Frequency-dependent fitness induces multistability in coevolutionary dynamics

Hinrich Arnoldt¹,²,*, Marc Timme¹ and Stefan Grosskinsky³

¹Network Dynamics Group, Max Planck Institute for Dynamics and Self-Organization (MPIDS), Bunsenstrasse 10, 37073 Göttingen, Germany
²Georg August University Göttingen, 37077 Göttingen, Germany
³Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, UK

Evolution is simultaneously driven by a number of processes such as mutation, competition and random sampling. Understanding which of these processes is dominating the collective evolutionary dynamics in dependence on system properties is a fundamental aim of theoretical research. Recent works quantitatively studied coevolutionary dynamics of competing species with a focus on linearly frequency-dependent interactions, derived from a game-theoretic viewpoint. However, several aspects of evolutionary dynamics, e.g. limited resources, may induce effectively nonlinear frequency dependencies. Here we study the impact of nonlinear frequency dependence on evolutionary dynamics in a model class that covers linear frequency dependence as a special case. We focus on the simplest non-trivial setting of two genotypes and analyse the co-action of nonlinear frequency dependence with asymmetric mutation rates. We find that their co-action may induce novel metastable states as well as stochastic switching dynamics between them. Our results reveal how the different mechanisms of mutation, selection and genetic drift contribute to the dynamics and the emergence of metastable states, suggesting that multistability is a generic feature in systems with frequency-dependent fitness.

Keywords: population dynamics; dynamic fitness; stochastic switching; multistability

1. INTRODUCTION

Selection, random genetic drift and mutations are the processes underlying Darwinian evolution. For a long time, population geneticists have analysed the dynamics in the simplest setting consisting of two genotypes evolving under these processes [1]. In those studies, a genotype represents an individual’s genetic make-up, completely determining all relevant properties of the individual. A key concept is the so-called fitness of a genotype that represents the selection pressure for the individuals. The fitness defines the expected number of offspring an individual will produce. Thus, selection acts on fitness differences preferring individuals with higher fitness over individuals with lower fitness. Usually it is assumed that individuals have fixed fitnesses defined by their genotype alone [1–3]. Yet, experimental studies have revealed that many natural systems exhibit frequency-dependent selection [4–6], which means that an individual’s fitness not only depends on its genotype, but also on its interactions with other individuals and hence on the frequency of the different genotypes in the population. Although such frequency-dependent selection had already been studied early by Crow & Kimura [3], only recently has it received more attention [7–14]. In these theoretical and computational studies, individuals’ interactions are represented by interaction matrices from game theory. This leads to a frequency dependence where the fitness depends directly on the interaction parameters in a linear way. However, fitness may depend on many diverse factors such as cooperation (i.e. individuals acting together to increase their fitness [12,13]) and resource competition, so that certain systems may exhibit frequency-dependent fitness that is nonlinear. For example, in experiments, certain hermaphrodites exhibit such nonlinear fitness-dependence [6]. To the best of our knowledge, the impact of such nonlinear dependencies on coevolutionary dynamics has not been investigated theoretically.

In this article, we show that nonlinear frequency dependence [6,15] may induce new stable configurations of the evolutionary dynamics. Furthermore, we study the impact of asymmetric mutation probabilities on the dynamics [16,17], which was also neglected in most models until now [7,9,13,14]. As in previous works on coevolutionary dynamics, we base our work on the Moran process in a non-spatial environment, which is a well-established model to study evolutionary dynamics and was already used in many applications [18]. The Moran process is a stochastic birth–death process that keeps the
population size constant [19]. Therefore, in a two-genotype model, the system becomes effectively one-dimensional, so that the dynamics may be described by a one-dimensional Markov chain with transition rates defined by the Moran process. We derive the stationary probability distribution of the system dynamics via its Fokker–Planck equation [20]. Sharp maxima of the distribution reveal metastable points of the dynamics and a multitude of such maxima lead to stochastic switching dynamics between multiple stable points.

The article is structured as follows. In §2, we introduce the model details, and in §3, we derive the Fokker–Planck equation describing the probabilistic dynamics of the population. Using this equation, we derive the stationary probability distribution that describes the long-time behaviour of the system. In §4, we analyse this probability distribution, which yields information about the impact of nonlinear frequency-dependent selection and of different mutation probabilities on the coevolutionary dynamics. In §5, we give a summary and discuss our results.

2. THE MODEL

Consider a population of \( N \) individuals evolving in a non-spatial environment, meaning that the population is well-mixed so that each individual interacts with all other individuals at all times. In this population, the individuals may assume one out of two genotypes \( A \) and \( B \). The population sizes \( k_A \) and \( k_B \) (\( k_A + k_B = N \)) evolve according to the time-continuous Moran process described in the following (cf. [19]). The number of individuals \( k_A \) of genotype \( A \) completely determines the state of the system as \( k_B = N - k_A \). At all times, the interactions of the individuals determine the actual (frequency-dependent) fitness, so that an individual’s fitness of genotype \( A \) or \( B \) is defined by a fitness function \( f_A(k_A) \) or \( f_B(k_A) \), respectively. The fitness functions \( f_A(k_A) \) and \( f_B(k_A) \) may be any functions with the only condition that \( f_A(k_A) \geq 0 \) for all \( k_A \in [0,N] \) as negative fitness is not defined. At rate \( f_A(k_A)k_A \), an individual of type \( A \) produces an identical offspring that may mutate to genotype \( B \) with probability \( \mu_{AB} \). This applies analogously to genotype \( B \). Then one individual of the population is chosen randomly to die, so that the population size \( N \) stays constant and the variables \( k_A \) change maximally by 1. This is the so-called Moran process, which was originally introduced for a population of two genotypes with fixed fitnesses and no mutations occurring (cf. [19]). However, this process is easily generalizable to more genotypes and frequency-dependent selection in the way described already (cf. [11,21]).

Note that in our model, the rate of reproduction of, e.g. type \( A \) is directly determined by the term \( k_Af_A(k_A) \) as, e.g. in [22]. In other models, the fitness is first normalized so that the rate of reproduction is given by \( k_Af_A(k_A)/f(k_A) \) [13], where

\[
\bar{f}(k_A) = \frac{1}{N} [k_Af_A(k_A) + (N-k_A)f_B(k_A)],
\]

is the population’s mean fitness. While in both of these models the events occur with the same probability, the times between the events differ by a common factor determined by the mean fitness (2.1). Thus, these models exhibit a quantitatively different time course, but the same event sequences.

Until now, usually linear functions have been considered for \( f(k_A) \). However, in many applications, nonlinear functions seem more appropriate [6,15]. For example, cooperation effects in game theory induce functions linearly increasing in \( k_A \), so that \( f_A(k_A) = 1 + a k_A \) [11,13]. This would mean that fitness increases infinitely with the population size of genotype \( A \). Yet, in any habitat there is only a limited amount of resources available for living. Therefore, fitness should decline when the population of one genotype becomes too large as all of the individuals will compete for the same resources. We conclude that a fitness function including these two effects has to contain a nonlinear factor, e.g.

\[
f_A(k_A) = 1 + a k_A - b k_A^2,
\]

where \( a > 0 \) and \( b > 0 \). Here we choose a quadratic nonlinear factor as a simple example. Naturally, all other nonlinear functions could be applicable as well.

Note that linearly increasing fitness functions such as \( f_A(k_A) = 1 + a k_A \) and \( f_B(k_A) = 1 + b k_A \) lead to quadratic relations for the mean fitness \( \bar{f}(k_A) \) (see equation (2.1)) of the population, which becomes

\[
\bar{f}(k_A) = 1 + b k_A + (a - b) \frac{k_A^2}{N}.
\]

The quadratic factor originates from the fact that the linear fitness function in (2.1) is multiplied by the number of individuals, leading to a quadratic dependence on \( k_A \). e.g. \( k_Af_A(k_A) = k_A(1 + a k_A) = k_A + a k_A^2 \). This factor enters the replicator equation in evolutionary game theory, thus leading to a quadratic fitness dependence in standard game-theoretic problems [11,22]. However, this does not imply nonlinear interactions and thus fitness effects such as presented in equation (2.2). Therefore, the fitness functions that can be generated by evolutionary game theory are a special case of the functions that occur in the theory presented here. Fitness functions as for example presented in equation (2.2) will—as we show in this article—lead to dynamics with more complex stability properties.

Usually, when analysing the dynamics of such a model system as described earlier, it was assumed either that no mutations occur at all \( (\mu_A = 0) \) [10,11,19] or that the mutation probabilities are equal \( (\mu_{AB} = \mu_{BA}) \) [13,14,21]. However, usually mutation probabilities can be highly diverse [16,17]. This has been considered in some examples [3], but until now the effect of different mutation probabilities in the introduced model system has not been studied systematically. Here, we explicitly study the effects caused by asymmetric mutation rates (see figure 2).

3. ANALYSIS

In the following, we assume the fitness of the individuals to be given by

\[
f_A(k) = 1 + g_A(k)
\]
Multistability in evolutionary dynamics  

H. Arnoldt et al.  

3389

and

\[ f_B(k) = 1 + g_B(k) \]  (3.1)

with the interaction functions \( g_A > -1 \) and \( g_B > -1 \).

The model system is effectively one-dimensional with the variable \( k = k_A = N - k_B \) completely describing the system state [12]. At each event \( k \) may change by at most \( \pm 1 \), depending only on the actual state of the system, so that the dynamics are Markovian. Let the transition \( k \to k + 1 \) occur with the rate \( r_k^+ \). Then this rate is determined by the fitnesses and mutation probabilities in the following way.

- Each individual in the population receives offspring with rate \( f_j(k) \) with \( j \in \{ A, B \} \). This means that genotype \( A \) receives one offspring with rate \( k f_A(k) \).
- Such an offspring increases the population \( A \) with probability \( 1 - \mu_{AB} \), and population \( B \) in case it mutates with probability \( \mu_{AB} \).
- One individual is chosen uniformly to die belonging to genotype \( B \) with probability \( (N - k)/(N + 1) \).
- Equivalently, the number of individuals of genotype \( A \) can increase if genotype \( B \) receives one offspring, which mutates to genotype \( A \) with probability \( \mu_{BA} \), and one individual of \( B \) is chosen to die.

Taken these processes together, the transition rate for \( k \to k + 1 \) is

\[ r_k^+ = (1 - \mu_{AB}) k f_A(k) \frac{N - k}{N + 1} + \mu_{BA} (N - k) f_B(k) \frac{N - k}{N + 1}. \]  (3.2)

Decreasing \( k \) to \( k - 1 \) is equivalent to increasing the number of individuals of genotype \( B \), so that the transition rate

\[ r_k^- = (1 - \mu_{BA}) (N - k) f_B(k) \frac{k}{N + 1} + \mu_{AB} f_A(k) \frac{k}{N + 1} \]  (3.3)

is obtained analogously. The master equation of the Markov chain is then given by

\[
\frac{dp_k(t)}{dt} = r_{k-1}^- p_{k-1}(t) - (r_k^- + r_k^+) p_k(t) \\
+ r_{k+1}^+ p_{k+1}(t), \quad k \in \{0, 1, 2, \ldots, N\}
\]  (3.4)

where we define \( r_{N+1}^- = r_N^- = 0 \). As also \( r_0^+ = r_{k_A}^+ = 0 \), the system has reflective boundaries at \( k = 0 \) and \( k = N \). If \( \mu_{AB} = 0 \) \( (\mu_{BA} = 0) \), then \( k_A = N \) \( (k_B = 0) \) is an absorbing state of the dynamics.

It is possible to derive the exact solution to such an equation, similar to the derivations in [23], which we do in appendix B. However, the obtained solutions are not very explicit and their implications are hard to grasp. It is thus more useful to advance to the Fokker–Planck equation [20] describing the system in the limit of large \( N \). As previous works have shown, this approximation already leads to very good results for moderate population sizes of \( N < 1000 \) (cf. [13]).

We use the transformation

\[ x = \frac{k}{N} \in [0, 1], \quad s = \frac{t}{N} \quad \text{and} \quad \bar{\mu}_j = \mu_j N, \]  (3.5)

and

\[ \rho(x, s) = \rho_{AB} (sN) N \quad \text{and} \quad \bar{\gamma}_j(x) = g_j(xN) N \]  (3.6)

yielding the Fokker–Planck equation. The scaling factors are derived in appendix A, where the derivation of the Fokker–Planck equation in the scaling limit \( N \to \infty \) is shown in detail (equation (A5)). This leads to

\[
\frac{dp(x, s)}{ds} = -\frac{\partial}{\partial x} \left[ \left( \bar{\mu} (1 - 2x) + \bar{\gamma}_A (x) \right) x (1 - x) - \bar{\gamma}_B (x) x (1 - x) - \Delta \bar{\mu} \rho(x, s) \right] + \frac{\partial^2}{\partial x^2} \left[ x (1 - x) \rho(x, s) \right],
\]

with the normalization condition

\[
\int_0^1 \rho(x, s) dx = 1.
\]  (3.8)

for all \( s \geq 0 \). Here \( \bar{\gamma}_A = N g_A(Nx) \) and \( \bar{\gamma}_B = N g_B(Nx) \) are the rescaled interaction functions defined in (3.6), \( \bar{\mu} = N (\mu_{AB} + \mu_{BA})/2 \) is the rescaled mean mutation rate and \( \Delta \bar{\mu} = N (\mu_{AB} - \mu_{BA})/2 \) is the rescaled mutation rate difference. These functions have to be rescaled to obtain a non-degenerate limit, the so-called weak selection regime.

Notice, that the drift term

\[ D^{(1)}(x) = \bar{\mu} (1 - 2x) + \bar{\gamma}_A (x) - \bar{\gamma}_B (x) x (1 - x) - \Delta \bar{\mu}, \]  (3.9)

contains three different effects. The fitness difference at each point \( x \) causes a selective drift, the mean mutation rate causes a selective drift directed to \( x = \frac{1}{2} \) and the mutation rate difference causes a one-directional drift towards one of the boundaries \( x = 0 \) or \( x = 1 \).

The diffusion term

\[ D^{(2)}(x) = x (1 - x), \]  (3.10)

reflects the undirected genetic drift.

If fitness differences are on a scale of \( \mathcal{O}(N^{-1}) \), then mutation, selection and genetic drift all act on the same scale [13,24]. Then the dynamics are characterized by an interplay of these different effects. On the other hand, in the strong selection regime—where fitness differences are on a scale \( \mathcal{O}(1) \)—genetic drift becomes negligible and the dynamics in the bulk of the system become deterministic for large \( N \) [25]. However, as the factor \( x (1 - x) \) vanishes for \( x \to 0 \) and \( x \to 1 \) the effects of mutations play an important role. Clearly, without mutations, \( x = 0 \) and \( x = 1 \) are absorbing states, but even in the strong selection regime, they become non-absorbing for any positive mutation rate. This is reflected in the form of the drift term \( D^{(1)}(x) \), as only the factors reflecting mutations \( \bar{\mu} (1 - 2x) - \Delta \bar{\mu} \) remain non-zero in the limits \( x \to 0 \) and \( x \to 1 \).
For a one-dimensional Fokker–Planck equation (3.7), the stationary distribution is [20]

$$\rho^*(x) = Ce^{-\varphi(x)},$$

(3.11)

with the potential

$$\varphi(x) = \ln(D^{(2)}(x)) - \int \frac{D^{(1)}(x)}{D^{(2)}(x)} \, dx,$$

(3.12)

which is defined up to a constant irrelevant to our calculations as it is cancelled by the normalization

$$C = \frac{1}{\int_0^1 e^{-\varphi(x)} \, dx}. $$

(3.13)

Thus, we obtain

$$\rho^*(x) = \ln(x(1-x))$$

$$-\int \left[ \frac{\mu(1-2x)}{x(1-x)} + \hat{g}_A(x) - \hat{g}_B(x) - \frac{\Delta \mu}{x(1-x)} \right] \, dx$$

$$= \ln(x(1-x)) - \mu \ln(x(1-x))$$

$$+ \Delta \mu \ln \left[ \frac{x}{1-x} \right] - \int [\hat{g}_A(x) - \hat{g}_B(x)] \, dx,$$

(3.14)

where we used

$$1 - 2x = \frac{\partial}{\partial x} [x(1-x)].$$

The stationary solution becomes

$$\rho^*(x) = C \frac{x(1-x)}{(1-x)^{\Delta \mu}}$$

$$\times \left[ \frac{x}{1-x} \right]^{-\Delta \mu} e^{\int [\hat{g}_A(x) - \hat{g}_B(x)] \, dx},$$

(3.15)

where the constant C defined in (3.13) has to be calculated numerically for given parameters.

This stationary distribution contains contributions from the earlier-mentioned four different effects.

— Genetic drift is reflected by $$(x(1-x))^{-\Delta \mu},$$ which diverges for $$x \in [0, 1]$$ at the boundaries.

— The mean mutation rate $$\mu > 0$$ causes the balancing term $$(x(1-x))^\mu.$$
the interval $[0, 1]$. For away from the maxima at height of the maximum at dashed line) and one stable state at $x = 0$. System parameters were $B$ advantage for genotype $B$ as shown in (b) (solid line: genotype $A$; dashed line: genotype $B$). Solid and dashed lines in (a) show the theoretical curves from equation (3.15), crosses the data from simulations with $N = 1000$ (cross symbols, $\Delta \bar{\mu} = 0$; plus symbols, $\Delta \bar{\mu} = -0.5$). Interaction parameters are $a_A = 0.077$, $b_A = 0.1$, $a_B = 0.2$ and $b_B = 0.15$ (see equation (4.1)) while mean mutation rate is $\bar{\mu} = 1$. (Online version in colour.)

Figure 3. The mutation rates strongly influence the system’s dynamics. (a) Stationary solutions of an example system with interaction functions as in equation (4.1) for three different symmetric ($A \Delta \bar{\mu} = 0$) mutation rates: $\bar{\mu} = 1$ (solid line), $\bar{\mu} = 0.1$ (long dashed line) and $\bar{\mu} = 0.01$ (short dashed line). This illustrates that for low mutation rates the dynamics stay close to the system’s edges for long times. Here, differences in the mutation rates only slightly affect the shape of the stationary distribution. On the other hand, (b) shows that for high mutation rates ($\bar{\mu} = 5$) the dynamics are drawn stronger towards the middle. Here, differences in the mutation rates have a stronger impact, as the three curves with $\Delta \bar{\mu} = 0$ (solid line), $\Delta \bar{\mu} = 2.5$ (long dashed line) and $\Delta \bar{\mu} = 4.5$ (short dashed line) demonstrate. The increasing $\Delta \bar{\mu}$ shifts the existing stable state towards the edge of the system until it vanishes. System parameters were $N = 1000$ and a symmetric interaction function according to equation (4.1) with one stable state at $x = 0.5$ with $a_A = a_B = -0.01$, $b_A = b_B = 0.005$. (Online version in colour.)

and for $x > 0.2$ genotype $A$ is less fit than $B$, thus decreasing in frequency (cf. figure 1b). The maximum of the stationary distribution is thus exactly at the point where $f_A(x) = f_B(x)$. For $x > 0.7$ again genotype $A$ is fitter than $B$, and thus the dynamics are driven towards $x = 1$ by selection as well as genetic drift. The mutational force induced by $\bar{\mu} = 0.5$ increases the height of the maximum at $x \approx 0.2$, driving the system away from the maxima at $x = 0$ and $x = 1$. So, in this example, genetic drift, mutation and selection all significantly influence the dynamics.

Let us now study the influence of the mutation rates in detail. Interestingly, for asymmetric mutation rates ($\Delta \bar{\mu} \neq 0$) the factor $(x/(1-x))^{-\Delta \bar{\mu}}$ always diverges in the interval $[0, 1]$. For $\Delta \bar{\mu} > 0$, it diverges for $x \to 0$; otherwise for $x \to 1$. This can cause the emergence of a maximum of the stationary distribution at $x = 0$ (or, respectively, $x = 1$). Figure 2 shows an example where owing to asymmetric mutation rates a maximum occurs at $x = 1$, when for $\Delta \bar{\mu} = 0$ there is an absolute minimum at $x = 1$ as this stable state minimizes the population’s mean fitness. This means that the system dynamics are mutation dominated near $x = 1$. Furthermore, the maximum located at $x = 0$ for $\Delta \bar{\mu} = 0$ is shifted for $\Delta \bar{\mu} > 0$ to values $x > 0$, causing a minimum of the stationary distribution at $x = 0$. Thus, in this example, both selection and the asymmetry in mutation probabilities are the driving forces of the system’s dynamics.

For low mutation rates, the population has a high tendency to become dominated by one of the two genotypes for long times, which is illustrated in figure 3a. Thus, the dynamics stay close to the edges $x = 0$ or $x = 1$ waiting for one of the few mutations to occur. For low overall mutation rates, differences in the mutation rates have only weak effects. High mutation rates cause the population to be drawn towards a mixture of both genotypes ($x \approx 0.5$). Even more, for high mutation rates the differences in the mutation rates $\Delta \bar{\mu}$ can have an important influence on the dynamics. For instance, figure 3b illustrates a shift of an existing stable point, which completely vanishes for high $\Delta \bar{\mu}$. We conclude that asymmetric mutation rates can cause the emergence of new stable states (as in figure 2), the disappearance of...
which takes the mean distance of the measured empirical distribution \( \bar{d} \) from the theoretical distribution for every point except \( k = 0 \) and \( k = N \) where the density \( \rho'(x) \) diverges. Here, the theoretical distribution is calculated by integrating the theoretical density over bins of the size 1/N around the points k/N. Figure 4b shows that the decay of \( \bar{d} \) decays with increasing N.

The decay of this quantity is dominated by the slow convergence at the domain boundaries \( k = 0 \) and \( k = N \). Therefore, we additionally defined the maximum distance measure

\[
\bar{d}_{\text{max}} := \max_{k\in[1,N-1]} \left\{ \left| \mathbf{\pi}_k - \int_{k/N-1/2N}^{k/N+1/2N} \rho'(x) \, dx \right| \right\},
\]

which gives the maximum distance between the measured probability distribution from the theoretical distribution for all points except \( k = 0 \) and \( k = N \) due to the same argument as above. Figure 4b shows that the maximum distance \( \bar{d}_{\text{max}} \) decreases with increasing N, however slower than the mean distance \( \bar{d} \) owing to the divergence of \( \rho(x) \) at the boundaries. Altogether we conclude that the theoretical curve fits the data very well already for \( N \gtrsim 1000 \). However, near the domain boundaries, special care has to be taken, as divergences of the theoretical curve lead to larger deviations.

Furthermore, as the Fokker–Planck approximation only holds for weak selection, we study the quality of the solution obtained through the Fokker–Planck equation in dependence of the selection strength. For this, we introduce a scaling factor \( \xi \) to the interaction functions, so that they become

\[
\tilde{g}_A(x) = \xi N(a_A x - b_A x^2)
\]

and

\[
\tilde{g}_B(x) = \xi N(a_B (1 - x) - b_B (1 - x)^2).
\]

We find that the stationary solution of the Fokker–Planck equation well approximates the solution of the Master equation

\[
p_k^\xi = \frac{\prod_{j=0}^{k-1} (r_j / r_{j+1})}{\sum_{l=0}^{k-1} \prod_{j=0}^{l-1} (r_j / r_{j+1})}
\]

for selection strengths up to \( \xi \) of the order of 1 (cf. figure 5). We derive the solution (4.6) of the Master equation in appendix B. Note that for the derivation of this solution, no approximation is necessary and hence fits data from simulations perfectly up to an error due to finite sampling in simulations. As figure 5 illustrates, for strong selection \( \xi \approx 1 \), the solution of the Fokker–Planck equation does not fit the exact solution of the Master equation perfectly, but still catches the overall trend of the dynamics. The error, quantified by the mean distance measure analogous to (4.3), increases with \( \xi \) as a power law both in comparison with the solution from the Master equation and with existing ones (as in figure 3), and shift existing stable states to new positions.

As we saw earlier, the fit of the theoretically calculated stationary distribution to simulation data is almost perfect for weak selection, i.e. the maximal fitness difference satisfies \( \Delta f_{\text{max}} \ll 1 \) (cf. [13]). To quantify the quality of the fit, we measured the stationary distribution with the same parameters as mentioned earlier for different population sizes \( N \). This is shown in figure 4, demonstrating that for increasing \( N \) the data get even closer to the theoretically obtained distribution. For this, we define the empirical distribution

\[
\pi_k := \frac{1}{T_{\text{meas}}} \sum_{t=0}^{T_{\text{meas}}} \delta(X_t + \tau_{\text{mix}}, k),
\]

where \( (X_t : t \geq 0) \) is the discrete process defined in equation (3.4). \( T_{\text{meas}} \) is a time large enough for the process to reach stationarity and depends on the system size, as does the measurement time \( T_{\text{meas}} \) which is specified in figure 4. Using this definition, we introduce a scaling factor \( \xi \) to the interaction functions, so that they become

\[
\tilde{g}_A(x) = \xi N(a_A x - b_A x^2)
\]

and

\[
\tilde{g}_B(x) = \xi N(a_B (1 - x) - b_B (1 - x)^2).
\]

We find that the stationary solution of the Fokker–Planck equation well approximates the solution of the Master equation

\[
p_k^\xi = \frac{\prod_{j=0}^{k-1} (r_j / r_{j+1})}{\sum_{l=0}^{k-1} \prod_{j=0}^{l-1} (r_j / r_{j+1})}
\]

for selection strengths up to \( \xi \) of the order of 1 (cf. figure 5). We derive the solution (4.6) of the Master equation in appendix B. Note that for the derivation of this solution, no approximation is necessary and hence fits data from simulations perfectly up to an error due to finite sampling in simulations. As figure 5 illustrates, for strong selection \( \xi \approx 1 \), the solution of the Fokker–Planck equation does not fit the exact solution of the Master equation perfectly, but still catches the overall trend of the dynamics. The error, quantified by the mean distance measure analogous to (4.3), increases with \( \xi \) as a power law both in comparison with the solution from the Master equation and with

![Figure 4](http://rsif.royalsocietypublishing.org/Downloaded from http://rsif.royalsocietypublishing.org/)
data from simulations (cf. figure 5d). All in all, for weak selection the Fokker–Planck approximation works well, while for strong selection, it does not perfectly predict the stationary distribution, but still reflects the overall trend of the dynamics. If this approximation is not satisfying, the direct solution (4.6) of the Master equation yields the exact distribution fitting the data for all selection strengths.

We have shown earlier that the dynamics of a two-genotype system can exhibit multiple stable points induced by nonlinear selection. Even more, there is no theoretical limit to the number of stable points in the system. Assume for example $\bar{\mu} = 0.5, \alpha_l = \alpha_r = -0.01$ and $b_l = b_r = 0.005$. The measured empirical distributions were obtained by simulating the system dynamics from an initial state drawn from $p^*(x)$ for a mixing time $T_{\text{mix}} = 10^5$ and then recording the density for a time $T_{\text{meas}} = 10^7$. (Online version in colour.)

Figure 6. Multiple stability dynamics for periodic interaction functions. (a) Theoretically computed stationary distribution from equation (3.15) (solid line) together with simulation data with $N = 1000$ (cross symbols) for the interaction functions given in (4.7). The computed stationary distribution diverges for $x \to 0$ and $x \to 1$ while the simulation data remain finite owing to the finite number of individuals. (b) Sample path that exhibits switching between the different maxima of the distribution. Parameters are $a = 30, \beta = 25, \bar{\mu} = 0.8$ and $\Delta\bar{\mu} = 0$. (Online version in colour.)
where we used periodic interaction functions
\[ g_A(x) = \alpha (1 + \sin(\beta x)) \]
and
\[ g_B(x) = \alpha (1 + \cos(\beta x)). \]

Although this is not a realistic interaction function in most applications, it demonstrates what is theoretically possible in the introduced system.

5. SUMMARY

In this article, we analysed a two-genotype system in a very general setting with (possibly) asymmetric mutation probabilities and nonlinear fitness functions [6,15] in finite populations. The underlying Moran process is a well-established model [1,12,13] to gain an understanding of the interplay of selection, mutation and genetic drift in evolutionary dynamics. However, the Moran process is studied mostly with symmetric mutation probabilities and at most linear interaction functions. We reasoned that neither need mutation probabilities be symmetric—as experiments have shown that mutation probabilities are often asymmetric [16]—nor can all interaction effects be described by linear interaction functions. For example, cooperation in game theory leads to an interaction function increasing linearly in the frequency of the cooperating genotype. Yet, in many applications, also cooperators in the end compete for the same type of resource that is limited. Therefore, owing to limited resources, a population being too large cannot be sustained, leading to a decrease in the fitness. There is no linear function that can reflect both of these effects at the same time.

We derived the Fokker–Planck equation describing the dynamics of the number of individuals \( k \) of genotype \( A \) in the limit of large population sizes \( N \). We quantified the quality of the Fokker–Planck approach for an example (cf. figure 4), where the difference of simulation data and theoretical solution became almost not detectable for population sizes larger than \( N \gtrsim 1000 \). Actually, if the system exhibits absorbing states, then the Fokker–Planck method does not work to study the corresponding quasi-stationary distributions. Instead, Wentzel–Kramers–Brillouin methods are more appropriate to describe the system dynamics as, for example, in [10], where fixation resulting from large fluctuations was studied. In our model system, no such absorbing states exist, as long as the mutation rates are positive \( (\mu \neq 0) \) and therefore the Fokker–Planck equation is appropriate to describe the system dynamics.

We identified the individual effects of selection, mean mutation rate and mutation difference as well as genetic drift and derived the stationary probability distribution as determined by the Fokker–Planck equation. Analysing the distribution, we found that asymmetries in the mutation probabilities may not only induce the shifting of existing stable points of the dynamics to new positions, but also lead to the emergence of new stable points. Thus, a genotype that has a selective disadvantage can anyway have a stable dynamical state where its individuals dominate the population owing to a higher mutational stability (figure 2). Further, we found that dynamic fitness leads to multiple stable points of the dynamics induced by selection and also genetic drift. We showed an example (figure 1) where three stable points exist, two caused by selection and one by genetic drift.

We conclude that frequency-dependent fitness together with asymmetric mutation rates induce complex evolutionary dynamics—in particular, if the interactions imply nonlinear fitness functions. Theoretically, there is no limit for the number of stable points that the dynamics can exhibit (figure 6). All in all, we interpret our results such that in real biological systems, multiple metastable equilibria may exist whenever species interact in a way complex enough to imply a fitness that nonlinearly depends on frequency. As a consequence, one species may exhibit a certain frequency for a long time before a sudden shift occurs and then a new frequency prevails. Such a change may thus occur even in the absence of changes of the environment; it may be induced as well by a stochastic switching from one metastable state to another, owing to complex inter-species interactions.

The Moran process is a standard tool to gain theoretical insights into experimental data [18]. Of course, we do not propose here that any experimental set-up may be exactly described by the Moran process. However, we think that it should be feasible to develop an experiment where two different mutants evolve with asymmetric mutation rates. To find an experimental set-up where the two genotypes also exhibit nonlinear fitness could however prove more difficult. Rather, our study is a theoretical study indicating that nonlinear fitness may be the cause of multiple stable states when observed in experimental data.

For further studies on systems with more genotypes, it should be useful to combine our considerations presented here with the work of Traulsen et al. [13], where an analysis of systems with more than two genotypes was carried out. Extending those results, it may be possible to gain a better understanding of the effects of nonlinear interactions for many different genotypes. Also, it may be interesting to study the effects of changing interactions, where the interactions change according to the system dynamics [26]. Thus, our study might serve as a promising starting point to investigate how nonlinear frequency dependencies impact evolutionary dynamics in complex environments.

We thank Steven Strogatz for fruitful discussions during project initiation and Stefan Eule for helpful technical discussions. Stefan Grosskinsky acknowledges support by EPSRC, grant no. EP/E501311/1.

APPENDIX A. THE FOKKER–PLANCK EQUATION

The master equation (3.4) may be transformed to a Fokker–Planck equation in the limit of large \( N \) [20]. We introduce the transformation

\[ x = \frac{k}{N}, \quad s = t F(N) \quad \text{and} \quad \bar{\mu}_y = \mu_0 G(N), \]

\[ \bar{\mu}_y = \mu_0 G(N), \]
We choose \( F(N) = 1/(N + 1) \) and \( G(N) = H(N) = N \) so that in the limit \( N \rightarrow \infty \) the terms stay finite. Further, we introduce the mean mutation rate \( \tilde{\mu} = N(\mu_{AB} + \mu_{BA})/2 \) and the mutation rate difference \( \Delta \tilde{\mu} = N(\mu_{AB} - \mu_{BA})/2 \). To not overload the notation, we drop the time argument \( s \) of \( \rho(x, s) \) in the following calculation. This leads to

\[
\begin{align*}
\frac{d\rho(x)}{ds} &= N^2 \left\{ -2x(1-x)\rho(x) + x_+(1-x_+)^2\rho(x_+) + x_-(1-x_-)^2\rho(x_-) \\
&+ \frac{\tilde{\mu}}{N} \left(-2x(1-x)^2\rho(x) + \frac{1}{2} (1-2x^2)^2\rho(x) + \frac{1}{2} (1-2x_-^2)^2\rho(x_-) \right) \right\} \\
&+ N \{ \tilde{g}_A(x_+)(1-x_-)\rho(x_-) - \tilde{g}_A(x_+)(1-x_-)\rho(x) \} \\
&+ \Delta \tilde{\mu} [x_+\rho(x_+) - x_-(1-x_-)\rho(x)] \\
&+ \frac{\tilde{\mu}}{N} \left( \tilde{g}_A(x_+)^2 + \tilde{g}_B(x_+)^2 - \tilde{g}_A(x_-)^2 + \tilde{g}_B(x_-)^2 + \tilde{g}_B(x)^2 - \tilde{g}_B(x_+)^2 - \tilde{g}_B(x_-)^2 \right) \rho(x) \\
&+ (\tilde{g}_A(x_+)(1-x_-)\rho(x) + \Delta \tilde{\mu} [\tilde{g}_A(x_+)(1-x_-)^2\rho(x_-) + \tilde{g}_B(x_+)^2 - \tilde{g}_B(x_-)^2]) \rho(x) \\
&+ \tilde{g}_B(x_+)(1-x_-)^2\rho(x) - (\tilde{g}_A(x_+)(x_+^2 - x_+^2) + \tilde{g}_B(x_-)(x_-^2 - x_-^2)) \rho(x) \\
&+ \tilde{g}_B(x_+)(1-x_-)^2\rho(x) - (\tilde{g}_A(x_+)(x^2 - x^2) - \tilde{g}_B(x_-)(1-x^-)^2) \rho(x) \\
&+ \tilde{g}_B(x_+)(1-x_-)^2\rho(x) - (\tilde{g}_A(x_+)(x_+^2 - x_+^2) - \tilde{g}_B(x_-)(x_-^2 - x_-^2)) \rho(x) \\
&+ \tilde{g}_B(x_+)(1-x_-)^2\rho(x) - (\tilde{g}_A(x_+)(x^2 - x^2) - \tilde{g}_B(x_-)(1-x^-)^2) \rho(x) \right\}.
\end{align*}
\]

In the limit \( N \rightarrow \infty \), the different terms with \( N^2 \) in front become second-order derivatives with respect to \( x \), while the other terms become first-order derivatives. The terms that have a \( 1/N \) factor vanish in the limit \( N \rightarrow \infty \) and thus the earlier-mentioned equation becomes

\[
\frac{d\rho(x)}{ds} = -\frac{\partial}{\partial x} [\tilde{\mu}(1-2x)\rho(x) + (\tilde{g}_A(x) - \tilde{g}_B(x))(1-x)\rho(x)] \\
- \frac{\partial^2}{\partial x^2} [x(1-x)\rho(x)] \\
+ \frac{\Delta \tilde{\mu}}{N} [x+\rho(x)] \\
+ \frac{\partial^2}{\partial x^2} [x(1-x)\rho(x)] \\
+ \frac{\partial^2}{\partial x^2} [x(1-x)\rho(x)] \\
+ \frac{\partial^2}{\partial x^2} [x(1-x)\rho(x)] \\
+ \frac{\partial^2}{\partial x^2} [x(1-x)\rho(x)] \\
+ \frac{\partial^2}{\partial x^2} [x(1-x)\rho(x)]
\]

which is the Focker–Planck-equation of the system.

**APPENDIX B. STATIONARY SOLUTION OF THE MASTER EQUATION**

We directly derive the stationary solution \( p_k^* \) of the master equation (3.4) using the detailed balance equation

\[
r_{k+1}^* p_k^* = r_k^* p_{k+1}^* \tag{B1}
\]

which applies to any chain with only nearest neighbour transitions [20] (cf. also [23]). Thus, by using the rewritten balance equation

\[
p_k^* = \frac{r_{k+1}^*}{r_k^*} p_{k+1}^* \tag{B2}
\]
We then obtain the exact stationary solution of the master equation

\[ p_k^* = p_0^* \prod_{j=0}^{k-1} \frac{r_{j+1}^+}{r_{j+1}^-}. \]  

(B3)

Finally, we may use the normalization condition

\[ \sum_{j=0}^{N} p_j^* = 1 \]  

(B4)

of the stationary distribution to eliminate the factor \( p_0^* \). We then obtain the exact stationary solution of the master equation

\[ p_k^* = \frac{\prod_{j=0}^{k-1} (r_{j+1}^+ / r_{j+1}^-)}{\sum_{j=0}^{N} \prod_{j=0}^{k-1} (r_{j+1}^+ / r_{j+1}^-)}, \]  

(B5)

which can be evaluated numerically for any transition rates \( r_j^+ \) and \( r_j^- \). For more details on the exact solution of the Master equation in the Moran process, see, for example, the work by Claussen & Traulsen [23].

REFERENCES


