Robustness and epistasis in mutation-selection models

Andrea Wolff and Joachim Krug

Institut für Theoretische Physik, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany
E-mail: awolff@thp.uni-koeln.de and krug@thp.uni-koeln.de

Received 6 January 2009
Accepted for publication 7 April 2009
Published 1 May 2009
Online at stacks.iop.org/PhysBio/6/036007

Abstract
We investigate the fitness advantage associated with the robustness of a phenotype against deleterious mutations using deterministic mutation-selection models of a quasispecies type equipped with a mesa-shaped fitness landscape. We obtain analytic results for the robustness effect which become exact in the limit of infinite sequence length. Thereby, we are able to clarify a seeming contradiction between recent rigorous work and an earlier heuristic treatment based on mapping to a Schrödinger equation. We exploit the quantum mechanical analogy to calculate a correction term for finite sequence lengths and verify our analytic results by numerical studies. In addition, we investigate the occurrence of an error threshold for a general class of epistatic landscapes and show that diminishing epistasis is a necessary but not sufficient condition for error threshold behaviour.

1. Introduction

In current evolutionary theory, the concept of robustness, referring to the invariance of the phenotype under perturbations, is of central importance [1, 2]. Here we address specifically mutational robustness, which we take to imply the stability of some biological function with respect to mutations away from the optimal genotype. To be precise, suppose the genotype is encoded by a sequence of length $L$ and the number of mismatches with respect to the optimal genotype is denoted by $k$. Robustness is then quantified by the maximum number of mismatches $k_0$, which can be tolerated before the fitness of the individual falls significantly below that of the optimal genotype at $k = 0$. This situation arises e.g. in the evolution of regulatory motifs, where the fitness is a function of the binding affinity to the regulatory protein [3, 4]. Assuming that the fitness is independent of $k$ for both $k < k_0$ and $k > k_0$, the fitness landscape has the shape of a mesa parametrized by its width $k_0$ and height $w_0$ [5].

We consider deterministic mutation-selection models of a quasispecies type, which describe the dynamics of large (effectively infinite) populations [6]. We analyse the stationary state of mutation-selection balance, focusing on the dependence of the population fitness on the parameters $k_0$ and $w_0$. This allows us to identify the conditions under which a broad fitness peak of a relatively low selective advantage outcompetes a higher but narrower peak [7], a phenomenon that has been referred to as the survival of the flattest [8–14].

Robustness in the sense described above is a special case of epistasis, which refers more generally to any nonlinear relationship between the number of mutations away from the optimal genotype and the corresponding fitness effect [15]. A simple way to parametrize epistasis is to let the loss of fitness vary with the number of mismatches as $k^\alpha$, such that the non-epistatic case $\alpha = 1$ separates regimes of synergistic ($\alpha > 1$) and diminishing ($\alpha < 1$) epistasis [16, 17]. An important problem in previous work on mutation-selection models has been to identify the conditions under which epistatic fitness landscapes display an error threshold, a term that refers to the discontinuous delocalization of the population from the vicinity of the fitness peak as the mutation rate is increased beyond a critical value [6, 16, 18]. Improving on earlier work that found that only landscapes with diminishing epistasis ($\alpha < 1$) have an error threshold, we derive here the more stringent condition $\alpha \leq 1/2$ on the epistasis exponent.
1.1. Organization of this paper

We base our work on two complementary analytic approaches. First, recent progress in the theory of mutation-selection models [5, 18–24] provides an expression for the population fitness in terms of a maximum principle (MP) that becomes exact when the limit \( L \to \infty \) is performed keeping the ratio \( x_0 = k_0/L \) fixed. Second, Gerland and Hwa (GH) [3] have used a continuum approximation to map the mutation-selection problem onto a one-dimensional Schrödinger equation which is then analysed with standard techniques.

Our work was initially motivated by the observation of a discrepancy between the two approaches: whereas the MP predicts that the selective advantage of a broad mesa should vanish when the limit \( L \to \infty \) is taken at fixed \( k_0 \), in the GH approach a finite selective advantage is retained in this limit, which depends on the absolute value of \( k_0 \) rather than on \( x_0 \). After introducing the model and briefly reviewing the results of the MP approach in section 2, we therefore provide a detailed discussion of the continuum approximation used by GH in section 3. We emphasize that it amounts to a harmonic approximation around \( k = L/2 \) and show how it can be improved in such a way that the results based on the MP are recovered.

The mapping to a Schrödinger equation is nevertheless useful, as it allows us to derive the leading correction to the population fitness for finite sequence length \( L \). As a consequence, we find excellent agreement between the analytic predictions and numerical solutions of the discrete mutation-selection equations. In section 4 we consider the selection transition in a two-peak landscape first studied by Schuster and Swetina [7, 25], in which the population shifts from a high, narrow fitness maximum to a lower but broader peak as the mutation rate is increased. The occurrence of this transition is an indicator for the superiority of robustness over fitness in certain parameter regimes. In section 5, we use the MP approach to derive the critical value of the epistasis exponent \( \alpha \) and verify our prediction by numerical calculations. Finally, some conclusions are presented in section 6. Details of the derivation of the improved continuum limit and the generalization to an arbitrary alphabet size can be found in the two appendices.

2. Mutation-selection models and the maximum principle

Mutation-selection models are commonly studied in two varieties, the Eigen model [26], in which mutations are coupled to reproduction events, and the Crow–Kimura model [27], in which mutation and reproduction occur in parallel. In fact the two models are very similar, as the Crow–Kimura model can be obtained from the Eigen model in the limit of weak selection and small mutation rates [28]. Alternatively, the Eigen model may be viewed as a discrete time model from which the Crow–Kimura model arises in a continuous time limit [6]. For the convenience of the reader, we sketch here the latter relation and then focus on the continuous time Crow–Kimura model in the remainder of this paper. Throughout we consider asexual populations of haploid individuals and encode the genotype by a binary sequence of fixed length.

In the discrete time Eigen model, mutations occur as copying errors during reproduction with probability \( \bar{\mu} \) per site and generation, and the Wrightian fitness \( W_\sigma \) of a sequence \( \sigma \) is a non-negative number that is proportional to the expected number of offsprings of an individual carrying that sequence. The population structure is then described by the fraction \( p_\sigma(t) \) of individuals with sequence \( \sigma \), which satisfies the evolution equation

\[
p_\sigma(t + \Delta t) = \frac{\sum_\sigma Q_{\sigma\sigma'} W_{\sigma'} p_{\sigma'}(t)}{\sum_\sigma W_\sigma p_\sigma(t)}.
\]  

(1)

Here, \( \Delta t \) is the generation time of the non-overlapping generations and the mutation probability \( Q_{\sigma\sigma'} \) is given by

\[
Q_{\sigma\sigma'} = \bar{\mu} d(\sigma, \sigma')(1 - \bar{\mu})^{-d(\sigma, \sigma')},
\]  

(2)

with \( d(\sigma, \sigma') \) being the Hamming distance between the two binary sequences \( \sigma \) and \( \sigma' \). In order to derive the Crow–Kimura model, we perform the limit \( \Delta t \to 0 \) and thus write \( W_\sigma \) in terms of the Malthusian fitness \( w_\sigma \), expanding to first order in \( \Delta t \to 0 \):

\[
W_\sigma = e^{w_\sigma \Delta t} \approx 1 + w_\sigma \Delta t.
\]  

(3)

Similarly, denoting the continuous time mutation rate per site by \( \mu \) and setting \( \bar{\mu} = \mu \Delta t \), for \( \Delta t \to 0 \) equation (2) reduces to

\[
Q_{\sigma\sigma'} \approx \delta_{\sigma\sigma'} + \eta_{\sigma\sigma'} \mu \Delta t
\]  

(4)

with

\[
\eta_{\sigma\sigma'} = \begin{cases} 0 & : d(\sigma, \sigma') > 1 \\ \mu & : d(\sigma, \sigma') = 1 \\ -\mu L & : d(\sigma, \sigma') = 0. \end{cases}
\]  

(5)

Finally, expanding also the change in the population distribution to first order in \( \Delta t \)

\[
p_\sigma(t + \Delta t) \approx p_\sigma(t) + \frac{dp_\sigma(t)}{dt} \Delta t,
\]  

(6)

we arrive at the Crow–Kimura model

\[
\frac{dp_\sigma(t)}{dt} = (w_\sigma - w) p_\sigma(t) + \sum_{\sigma'} \eta_{\sigma\sigma'} p_{\sigma'}(t).
\]  

(7)

Note that the discrete time model (1) is invariant under multiplicative shifts \( W_\sigma \to C W_\sigma \) of the Wrightian fitness, whereas the continuous time model (7) is invariant under additive shifts of the Malthusian fitness, \( w_\sigma \to w_\sigma + C \). In particular, a lethal genotype would be described by \( W_\sigma = 0 \) in the first case but by \( w_\sigma = -\infty \) in the second case. Lethal mutations will be not be considered in this paper.

We now assume the Malthusian fitness to be a function \( w_k \) only of the Hamming distance \( k \) to the optimal sequence \( \sigma_0 \) at \( k = 0 \). Then (7) transforms into the set of equations

\[
\frac{dP_k}{dt} = (w_k - w) P_k + \mu (k + 1) P_{k+1} + \mu (L - k + 1) P_{k-1} - \mu L P_k,
\]  

(8)

for the population fraction

\[
P_k = \sum_{\sigma \ : d(\sigma, \sigma_0) = k} p_\sigma
\]  

(9)
in the Hamming class \( k \). Here, \( 1 \leq k \leq L - 1 \) and obvious modifications of (8) are implied for \( k = 0 \) and \( k = L \). The nonlinearity introduced into (8) by the mean fitness \( \bar{w}(t) = \sum w_i P_i(t) \) can be eliminated by passing to unnormalized population variables [6, 29]. At long times, the population distribution therefore approximates the principal eigenvector \( P^e \) of the linear dynamics, which is the solution of the eigenvalue problem

\[
\Lambda P^e_k = (w_k - \mu L) P^e_k + \mu (k + 1) P^e_{k+1} + \mu (L - k + 1) P^e_{k-1} \tag{10}
\]

with the maximal eigenvalue \( \Lambda \). This eigenvalue is equal to the long-time limit of the mean population fitness \( \bar{w} \), and it is the main quantity of interest in this paper. Depending on the context we will refer to \( \Lambda \) as the mean population fitness, the population growth rate, the principal eigenvalue of the mutation-selection matrix defined by (10) or the ground-state energy of the corresponding quantum mechanical problem, to be defined in subsection 3.2.

A considerable body of work has been devoted to the solution of (10) for large \( L \). In order to obtain nontrivial behaviour in the limit \( L \to \infty \), it is necessary to either scale the mutation rate \( \gamma \) or the fitness \( \gamma \). We adopt here the first choice and take \( L \to \infty, \mu \to 0 \) with \( \gamma = \mu L \) fixed. If, in addition, the fitness landscape \( w_k \) is assumed to depend only on the relative number of mismatches, such that

\[
w_k = f(x), \quad x = k/L, \tag{11}
\]

then the principal eigenvalue in (10) is given, for \( L \to \infty \), by the solution of a one-dimensional variational problem as [18–23]

\[
\Lambda = \max_{x \in [0,1]} \left\{ f(x) - \gamma \left[ 1 - 2 \sqrt{x(1-x)} \right] \right\}. \tag{12}
\]

See subsection 3.5 and appendix A for a heuristic derivation and appendix B for the generalization to an arbitrary alphabet size. Moreover, if \( f(x) \) is differentiable the leading order correction to (12) takes the form [21, 22]

\[
\Delta \Lambda = \frac{\gamma}{2L \sqrt{x_c - x_c^2}} \left[ 1 - \sqrt{1 - 2f''(x_c)(x_c - x_c^2)^{-3/2}} / \gamma \right], \tag{13}
\]

where \( x_c \) is the value at which the maximum in (12) is attained.

A similar maximum principle holds for the largest eigenvalue \( \bar{\Lambda} \) of the linear operator acting on the right-hand side of the discrete time Eigen model (1). When the Wrightian fitness is a function of the relative number of mismatches, \( W_k = F(k/L) \), one finds that [20, 21, 23]

\[
\bar{\Lambda} = \max_{x \in [0,1]} \left\{ \bar{F}(x) e^{-\sqrt{1-2\sqrt{x(1-x)}}} \right\}, \tag{14}
\]

where \( \bar{\gamma} = \mu L \) is kept fixed as \( L \to \infty \). Note that this is precisely what one would expect on the basis of the relation between the two models discussed above.

In the first part of this paper, we will focus on mesa landscapes of the form

\[
w_k = \begin{cases} w_0 > 0 & : 0 \leq k \leq k_0 \\ 0 & : k > k_0 \end{cases}, \tag{15}
\]

where \( w_0 \) is the selective advantage of the functional phenotype and \( k_0 \) denotes the number of tolerable mismatches. Within

\begin{center}
\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Equilibrium population distribution \( P^e_k \) for a mesa-shaped fitness function with plateau width \( k_0 = 20 \) and sequence length \( L = 100 \) and two different mutation rates. For the localized population distribution \( P^e_{k,loc} \), the mutation rate per sequence is \( \gamma = 1 \); for the delocalized distribution \( P^e_{k,deloc} \), the mutation rate is \( \gamma = 5 \). The vertical axis on the left displays the fitness (dashed line) and the axis on the right displays the population distributions \( P^e_k \). The localized population resides near the edge of the plateau, because as a consequence of the mapping from the sequence space to Hamming space, the number of available sequences increases strongly with increasing \( k \). The delocalized population is binomially distributed in the Hamming space, which corresponds to a uniform distribution in the sequence space.}
\end{figure}
\end{center}

the class of scaling landscapes defined by (11), this is realized by setting

\[
f(x) = w_0 \Theta(x_0 - x), \tag{16}\]

where \( \theta \) is the Heaviside step function and \( x_0 = k_0/L \). Provided \( x_0 < 1/2 \), the application of the maximum principle (12) yields

\[
\Lambda = \begin{cases} w_0 - \gamma (1 - 2 \sqrt{x_0 (1 - x_0)}) & : w_0 > w_0^e \\ \gamma (1 - 2 \sqrt{x_0 (1 - x_0)}) & : w_0 < w_0^e \end{cases} \tag{17}
\]

The value \( w_0^e \) of the selective advantage marks the location of the error threshold at which the population delocalizes from the fitness peak and the location \( x_c \) of the maximum in (12) jumps from \( x_c = x_0 \) to \( x_c = 1/2 \). Figure 1 shows the form of the stationary population distribution \( P^e_k \) for this landscape in the localized \((w_0 > w_0^e)\) and delocalized \((w_0 < w_0^e)\) phases. Expression (13) for the finite sequence length correction is clearly not applicable to the discontinuous mesa landscape (16). In fact, we will show below that the leading order correction \( \Delta \Lambda \) is of order \( L^{-2/3} \) or \( L^{-1/2} \) rather than \( L^{-1} \) in this case.

3. **Continuum limit in the Hamming space**

3.1. **Derivation and status**

A natural approach to analysing (8) and (10) for large \( L \) is to perform a continuum limit in the Hamming class index \( k \). To this end, we introduce \( \epsilon = 1/L \) as a small parameter and replace the population variable \( P_k \) by a function

\[
\phi(x) = \lim_{L \to \infty} P_{\Delta L}. \tag{18}
\]
The current electron is taken to be of the general form (11). Expanding the finite differences in (10) to second order in $\epsilon$ then yields the stationary drift-diffusion equation

\[ f \phi - \epsilon \gamma \frac{d}{dx} (1 - 2x) \phi + \frac{\epsilon^2 \gamma}{2} \frac{d^2}{dx^2} \phi = \Lambda \phi. \]  

(19)

This is identical to the equation obtained by GH [3], who however write it in terms of the unscaled variable $k = Lx$.

Before proceeding with the analysis of (19), some remarks concerning the accuracy of the second-order expansion are appropriate. In the absence of selection ($f = 0$), the principal eigenvalue in (19) is readily seen to be $\Lambda = 0$ and the corresponding (right) eigenfunction is a Gaussian centred at $x = 1/2$.

\[ \phi_0(x) \sim \exp[-(1 - 2x)^2 / 2\epsilon]. \]  

(20)

This is just the central limit approximation to the binomial distribution

\[ P_k = 2^{-L} \binom{L}{k}, \]  

(21)

which solves (10) for $w_k = 0$ and $\Lambda = 0$. It is well known that the central limit approximation of (21) is accurate in a region of size $\sqrt{L}$ around $k = L/2$, but becomes imprecise for deviations of order $L$. An improved approximation is provided by the theory of large deviations [30], in which the ansatz

\[ P_k \sim \exp[-L u(x)] \]  

(22)

is made to obtain an expression for the large deviation function $u(x)$. In the context of mutation-selection models, this approach has recently been introduced by Saakian [22], who showed that it allows one to derive the exact relation (12) in a relatively straightforward manner (see appendix A). Equivalent results can be obtained by continuing the expansion in (19) to all orders in $\epsilon$ and treating the resulting equation in a WKB-type approximation, which essentially corresponds to the ansatz (22) (see [24]).

We conclude that the approximation (19) can be expected to be quantitatively accurate only near the centre $x = 1/2$ of the Hamming space. We will nevertheless adhere to this approximation in the following three subsections, because it allows us to make contact with the work of GH and to formulate the eigenvalue problem (10) in the familiar language of one-dimensional quantum mechanics. In subsection 3.3, we will then show how to go beyond the second-order approximation.

### 3.2. Mapping to a one-dimensional Schrödinger equation

The key step in reducing (19) to a standard form is to symmetrize the linear operator on the left-hand side, thus eliminating the first-order drift term. This can be achieved by the transformation

\[ \phi(x) = \sqrt{\phi_0(x)} \psi(x), \]  

(23)

with $\phi_0(x)$ from (20), which leads to the stationary Schrödinger equation

\[ -\frac{\epsilon^2 \gamma}{2} \frac{d^2}{dx^2} \psi + V(x) \psi = -(\Lambda - \epsilon \gamma) \psi \]  

(24)

with the effective potential

\[ V(x) = \frac{\gamma}{2}(1 - 2x)^2 - f(x). \]  

(25)

The latter consists of the superposition of a harmonic oscillator centred around $x = 1/2$ with the (negative) fitness landscape. As pointed out in [5], the inverse sequence length $\epsilon$ plays the role of Planck's constant $\hbar$, which implies that the case of interest corresponds to the semiclassical limit of the quantum mechanical problem. In particular, for $\epsilon \to 0$ the ground-state energy $-\Lambda$ becomes equal to the minimum of the effective potential. We thus arrive at the variational principle

\[ \Lambda = \max_{x \in [0,1]} \left[ f(x) - \frac{\gamma}{2}(1 - 2x)^2 \right], \]  

(26)

which is precisely the harmonic approximation (in the sense of a quadratic expansion around $x = 1/2$) of the exact relation (12). In this perspective, the error threshold corresponds to a shift between different local minima of $V(x)$, which become degenerate at the transition point. The transition is generally of first order, in the sense that the location $x_c$ of the global minimum jumps discontinuously. Within the harmonic approximation, the transition occurs at

\[ w_0 = \frac{\gamma}{2}(1 - 2x_0)^2 \approx \frac{\gamma}{2} \left( 1 - \frac{4k_0}{L} \right) \]  

(27)

for $x_0 = k_0/L \ll 1$.

#### 3.3. Corrections for finite sequence length

For small but finite $\epsilon$, corrections to the classical limit (26) have to be taken into account. With the help of the quantum mechanical picture, this can be achieved quite easily. If $f(x)$ is smooth, the ground-state wavefunction is localized near the minimum $x_c$ of the effective potential, and the shift in the ground-state energy can be computed by replacing $V(x)$ by a harmonic well:

\[ V(x) \approx V(x_c) + \frac{1}{2} V''(x_c)(x - x_c)^2 \]  

\[ = V(x_c) + \frac{1}{2} [4\gamma - f''(x_c)](x - x_c)^2. \]  

(28)

Identifying $1/\gamma$ with the mass $m$ of the quantum particle (compare to (24)), we see that this corresponds to a harmonic oscillator of frequency $\omega = 2\gamma / \sqrt{1 - f''(x_c)/4\gamma}$. The ground-state energy $\epsilon \omega/2$, together with the shift $\epsilon \gamma$ on the right-hand side of (24), thus gives rise to the leading order correction

\[ \Delta \Lambda = \frac{\gamma}{L} \left[ 1 - \sqrt{1 - f''(x_c)/4\gamma} \right], \]  

(29)

which coincides with (13) evaluated for $x_c \approx 1/2$. Similarly, the width of the wavefunction is given by

\[ \xi = \sqrt{\frac{\gamma \epsilon}{2m}} \sqrt{\frac{1}{8\sqrt{1 - f''(x_c)/4\gamma}}} \frac{1}{1/2}. \]  

(30)

In the case of the mesa landscape (16), the potential near $x_c = x_0$ consists of a linear ramp of slope

\[ -a = V'(x_0) = 2\gamma(2x_0 - 1) < 0 \]  

(31)

followed by a jump of height $w_0$. For small $\epsilon$, the jump can be considered as effectively infinite (as the

1 Note that, because of the factor $\sqrt{\Lambda}$ in (23), this is not equal to the width of the stationary population distribution.
the exponential factor $e^{\psi(x)}$ point at which the decay of the wavefunction $\gamma/L$ in considerations to be of the order $\bar{\hbar}/x$. The scaling $\Delta \sim L^{-2/3}$ was already noted in [5]. The width of the wavefunction can be estimated to be of the order $\xi \sim (\hbar^2/ma)^{1/3} \sim \epsilon^{2/3}$ in this case.

### 3.4. Comparison to the approach of Gerland and Hwa

We are now prepared to make contact with the approach of GH [3]. Assuming from the outset that the maximal number of mismatches is small compared to the sequence length, $1 \ll k_0 \ll L$, they neglect the contribution $2x$ in the drift term on the left-hand side of (19). The linear operator can then be symmetrized by the transformation

$$\phi(x) = e^{\psi(x)},$$

(34)

which is obtained from (23) by neglecting the terms quadratic in $x$ in $\phi_0$. This leads to a Schrödinger problem similar to (24), but with a potential that differs from $-f(x)$ only by the constant term $\gamma/2$. The error threshold is determined by the point at which the decay of the wavefunction $\psi(x)$ matches the exponential factor $e^{\psi(x)}$ in (34), such that $\phi(x)$ ceases to be normalizable. For $k_0 \gg 1$, the location of the transition is found by GH to be

$$w_0^\gamma = \frac{\gamma}{2} \left(1 + \frac{\pi^2}{k_0^2}\right),$$

(35)

which depends on the absolute number of mismatches $k_0$, but is independent of $L$.

To reconcile this with the result (27), we note that the semiclassical approximation must break down when the width of the semiclassical wavefunction, as estimated in subsection 3.3, becomes comparable to the width of the potential well provided by the fitness function. For the discontinuous mesa landscape, this occurs when

$$\xi \sim \epsilon^{2/3} \sim x_0 = \epsilon k_0 \Rightarrow k_0 \sim \epsilon^{-1/3} = L^{1/3}.$$

(36)

For a mesa that is shorter than $L^{1/3}$, the energy of the wavefunction is determined by its confinement on the scale $x_0$, and it can be estimated from standard quantum mechanical considerations to be of the order of $\hbar^2/(mx_0^2) \sim \gamma \epsilon^2/x_0^3 \sim \gamma/k_0^2$. For $k_0 \ll L^{1/3}$, this supersedes the contribution $\sim k_0/L$ on the right-hand side of (27). We conclude, therefore, that the leading ‘quantum’ correction to the ‘classical’ eigenvalue $\Lambda = w_0 - \gamma/2$ is a negative contribution proportional to $\gamma/k_0^2$, which leads to a corresponding positive shift in $w_0^\gamma$, in qualitative agreement with (35). For smooth fitness landscapes the breakdown of the semiclassical regime already occurs at $k_0 \sim L^{1/2}$, but the condition for the confinement energy contribution $\gamma/k_0^2$ to dominate the $k_0/L$ term in (27) still reads as $k_0 \ll L^{1/3}$.

This requires $\psi$ to decay on a scale of order unity in unscaled coordinates at the transition, which is actually inconsistent with the assumption of slow variation on the scale of the Hamming class index $k$ that underlies the continuum approximation.

### 3.5. Beyond the harmonic approximation

So far, we have worked in the harmonic approximation around $x = 1/2$, which breaks down near the boundaries $x = 0$ and $x = 1$. However, to access the regime $1 \ll k_0 \ll L$ considered by GH, an accurate treatment of the region of small $x \ll 1$ is clearly necessary. In Appendix A, we show how the quantum mechanical treatment can be extended such that it becomes quantitatively valid over the whole interval $0 \leq x \leq 1$. Based on the considerations of [22], we arrive at the modified Schrödinger equation

$$-\epsilon^2 \sqrt{x(1-x)} \frac{d^2}{dx^2} \psi + \left[\gamma(1 - 2\sqrt{x(1-x)}) - f(x)\right] \psi = -\Lambda \psi,$$

(37)

which differs from (24) in two respects. First, the potential (25) is replaced by

$$V_{\text{full}}(x) = \gamma(1 - 2\sqrt{x(1-x)}) - f(x).$$

(38)

In the asymptotic limit $\epsilon \to 0$ the principal eigenvalue is given by minimizing $V_{\text{full}}$, which exactly recovers the maximum principle (12). Second, the mass of the quantum particle described by (37) becomes position dependent:

$$m(x) \approx (2\sqrt{x(1-x)})^{-1},$$

(39)

which replaces the simple identification $m \approx 1/\gamma$ in the harmonic case. Inserting (38) and (39) into expression (32) for the finite sequence length correction yields

$$\Delta \Lambda = 2^{-1/3} \epsilon^{2/3} \gamma(1 - 2x_0)^{2/3} \sim [x_0(1 - x_0)]^{-1/6}.$$  

(40)

For fixed $x_0$ this still scales as $\epsilon^{2/3} = L^{-2/3}$, but when taking $L \to \infty$ at fixed $k_0$, such that $x_0 \to 0$, we find instead that

$$\Delta \Lambda \to 2^{-1/3} \epsilon^{2/3} \gamma x_0^{-1/6} \epsilon^{2/3} = 2^{-1/3} \epsilon^{2/3} \gamma k_0^{-1/6} L^{-1/2}.$$  

(41)

We next revisit the considerations of subsection 3.4. The width of the ground-state wavefunction is of order $\xi \sim (\hbar^2/ma)^{1/3}$, where both $m$ and $a$ now diverge as $x_0^{-1/2}$ for $x_0 \to 0$. Consequently, (35) is replaced by

$$\xi \sim (\hbar^2/ma)^{1/3} = \epsilon k_0^{1/3},$$

(42)

and we see that the condition $\xi \gg k_0$ for the breakdown of the semiclassical approximation is never satisfied. We conclude that the ‘quantum confinement’ regime discussed in subsection 3.4 in fact does not exist, and hence the improved semiclassical expression (40) for the finite sequence length correction is expected to remain valid for all $k_0$ and all $L$, provided that $k_0, L \gg 1$.

### 3.6. Numerical results

To test the analytical predictions derived in the preceding subsections, we have carried out a detailed numerical study of the dependence of $\Lambda$ and $w_0^\gamma$ on $k_0$, $L$, and $\gamma$. In figure 2, we show two examples for the dependence of $\Lambda$ on the plateau width $k_0$. The prediction of the asymptotic maximum principle (12) reproduces the qualitative behaviour of the numerical data but significantly overestimates the value of $\Lambda$. The $L^{-2/3}$ finite sequence length correction (32) derived in the harmonic approximation improves the comparison, but
length is been obtained by (numerical) calculation of the largest eigenvalue of the matrix defined by equation (10). The term (including the position-dependent mass) provides the best agreement with the numerics. The numerical values of the growth rate have been obtained by (numerical) calculation of the largest eigenvalue of the matrix defined by equation (10).

The solution of the maximum principle together with the $L^{-1/2}$ correction term (including the position-dependent mass) provides the best agreement with the numerics. The numerical values of the growth rate have been obtained by (numerical) calculation of the largest eigenvalue of the matrix defined by equation (10).

Figure 2. Growth rate $\Lambda$ as a function of the plateau width $k_0$ for two values of the plateau height, $w_0 = 0.5$ and $w_0 = 0.95$. The sequence length is $L = 100$ and the mutation rate per sequence is $\gamma = 1$. The solution of the maximum principle together with the $L^{-1/2}$ correction term (including the position-dependent mass) provides the best agreement with the numerics. The numerical values of the growth rate have been obtained by (numerical) calculation of the largest eigenvalue of the matrix defined by equation (10).

Figure 3. Critical plateau height as a function of the plateau width $k_0$. The sequence length is $L = 500$ and the mutation rate per sequence is $\gamma = 1$. The solution of the maximum principle together with the $L^{-1/2}$ correction term provides the best agreement with the numerics. With increasing $x_0$, the $L^{-1/3}$ and $L^{-1/2}$ corrections approach each other. The numerical values have been obtained via calculation of the average magnetization $M$ of the population and determining the plateau height, where $M$ jumps from a finite value to zero (see subsection 4.1 for a detailed discussion of $M$). The slight modulation of the red solid line arises from the finite numerical resolution of this procedure.

Quantitative agreement is achieved only using the refined expression (40), which is proportional to $L^{-1/2}$. Figure 3 shows a similar comparison for the critical plateau height $w_c^0$. Here the prediction (35) of GH is also included and seen to match the numerical outcome only poorly, whereas the MP result with the finite sequence length correction (40) produces excellent agreement. Finally, in the left picture of figure 4 we verify that the correction $\Lambda \Lambda$ indeed varies as $L^{-2/3}$ when $L$ is increased at fixed relative plateau width $x_0$. The right figure shows the corresponding $L^{-1/2}$ dependence for fixed absolute plateau width $k_0$.

4. Fitness landscapes with competing peaks

4.1. The selection transition

Since we have validated the analytical results of section 3 via numerical studies, we can now apply the analytical theory to the phenomenon of the survival of the flattest as explained in section 1. To be specific, we want to find out whether a broad plateau outcompetes a smaller but higher one even in the limit of long sequences.

In the literature, this question has already been discussed to some extent by Schuster and Swetina [7]. The question can be answered by investigating a fitness landscape consisting of two fitness plateaus with fixed absolute widths at the opposing ends of the Hamming space (see figure 5). As was shown in [7], for small $\mu$ the interference between the two plateaus is negligible when they are separated by a few mutational distances. The stationary state of the system is therefore to a very good approximation determined by the comparison between the population growth rates associated with each of the two plateaus in isolation.

Observing the centre of mass of the population as a function of the mutation rate and for fixed sequence length, we find two types of transitions. The first one is a jump of the population from the higher to the broader plateau, which we will refer to as the selection transition [25] taking place at a mutation rate $\mu_s$. The second transition is the well-known error threshold taking place at $\mu_{tr}$, where the population becomes uniformly spread in the sequence space. To analyse these transitions, an order parameter is needed. A convenient quantity is the population-averaged ‘magnetization’ defined by

$$M = 1 - 2\langle x \rangle \in [-1, 1] \quad \text{with} \quad \langle x \rangle = \frac{1}{L} \sum_{k=0}^{L} k P_k^*.$$  (43)

If the whole population consists only of master sequences, the magnetization is $M = 1$. If only the inverse master sequence is present, the magnetization becomes $M = -1$. For a uniform distribution in the sequence space (delocalized population), the magnetization is $M = 0$. Thus, we can distinguish the qualitatively different states of the population in the two-plateau landscape by considering the population-averaged magnetization $M$ as a function of $\mu$.

As can be seen from figure 6, the selection transition (the jump between the two plateaus) is sharp even for finite sequence lengths, whereas the error threshold is a continuous

$$M = 0 \quad \text{holds in fact for any distribution symmetric to } x = 1/2.$$  But since here such distributions do not arise, we may identify $M = 0$ with a delocalized population.
transition for finite sequence length and only becomes sharp in the limit of infinite sequence length [25]. With growing sequence length the two critical mutation rates $\mu_s$ and $\mu_t$ become smaller and also approach each other until, at a critical sequence length $L^*$, the selection transition completely disappears. For sequences longer than $L^*$, the population delocalizes directly from the high, narrow peak and the low, broad plateau is never substantially populated.

With the help of the maximum principle, this surprising behaviour can be easily understood. Using (12), the selection threshold is obtained by equating the population mean fitness for the two competing plateaus, which yields

$$
\mu_p = \frac{w_0 - w_1}{2(\sqrt{k_1L - k_1^2} - \sqrt{k_0L - k_0^2})} \approx \frac{w_0 - w_1}{2(\sqrt{k_1} - \sqrt{k_0})} L^{-1/2},
$$

(44)

where $k_0, k_1 \ll L$ has been assumed in the last step. On the other hand, the error threshold $\mu_{tr}^{(i)}$ associated with plateau $i = 0, 1$ is determined by the vanishing of the corresponding principal eigenvalue $\lambda_i$, which gives

$$
\mu_{tr}^{(i)} = \frac{w_i}{L(1 - 2\sqrt{k_iL - k_i^2})} \approx \frac{w_i}{1 - 2\sqrt{k_i/L}} L^{-1}.
$$

(45)

The different scaling of the two types of thresholds with sequence length implies that for large $L$, the error threshold of the higher peak is encountered before the selection threshold, which therefore is no longer observable. The critical sequence length $L^*$ where the selection transition vanishes can be estimated by equating the approximate expressions (44) and (45), which yield

$$
L^* \approx \frac{4(w_0\sqrt{k_1} - w_1\sqrt{k_0})^2}{(w_0 - w_1)^2}.
$$

(46)

Following [7, 25], in our numerical work we have considered short plateaus, $k_0 = 1$ and $k_1 = 2$, with relative fitness values $w_1/w_0 = 0.9$, for which (46) gives $L^* \approx 106$. Comparison with the numerical values for the selection and error thresholds in figure 7 shows that this significantly underestimates the value of $L^*$; moreover, the agreement between theory and numerics is not substantially improved by using the full expressions for the principal eigenvalues $\lambda_0$ and $\lambda_1$, including the $L^{-1/2}$ correction derived in subsection 3.5. This is not surprising, as the continuum approach developed in section 3 cannot be expected to be quantitatively accurate for plateau sizes of order unity.

For completeness we mention that for plateau widths scaling with the sequence length (such that $x_0 = k_0/L$ and $x_1 = k_1/L$ are kept fixed as $L \to \infty$), the selection transition is maintained at a fixed value of $\gamma$ [20].

4.2. The ancestral distribution

In addition to the equilibrium population distribution $P^*_k$ attained at long times, we can also consider the ancestral distribution, the equilibrium distribution of the backward time process, as introduced by Baake and collaborators [18, 23]. The ancestral distribution $a_k$ gives information on the origin of the equilibrium population and is obtained as the product of the right eigenvector $P^*_k$ and the left eigenvector $P_{kk}^{**}$ of the mutation-selection matrix defined through (10), $a_k \sim P^*_k \cdot P_{kk}^{**}$.

For the fitness landscape with two competing peaks, we find an ancestral population that is either located on one of the plateaus or uniformly distributed in the sequence space (figure 8). The transitions between these states are all of first order. The continuous character of the error threshold transition of the equilibrium population distribution, as opposed to the discontinuous transition of the ancestral distribution, can be explained by the growing mutational pressure affecting the population on the plateau and driving
it toward the middle of the Hamming space. Mutations cause the population to ‘leak out’ from the plateau. Nevertheless, the individuals maintaining the population and compensating for the mutational loss are those with the highest fitness, which are located on the plateau and make up the ancestral distribution.

Before closing the analysis of plateau-shaped fitness landscapes, we want to mention the connection between our description and the popular language of Ising chains or semi-infinite Ising models [32, 33]. In the Ising picture, the ancestral distribution becomes the bulk distribution on a semi-infinite two-dimensional (spatial or spatio-temporal) lattice, and the equilibrium distribution becomes the distribution in the surface layer. This analogy can be seen very clearly in the paper by Tarazona [25], where the different orders of the transitions in the two distributions are explained in terms of surface wetting.

5. Epistasis and the error threshold

So far, we have discussed the robustness of phenotypes using plateau-shaped fitness landscapes, which are a special case of the class of epistatic fitness functions. We now want to discuss the latter in a more general framework. Epistasis describes the non-linear dependence of the fitness function on the number of mismatches \( k \) [15]. Every additional mismatch is penalized harder (synergistic epistasis of deleterious mutations) or less hard (diminishing epistasis) than the previous one. Here we address the effect of epistasis on the existence of an error threshold, which is defined for our purposes as a singularity in the dependence of the population mean fitness on the mutation rate. In general (but not always, see below), such a singularity is associated with a discontinuous jump in the location of the most populated genotype.

Following Wiehe [16], we consider the class of epistatic fitness functions

\[
w_k = w_0 - bk^\alpha,
\]

(47)

where \( k \) is again the Hamming distance to the master sequence and \( b > 0 \). The epistasis exponent \( \alpha \) takes the value \( \alpha = 1 \) in the non-epistatic case, while \( \alpha > 1 \) and \( \alpha < 1 \) produce landscapes with synergistic and diminishing epistasis, respectively. For \( \alpha \to 0 \), (47) reduces to the sharp peak landscape \( w_k = w_0 - b(1 - \delta_k) \). It is well known that an error threshold exists for \( \alpha \to 0 \), but not for \( \alpha = 1 \) [6]. Neglecting backward mutations, Wiehe argued in [16] that an error threshold emerges whenever \( \alpha < 1 \). In the following we show that, based on the maximum principle (12), the critical value of the epistasis exponent below which an error threshold develops is in fact \( \alpha = 1/2 \). This result can also be inferred from [18], where a general criterion for the existence of an error threshold is given for a larger class of mutation schemes; for unidirectional mutations, this criterion coincides with Wiehe’s result.

As before, we work in the scaling limit \( L \to \infty \) and \( \mu \to 0 \) with the mutation rate per sequence \( \gamma = \mu L = \text{const.} \) In order to cast (47) into form (11) required for the application of the maximum principle, we write

\[
w_k = f(x) = w_0 - \tilde{b}x^\alpha, \quad \text{with} \quad x = k/L, \quad \tilde{b} = bL^\alpha,
\]

(48)
and the limit \( L \to \infty \) should be combined with \( b \to 0 \), such that \( \tilde{b} = \text{const} \). Since \( b \) can be interpreted as an effective selection coefficient, we thus consider a situation where both the mutation rate (per site) and the selection forces are small. Applying the maximum principle (12) to this landscape, the mean fitness \( \Lambda \) of the population in the equilibrium state is given by

\[
\Lambda = \max_{x \in [0,1]} \left\{ w_0 - \tilde{b} x^a - \gamma \left[ 1 - 2 \sqrt{x(1-x)} \right] \right\} \equiv \max_{x \in [0,1]} \lambda(x),
\]

where \( \lambda(x) \) is the function inside the curly brackets.

To find the condition under which the maximum is attained inside the interval \( x \in (0,1) \), we set \( d\lambda/dx = 0 \), yielding the condition

\[
\frac{\gamma}{\tilde{b}} (1 - 2x) = ax^{a-1/2} \sqrt{1-x}.
\]

(50)

For \( a > 1/2 \) the right-hand side is a convex function which vanishes at \( x = 0,1 \), with an infinite slope at \( x = 1 \). As a consequence, there always exists a unique solution \( x_c \in (0,1) \) for any value of \( \gamma/\tilde{b} \), which describes the location of the population for \( L \to \infty \). The location varies smoothly from \( x_c = 0 \) for \( \gamma/\tilde{b} \to 0 \) to \( x_c \to 1/2 \) for \( \gamma/\tilde{b} \to \infty \), and there is no error threshold. However, for \( a < 1/2 \) the right-hand side diverges at \( x = 0 \), and there is no solution for small \( \gamma/\tilde{b} \). The function \( \lambda(x) \) then monotonically decreases, which implies that the maximum in (49) is located at the boundary point \( x = 0 \) over a finite interval of \( \gamma/\tilde{b} \). Increasing \( \gamma/\tilde{b} \) the function \( \lambda(x) \) develops a local maximum, which eventually exceeds the boundary value \( \lambda(0) = w_0 - \gamma \). At this point, the population discontinuously delocalizes to an interior point \( x_c \in (0,1) \). The error threshold condition is of the form \( \gamma/\tilde{b} = g(\alpha) \), which translates into the expression

\[
\mu_u = \frac{\gamma_u}{L} = \frac{\tilde{b} g(\alpha)}{L} = g(\alpha) b L^{a-1},
\]

(51)

for the critical mutation rate \( \mu_u \). The same scaling was obtained in [16]. In the sharp peak limit \( a \to 0 \) the threshold occurs at \( \gamma/\tilde{b} = \gamma/\tilde{b} = 1 \), which implies that \( g(0) = 1 \). On the other hand, for \( a = 1/2 \) the expansion of \( \lambda(x) \) near \( x = 0 \) reads as

\[
\lambda(x) \approx w_0 - (\tilde{b} - 2\gamma) x^{1/2} - \gamma x^{3/2},
\]

(52)

which shows that \( g(1/2) = 1/2 \). For \( \gamma/\tilde{b} > 1/2 \) an interior maximum appears at \( x_c = (2 - \tilde{b}/\gamma)/3 \), which moves continuously away from \( x = 0 \). In the language of phase transitions, \( \alpha = 1/2 \) can thus be viewed as a critical endpoint terminating the line of discontinuous phase transitions that occur for \( \alpha < 1/2 \).

Our predictions are fully confirmed by numerical calculations for finite sequence length. Figure 9 illustrates, for \( \alpha = 1/3 < 1/2 \), how the maximum of the stationary population distribution \( P_k^* \) shifts from the boundary \( k = 0 \) to an interior point as the error threshold is crossed. In figure 10, we demonstrate the existence of an error threshold for \( \alpha < 1/2 \) and its absence for \( \alpha > 1/2 \) by showing the behaviour of the magnetization \( M \) as a function of \( \gamma \) for two different cases. The magnetization displays a non-analytic jump for \( \alpha < 1/2 \) and
Calculations have been done for a sequence length of $L$ discontinuous jump, whereas for $\alpha$ terminates in a second-order phase transition at $x_c > 0$ (above the line) to a delocalized state the error threshold as function have been studied previously in the context of regulatory motifs describing the mutational robustness of phenotypes, which landscapes, with particular emphasis on mesa landscapes In this paper we discussed the properties of epistatic fitness functions varying along a single direction in the space of genotypic sequences. However, such fitness functions represent only the simplest kind of epistasis; more complex forms, such as multidimensional epistasis and sign epistasis are likely to be prevalent in most biological applications. Future work therefore needs to develop new concepts to quantify general forms of epistatic interactions and to explore their evolutionary consequences.

Figure 10. Magnetization as a function of the mutation rate for the fitness landscape (47) with epistasis exponents $\alpha = 0.4$ and $\alpha = 0.52$, respectively. For $\alpha = 0.4$ the magnetization undergoes a discontinuous jump, whereas for $\alpha = 0.52$ it changes smoothly. Calculations have been done for a sequence length of $L = 500$.

Figure 11. Numerically determined phase diagram for the epistatic fitness landscape (47). At the thick line, the population undergoes a first-order phase transition from a state localized at $x_c = 0$ (below the line) to a delocalized state $x_c > 0$ (above the line). This line terminates in a second-order phase transition at $\alpha = 1/2$. The deviation from the prediction $\gamma/b = g(1/2) = 1/2$ at $\alpha = 1/2$ is due to finite sequence length corrections. For all larger values of the epistasis exponent, $\alpha > 1/2$, the population changes smoothly. Calculations have been performed for a sequence length of $L = 750$. The stair-like character of the line is only due to the numerical step size.

6. Conclusions

In this paper we discussed the properties of epistatic fitness landscapes, with particular emphasis on mesa landscapes describing the mutational robustness of phenotypes, which have been studied previously in the context of regulatory motifs as a population evolution model we used the continuous time Crow–Kimura model, which is a quasispecies model for asexual and haploid organisms, and analysed its stationary states for sequences consisting of two letters. As explained in appendix B, it is straightforward to generalize our results to sequence alphabets of a general size.

We reviewed two existing approaches to this problem and explained the discrepancy between their predictions by extending the approach of Gerland and Hwa beyond the harmonic approximation. Based on a quantum mechanical analogy we derived a novel finite sequence length correction term to the maximum principle of [18], which significantly improves the agreement with numerical calculations. Our central result is that the relative number of tolerable mismatches $x_0 = k_0/L$ is the relevant parameter for the fitness effect of mutational robustness, and we provide accurate formulae for its quantitative evaluation. As a consequence, we showed that the selection transition first described by Schuster and Swetina (7) disappears for long sequences.

Finally, in section 5, we discussed more general forms of epistatic fitness landscapes with regard to the existence of an error threshold. Based on the results of [18, 23], we improved on earlier work [16] and showed that diminishing epistasis ($\alpha < 1$ in the fitness function (47)) is not a sufficient condition for an error threshold to occur.

In conclusion, the theory of mutation-selection models has reached a high degree of sophistication, and precise quantitative predictions are possible for one-dimensional fitness functions varying along a single direction in the space of genotypic sequences. However, such fitness functions represent only the simplest kind of epistasis; more complex forms, such as multidimensional epistasis and sign epistasis, are likely to be prevalent in most biological applications. Future work therefore needs to develop new concepts to quantify general forms of epistatic interactions and to explore their evolutionary consequences.

Acknowledgments

We are grateful to Alexander Altland, Ellen Baake, Ulrich Gerland, Kavita Jain, Luca Peliti and David Saakian for useful discussions. This work was supported by DFG within the Bonn Cologne Graduate School of Physics and Astronomy, SFB 680 Molecular basis of evolutionary innovations and SFB/TR12 Symmetries and Universality in Mesoscopic Systems.

Appendix A. The large deviations approach

We start by symmetrizing the eigenvalue problem (10). The discrete analogue of the transformation (23) is

$$Q_k = \left(\frac{L}{2}\right)^{1/2} P_k^\star,$$

which leads to

$$\Lambda Q_k = (w_k - \gamma)Q_k + \mu\sqrt{(L - k)(k + 1)}Q_{k+1}$$

$$+ \mu\sqrt{(L - k + 1)k}Q_{k-1}.$$  

(A.2)

Following [22], we now perform the continuum limit by making a large deviations ansatz for $Q_k$:

$$Q_k = Q_{1L} = \psi(x) = \exp[-\epsilon^{-1}u(x)].$$

(A.3)
with \( \epsilon = 1/L \). Inserting this into (A.2), one finds
\[
(\Lambda - f(x) + \gamma) \psi = 2\sqrt{\gamma}/\sqrt{1-x} \cosh[u']\psi. \tag{A.4}
\]
Cancelling \( \psi \) on both sides yields a Hamilton–Jacobi equation for the ‘action’ \( u(x) \), with \( u' = du/\mathrm{d}x \) playing the role of a canonical momentum [22]. In order to cast (A.4) into the form of a Schrödinger equation, we expand the momentum-dependent factor to quadratic order, \( \cosh(u') \approx 1 + (u')^2/2 \), and make use of the relation
\[
(u')^2 = \epsilon^2/\sqrt{\gamma}(1 + (u')^2/2), \tag{A.5}
\]
which follows from (A.3) to leading order in \( \epsilon \). Inserting this into (A.4) results in (37).

Appendix B. General alphabet size

Here we show how our results generalize to the case where the symbols in the genetic sequence are taken from an alphabet of \( A > 2 \) letters (for nucleotide sequences, \( A = 4 \)). We assume a uniform point mutation rate \( \mu \), connecting any two of the possible states of a site in the sequence and a fitness function \( w_0 \) that depends on the relative number of mismatches according to (11). It is then straightforward to see that the basic eigenvalue problem (10) generalizes to
\[
(\Lambda - w_0)P^*_k = (\Lambda - 1)\gamma \left[ \frac{k + 1}{A - 1} P^*_k + (L - k + 1)P^*_{k-1} \right] - \left[ L - k + \frac{k}{A - 1} \right] P^*_k. \tag{B.1}
\]
Applying the results of [19, 36, 37] to this problem we find that, asymptotically for large \( L \), the principal eigenvalue is given by the maximum principle
\[
\Lambda = \max_{x \in [0,1]} \left\{ f(x) - (\Lambda - 1)\gamma \left[ 1 - \frac{(A - 2)x}{A - 1} \right] \right\}. \tag{B.2}
\]
For the case of the mesa landscape (15), this implies that the population is localized near the edge \( k = k_0 \) of the fitness plateau for \( w_0 > w_0^* \) with
\[
w_0^* = \gamma/(\Lambda - 1) \left[ 1 - \frac{(A - 2)x_0}{A - 1} \right] - \frac{2\sqrt{x_0(1-x_0)}}{\sqrt{A - 1}}. \tag{B.3}
\]
The right-hand side is a monotonically decreasing function of \( x_0 \) which vanishes at \( x_0 = 1 - 1/A \). For fixed \( x_0 \), it is an increasing function of \( \Lambda \).

Glossary

**Epistasis.** Nonlinear dependence of the fitness effect on the number of mutations, when all mutations are either beneficial (increasing fitness) or deleterious (decreasing fitness). Sign epistasis implies that the sign (beneficial or deleterious) of a mutation depends on the presence of other mutations.

**Error threshold.** Non-analytic behaviour of the mean population fitness and the stationary population distribution upon the variation of model parameters, such as the mutation rate, the height or the width of the fitness plateau. For fitness landscapes which are peaked around an optimal genotype \( x_0 \), the stationary population distribution delocalizes from the peak when the error threshold is crossed.

**Hamming space.** The one-dimensional space of Hamming classes encompassing all sequences with the same number of mismatches with respect to a reference genotype, usually the genotype of optimal fitness. For the case of binary sequences, the number of sequences in the Hamming class \( k \) is \( \binom{x}{A} \). The Hamming distance between two sequences is equal to the number of letters in which they differ.

**Maximum principle.** The one-dimensional variational principle (12) for the principal eigenvalue of the mutation-selection operator. It gives the exact eigenvalue in the limit \( L \to \infty \) for fitness functions that depend only on the Hamming class index \( k \).

**Schrödinger equation.** The stationary, one-dimensional Schrödinger equation for the quantum-mechanical wavefunction \( \psi(x) \) describing a particle of mass \( m \) in a potential \( V(x) \) reads as
\[
-h^2/2m \frac{d^2\psi}{dx^2} + V(x)\psi = E\psi.
\]
This is an eigenvalue problem for the energy \( E \) of the particle. The ground state is the solution of the lowest energy. The harmonic oscillator potential mentioned in subsection 3.3 is given by \( V(x) = \frac{1}{2}m\omega^2x^2 \), where \( \omega \) is the oscillator frequency. The classical limit amounts to taking Planck’s constant \( h \to 0 \), and semiclassical theory is concerned with the study of the solutions of the Schrödinger equation when \( h \) is small but nonzero.

References

[1] de Visser J A G M et al 2003 Perspective: evolution and detection of genetic robustness *Evolution* 57 1959–72
[2] Krakauer D C and Plotkin J B 2002 Redundancy, antiredundancy, and the robustness of genomes *Proc. Natl Acad. Sci.* 99 1405–9
[3] Gerland U and Hwa T 2002 On the selection and evolution of regulatory DNA motifs *J. Mol. Evol.* 55 386–400
[4] Berg J, Willmann S and Lässig M 2004 Adaptive evolution of transcription factor binding sites *BMC Evol. Biol.* 4 42
[5] Peliti L 2002 Quasispecies evolution in general mean-field landscapes *Europhys. Lett.* 77 745
[6] Jain K and Krug J 2007 Adaptation in simple and complex fitness landscapes *Structural Approaches to Sequence Evolution: Molecules, Networks and Populations* ed U Bastolla, M Porto, H E Roman and M Vendruscolo (Berlin: Springer) pp 299–339
[7] Schuster P and Swetina J 1988 Stationary mutant distributions and evolutionary optimization *Bull. Math. Biol.* 50 635–60
[8] Wilke C O, Wang J L, Ofria C, Lenski R E and Adami C 2001 Evolution of digital organisms at high mutation rates leads to survival of the flattest *Nature* 412 531–3
[9] Wilke C 2001 Selection for fitness versus selection for robustness in RNA secondary structure folding *Evolution* 55 2412
[10] Tannenbaum E and Shakhnovich E 2004 Solution of the quasispecies model for an arbitrary gene network *Phys. Rev.* E 70 021903
[11] Codoner F M, Daros J-A, Solé R V and Elena S F 2006 The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens PLoS Pathogens 2 e136

[12] Sanjuan R, Cuevas J M, Furio V, Holmes E C and Moya A 2007 Selection for robustness in mutagenized RNA viruses PLoS Genet. 3 e93

[13] Rolland M, Brander C, Nickle D C, Herbeck J T, Gottlieb G S, Campbell M S, Maust B S and Mullins J I 2007 HIV-1 over time: fitness loss or robustness gain? Nature Rev. Microbiol. 5 1–2

[14] Sardanyés J, Elena S F and Solé R V 2008 Simple quasispecies models for the survival-of-the-flattest effect: the role of space J. Theor. Biol. 250 560–8

[15] Phillips P C, Otto S P and Whitlock M C 2000 Beyond the average: the evolutionary importance of gene interactions and variability of epistatic effects Epistasis and the Evolutionary Process ed J B Wolf, E D Brodie III and M J Wade (Oxford: Oxford University Press)

[16] Wiehe T 1997 Model dependency of error thresholds: the role of fitness functions and contrasts between the finite and infinite sites models Genet. Res. Camb. 69 127–36

[17] Jain K 2008 Loss of least-loaded class in asexual populations due to drift and epistasis Genetics 179 2125

[18] Hermisson J, Redner O, Wagner H and Baake E 2002 Mutation-selection balance: ancestry, load, and maximum principle Theor. Pop. Biol. 62 9–46

[19] Baake E, Baake M, Bovier A and Klein M 2005 An asymptotic maximum principle for essentially linear evolution models J. Math. Biol. 50 83–114

[20] Saakian D B and Hu C-K 2006 Exact solution of the Eigen model with general fitness functions and degradation rates Proc. Natl Acad. Sci. 103 4935–9

[21] Park J M and Deem M W 2006 Schwinger boson formulation and solution of the Crow–Kimura and Eigen models of quasispecies theory J. Stat. Phys. 125 975

[22] Saakian D B 2007 A new method for the solution of model of biological evolution: derivation of exact steady-state distributions J. Stat. Phys. 128 781–98

[23] Baake E and Georgii H-O 2007 Mutation, selection, and ancestry in branching models: a variational approach J. Math. Biol. 54 257

[24] Sato K and Kaneko K 2007 Evolution equation of phenotype distribution: general formulation and application to error catastrophe Phys. Rev. E 75 061909

[25] Tarazona P 1992 Error thresholds for molecular quasispecies as phase transitions: from simple landscapes to spin-glass models Phys. Rev. A 45 6038

[26] Eigen M 1971 Selforganization of matter and the evolution of biological macromolecules Naturwissenschaften 58 465

[27] Crow J F and Kimura M 1970 An Introduction to Population Genetic Theory (New York: Harper and Row)

[28] Hofbauer J 1985 The selection mutation equation J. Math. Biol. 23 41–53

[29] Thompson C J and McBride J L 1974 On Eigen’s theory of the self-organization of matter and the evolution of biological macromolecules Math. Biosci. 21 127

[30] Sornette D 2000 Critical Phenomena in Natural Sciences (Berlin: Springer)

[31] Rollnik H 1995 Quantentheorie I (Wiesbaden: Vieweg)

[32] Leuthäusser I 1987 Statistical mechanics of Eigen’s evolution model J. Stat. Phys. 48 343

[33] Baake E and Wagner H 2001 Mutation-selection models solved exactly with methods from statistical mechanics Genet. Res. 78 93

[34] Kondrashov F A and Kondrashov A S 2001 Multidimensional epistasis and the disadvantage of sex Proc. Natl Acad. Sci. 98 12089–92

[35] Weinreich D M, Watson R A and Chao L 2005 Perspective: sign epistasis and genetic constraint on evolutionary trajectories Evolution 59 1165–74

[36] Garske T and Grimm U 2004 A maximum principle for the mutation-selection equilibrium of nucleotide sequences Bull. Math. Biol. 66 397

[37] Garske T and Grimm U 2004 Maximum principle and mutation thresholds for four-letter sequence evolution J. Stat. Mech. P07007