa_2] &= \tilde{\mathbb{E}}_0[M^{(1)}\kappa_3] = \frac{1}{3}\tilde{\mathbb{E}}_0[\kappa_4] = -\frac{2}{3}\mu^2 \\ \tilde{\mathbb{E}}_0[M^{(3)}\kappa_2] &= \tilde{\mathbb{E}}_0[M^{(2)}\kappa_3] = \frac{1}{3}\tilde{\mathbb{E}}_0[M^{(1)}\kappa_4] = \frac{1}{30}\tilde{\mathbb{E}}_0[\kappa_5 - 12\kappa_2\kappa_3] = (2q-1)(\frac{\mu^2}{3} + \frac{\mu}{20}).\end{aligned}$$

Recall that one of the hurdles to overcome in studying the behaviour of (1) is that the moment equations for P are not closed. Using the Girsanov transform, we are able to resolve this problem and write the rate of adaptation in terms of the moments of the *neutral* process, which *are* closed and can be solved, at least theoretically. The first few terms of the expansion (7) for the rate of adaptation are as follows:

$$(2q-1)\mu + s\mu + s^2(2q-1)\frac{\mu}{3} - s^3\frac{2\mu^2}{3} + s^4(2q-1)(\frac{\mu^2}{9} + \frac{\mu}{60}) + \dots$$

5 Future work: Incorporating Recombination

So far, we have only been concerned with asexual populations. A logical next step is to extend to sexual populations, which undergo recombination. As explained as early as 1889, by Weismann, sex does not increase the mean fitness directly (indeed it is costly) but it increases the variance in fitness upon which natural selection acts. Using the Girsanov transform, we hope to obtain an expansion formula for the rate of adaptation for a sexual population and compare that to the rate of adaptation of an asexual population to actually *quantify* the effect of recombination on the rate of adaptation. We will use single-crossover models presented in Baake & Baake (2003) as a basis. More concretely, we imagine chromosomes as a linear arrangement of n sites, thus the state space becomes $(k_1, \dots, k_n)$, where each k_α denotes the fitness class of site α. We will restrict to the simplest case involving only two sites and thus one crossover point. A recombination event randomly picks two individuals i and j and then replaces individual i with an individual of type (i_1, j_2). The effect of recombination should be to add a drift term to (1), if one scales things properly, so that one obtains the following system of SDE's for $P_{k,l}$, the proportion of individuals with k mutations at locus 1 and l mutations at locus 2:

$$\begin{aligned} dP_{k,l} &= [\mu_1(q_1 P_{k-1,l} - P_{k,l} + (1-q_1)P_{k+1,l}) + \mu_2(q_2 P_{k,l-1} - P_{k,l} + (1-q_2)P_{k,l+1}) \\ &\quad + s(k+l-m(P))P_k + \rho((P_{k1}+P_{k2})(P_{1l}+P_{2l}) - P_{kl})]\, dt + \sum_{k',l' \in \mathbb{Z}} \sqrt{P_{kl}P_{k'l'}}\, dW_{kl,k'l'}, \end{aligned}$$

where ρ is the scaled recombination rate, each mutation on either site has fitness effects $\pm s$, and $W_{kl,k'l'}$'s are independent Brownian motions such that $W_{kl,k'l'} = -W_{k'l',kl}$. Just as in the asexual case, solutions of the above SDE system can be obtained via the Girsanov transform from those of a system without selection, which should lead to an expansion formula analogous to (7).

References

[1] *N. H. Barton. Linkage and the limits to natural selection.* Genetics*, 140:821–841, 1995.*

[2] *E. Brunet, I. Rouzine, and C. Wilke. The stochastic edge in adaptive evolution.* Genetics*, 179:603–620, 2008.*

[3] *C. Cuthbertson, A. M. Etheridge, and Feng Yu. Fixation probability for competing selective sweeps.* Preprint, submitted to Electronic Journal of Probability, arXiv:0812.0104*, 2010.*

[4] *Donald A. Dawson. Measure-valued Markov processes. In* École d'Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of* Lecture Notes in Math.*, pages 1–260. Springer, Berlin, 1993.*

[5] *M. M. Desai and D. S. Fisher. Beneficial mutation-selection balance and the effect of linkage on positive selection.* Genetics*, 176:1759–98, 2007.*

[6] *R. A. Fisher.* The Genetical Theory of Natural Selection*. Clarendon Press, Oxford, 1930.*

[7] *P. J. Gerrish and R. E. Lenski. The fate of competing beneficial mutations in an asexual population.* Genetica*, 102/103:127–144, 1998.*

[8] Paul Higgs and Glenn Woodcock. The accumulation of mutations in asexual populations, and the structure of genealogical trees in the presence of selection. J. Math. Biol.*, 33:677–702, 1995.*

[9] D. Nualart. The Malliavin Calculus and Related Topics. *Probability and Its Applications. Springer, 1995.*

[10] Su-chan Park, Damien Simon, and Joachim Krug. The speed of evolution in large asexual populations. Journal of Statistical Physics*, 138:381–410, 2010.*

[11] I. Rouzine, J. Wakeley, and J. M. Coffin. The solitary wave of asexual evolution. Proceedings of the National Academy of Sciences*, 100(2):587–592, 2003.*

[12] Tokuzo Shiga. Wandering phenomena in infinite-allelic diffusion models. Adv. in Appl. Probab.*, 14(3):457–483, 1982.*

[13] A. Weismann. The significance of sexual reproduction in the theory of natural selection. In Essays upon heredity and kindred biological problems*, pages 251–332. Clarendon Press, Oxford, 1889.*

[14] Claus O. Wilke. The speed of adaptation in large asexual populations. Genetics*, 167:2045–2054, 2004.*

[15] Feng Yu and A. M. Etheridge. Rate of adaptation of large populations. In Evolutionary Biology from Concept to Application. *Springer, 2008.*

[16] Feng Yu and A. M. Etheridge. The fixation probability of two competing beneficial mutations. Theor. Popul. Biol.*, 78:36–45, 2010.*

[17] Feng Yu, A. M. Etheridge, and C. Cuthbertson. Asymptotic behaviour of the rate of adaptation. Annals of Applied Probability*, 20:978–1004, 2010.*