6. Finite Population Dynamics

As we saw in section 4, the fitness distribution state space for a finite population is given by a rectangular lattice Λ_M over the $N+1$ dimensional simplex Λ_∞ with lattice spacing $1/M$. As in the infinite population case, applying the generation operator $\mathbf{G}$ to a fitness distribution $\vec{P} \in \Lambda_M$ gives the *expected* fitness distribution $\langle \vec{P} \rangle$ at the next time step. Viewed in a slightly different way, $\mathbf{G}_i(\vec{P}) = \langle P_i \rangle$ is the probability that if one string is selected from $\vec{P}$ and mutated, it will turn out to have fitness i. Since the new population is created from M of these selections and mutations, it is clear that the new population is a sample of size M of the distribution $\langle \vec{P} \rangle$. In the infinite population case the *expected* distribution $\langle \vec{P} \rangle$ is always attained; there are no fluctuations in the $\langle P_i \rangle$ in this limit. For a finite population of size M, the finite sampling from $\langle \vec{P} \rangle$ leads to stochasticity in the dynamics. At each time step the population that is actually attained will fluctuate around the expected distribution $\langle \vec{P} \rangle$. We therefore see that the flow in state space for a finite population is given by the same flow operator $\mathbf{G}$ but now over a discrete space and with noise added due to the population's finiteness. It is important to note that the finite population sampling noise added to the operation of $\mathbf{G}$ takes into account the combined effect of the fluctuations in both mutation and selection. Normally, one would consider the sampling fluctuations in mutation and selection to be two distinct sources of fluctuations. In combining the selection and mutation effects into the operator $\mathbf{G}$, all fluctuations are due to the finite population sampling applied to the operation of $\mathbf{G}$.

In general, the probability $\Pr[\vec{P}^n \to \vec{P}^m]$ that a fitness distribution $\vec{P}^n = (n_0, n_1, \ldots, n_N)/M$ will go to a fitness distribution $\vec{P}^m = (m_0, m_1, \ldots, m_N)/M$ under mutation and selection is given by a multinomial sampling distribution with mean $\mathbf{G}(\vec{P}^n)$:

$$\Pr[\vec{P}^n \to \vec{P}^m] = M! \prod_{i=0}^{N} \frac{[\mathbf{G}_i(\vec{P}^n)]^{m_i}}{m_i!}, \tag{6.1}$$

where $\mathbf{G}_i(\vec{P}^n)$ is the expected proportion of individuals with fitness i at the next generation. The multinomial distribution is nothing more than the distribution of a random sample of size M of the expected distribution $\mathbf{G}(\vec{P})$. In Appendix B this distribution is used to prove that for large populations the finite population dynamics approaches the infinite population dynamics arbitrarily closely.

6.1 Fitness Epochs

We will now turn to the fitness epochs that occur for finite population size. As we have seen, $\mathbf{G}(\vec{P})$ gives the expected distribution at the next time step if the current distribution is $\vec{P}$. The expected flow $\langle d\vec{P} \rangle$ in fitness distribution space is given by $\langle d\vec{P} \rangle = \mathbf{G}(\vec{P}) - \vec{P}$. Now let us consider what happens if the absolute value $|\langle dP_i \rangle|$ of a certain component $\langle dP_i \rangle$ of the expected flow is much smaller than the lattice spacing; that is, if $|\langle dP_i \rangle| \ll 1/M$. The fitness distribution $\vec{P}$ for a finite population can only be on the lattice points of Λ_M, since the actual difference dP_i can only be $\ldots -1/M\ 0,\ 1/M,\ 2/M \ldots$. Thus, we expect the actual difference dP_i to be 0 most of the time; this means that the actual component P_i will likely not change for some time. For example, in the previous case with $N = 3$, $K = 4$, and $q = 0.01$, but with finite $M = 100$, we find that if $P_3 = 0$, the dynamics will push P_0, P_1, and P_2 into a region where

$$|\langle dP_3 \rangle| \approx 6.5 * 10^{-4} \ll \frac{1}{M} = 0.01. \tag{6.2}$$

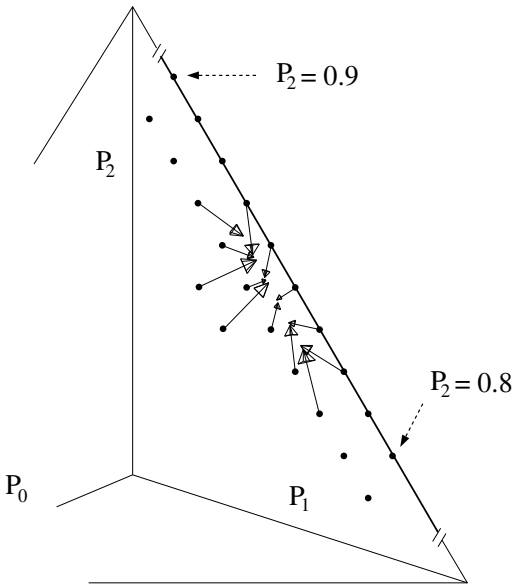

Figure 5: Face view of the flow $\langle d\vec{P} \rangle$ within the surface $P_3 = 0$. The flow of fitness distributions centers around an area where $P_2 \approx 0.85$, $P_1 \approx 0.14$ and $P_0 \approx 0.01$. This region is shown substantially enlarged. Parameters are set to $N = 3$, $K = 4$, $q = 0.01$, and $M = 100$. The dots indicate the allowed finite populations on a portion of Λ_M.

Since $\langle dP_3 \rangle$ is so small[7] compared to $1/M$, we expect the distribution to stabilize itself in that region for some time. Figures 5 and 6 illustrate this behavior. Figure 5 shows the flow at points in the plane $P_3 = 0$ as arrows whose length is the value of $|\langle d\vec{P} \rangle|$. We see how the flow stabilizes the dynamics in an area on the surface where $\vec{P} = (P_0, P_1, P_2, P_3) \approx (0.01, 0.14, 0.85, 0.0)$. Figure 6 shows a side view, perpendicular to the $P_3 = 0$ plane. The arrows show the expected change in the component dP_3 off the $P_3 = 0$ surface. We see that the flow off the surface is so small that the dynamics is likely to remain on that surface for some time. This region on the $P_3 = 0$ surface is a likely place for a fitness epoch to occur.

We would like to have a way to determine *where* metastable regions like the one shown in figure 6 are to be found in the state space. In section 5.5 we saw that all eigenvectors of the operator $\tilde{\mathbf{G}}$ correspond to fixed points of $\mathbf{G}$. It is therefore natural to assume that the metastable regions of the finite population dynamics are to be found in the vicinities of the eigenvectors $\vec{V}^i$. We saw, in addition, that the average fitness in the population in the neighborhood of the eigenvectors $\vec{V}^i$ is equal to the eigenvalues g_i of the operator $\tilde{\mathbf{G}}$. We therefore expect the population to have mean fitness close to g_i during fitness epoch i.

6.2 Predicted Epoch Levels

This section assesses under what circumstances we can predict the mean fitness of the population during an epoch and lists additional aspects of the empirical GA behavior that will be explained in subsequent sections.

Figures 7(a) through 7(i) show the same runs as those in figures 1(a) through 1(i) together with the theoretical predictions of the fitness levels (horizontal lines) at which we expect the

[7] The expression for $\langle dP_3 \rangle$ can be obtained by plugging the epoch center $\vec{P} = (0.01, 0.14, 0.85, 0.0)$ into $\langle d\vec{P} \rangle = \mathbf{G}(\vec{P}) - \vec{P}$. Epoch centers are defined in section 6.3.

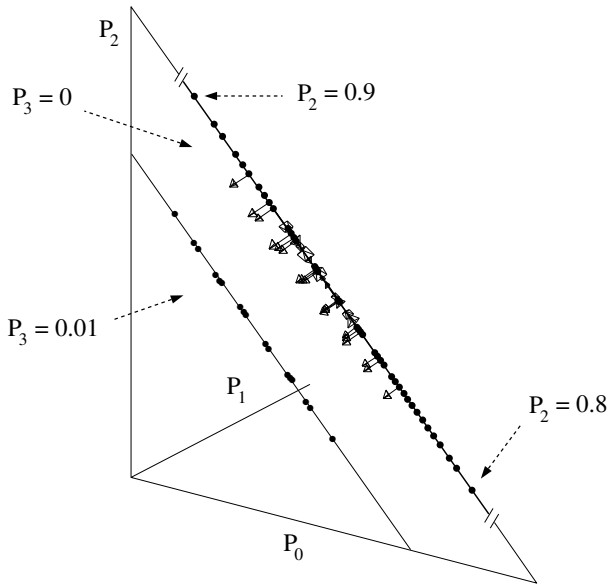

Figure 6: Side view of the simplex perpendicular to the surfaces $P_3 = 0$ and $P_3 = 0.01$, which are shown as the upper and lower diagonal solid lines, respectively. The dots indicate the allowed finite populations in Λ_M. The expected change $\langle dP_3 \rangle$ (arrows) is so small that the dynamics is likely to stay on the surface $P_3 = 0$ for some time. Note that the arrows emanating from the $P_3 = 0$ surface have been magnified five times to make them visible. Parameters are $N = 3$, $K = 4$, $q = 0.01$, and $M = 100$.

epochs to occur. The levels were calculated numerically by determining the eigenvalues g_i of $\tilde{\mathbf{G}}$ for the different parameter settings of each run. Recall that the fitness levels are not a function of the population size M; they are determined by the infinite population dynamics.

Caption for figure 7: Average fitness over time for nine runs of the Royal Road GA with different parameter settings together with the theoretical predictions (horizontal lines) for the epoch levels. See figures 1(a) through 1(i) for the corresponding parameter settings for each run. Although there are N possible fitness levels in each case, for the sake of clarity only subsets are shown. Figures 7(a) and 7(d) show epoch levels g_3 through g_{10} and figure 7(g) shows levels g_4 through g_{10}. Note that the theoretical levels are the same for these three runs since the epoch levels don't depend on population size. Figure 7(b) shows predicted levels g_9 through g_{20}. Only the asymptotic epoch g_{20} can be seen to occur in this run. The same epoch levels g_9 through g_{20} are shown in figure 7(c). In run 7(c) epochs can be seen to occur at g_{14}, g_{17}, and g_{20}. Figure 7(e) shows epoch levels g_5 through g_{10}. No epochs can be seen to occur for these parameter settings. Figure 7(f) shows the last two of the epoch levels g_9 and g_{10} as dashed lines. The solid lines in 7(f) show the upper bounds on these epoch levels (see text for discussion). Epoch levels g_3 through g_{10} are shown in figure 7(h) in which no clear epochs can be distinguished. Although, one can see time intervals for which the average fitness fluctuates around g_5. Finally, figure 7(i) shows all epoch levels g_1 through g_4. In all of these figures, the epoch levels g_i were obtained by numerically solving for the eigenvalues of $\tilde{\mathbf{G}}$ for each of the parameter settings. In section 6.4 we calculate simple analytical expressions for these epoch levels.

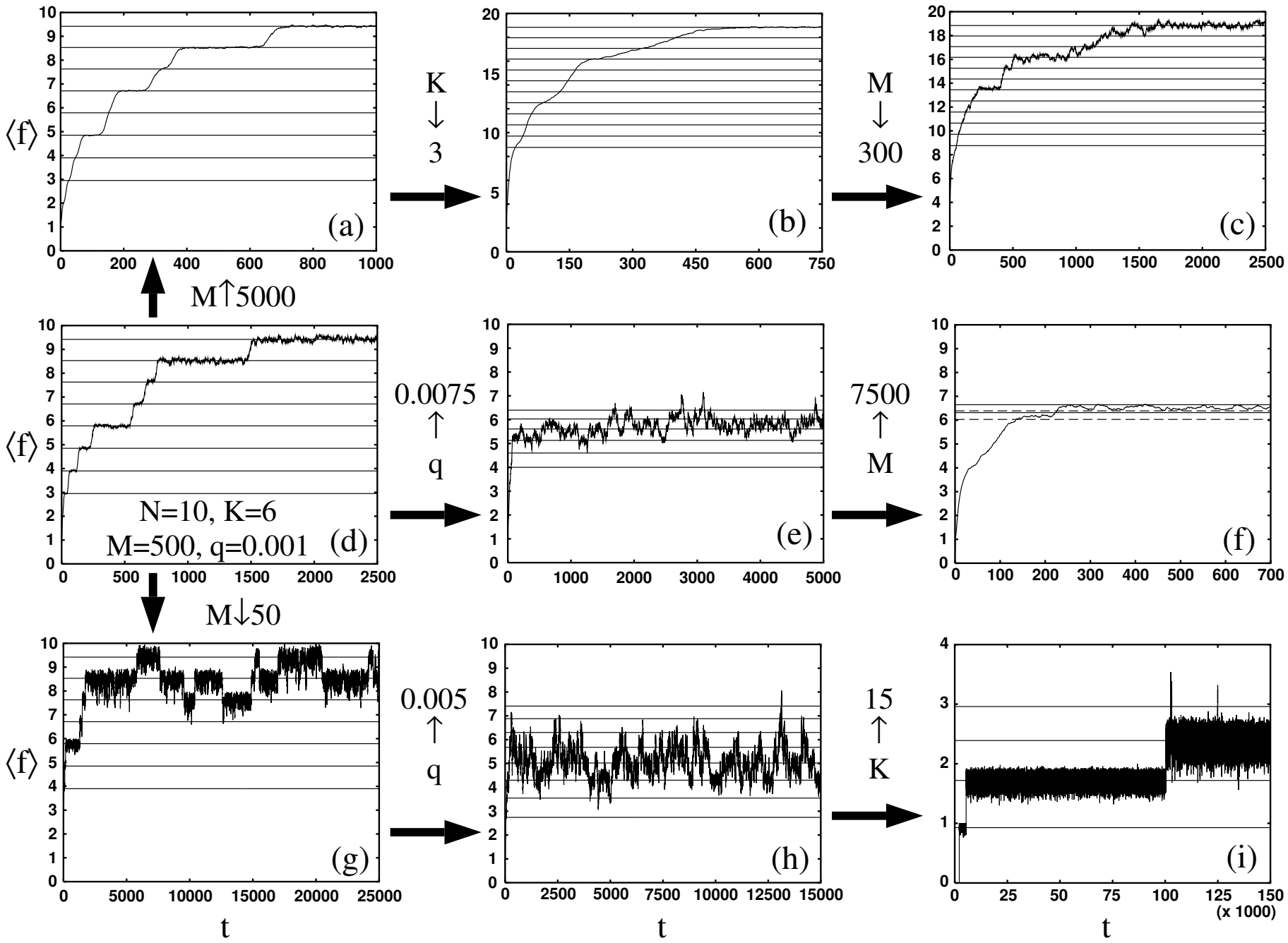

Figure 7: For caption see text.

In runs 7(a), 7(d), and 7(g) (first column of figure 7), the theory correctly predicts the fitness levels at which epochs occur. The variation of epoch durations across fitness levels and across the runs, sizes of fitness fluctuations in the epochs, and the intermittent behavior of run 7(g) need to be explained.

For the runs in the middle column (figures 7(b), 7(e), and 7(h)) we have plotted the theoretical fitness level predictions in a band of fitness values around which most of the behavior takes place. The theoretical predictions of the fitness levels correspond in only a limited way to the GA's behavior for these parameter settings. In run 7(b) the predicted highest level matches the empirical asymptotic fitness level. This is, of course, in accord with the analysis. For large populations the behavior approaches the infinite population dynamics we described in section 5; epochs do not occur. But we must explain why the epochs appear in run 7(a) but disappear in run 7(b) as a result of decreasing the block size K to 3.

For run 7(e) we have plotted only the predicted fitness levels g_5 through g_{10}. The low value of g_{10} explains why the average fitness stays low throughout the run. We must explain why the average fitness does not stabilize onto the highest fitness level g_{10}, but instead fluctuates in a band between g_6 and g_{10}.

In run 7(h) we have plotted the theoretical fitness levels of epochs 3 to 10. Again, the low value of g_{10} explains the apparent ceiling on the average fitness. We must explain why the average fitness does not stabilize onto distribution V^{10} at fitness g_{10} and why the fluctuation amplitudes are larger than those in 7(e).

The theoretical epoch fitnesses plotted in run 7(c) correctly predict the levels of the epochs that can be distinguished in this case as well as the asymptotic fitness level.

In run 7(f), however, it turns out that we see for the first time the consequences of the breakdown of our approximation of the probability A to create a block. The dashed lines show the predicted last two epoch levels using the analysis of the foregoing sections. They clearly underestimate the levels at which the epochs occur. Recall that in calculating the probability A to align a block we assumed that all unaligned blocks have never been aligned before. This approximation breaks down for the parameters used in this case. With this run's high mutation rate, $q = 0.0075$, many blocks in $\mathcal{P}_h$ are destroyed through mutation. Therefore, the blocks in the fitness class just below $\mathcal{P}_h$ have a relatively large proportion of their bits set. This means that the probability to align a block is much higher for these strings than we assume it to be. Our prediction should thus be seen as a lower bound. In section 5.1 we obtained an upper bound of $A = q(1-q)^{K-1}$ by assuming all blocks have $K-1$ of the K bits set. We plotted the upper bounds for the 9^{th} and 10^{th} epochs as solid lines in run 7(f). These upper and lower bounds still give quite accurate predictions for the observed fitness levels.

Finally, in run 7(i) the theoretical values correctly predict the epoch levels. The large fitness fluctuations within the epochs of this run remain to be explained.

To understand some of the remaining phenomena, we next analyze in more detail the structure of the epochal dynamics in the simplex.

6.3 Epochal Dynamics as the Evolutionary Unfolding of the State Space

We saw earlier that of all $\tilde{\mathbf{G}}$'s eigenvectors $\vec{V}^i$, only the principal eigenvector $\vec{V}^N$ is positive definite and thus interpretable as a fitness distribution. This means that the nonprincipal eigenvectors are not points in the simplex Λ_∞. The fitness epochs could only occur at fitness

distributions that are *close* but not at the eigenvectors $\vec{V}^i$. Let the term "epoch center" refer to the average position in Λ_M of the fitness distribution during an epoch. We now present a way to exactly obtain the epoch centers for each epoch.

Intuitively, the N epochs come about because there are N different possible blocks to align. In the first epoch, some individuals have one aligned block (fitness 1) and some have no aligned blocks (fitness 0) but none has more than one block aligned. In the second epoch, some individuals have two aligned blocks, some have one, and some have none. Whenever a new block is aligned and spreads through the population, the current fitness epoch becomes unstable and the population moves up to the next epoch. In terms of the dynamics in the simplex, the population starts on a low-dimensional subset of the simplex and finds its way to higher dimensions step by step. Initially (for small populations), there are typically only individuals with fitness 0 or 1. The fitness distribution then stabilizes somewhere on the line $P_0 + P_1 = 1, P_i \geq 0$. After a new, second aligned block is discovered the population moves off that line and onto the plane $P_0 + P_1 + P_2 = 1, P_i \geq 0$ and stabilizes there. When the third block is found the population moves into the three-dimensional space $P_0 + P_1 + P_2 + P_3 = 1, P_i \geq 0$. The evolution proceeds in this incremental fashion until all blocks are discovered or, as we will discuss later on, until the population cannot stabilize within successively higher-dimensional subspaces.

This rough picture illustrates how the subsimplices unfold dimension by dimension through the evolutionary search. When the population is in the n^{th} epoch it will move into the $(n+1)$-dimensional subsimplex with

$$\sum_{i=0}^{n} P_i = 1, P_i \geq 0. \tag{6.3}$$

This points to a way of calculating the actual epoch centers. While a population is in the n^{th} epoch, the selection and mutation operators, by definition, act only on strings s with $f(s) \leq n$ and produce new strings s' with $f(s') \leq n$. We can find the dynamics of the GA on this subspace by projecting $\mathbf{G}$ onto this subspace to form a restricted operator $\mathbf{G}^n$:

$$\mathbf{G}^n_{ij} = \begin{cases} \mathbf{G}_{ij} & \text{if } i, j \leq n \\ 0 & \text{otherwise.} \end{cases} \tag{6.4}$$

By acting on the fitness distributions with this restricted operator we can find the flow of the population in the n-dimensional subspace of the full simplex. The center $\vec{P}^n$ of the n^{th} fitness epoch is given by the principal eigenvector $\vec{P}^n$ of the restricted linearized operator $\tilde{\mathbf{G}}^n$. That is,

$$\tilde{\mathbf{G}}^n \cdot \vec{P}^n = e_n \vec{P}^n, \tag{6.5}$$

where e_n is $\tilde{\mathbf{G}}^n$'s principal eigenvalue. In short, by restricting $\mathbf{G}$ to the subsimplex of dimension n we can obtain the exact center of the n^{th} fitness epoch.

6.4 Eigensystem of the Restricted Generation Operators

We will now derive expressions for the different epoch centers $\vec{P}^n$ and the associated average fitness f_n in each epoch for small mutation rates q. For notational simplicity we will refer to the restricted operators $\tilde{\mathbf{G}}^n$ as $\tilde{\mathbf{G}}$ when calculating e_n and $\vec{P}^n$ in the following.

We first expand $\tilde{\mathbf{G}}$ to first order in q. Starting from the definition of $\mathbf{G}$ in terms of $\mathbf{M}$ and $\mathbf{S}$, we obtain

$$\tilde{\mathbf{G}}_{ij} = j\left[\delta_{ij}\left(1 - q(A_1(N-j) + Kj)\right) + \delta_{(i-1)j}A_1(N-j)q + \delta_{(i+1)j}Kjq\right] + \mathcal{O}(q^2). \tag{6.6}$$

where $A_1 q$ is the probability A to align a block to first-order in q (from equation 5.3):

$$A = A_1 q + \mathcal{O}(q^2) = \frac{K}{2^K - 1}q + \mathcal{O}(q^2) \tag{6.7}$$

Formally, we can split equation 6.6 into a term that is zeroth order in q and a term that is first order in q:

$$\tilde{\mathbf{G}} = \tilde{\mathbf{H}}^{\mathbf{0}} + q\tilde{\mathbf{H}}^{\mathbf{1}} + \mathcal{O}(q^2), \tag{6.8}$$

where $\tilde{\mathbf{H}}^{\mathbf{0}}_{ij} = j\delta_{ij}$. In the limit $q \to 0$, $e_n = n$ and $P_i^n = \delta_{in}$; that is, all strings have fitness n during the n^{th} epoch in this limit. For nonzero q we can expand $\vec{P}^n$ and e_n to first order in q:

$$P_i^n \equiv \delta_{in} + q\Delta_i^1 \tag{6.9}$$

and

$$e_n = n + qe_n^1. \tag{6.10}$$

Using the eigenvalue equation

$$\tilde{\mathbf{G}} \cdot \vec{P}^n = e_n \vec{P}^n, \tag{6.11}$$

and equating coefficients of equal powers in q, we obtain the n^{th} and $(n-1)^{st}$ components of $\vec{P}^n$:

$$P_n^n = (1 - n^2 Kq) \text{ and } P_{n-1}^n = n^2 Kq. \tag{6.12}$$

All other components are 0 to first order in q. For e_n we find

$$e_n = n - \left[n^2 K + n(N-n)A_1\right] q. \tag{6.13}$$

For the average fitness

$$f_n = \sum_{i=0}^{n} iP_i^n, \tag{6.14}$$

we obtain

$$f_n = n - n^2 Kq. \tag{6.15}$$

Notably, the epoch fitness levels can also be approximated in a more straightforward way that gives an insightful result, if we assume that the number of fitness n strings generated by block-aligning mutations of lower fitness strings is negligible. This assumes that the proportion P_n^n of strings in the highest fitness class is kept constant by a balance between block destroying mutations and selection. We then simply find that under selection and mutation:

$$P_n^n \to \frac{n}{f_n}P_n^n \to P_n^n \frac{n}{f_n}(1-q)^{nK}, \tag{6.16}$$

where the last factor $(1-q)^{nK}$ arises because only the strings that do not mutate the nK bits in their aligned blocks remain in fitness class n. From the above equation it follows that:

$$f_n = n(1-q)^{nK}. \tag{6.17}$$

This equation nicely shows that the average fitness in epoch n is proportional to the probability of all (nK) defining bits replicating without mutations. Note that equation 6.15 follows immediately to first order in q.

While the population resides in a fitness epoch there is a (metastable) equilibrium distribution of the population over the different fitness classes. We see that for very small q almost all individuals in the population are in the current highest fitness class, which is n for the n^{th} epoch. The higher the mutation rate, the more individuals will have lower fitness and accordingly the average fitness of the population decreases as q increases. Also, for higher epochs (larger n) the proportion of individuals in the current highest fitness class decreases with the square of n. This is also the case for the asymptotic fitness distribution $\vec{P}^N = \vec{V}^N$. The main difference between the metastable fitness distributions and the asymptotic distribution $\vec{V}^N$ is that the epoch distributions can become unstable when a new block is found and spreads through the population. Until this happens there is a metastable equilibrium between the effects of mutation and selection on the strings.

We noted earlier that the asymptotic distribution $\vec{V}^N$ is the phenotypic analogue of a quasispecies of replicating molecules or genomes. In light of the results above it is natural to extend this notion of quasispecies to include *metastable* (phenotypic) quasispecies that occur at earlier fitness epochs. Since in general the state space of molecular or genetic sequences is vastly larger than any realistic population size, we also expect finite-population-induced fitness epochs to occur in molecular evolution. During the evolution of a population there are large proportions of the sequence space that the population has never visited which might contain higher fitness strings than the current fittest genotypes. As already pointed out by Eigen and his colleagues [12, 32], the population will therefore reside in a metastable quasispecies distribution (or "metaspecies") until a mutation discovers one of the higher fitness strings. In general, though, the mechanism commonly proposed for metastability, e.g., [32], is that of a metastable quasispecies that is *localized* on a fitness peak in sequence space and that remains stable until one of the mutants at the edge of the quasispecies cloud finds a higher fitness peak. In this view, metastability is caused by local optima in the fitness landscape. The population will remain stable until one of the mutants crosses the fitness valley to a higher peak. In other words, the population has to cross a "fitness barrier" towards a higher peak.

The fitness-barrier mechanism is in sharp contrast to the mechanism presented here. There are no local optima in Royal Road fitness landscape. Rather, there are large connected subspaces of strings with equal fitness. Let S_n denote the subspace of strings with fitness n. The metastable distributions $\vec{P}^n$ are *not* strictly localized in sequence space but diffuse around randomly in the subspace S_n—a "subbasin" of the transient evolutionary dynamics—until one of the individuals discovers a rare "portal" to the subspace S_{n+1} by aligning a new block.[8] Therefore, the metastability in this view is the result of an "entropic barrier" that has to be crossed by the population. As a consequence, the evolutionary dynamics naturally splits into a neutral and an adaptive regime. The population spends most of its time randomly diffusing in the subbasin, or "neutral network", S_n of strings with equal fitness until one mutant finds one of the rare portal connections to S_{n+1}, after which the population adapts by moving into this new subbasin. In [25] this entropic view of evolution was also advocated in the context of molecular evolution.

[8]The nested-simplex state-space architecture also describes the transient behavior of diffusive annihilating particles, from which we have borrowed the subbasin and portal terminology [6]. In the evolutionary setting the simplices unfold; in the particle case they collapse, due to annihilation.

6.5 Crossover

So far we have restricted our attention to a mutation-only GA while, in fact, most applications of GAs employ some form of crossover—an operator that combines portions of parental chromosomes to produce offspring. It might come as somewhat of a surprise, then, that the epoch levels for the GA including crossover are empirically found to be the same as the epoch levels for the GA without crossover. Figure 8 shows the results of a GA run with parameters $N = 10$, $K = 6$, $M = 100$, and $q = 0.001$ and single-point crossover probability 1.0, together with the theoretical predictions of the epoch levels f_8, f_9, and f_{10}. Pairs of parents are selected in proportion to fitness from the old population and are always crossed over at a single point to produce two strings for the new generation. From figure 8 it is clear that the same theoretical

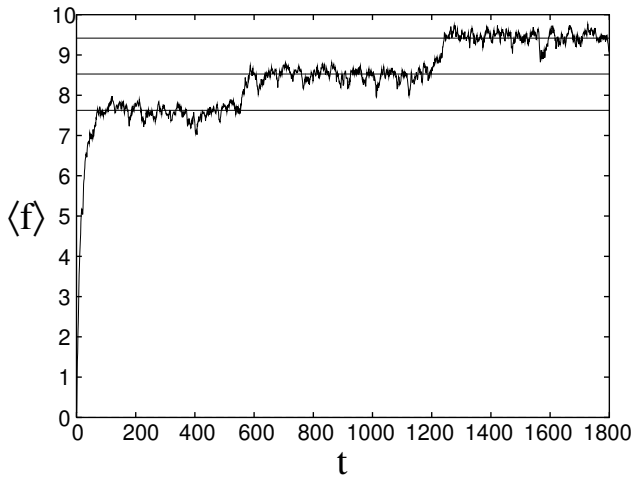

Figure 8: A run of the GA with single-point crossover, displaying the average fitness (solid curve) as a function of generation, together with the theoretically predicted epoch fitness levels (solid lines) for the 8^{th}, 9^{th}, and 10^{th} epochs. The parameters for this run are $N = 10$, $K = 6$, $q = 0.001$, and $M = 100$. Cf. figures 1(d) and 7(d).

predictions for the epoch levels of the GA without crossover also correctly predict the epoch levels for the GA with crossover. This can be understood in the following way.

When the population resides in the n^{th} epoch all individuals in the population have fitness n or less. An innovation occurs when one or more individuals with fitness higher than n are discovered. It is unlikely that more than one such individual is discovered in the same generation. Therefore, it is almost always the case that one individual is the founder of a new epoch. When the population has moved up to epoch $n + 1$ all the strings with fitness $n + 1$ in the population are descendants of that founder string and thus share its aligned blocks (as well as other bits). This means that the $n + 1$ aligned blocks in each string with fitness $n + 1$ will be in corresponding positions in each string. If the population is dominated during the epoch by strings of fitness $n + 1$, crossover on pairs of selected strings will generally not break up the aligned blocks. Blocks that are destroyed are mostly destroyed through mutations, *not* through crossover. Because of this, the balance between block destruction and selection is hardly affected and the epochs occur at the same fitness levels as in the case without crossover.

However, at the very beginning of a run, the population will not be converged and aligned blocks will be in different positions along the strings. Therefore, in contrast to the mutation-only GA, crossover *can* combine different blocks from different strings at the start of the run to create

higher fitness strings. Hence, the number of initial epochs expressed in runs with crossover is smaller than the number of epochs in the runs without crossover. Once crossover has put together the best string that can be produced from the aligned blocks in the initial population, the population will converge onto copies of that string. From that point on crossover will act as a "localized" mutation operator. It will introduce mixing only in the bits of the unaligned blocks, since all the bits in the aligned blocks are shared by the strings in the population.

This illustrates again that phrases such as "crossover combines building blocks" or "crossover introduces nonlocal mixing", are meaningless without precise specification of *all* other components of the dynamics. For crossover to be able to perform such actions as referred to above, it is necessary that *different* useful building blocks be present in the population at the same time and that generally the population is genetically diverse. We have argued that very early in the run, due to the randomness of the initial population, crossover can indeed combine building blocks. However, as soon as the epochal behavior sets in, all aligned blocks occur in corresponding positions in the strings and it is highly unlikely that different building blocks will occur in the population at the simultaneously ever again.

A similar argument holds for the nonlocal mixing that crossover is purported to introduce. For the random initial population this will certainly be the case, but as soon as the epochal dynamics sets in crossover can only introduce nonlocal mixing in the unaligned blocks. This nonlocal mixing is furthermore restricted by the diversity that *mutation* has introduced in bits of the unaligned blocks. In summary, for systems that show mainly epochal evolution, it seems unlikely that crossover's "combining building blocks" or "introducing nonlocal mixing" are important contributions to the search dynamics.

6.6 Stable and Unstable Manifolds

It is clear from figure 1 that not all of the N possible fitness epochs are visited for a given GA run; later epochs tend to be visited in more runs than earlier epochs. Later epochs also tend to have longer durations than earlier epochs in a given run. In addition, for higher mutation rates, epochs appear less distinct and innovations less steep. Since, for small mutation rate q, the metastable fitness distributions in which a finite population can get trapped are close to the fixed points of $\mathbf{G}$, we can obtain a qualitative understanding of these features by analyzing the local stability of these fixed points.

We will analyze the topological structure in Λ_∞ that determines the global stability of these metastable states by looking at the local stability of the fixed points themselves. The local stability around the fixed points is determined by the Jacobian matrix $\mathbf{DG}$ at each fixed point, since it gives the first-order approximation of the dynamics in the vicinity of the fixed points. That is, around a fixed point $\vec{V}$ we have:

$$\mathbf{G}(\vec{V} + \vec{\epsilon}) = \vec{V} + \mathbf{DG} \cdot \vec{\epsilon}, \tag{6.18}$$

where $\vec{\epsilon}$ is a small deviation vector. Consider the Jacobian matrix at the fixed point $\vec{V}^n$ of the n^{th} fitness epoch. Using the basic definitions, $\tilde{\mathbf{G}} = \mathbf{M} \cdot \tilde{\mathbf{S}}$, and the fact that $\vec{V}^n$ is a fixed point $\tilde{\mathbf{G}} \cdot \vec{V}^n = g_n \vec{V}^n$, we find that

$$\mathbf{DG}_{ij}(\vec{V}^n) = \left[\frac{\partial \mathbf{G}_i(\vec{P})}{\partial P_j} \right]_{\vec{P}=\vec{V}^n} = \frac{\tilde{\mathbf{G}}_{ij} - jV_i^n}{g_n}. \tag{6.19}$$

The local stable and unstable manifolds of this fixed point can be obtained by solving for the eigenvectors of this matrix. Eigenvectors with eigenvalues $\lambda_i > 1$ give the directions in which this fixed point is unstable and eigenvectors with eigenvalues $\lambda_i < 1$ give the directions in which the fixed point is stable. This follows from equation 6.18, since the deviations grow and shrink for eigenvalues greater or less than 1, respectively. By inspection of equation 6.19, it is easy to see that the eigenvectors $\vec{U}^i$ of the Jacobian matrix $\mathbf{DG}(\vec{V}^n)$ are given by the differences between the eigenvectors $\vec{V}^i$ and the eigenvector $\vec{V}^n$:

$$\vec{U}^i = \vec{V}^i - \vec{V}^n. \tag{6.20}$$

Substituting these vectors into the eigenvalue equation, we find that

$$\mathbf{DG}(\vec{V}^n) \cdot \vec{U}^i = \frac{g_i}{g_n}\, \vec{U}^i. \tag{6.21}$$

Thus, the eigenvalues λ_i^n of this fixed point are given by[9]

$$\lambda_i^n = \frac{g_i}{g_n}. \tag{6.22}$$

This means that, since the g_i are ordered in increasing magnitude, the fixed point $\vec{V}^i$ has i stable directions and $N - i$ unstable directions. This is intuitively clear from the fact that in the i^{th} epoch the subsimplex that is already *discovered* by the population is i-dimensional (corresponding to the i aligned blocks) and the undiscovered part of the simplex is $(N - i)$-dimensional (corresponding to the $N-i$ yet-to-be-aligned blocks). Furthermore, we see that the strength of the Jacobian eigenvalues λ_i^n is determined by the ratios of the different generation operator eigenvalues g_i with respect to the eigenvalue g_n of the epoch under consideration. As we have seen, to a high approximation the generation operator eigenvalues are equal to the average fitnesses in the different epochs. So we see that the eigenvalues of the Jacobian matrix are simply determined by the ratios of average fitnesses in the different epochs.

Thus, for small mutation rate q we can use equation 6.17 to obtain analytical expressions for the eigenvalues of $\mathbf{DG}$ in the n^{th} fitness epoch. These are

$$\lambda_i^n = \frac{i}{n}(1-q)^{(i-n)K}\ , 0 \le i \le N\ . \tag{6.23}$$

The relative sizes of these eigenvalues as a function of i, n, K, and q control the dynamical features of the different fitness epochs.

Now we can qualitatively explain the first of our observations, namely, that the higher fitness epochs are longer. The population will remain in an epoch until it finds, and spreads into, one of the undiscovered dimensions of the simplex. Since the later epochs have more stable dimensions and fewer unstable ones, the population is less likely to find one of the unstable dimensions and the epochs will be longer. This is related to the simple fact that it is easier for mutations to align one or more new blocks when there are more blocks still unaligned.

We have also observed that later epochs typically appear in more runs than earlier epochs. The first and obvious reason for this is that the initial population is likely to contain strings

[9] $\vec{U}^n$ is the null-vector which, of course, has eigenvalue 0. It does not, however, give an independent direction in the simplex.

with different fitness values and that the first epoch n that will be visited corresponds to the fitness n of the highest fitness strings in the initial population. Let $\Pr(f < n)$ denote the probability that *none* of the M individuals of the initial population $\vec{P}(0)$ has a fitness of n or higher:

$$\Pr(f < n) = \Pr(P_i = 0,\ i \geq n). \tag{6.24}$$

We then have

$$\Pr(f < n) = \left[\sum_{i=0}^{n-1} \binom{N}{i} 2^{-Ki} \left(1 - 2^{-K}\right)^{N-i}\right]^M . \tag{6.25}$$

It turns out that as n increases this probability jumps up sharply from almost 0 to almost 1 at some value of n (provided that K is not too small). If n_f is the first value for which $\Pr(f < n) \approx 1$ then it is likely that the first epoch to appear will be the $(n_f - 1)^{th}$ epoch, since strings with fitness $n_f - 1$ are likely to occur in the initial population and strings with fitness n_f are very unlikely to occur in the initial population. Furthermore, when one or more new blocks are discovered in epoch n, there are $N - n$ possible unstable dimensions into which the population can move out of the subsimplex. Each of these dimensions corresponds to a portal through which the next epoch can be visited. From this it follows that higher epochs with a larger number of attracting dimensions are more likely to be visited.

Next we consider the steepness of the innovations. The steepness of an innovation out of epoch n is related to the size of the eigenvalues λ_i^n for the unstable dimensions $i > n$. The larger these eigenvalues, the more quickly the population will move away from the metastable region once an unstable dimension is explored. From equation 6.23 we see that for larger q the eigenvalues λ_i^n, given by the fitness ratios f_i/f_n, approach 1. This means that the innovations become less steep for larger values of q, which is indeed what we have observed.

From the fact that the eigenvalues λ_i^n approach 1 for larger q, it also follows that the fitness fluctuations in an epoch increase with q. This is because the smaller the stable-direction eigenvalues $\{\lambda_i^n : i < n\}$, the more strongly the population is restored to the epoch center after a fluctuation. So if these eigenvalues increase toward 1, the fitness fluctuations increase as well. From the fact that the stable-direction eigenvalues $\{\lambda_i^n : i < n\}$ are larger and so closer to 1 for higher n, we also see that later epochs have larger fitness fluctuations than early ones (e.g., see figure 1(d)). We will discuss the size of the fitness fluctuations in more detail in section 6.8 below.

Finally, we see that the unstable-direction eigenvalues $\{\lambda_i^n : i > n\}$ for large n are smaller than those for small n. From this, it follows that later innovations are, in general, less steep than early ones (e.g., see figures 1(a) and 1(d)). Following up this line of qualitative analysis, in the next section we estimate the innovation times quantitatively from the values of λ_i^n.

Recapitulating, we have argued that the main qualitative features of epochal evolution can be understood in terms of the eigenvalues λ_i^n of the Jacobian matrix in the vicinity of the n^{th} epoch. In the following sections we will use these eigenvalues to make more quantitative predictions about the epochal behavior.

6.7 Innovation Durations

We now estimate the expected time for an innovation from the n^{th} to the $(n+1)^{st}$ epoch to take place. During the n^{th} epoch the proportion of individuals with fitness $n+1$ is 0. When some individual finds a new aligned block it will either start spreading through the population

or it will be lost through a sampling fluctuation. (See figure 1 for examples of this loss.) We are interested in the time it takes a string with the new aligned block to fully spread through the population *once* the string appears and continues to spread through the population. (For convenience in this section we take the time of this first appearance to be $t = 0$.) Thus, we will assume that the effects of the finite population sampling noise can be neglected during the innovation.

Thus, we refer to the time at which the epoch $n + 1$ begins as $t = 0$. At this time, the number of individuals with fitness $n + 1$ is 1; that is, $P_{n+1}(0) = 1/M$. From that time on, P_{n+1} increases until it reaches the equilibrium value P_{n+1}^{n+1} of the $(n + 1)^{st}$ epoch. In going from the n^{th} to the $(n + 1)^{st}$ epoch, the population moves in the direction of the Jacobian eigenvector $\vec{U}^{n+1} = \vec{P}^{n+1} - \vec{P}^{n}$. Thus, during the innovation from the n^{th} to the $(n + 1)^{st}$ epoch, we can write the fitness distribution $\vec{P}(t)$ of the population as a linear combination of the n^{th} epoch eigenvector $\vec{P}^{n}$ and the $(n + 1)^{st}$ epoch eigenvector $\vec{P}^{n+1}$:

$$\vec{P}(t) = (1 - \alpha(t))\vec{P}^{n} + \alpha(t)\vec{P}^{n+1}, \tag{6.26}$$

with the initial condition $\alpha(0) = 1/M$. One should interpret α as giving the decomposition of the fitness distribution during the innovation in terms of the lower and higher epoch distributions. During the innovation the balance shifts from the lower epoch $\alpha \approx 0$ to the higher epoch $\alpha \approx 1$.

Applying the generation operator to the fitness distribution in equation 6.26 and using the fact that the epochs correspond to eigenvectors of the generation operator, we find for the evolution of $\alpha(t)$ that

$$\alpha(t + 1) = \frac{f_{n+1}}{f_{n+1}\alpha(t) + (1 - \alpha(t))f_n}\alpha(t). \tag{6.27}$$

Assuming that $\alpha(t)$ changes slowly and smoothly as a function of time, we can deduce a differential equation from the above equation:

$$\frac{d\alpha}{dt} \approx \alpha(t + 1) - \alpha(t) = \gamma_n \alpha \frac{1 - \alpha}{\gamma_n \alpha + 1}, \tag{6.28}$$

where

$$\gamma_n = \frac{f_{n+1} - f_n}{f_n} \tag{6.29}$$

is the relative increase in fitness from the n^{th} to the $(n + 1)^{st}$ epoch. It is possible to solve equation 6.28 analytically to obtain t as a function of α:

$$t = \frac{1}{\gamma_n}\left[\log(M\alpha) - (1 + \gamma_n)\log\frac{M(1 - \alpha)}{M - 1}\right], \tag{6.30}$$

where we have used the boundary condition $\alpha(0) = 1/M$. In general it is impossible to invert equation 6.30 analytically to obtain α as a function of t. Nonetheless, we can use equation 6.30 to obtain an estimate of the duration t_n of the innovation from the n^{th} to the $(n + 1)^{st}$ epoch. We consider the innovation to have ended once α has reached the level $1 - 1/M$. Since we treat α as a continuous variable, it will approach 1 only asymptotically. For the finite population case we truncate this continuous dynamics at $\alpha = 1 - 1/M$ and solve for t_n, finding that

$$t_n = \frac{2 + \gamma_n}{\gamma_n}\log[M - 1]. \tag{6.31}$$

This equation has a simple interpretation. γ_n gives the relative fitness increase from the n^{th} to the $(n+1)^{st}$ epoch. The innovation duration is roughly inversely proportional to this relative fitness increase. The innovation duration, in addition, is proportional to the logarithm of the population size. The logarithm of the population size apparently controls the "inertia" of the distribution with respect to innovations. Certainly, a single more-fit string will take longer to dominate a larger population. And since replication is exponential in time, one obtains a logarithmic dependence on M.

Expanding γ_n to first order in the mutation rate q and setting $\log[M-1] \approx \log[M]$ we find that

$$t_n = [1 + 2n + 2Kn(n+1)q] \log M + \mathcal{O}(q^2). \tag{6.32}$$

We see that the innovation duration is proportional to the epoch level and the logarithm of the population size. Furthermore, it is clear that increasing the mutation rate makes the innovations less steep.

We can also approximate $\alpha(t)$ from equation 6.28 by neglecting the small term $\gamma_n \alpha$ in the denominator.[10] Equation 6.28 then turns into the well known logistic growth equation, which can be solved analytically:

$$\alpha(t) = \frac{\exp(\gamma_n t)}{\exp(\gamma_n t) + M - 1}. \tag{6.33}$$

Again, the speed of the innovation is controlled by the relative fitness increase γ_n from the n^{th} to the $(n+1)^{st}$ epoch. The average fitness $\langle f(t) \rangle$ during the innovation is proportional to $\alpha(t)$:

$$\langle f(t) \rangle = f_n + (f_{n+1} - f_n)\alpha(t). \tag{6.34}$$

Figures 9(a) and 9(b) plot the theoretical predictions for the average fitness (thick lines) during the innovations between the third and fourth and eighth and ninth epochs, respectively, for the parameter settings $N = 10$, $K = 6$, $q = 0.001$, and $M = 500$. The thin lines in figures 9(a) and 9(b) give some examples of the empirically observed innovations from runs of the GA. Figure 9(a) shows the innovation from just one run with these parameter settings. It is clear that the logistic growth approximation of the innovation gives an accurate prediction of the shape and length of this innovation. Figure 9(b) shows the innovation between the eight and ninth epoch from three different runs. At higher epochs there are large fitness fluctuations[11] and the exact shape of the innovations differs from run to run. Still the theory accurately predicts the *average* shape of the innovation as can be seen from figure 9(b). The predicted innovation durations in the above cases are $t_3 \approx 32$ and $t_8 \approx 98$, respectively.

6.8 Fitness Fluctuations

The finite population dynamics in and around the epochs is, of course, essentially stochastic and cannot be modeled by deterministic dynamical equations. From here on, we therefore have to turn to stochastic equations describing the evolution of a probability distribution $\Pr[\vec{\epsilon}, t]$ which gives the probability that the population is a distance $\vec{\epsilon}$ away from some epoch center $\vec{P}^n$ at time t. We will model the evolution of $\Pr[\vec{\epsilon}, t]$ by means of a diffusion (or Fokker-Planck) equation, where the *average* change per generation of $\vec{\epsilon}$ is determined by the Jacobian

[10] Since $\gamma_n \approx 1/n$ and $\alpha < 1$ this approximation is justified, especially for later epochs and for the small α at the start of the innovation.

[11] This will be discussed in the next section.

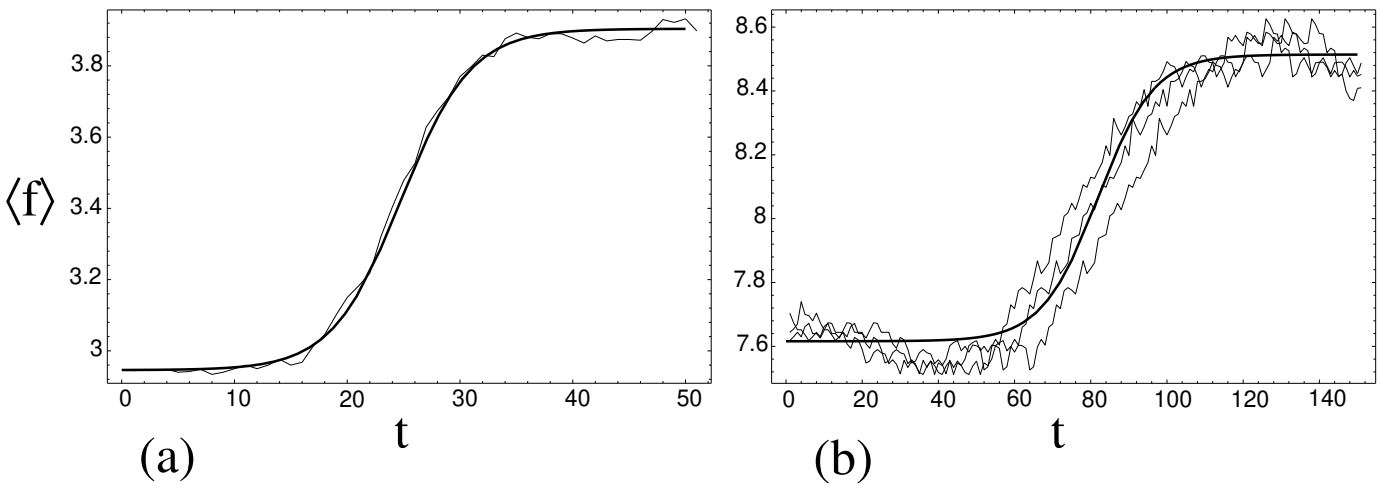

Figure 9: Innovation between the third and fourth (a) and eight and ninth (b) epochs for GA runs with parameters $N = 10$, $K = 6$, $q = 0.001$, and $M = 500$. (See figure 1(d).) The thick lines give the theoretical predictions of the innovation curves. The thinner lines show examples of these innovations from runs of the GA.

eigenvalues λ_i^n of epoch n and the fluctuations due to the finite population sampling occur as a Gaussian-noise diffusion. The idea of using continuous-variable diffusion models, like the Fokker-Planck equation, to solve for the stochastic dynamics of gene frequencies in a population was developed by the population geneticist Kimura, see [27]. This type of stochastic model assumes that gene frequencies, such as P_i or ϵ_i, can be approximated by real variables and that the gene frequencies change slowly over a single time step. (See [22] for a recent overview of the validity of the diffusion models in this context.)

Using the average fitness levels in each epoch, we can now estimate the size of the fluctuations about those levels. During a fitness epoch the population jumps around on a set of different lattice points in Λ_M surrounding the epoch's center $\vec{P}^n$. From the Jacobian matrix $\mathbf{DG}(\vec{P}^n)$ we can derive the form of the flow in the vicinity of the fixed point. The difference vectors $\vec{U}^i = \vec{P}^i - \vec{P}^n$ between the different epoch locations and the location of the epoch of interest form a natural basis for modeling the fluctuations around the epoch center. We will therefore expand the fluctuation dynamics in terms of the Jacobian eigenbasis.

First, assume that the population fitness distribution $\vec{P}$ resides near the epoch center $\vec{P}^n$:

$$\vec{P} = \vec{P}^n + \sum_{i=0}^{n-1} \epsilon_i \vec{U}^i, \tag{6.35}$$

where $\vec{\epsilon} = (\epsilon_0, \ldots, \epsilon_{n-1})$ is a small deviation vector. In one time step, the vector $\vec{P}$ is expected to go to $\langle \vec{P}' \rangle$, which is given by

$$\langle \vec{P}' \rangle = \mathbf{G}(\vec{P}) = \vec{P}^n + \sum_{i=0}^{n-1} \frac{f_i}{f_n} \epsilon_i \vec{U}^i. \tag{6.36}$$

That is, the deviation $\vec{\epsilon}$ is expected to scale down by a factor of f_i/f_n in each direction $\vec{U}^i$. From the above equation we can now calculate the expected change $\langle d\epsilon_i \rangle$ of the deviation in

direction $\vec{U}^i$:

$$\langle d\epsilon_i \rangle = \left(\frac{f_i}{f_n} - 1 \right) \epsilon_i. \tag{6.37}$$

To simplify notation, we define the relative fitness decrease of epoch i with respect to epoch n:

$$\mu_i = \frac{f_n - f_i}{f_n}. \tag{6.38}$$

Then equation 6.37 takes on a simple form:

$$\langle d\epsilon_i \rangle = -\mu_i \epsilon_i. \tag{6.39}$$

Thus, fluctuations in direction $\vec{U}^i$ die off exponentially *on average* with rate μ_i. Note that this rate is smallest in the direction of the higher-fitness epochs.

Equation 6.39 gives the *expected* change of ϵ_i around the epoch center $\vec{P}^n$. Of course, there will be fluctuations in this change due to finite-size sampling. In equation 6.1 we saw that the probability for going from state $\vec{P}$ to state $\vec{P}'$ is a multinomial with mean $\mathbf{G}(\vec{P})$. From this we can derive expressions for the expected second moments of the change in the fitness distribution:

$$\langle dP_i dP_j \rangle = \frac{P_i(\delta_{ij} - P_j)}{M}. \tag{6.40}$$

From this we can derive the second moments of the change in the fluctuation vector $\langle d\epsilon_i d\epsilon_j \rangle$. To this end we transform the components P_i to the basis of the epoch centers $\vec{P}^i$. We first define the similarity transformation matrix $\mathbf{R}_{ij} = P_i^j$. Using its inverse, we can calculate the change of fluctuations in direction $\vec{U}^i$:

$$\epsilon_i = \sum_{i=0}^{n-1} \mathbf{R}_{ij}^{-1} P_j. \tag{6.41}$$

Note that the sum goes from 0 to $n-1$ and not up to n, since the component ϵ_n is not independent of the others. (It is determined from the requirement that the fitness distribution vector be normalized.) Using equations 6.40 and 6.41 the second moments of the change in the fluctuations are given by

$$\langle d\epsilon_i d\epsilon_j \rangle = \sum_{k,m=0}^{n-1} \mathbf{R}_{ik}^{-1} \mathbf{R}_{jm}^{-1} \frac{P_k(\delta_{km} - P_m)}{M}. \tag{6.42}$$

Assuming that the fluctuations ϵ_i are small compared to the epoch center components $\vec{P}_i^n$, we can approximate $\vec{P}_i$ by $\vec{P}_i^n$ in the above formula and find

$$\langle d\epsilon_i d\epsilon_j \rangle = \sum_{k,m=0}^{n-1} \mathbf{R}_{ik}^{-1} \mathbf{R}_{jm}^{-1} \frac{P_k^n(\delta_{km} - P_m^n)}{M} \equiv \frac{\mathbf{B}_{ij}}{M}, \tag{6.43}$$

where the components $\mathbf{B}_{ij}$ depend only on the location of the current epoch center $\vec{P}^n$.

So that we can use the Jacobian matrix approximation to the dynamics, we assume that the population size is large enough to keep the fluctuations localized in the area around the epoch

center. Then we can use equations 6.39 and 6.43 to approximate the stable limit distribution $\Pr[\vec{\epsilon}\,]$ of the fluctuations that occur while the population resides in epoch n. The distribution $\Pr[\vec{\epsilon}, t]$ gives the probability of finding the population at a deviation $\vec{\epsilon}$ from the epoch center at any particular time t during the epoch. This distribution can be obtained by solving the multivariate Fokker-Planck equation associated with the drift term of equation 6.39 and the diffusion term given by equation 6.43. This is

$$\frac{\partial \Pr[\vec{\epsilon}, t]}{\partial t} = -\sum_i \frac{\partial \langle d\epsilon_i \rangle \Pr[\vec{\epsilon}, t]}{\partial \epsilon_i} + \frac{1}{2} \sum_{i,j} \frac{\partial^2 \langle d\epsilon_i d\epsilon_j \rangle \Pr[\vec{\epsilon}, t]}{\partial \epsilon_i \partial \epsilon_j}. \tag{6.44}$$

As $t \to \infty$, the asymptotic (stationary) solution of the above equation is a multi-dimensional Gaussian peak around $\vec{\epsilon} = 0$ for the case of constant $\langle d\epsilon_i d\epsilon_j \rangle$ [43]. This is given by

$$\Pr[\vec{\epsilon}\,] = \frac{1}{\sqrt{(2\pi)^n Det[\mathbf{C}]}} \exp\left[-\frac{1}{2} \sum_{i,j=0}^{n-1} \epsilon_i \mathbf{C}_{ij}^{-1} \epsilon_j \right], \tag{6.45}$$

where the matrix $\mathbf{C}$ determines the second moments of the distribution. It is given by

$$\mathbf{C}_{ij} = \langle \epsilon_i \epsilon_j \rangle = \frac{\mathbf{B}_{ij}}{M(\mu_i + \mu_j)}. \tag{6.46}$$

Of course, the means are all 0:

$$\langle \epsilon_i \rangle = 0. \tag{6.47}$$

Using equation 6.45 we can solve for the expected fitness fluctuations during an epoch. With some algebra we find that the fitness variance $\mathrm{Var}[f]$ in epoch n is given to first order in q by

$$\mathrm{Var}[f] = \frac{K n^3 q}{2M}. \tag{6.48}$$

This shows rather transparently that the fluctuations scale inversely with the population size M and proportionally to the cube of the epoch number n.

Figures 10, 11, and 12 show the above analytical prediction of the fitness fluctuations during various epochs. The grey bands show the average fitness in the epochs plus and minus two standard deviations, given by $\sigma_n = \sqrt{\mathrm{Var}[f_n]}$. Figure 10 shows the fitness fluctuations for our canonical parameter settings $N = 10$, $K = 6$, $q = 0.001$, and $M = 500$. Figure 11 shows the fitness fluctuations for the same parameter settings, only with an increased mutation rate of $q = 0.002$. By comparing these two figures, it is clear that increased mutation rate increases the size of the fitness fluctuations during the epochs. As predicted by the analysis, the fitness fluctuations increase by a factor of roughly $\sqrt{2}$. Finally, figure 12 has the same parameter settings as figure 10, except for a lower population size of $M = 50$. As predicted, the fitness fluctuations increase by a factor of approximately $\sqrt{10}$.

6.9 Destabilizing Fluctuations and the Error Threshold

We saw that for high-fitness epochs the proportion of individuals in the highest fitness class decreases with the square of the epoch number n:

$$P_n^n = (1 - n^2 K q) + \mathcal{O}(q^2). \tag{6.49}$$

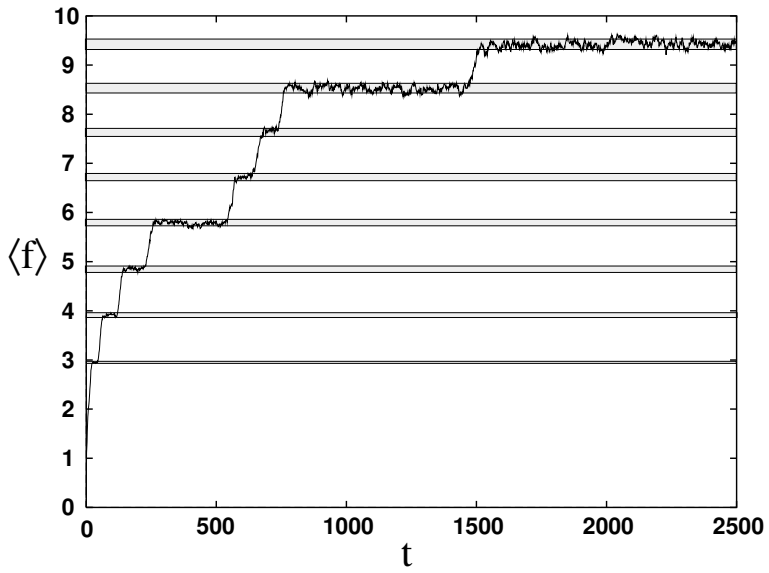

Figure 10: Predicted size of the epoch fitness fluctuations to two standard deviations, given by the grey bands. The parameters for this run are $N = 10$, $K = 6$, $q = 0.001$, and $M = 500$. This is the same run as was plotted in figure 1(d).

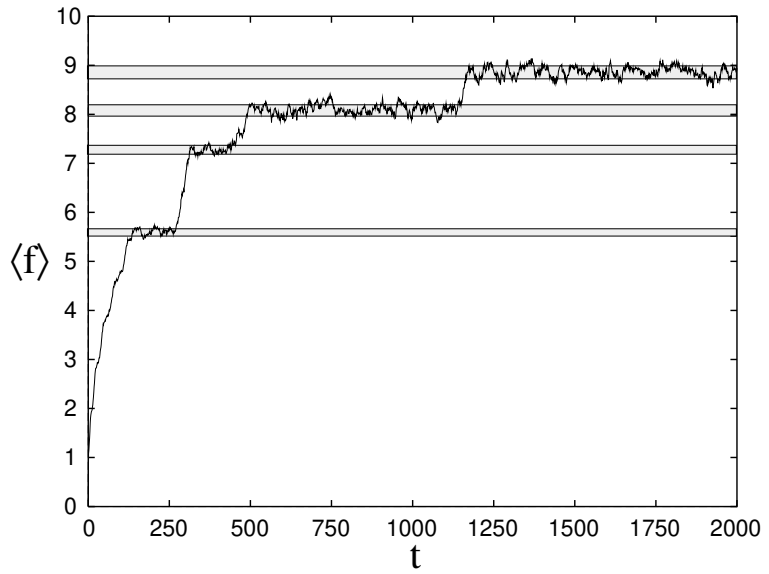

Figure 11: Predicted size of the fitness fluctuations during the epochs, plotted up to two standard deviations using the grey bands. The parameters for this run are $N = 10$, $K = 6$, $q = 0.002$, and $M = 500$. Cf. figure 10.

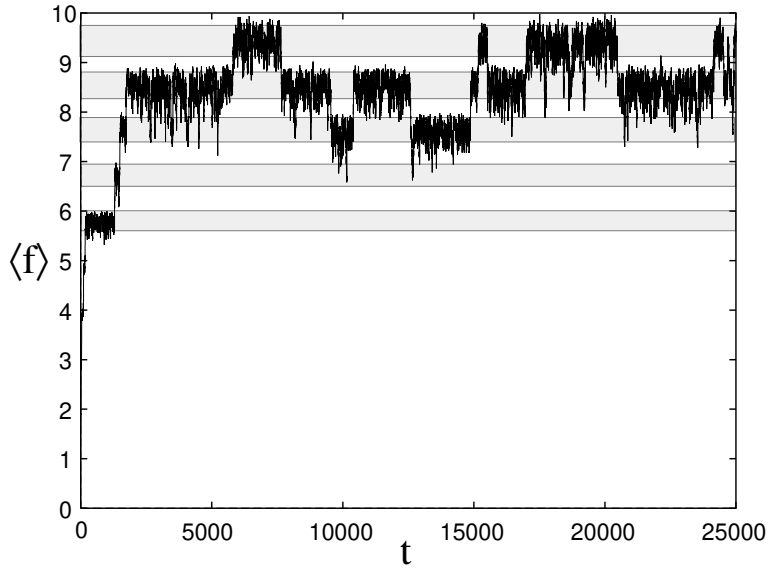

Figure 12: Predicted size of the fitness fluctuations during the epochs, plotted up to two standard deviations using the grey bands. The parameters for this run are $N = 10$, $K = 6$, $q = 0.001$, and $M = 50$. This is the same run as was plotted in figure 1(g).)

For high-fitness epochs, in the large N case when there are many possible epochs, the proportion P_n^n of strings in the highest fitness class can eventually become so small that there is an appreciable chance that all individuals in the highest fitness class will be lost through a sampling fluctuation. When this happens, the fitness distribution will fall back to the distribution $\vec{P}^{n-1}$ of epoch $n-1$ just below. If, after some time, the n^{th} block is rediscovered and spreads through the population, the distribution will move up again to the n^{th} epoch. This is exactly the process that causes the intermittent epochal behavior seen in figure 1(g).

We can obtain a rough magnitude estimate for the epoch at which one begins to observe this intermittency. Assume that the population resides at the epoch center. In generating the population for the next time step, each string has a probability P_n^n to be of fitness class n. Therefore, using equation 6.12, the probability $\Pr(f \neq n)$ that none of the strings will be in class n is given by

$$\Pr(f \neq n) = (1 - P_n^n)^M \approx (n^2 Kq)^M. \tag{6.50}$$

Demanding that this probability is of the order of (say) 1%, we can calculate for which epoch level n this condition is satisfied as a function of K, q, and M. For the case of $N = 10$, $K = 6$, $q = 0.001$, and $M = 50$, we find $n \approx 12$. From figure 12, we see that by epochs 9 and 10 the intermittency has set in. The crude estimate given above apparently underestimates the fluctuations in the proportion of individuals in the highest fitness class P_n. Using an analysis similar to that in the last section, we will now calculate more precisely the average time —the "destabilization" time— that the population spends in epoch n until all individuals with fitness n are lost through a sampling fluctuation.

To calculate the average destabilization time, we consider the dynamics of the fluctuations ϵ_n in the component P_n of the fitness distribution near the epoch center. That is, let

$$P_n = P_n^n + \epsilon_n. \tag{6.51}$$

The deviation ϵ_n will fluctuate up and down while the population resides in the epoch. It can

become at most $\epsilon_n = 1 - P_n^n$, corresponding to all strings fitness n, and when it becomes $-P_n^n$ all individuals with fitness n are lost and the epoch becomes unstable. We therefore want to calculate the average time it takes until $\epsilon_n = -P_n^n$ for the first time.

The deviation ϵ_n just compensates the sum of the deviations ϵ_i in the eigenvector directions $\vec{U}^i$ of the Jacobian[12]. That is,

$$\epsilon_n = -P_n^n \sum_{i=0}^{n-1} \epsilon_i. \tag{6.52}$$

After one time step the component P_n will on average move to

$$\langle P_n' \rangle = P_n^n - P_n^n \sum_{i=0}^{n-1} \frac{f_i}{f_n} \epsilon_i \equiv P_n^n + \lambda \epsilon_n, \tag{6.53}$$

where we have defined λ as the factor by which the fluctuation ϵ_n is scaled down. Of course, in general λ depends on the particular distribution of the fluctuations ϵ_i over the different directions. In the previous section we saw that the fluctuations ϵ_i during an epoch are all approximately normally distributed around the epoch center. As an approximation, we will assume that the fluctuation in direction i is proportional to the variance $\langle \epsilon_i^2 \rangle$. We then obtain

$$\lambda = \frac{\sum_{i=0}^{n-1} f_i \langle \epsilon_i^2 \rangle}{f_n \sum_{j=0}^{n-1} \langle \epsilon_j^2 \rangle}. \tag{6.54}$$

With the above definition of the average scale factor, we can obtain the expected change in the fluctuation:

$$\langle d\epsilon_n \rangle = -(1 - \lambda)\epsilon_n \equiv -\mu \epsilon_n, \tag{6.55}$$

where the second equality defines the coefficient μ. We can again approximate the expected change in the square of the deviation by the size of the sampling fluctuations at the epoch center:

$$\langle (d\epsilon_n)^2 \rangle = \langle (dP_n)^2 \rangle \approx \frac{P_n^n (1 - P_n^n)}{M}. \tag{6.56}$$

Since the above (diffusion) term is again a constant, the problem reduces to that of the first passage time of a homogeneously diffusing particle in a potential field (see [21]). The solution for the average time $T(0)$ for the fluctuation to reach $\epsilon_n = -P_n^n$ for the first time, given that the process starts with a fluctuation $\epsilon = 0$, is given by

$$T(0) = \frac{M P_n^n}{1 - P_n^n} + \frac{\pi}{2\mu} \mathrm{erfi}\left[\sqrt{\frac{M \mu P_n^n}{1 - P_n^n}} \right] \mathrm{erf}\left[\sqrt{\frac{M \mu (1 - P_n^n)}{P_n^n}} \right], \tag{6.57}$$

where $\mathrm{erf}(x)$ is the error function and $\mathrm{erfi}(x) = \mathrm{erf}(ix)/i$ is the imaginary error function. Similar waiting time distributions for evolutionary processes were derived by Kimura [29].

Table 1 shows the average times $T_n(0)$, starting from the epoch center $\epsilon = 0$, for some of the epochs of figure 1(g) with $N = 10$, $K = 6$, $q = 0.001$, and $M = 50$. Comparing these average destabilization times with figure 1(g) we see that they give reasonable predictions of the average epoch stability times. The above numbers should be seen as an "order of magnitude" estimate. They are rather sensitive to the exact value of P_n^n and since our calculation of P_n^n essentially involves approximations, so do the above destabilization times. They do, however,

$\mathbf{T_6(0)}$	$\mathbf{T_7(0)}$	$\mathbf{T_8(0)}$	$\mathbf{T_9(0)}$	$\mathbf{T_{10}(0)}$
3×10^{13}	4.8×10^8	9.9×10^5	2.5×10^3	2600

Table 1: Average destabilization times $T_n(0)$ for some epochs ($n = 6 - 10$) in the Royal Road GA with $N = 10$, $K = 6$, $q = 0.001$, and $M = 50$. These were the parameters used for the run plotted in figure 1(g).

$\mathbf{T_3(0)}$	$\mathbf{T_4(0)}$	$\mathbf{T_5(0)}$	$\mathbf{T_6(0)}$
1.7×10^{22}	5.5×10^8	21×10^3	375

$\mathbf{T_7(0)}$	$\mathbf{T_8(0)}$	$\mathbf{T_9(0)}$	$\mathbf{T_{10}(0)}$
65	28	15	9

Table 2: Average destabilization times $T_n(0)$ for some epochs ($n = 3 - 10$) in the Royal Road GA with $N = 10$, $K = 6$, $q = 0.005$, and $M = 50$. These were the parameters used for the run plotted in figure 1(h).

nicely explain the occurrence of the intermittent behavior seen around epochs 9 and 10 in this run.

Table 2 shows the average destabilization times for some of the epochs of figure 1(h), which has an increased mutation rate of $q = 0.005$. These predictions demonstrate why the fitness of the population gets trapped in a band that is set by the average fitnesses of the fifth and eighth epochs. Epoch 8 destabilizes so quickly that the population has almost no chance to find a ninth aligned block. Even if it could, it would not have time to stabilize on that epoch, since the innovations duration itself is longer than the destabilization time. The destabilization times are so short compared to innovation durations in this case that epochs are very hard to distinguish, if they can be distinguished at all. For these parameters the dynamics is almost completely governed by the fluctuations of the finite population sampling. Selection and mutation cannot stabilize the population against these sampling fluctuations, though they do set the bounds on this range. We find $f_5 - 2\sigma_5 \approx 4$ and $f_8 + 2\sigma_8 \approx 7$ which exactly matches the band within which the fitness fluctuates in figure 1(h). For fitnesses below 4 selection pushes the population up against the sampling fluctuations. For fitnesses above 7 the mutations push the population down against the sampling fluctuations. The same mechanism is at work in figure 1(e). Here, too, the size of the fluctuation band is explained by the above analysis. For $N = 10$, $K = 6$, $q = 0.0075$, and $M = 500$ (run 1(e)) we find that epoch 7 is still stable for 48×10^3 time steps, epoch 8 for about 226 time steps, epoch 9 for 35 time steps, and finally the tenth epoch is stable only for about 13 time steps on average. We find for the fluctuation band $f_7 - 2\sigma_7 \approx 4.7$ and $f_9 + 2\sigma_9 \approx 7.1$ which again explains the data from figure 1(e).

Destabilization is very closely related to the so-called "error threshold" of self-replicating molecules in the theory of molecular evolution [42]. It was found that when the size of the genome and the selection pressure are kept constant and the mutation rate is increased, there is a sharp transition from a regime where the most fit genotype is always in the population to a regime where it will almost always be lost. This transition point is referred to as the "error threshold". In a complementary way, for a fixed mutation rate, there will be an equivalent "size threshold" on increasing genome length since mutations *anywhere* in the genome will reduce its fitness. Above a certain critical size the mutations will out-compete selection and the fittest

[12]Note that $U_n^i = -P_n^n$ since $P_n^i = 0$ for all $i < n$.

genotype will be lost.

This is analogous to what happens in the Royal Road GA. Under constant mutation rate there is a certain upper limit on the number of aligned blocks that selection can keep in the population. In the above cases this threshold occurred around 9 and around 5 blocks for runs 1(g) and 1(h), respectively. In the region of the critical number of blocks this leads to intermittent behavior in the average fitness. The population "hops" between different epochs when either the highest fitness string is lost through a fluctuation or a new block is (re)discovered and spreads through the population. When the intermittency time becomes shorter than the innovation durations, epochal behavior disappears, and the population seems to fluctuate randomly in a wide band that encompasses several epoch levels.

We see that there is a functional genome-length $L_n = nK$ versus mutation rate threshold in this GA. Our analysis also establishes a population size M versus mutation rate q error threshold. This suggests that there is a critical error-threshold surface in the three dimensional phase space spanned by the parameters M, q, and L_n.

7. Epoch Durations

We will now turn to the most important feature of Royal Road GA behavior that remains to be addressed—namely, the average length of the fitness epochs. Until now, almost all behavioral features could be understood in terms of the epoch fitness levels f_n and the epoch centers $\vec{P}^n$. We will investigate to what extent the same analysis can be used to predict the average epoch durations.

The ending of an epoch has two phases. First, a string has to be created of higher fitness than currently exists in the population. Thus, if the population resides in epoch n, a string of at least fitness $n+1$ has to be created by mutation. Second, this string has to be able to spread through the population if the population is to leave epoch n. Especially for higher-fitness epochs, where the relative fitness increase of new strings with respect to the old strings becomes small (i.e., proportional to $1/n$), it is likely that the best string will be lost through a sampling fluctuation before it gets a chance to spread through the population as was first shown by Fisher [18].

7.1 Creation of a Higher Fitness String

We will first calculate the probability that a string with fitness $n+1$ or higher is created while the population resides in the n^{th} epoch. During the n^{th} epoch the fitness distribution is given by $\vec{P}^n$ on average and the population resides in the n-dimensional subsimplex. The probability for the population to remain in the subsimplex over one generation is the chance that all individuals of the new generation have fitness smaller or equal to n. When an individual is selected and mutated for the new generation, the probability $\Pr[in]$ that it will have a fitness $i \leq n$ and thus remain within the subsimplex is given by:

$$\Pr[in] = \sum_{i=0}^{n} \mathbf{G}_i(\vec{P}^n) = \sum_{i=0}^{n} \mathbf{G}_i^n(\vec{P}^n) = \sum_i \frac{e_n}{f_n} \vec{P}_i^n = \frac{e_n}{f_n} \ . \tag{7.1}$$

The first equality follows from the fact that, given the epoch distribution $\vec{P}^n$, the probability to create a string of fitness i is determined by the ith component of the generation operator

acting on $\vec{P}^n$. The second equality notes that by restricting ourselves to $i \leq n$, the component i of the generation operator acting on $\vec{P}^n$ is equal to the restricted operator component $\mathbf{G}_i^n$. The third equality uses the fact that $\vec{P}^n$ is an eigenvector of the linearized restricted operator $\tilde{\mathbf{G}}^n$ with eigenvalue e_n. The final equality uses the fact that $\vec{P}^n$ is normalized to one.

The probability that *all* M individuals remain in the subsimplex is given by $\Pr[in]^M$. The probability $\Pr[out]$ that one or more individuals have jumped out of the subsimplex, by creating a string with fitness greater than n, is

$$\Pr[out] = 1 - \left[\frac{e_n}{f_n}\right]^M. \tag{7.2}$$

We will assume that the population resides at the epoch center and therefore has the same probability $\Pr[out]$ at each time step to jump out of the epoch. The expected number of time steps τ_n until the population jumps out of the epoch n subsimplex is then given by

$$\tau_n = \frac{1}{Pr[out]} = \frac{f_n^M}{f_n^M - e_n^M}. \tag{7.3}$$

For small q or large K, the probability A to align a block is small and we can expand to leading order in A to find:

$$\tau_n = \frac{1}{M(N-n)A}. \tag{7.4}$$

Thus, the expected number of generations τ_n to jump out of epoch n is inversely proportional to the probability A to align a block, the population size M, and the number of unaligned blocks $N-n$. We now investigate the probability that the new higher-fitness string will spread through the population.

7.2 Takeover of the Population by a Higher Fitness String

When a string of fitness $(n+1)$ is created, the initial proportion P_{n+1} of such strings is $1/M$.

Using the results from section 6.7 we see that the expected change $\langle dP_{n+1} \rangle$ per time step is given by

$$\langle dP_{n+1} \rangle = \frac{f_{n+1} - f_n}{f_n} P_{n+1} = \gamma_n P_{n+1}, \tag{7.5}$$

for small P_{n+1}. The second moment of the change dP_{n+1} is given by the sampling fluctuations:

$$\langle (dP_{n+1})^2 \rangle = \frac{P_{n+1}(1 - P_{n+1})}{M}. \tag{7.6}$$

Assuming that the change per time step is dominated by these first two moments, we can solve for the probability $\pi(p)$ that the new higher-fitness string will spread through the population and eventually reach proportion P_{n+1}^{n+1} , given that it initially has proportion p. To solve for $\pi(p)$ we use the backward Fokker-Planck equation:

$$\frac{\partial \pi(p,t)}{\partial t} = \langle dP_{n+1} \rangle \frac{\partial \pi(p,t)}{\partial p} + \frac{\langle (dP_{n+1})^2 \rangle}{2} \frac{\partial^2 \pi(p,t)}{\partial p^2}, \tag{7.7}$$

where $\pi(p,t)$ is the probability that the higher fitness string will have reached P_{n+1}^{n+1} by time t. The probability $\pi(p)$ that the mutant will spread, is given by the limit of $\pi(p,t)$ as t goes to infinity. This calculation was first done by Kimura in 1962 [26] in the context of the drift of the frequency of a certain genotype in a population. The solution is

$$\pi(p) = \frac{\int_0^p G(x)dx}{\int_0^{P_{n+1}^{n+1}} G(x)dx}, \tag{7.8}$$

where, for our case, the function $G(x)$ is given by[13]

$$G(x) = (1-x)^{2M\gamma_n}. \tag{7.9}$$

Performing the integral, we obtain

$$\pi_n \equiv \pi(\frac{1}{M}) = \frac{1-\left(1-\frac{1}{M}\right)^{2M\gamma_n+1}}{1-\left(1-P_{n+1}^{n+1}\right)^{2M\gamma_n+1}} \approx 1-e^{-2\gamma_n}, \tag{7.10}$$

where we have set the initial proportion $p = 1/M$. The approximation on the right-hand side holds only for large population sizes. Equation 7.10 tells us that the population has to find a better string $1/\pi_n$ times on average before it finally moves from the n^{th} to the $(n+1)^{st}$ epoch. Therefore, the total average time T_n the population spends in epoch n is

$$T_n = \frac{\tau_n}{\pi_n} = \frac{f_n^M}{(f_n^M - e_n^M)(1-\exp(-2\gamma_n))}. \tag{7.11}$$

For small q or large K this becomes

$$T_n = \frac{1}{M(N-n)A\left[1-\exp(-2/n)\right]}. \tag{7.12}$$

As we found in section 6.8, the fitness fluctuates around f_n during the n^{th} epoch in an approximately Gaussian way with the standard deviation given by

$$\sigma_n = \sqrt{\frac{Kn^3q}{2M}}. \tag{7.13}$$

We thus find the average number of time steps $T_n(f)$ that the population has fitness f during epoch n is given by

$$T_n(f) = \frac{T_n}{\sqrt{2\pi}\sigma_n} \exp\left[-\frac{1}{2}\left(\frac{f-f_n}{\sigma_n}\right)^2\right]. \tag{7.14}$$

We performed experiments to test these theoretical predictions by accumulating histograms of the average number of time steps $\langle T\rangle$ that the population has fitness f during a run, averaged over a large number of runs. Figure 13 shows the results of such an experiment for the parameter setting $N = 10$, $K = 6$, $M = 500$, and $q = 0.001$ of run 1(d) together with the theoretical predictions for the peaks at these parameter settings. The inset plot shows a magnification of

[13]This function is an essential quantity in the diffusion equation method. It's obtained by taking the exponential of the integral of the ratio $\langle dx\rangle/\langle dx^2\rangle$.

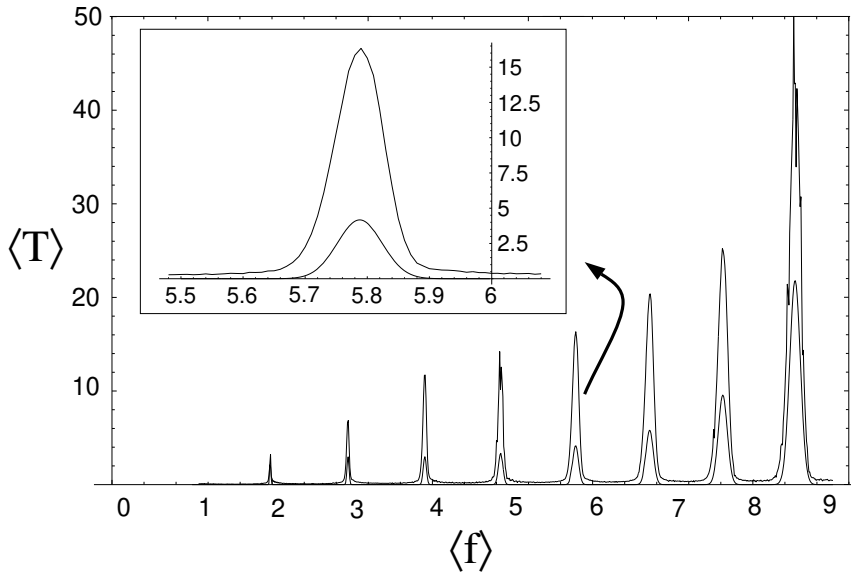

Figure 13: Empirical (upper curve) and theoretical (lower curve) fitness histograms. The horizontal axis shows the average fitness and the vertical axis shows the average number of time steps the population has that average fitness during a GA run, averaged over 500 runs. The parameters for this run are $N = 10$, $K = 6$, $q = 0.001$, and $M = 500$. The inset plot shows a magnification of the peak at the 6^{th} epoch.

the peak at the 6^{th} epoch. The upper curves are from the experiment and the lower curves plot the theoretical predictions (equation 7.14). As can be clearly seen from the figure the theory substantially underestimates the average lengths of the epochs found at this parameter setting. As shown before, the widths and locations of the peaks are correctly predicted by the theory. The empirically observed averages are offset vertically from the theoretical predictions outside of the peak region since the theoretical curve does not take into account the time the population spends in the innovations. This does not account, however, for the fact that the predicted peaks are about a factor of 6 too small.

Figure 14 shows the results for the parameter setting $N = 20$, $K = 3$, $M = 300$, and $q = 0.001$ of run 1(c) together with the theoretical predictions of the epoch duration peaks for these parameter settings. The inset plot shows a magnification of the peaks around the 16^{th}, 17^{th}, and 18^{th} epochs. Again it is clear that the theory underestimates the epoch durations; here it does so by a factor of approximately 3. Nonetheless, it is instructive to compare the above histogram with run 1(c). From that figure it is very hard to say how many epochs occur and what their durations and exact locations are. The fitness histogram of figure 14 clearly shows that the epochs are *still there* in the GA dynamics. They are reflected in the peaks in the plots. With this observation in mind, we propose to *define* the existence of an epoch in the dynamics by the appearance of a peak in the average-time histogram. In the empirical curve in figure 14 at least 10 peaks can be counted. It would have been very hard to get such a clear view of the epochal behavior from plots such as those in figure 1(c). Averaging the average fitness at each time step from a large number of runs like that in figure 1(c) only makes detecting epochs more difficult. Since the onset of each epoch shifts in time from run to run, averaging the fitness values at each time step over many runs completely washes out any trace of epochal dynamics.

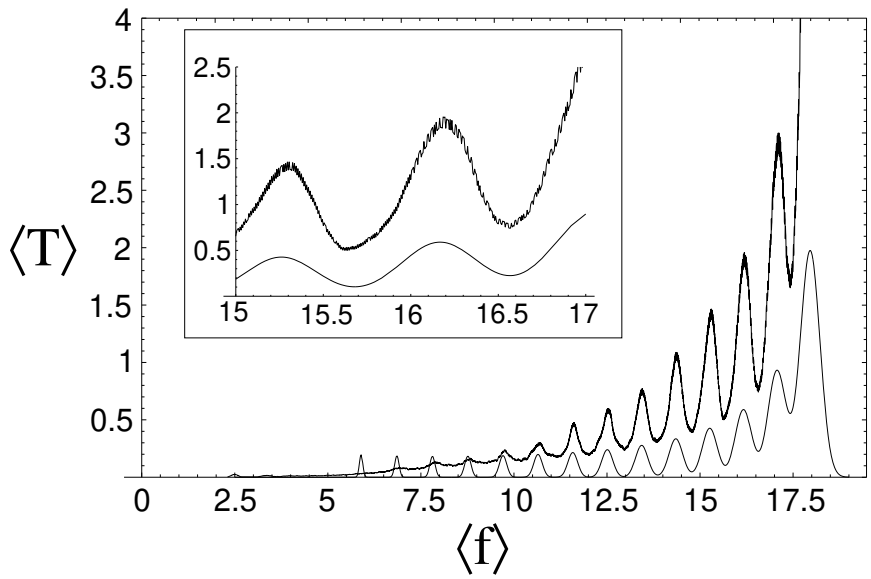

Figure 14: Empirical (upper curve) and theoretical (lower curve) fitness histograms. The horizontal axis shows the average fitness and the vertical axis shows the average number of time steps the population has that average fitness during the run, averaged over 500 runs. The parameters for this run are $N = 20$, $K = 3$, $q = 0.001$, and $M = 300$. The inset plot shows a magnification of the peaks at the 16^{th}, 17^{th}, and 18^{th} epochs.

We also see from figure 14 that the distributions associated with each epoch start overlapping at these parameters. This is, of course, the reason why the epochs are hardly discernible in figure 1(c). In obtaining the theoretical curve for figure 14 we summed the contributions of the peaks corresponding to the 6^{th} through 19^{th} epochs. It is clear from the figure that the peaks of the 6^{th}, 7^{th} and 8^{th} epochs do not actually occur in the behavior. The reason for this is that in calculating the peaks for the epochs we assumed that the population starts in the epoch center. In the actual behavior seen in the run the population never reaches these epoch centers.

Why do the theoretical predictions of the epoch durations markedly underestimate the actual epoch lengths in the behavior? As noted in section 6.5 a new epoch is almost always founded by a single individual. This means that at the start of the epoch all strings of the current highest fitness class are essentially the same, the population is highly converged. It is obvious in this case that assuming all unaligned blocks are statistically independent breaks down, because all unaligned blocks in the strings of the current highest fitness class are almost identical at the start of the epoch. The length of the epoch is very sensitive to the number of bits set to 1 in the unaligned blocks of the epoch's founding string. The fact that all unaligned blocks in the population are the *same* at the start of an epoch will cause the epoch to be longer on average than it would be if all unaligned blocks were independent. A more subtle but unfortunately poorly understood effect, as pointed out in [24], is the fact that finite population sampling in general prohibits the unaligned blocks in the population from becoming completely independent. That is, sampling causes the strings in the population to remain correlated for all time thereby effectively reducing the proportions of unaligned blocks that have a high fraction of bits set to 1. These factors lead to the theory's underestimation of the epoch durations. We are currently studying the exact dynamics of a finite population of strings searching for a new block while it resides in an epoch starting from a completely converged population at the start of the epoch. The results, which will be presented elsewhere, lead to greatly improved epoch

duration predictions.

8. Discussion

We have seen how most of the behavioral features of the Royal Road genetic algorithm—the appearance and disappearance of epochs, the structure of the innovations, the variation in fitness levels and their fluctuations, and the epoch durations—can be understood in terms of the properties of the infinite-population generation operator **G**. The analysis showed how the basic balance of evolutionary "forces"—ordering due to selection, increased diversity and aligned-block creation and destruction due to mutation, discreteness of the state space due to the finite population, and stochasticity from finite population sampling—competed and cooperated to produce a wide range of phenomena. Some of the trade-offs between these pressures, as controlled by the GA parameters, were nonmonotonic and occasionally counterintuitive. As a concise summary of these trade-offs, table 3 presents an overview of the major analytical results we obtained for the different dynamical quantities, for small mutation rates q.

$f_n = n(1-q)^{nK}$ (6.17)	Epoch fitness
$\sigma_n^2 = Kn^3 q/2M$ (6.48)	Epoch fitness fluctuations
$P_n^n = 1 - n^2 Kq$ and $P_{n-1}^n = n^2 Kq$ (6.12)	Epoch population
$\lambda_i^n = f_i/f_n = i(1-q)^{(i-n)K}/n$ (6.23)	Epoch stability
$T_n = [M(N-n)A(1-\exp(-2/n))]^{-1}$ (7.12)	Epoch duration
$t_n = [1 + 2n + 2Kn(n+1)q] \log M$ (6.32)	Innovation duration

Table 3: Low mutation behavior of the Royal Road genetic algorithm: an overview of the analytical results for small mutation rates q. n denotes the epoch number, K is the number of bits in a block, N the total number of blocks, q is the mutation rate, and M is the population size.

8.1 Low Mutation Rate Results

The first line shows the average fitness f_n of the n^{th} epoch. The fitness is decreased from n by a factor that drops geometrically as a function of the number of defining (aligned-block) bits (nK) of the epoch.

The second line gives the variance σ_n^2 of an epoch's average fitness fluctuation. The fluctuation amplitudes (σ_n) are proportional to the epoch level n to the power 3/2 and inversely proportional to the square root of the population size M. They are also proportional to the square root of the block size K.

The third line shows that the proportion P_n^n of individuals in the highest fitness class drops proportional to the block size K and with the square of the epoch number n. Likewise it

shows that the proportion P^n_{n-1} of strings in the $(n-1)^{th}$ fitness class also grows with the same coefficient.

The fourth line gives the eigenvalues λ^n_i of the Jacobian matrix around the n^{th} epoch center. These eigenvalues determine the bulk of the epoch's stability. In particular, they control the epoch's innovation and fluctuation dynamics. They can be simply expressed in terms of the relative sizes of the epoch fitness levels f_i.

The fifth line shows the theoretical predictions of the average epoch duration T_n of epoch n to leading order in q. We have seen that these predictions underestimate the average epoch duration. The expression is included for completeness. The theoretical epoch duration is inversely proportional to the probability to create a block A, the population size M, and a factor that depends on the probability, $1-\exp(-2/n)$, that a fitter string will spread in the population. Note, that since $1/A$ is proportional to 2^K, the epoch duration increases exponentially with the block length K. This explains why epochal behavior is mainly seen for large blocks.

Finally, the last line shows the average time t_n taken for the innovation from the n^{th} to the $(n+1)^{st}$ epoch. The result shows that the innovation time is proportional to $2n$ and to the logarithm of the population size M. It also shows that increasing mutation decreases the steepness of the innovations roughly in proportion to the square of the epoch number n, the size of the blocks K, and the logarithm of the population size M.

In focusing our attention on fitness distributions we assumed that the exact inner structure of the unaligned blocks is unimportant and can be taken as random. As we demonstrated, this maximum entropy assumption breaks down for the calculation of the average epoch durations. In general, analyzing GA behavior solely in terms of fitness distributions will work if strings within the same fitness class act similarly under the GA dynamics. In cases where this simplification does not work one must include additional order parameters to the describe the projected state of the microscopic system. In the Royal Road GA a number of alternatives come to mind. These include using a distribution of the number of 1-bits contained in the unaligned blocks or an order parameter that describes the convergence of the bits in the unaligned blocks. If the number of fitness classes or the number of "order parameters" in general becomes too large, the analysis, though still appropriate in principle, will break down from a practical point of view. The generation operator $\mathbf{G}$ could simply acquire too many components for it to be theoretically or even numerically analyzed. However, under some circumstances, such as when fitness is determined in some noisy way, different fitness classes may be grouped together to obtain a low-dimensional $\mathbf{G}$. The resulting coarse-grained fitness "landscape" itself may be rather complicated, but as long as it is known, our analysis can also be performed even in these cases. More interestingly, our analysis suggests that it might be enough to determine certain statistics of the fitness landscape to be able to predict the population dynamics. For instance, the fitness levels of the epochs depend only on the number of "defining bits" of the epoch's fitness class (which is roughly the logarithm of the number of strings in the fitness class).

8.2 Metastability, Unfolding, and Landscapes

The main result of the preceding analysis is our explanation of epochal evolution as an interplay between the infinite-population flow given by $\mathbf{G}$ and the coarse-graining of the state space due to finite population size. Note that this mechanism for metastability is quite general and applies outside of evolutionary search behavior. A large number of dynamical systems in nature, as well as evolutionary computation in general, are stochastic dynamical systems in which a large

set of identical subsystems evolve through a state space, in parallel, and under the influence of one another. Macroscopic states for these systems are often defined in terms of the first moments of the distributions over the state variables of the components. A commonly observed qualitative behavior is that the mean of some state variable alternates between periods of stasis and sudden change. We would like to suggest that this sort of punctuated equilibrium behavior [23] can be explained in terms of the simple mechanism presented here. In the limit of an infinite number of subsystems the global dynamics is often much more tractable. Solving for the "flow" through the appropriate state space in the limit of an infinite number of subsystems can be used to identify state space regions where the flow is weakest. Then, in the case of a finite number of subsystems, we expect the dynamics to get trapped in the weak-flow regions. The behavioral features of the finite-size dynamics can be almost completely understood in terms of the eigenvalues and eigenvectors of the infinite-population flow operator.

The behavior of evolutionary search algorithms is often informally described as moving along a "fitness landscape" directly defined by a fitness function. It is clear from both our experiments and our analysis that this geographic metaphor, originally due to Wright [46], can be misleading. The fitness function is only a partial determinant of the dynamics. Even with a fixed fitness function, population size, mutation rate, and other parameters of the search process can radically alter the population dynamical behavior of the system, revealing or hiding much of the structure in the fitness landscape. Moreover, significant features of the population dynamics, such as the metastability and destabilization of epochs, are endogenous in the sense that they cannot be understood from a naive analysis of the landscape alone.

Let us recall that for typical evolutionary problems the genetic sequence spaces explored by populations are vast. In our case of bit strings with modest length $L = 60$, the size of the sequence space is already $2^L \approx 10^{18}$. It is clear that at any point in time the population can occupy only a minute proportion of the sequence space. It is therefore logical to assume that the population in principle could become trapped in certain regions of this state space. It has become fashionable to assume that the fitness functions over the sequence spaces that occur for typical evolutionary search problems can be modeled as "rugged landscapes" [36]. These rugged landscapes have wildly fluctuating fitness values even on very small genome-variation scales and possess a large number of local fitness optima. It has been assumed that after some time, the population is likely to be found *localized* around some suboptimal fitness peak in the sequence space. This notion of locality is also common in the analysis of quasispecies in the molecular evolution theory [12]. There, the population consists of a cloud of mutants around a local fitness optimum called the "wild type". A balance exists between the forces of mutation and crossover that makes the population diffuse away from the peak and the force of selection that tends to restore the population towards the peak. Within this view it becomes natural to associate metastable evolutionary behavior with hopping between these local fitness optima. The population stabilizes at the local peak until one of the mutants at the outer edge of the quasispecies cloud discovers a new and higher fitness peak. The population is metastable until a mutant crosses a "valley" of lower fitness towards a higher fitness peak. The local fitness optimum therefore presents a fitness barrier that has to be crossed by an individual of the population. The height of this barrier is determined by some measure of the "steepness" of the local fitness peak.

This view of metastability is in sharp contrast with the mechanism presented in this paper. Kimura was the first to advocate that many of the point mutations that occur on the molecular level are neutral with respect to fitness [28]. There is often a large degeneracy between the

genotype and the fitness or functionality of its phenotype—many genotypes lead to the same or similar fitness. This means that although the fitness landscape might be rugged, there are always neutral ridges along which the genotype can move without affecting fitness. In some cases local optima might disappear completely from the fitness landscape, as in the Royal Road fitness function. As we discussed in section 6.4, if we let S_n denote that subspace of all sequences containing strings of fitness n, then there are always points in S_n that are only a single mutation away from S_{n+1}. In any particular time interval, the population is not likely to be localized in sequence space, but instead diffuses randomly within subspace S_n until one of its mutants moves into the higher-fitness subspace S_{n+1}. Metastability occurs here on the phenotypic level of fitness or functionality. As we have seen, the fitness distribution stabilizes on an epoch center $\vec{P}^n$, while the best individuals in the population randomly diffuse through the subspace S_n. The fitness distribution remains stable until one of the mutants moves into S_{n+1}. The time that this takes depends on the relative number of points in S_n that connect to the subspace S_{n+1}. In general, there is only a very small proportion of such "portal" genotypes in S_n. (In our case the proportion is on the order of 2^{-K}.) The consequence is that the metastability here is due to an entropic barrier—in which large fitness-neutral volumes must be traversed. The entropy of the population in S_n has to increase until almost all points in S_n have been visited by the population and a connection to S_{n+1} is discovered. Thus, increased phenotypic sophistication is reached by passing through conditions of increased genotypic disorder. This mechanism is quite different from the metastable behavior in sequence space due to local fitness barriers. We believe that the kind of entropy-induced metastable behavior described above is very common in evolutionary search. It will occur as long as there is a large degeneracy between the genotype and its actual functionality.

8.3 Future Work

We are currently studying in more detail the process of aligning blocks from a converged population. We hope that this will eventually enable us to predict the distribution of epoch durations in these entropic metastable states. In addition, we plan to include in the analysis other aspects of evolutionary systems such as crossover, geographically distributed populations with migration, and noise in the fitness function. We are currently studying more general fitness functions to see how the metastable behavior generalizes to those that have both entropic barriers and local optima and, moreover, to determine which macroscopic variables defined on fitness landscapes are most informative in predicting actual population dynamical behavior.

We are also investigating epochal evolution from an information and computation theoretic point of view [7]. Selection can be seen as installing structural information from the fitness function into the genotypes. Mutation and crossover, in contrast, can be seen as randomizing the bits in the genetic representation and thereby destroying the information that selection has stored in them, while providing a necessary source of genetic novelty. The genetic representation itself—the string length and the strings' blockiness in our simple GA—also imposes restrictions on how much of the fitness function's structure can be stored in the gene pool.

It is already evident from the preceding investigations that epochal dynamics leads to interesting informational behaviors of the population. At an innovation, for instance, the whole population becomes converged. This means that essentially *all* bits of the founder string are stored in the population, not just the "functional" bits in the aligned block, but also the arbitrary bits this founder string happens to have in the unaligned blocks. (This kind of phenomenon has also been called "hitchhiking" in the GA literature [34].) Thus, at an innovation,

more raw information is stored in the population than is necessary for improved fitness. During the epoch the unaligned blocks start to diversify again and the founder string information originally stored in these unaligned blocks is destroyed. This thermalization process increases the population entropy, but is actually a prerequisite for the search being able to find increasingly better genotypes. We shall present this complementary thermodynamic analysis elsewhere.

Eventually, we would like to extend the analysis to evolutionary processes in which there is a nontrivial mapping between the genotype and the phenotype, such as found in our other work on a genetic algorithm that evolves cellular automata [8, 9], work that originally motivated the preceding investigations. We presume that entropic metastability will also be observed in these more complex adaptive systems as long as there is enough redundancy between a genotype and its phenotype.

9. Acknowledgments

The authors thank Lee Altenberg, Marc Feldman, Martijn Huynen, Mark Newman, Richard Palmer, Christian Reidys, and Andreas Wagner for helpful discussions. This work was supported at the Santa Fe Institute by NSF IRI-9320200, DOE DE-FG03-94ER25231, and ONR N00014-95-1-0975, and at UC Berkeley under ONR N00014-95-1-0524.

A. Uniqueness of the Asymptotic Fitness Distribution

We will show that the matrix $\tilde{\mathbf{G}}$ has only one positive definite eigenvector as was claimed in section 5.5. Since individuals with fitness 0 have probability 0 of being selected we have $\mathbf{S}_{i0} = 0$ and $\tilde{\mathbf{G}}_{i0} = 0$. Because there is positive probability for any nonzero-fitness string to be selected and since, with positive probability, mutation can take a string with fitness j into a string with fitness i for any j and i, we have

$$\tilde{\mathbf{G}}_{ij} > 0,\ j > 0. \tag{A.1}$$

Define a new matrix $\mathbf{H}$ that is the restriction of $\tilde{\mathbf{G}}$ to the positive fitness subspace. This is given by

$$\mathbf{H}_{ij} = \tilde{\mathbf{G}}_{ij},\ i, j > 0. \tag{A.2}$$

In addition, define the vector $\vec{Q}$ to be the N-dimensional nonzero fitness projection of the $(N+1)$-dimensional fitness distribution $\vec{P}$ to be

$$Q_i = P_i\ ,\ i > 0. \tag{A.3}$$

We can now turn the eigenvector equation

$$\tilde{\mathbf{G}} \cdot \vec{P} = f\vec{P} \tag{A.4}$$

into an eigenvector equation for $\mathbf{H}$,

$$\mathbf{H} \cdot \vec{Q} = f\vec{Q}, \tag{A.5}$$

and into an equation for the zeroth component P_0,

$$P_0 = f^{-1} \sum_{j=1}^{N} \tilde{\mathbf{G}}_{0j} P_j. \tag{A.6}$$

Since $\mathbf{H}$ is a positive definite matrix, Perron's theorem applies: $\mathbf{H}$ has a unique positive definite eigenvector, the eigenvalue of which is larger than all other eigenvalues. This means that $\mathbf{H}$ has a unique positive eigenvector $\vec{Q}^{max}$ with maximal eigenvalue f_{max}. Since all components $\tilde{\mathbf{G}}_{0j}$ are positive we have from equation A.6 that P_0 is positive. Specifically, we have

$$P_0 = \frac{\sum_{j=1}^{N} \tilde{\mathbf{G}}_{0j} Q_j^{max}}{f_{max}} \tag{A.7}$$

and so $\vec{P} = (P_0, Q_1, \ldots, Q_N)$ is the unique eigenvector of $\tilde{\mathbf{G}}$ with maximal eigenvalue. Since all other eigenvectors of $\mathbf{H}$ have at least one negative component, the above eigenvector $\vec{P}$ is the unique positive definite eigenvector of $\mathbf{G}$ and therefore $\vec{P}$ is the only eigenvector of $\mathbf{G}$ that can be interpreted as a fitness distribution.

B. Finite Population Dynamics Convergence in the Infinite Population Limit

We will show that as the population size increases, the finite population dynamics approaches arbitrarily closely the infinite population dynamics for any finite number of time steps T. The proof presented here is a more elaborate version of a proof that was outlined in [35]. (Useful mathematical background can be found in [17].) Note that we will prove that the finite population dynamics converges towards the infinite population dynamics as given by $\mathbf{G}$, *provided* that $\mathbf{G}$ accurately describes the dynamics of fitness distributions for an infinite population.

In the infinite population limit the dynamics is deterministic. Let $\vec{I}(t)$ denote the fitness distribution at time t in the infinite-population limit. For the finite population the dynamics in fitness distribution space is stochastic. Let $\vec{P}(t)$ denote the finite population fitness distribution at time t. We define the distance between the i^{th} component of the finite population fitness distribution and the infinite population fitness distribution at time t to be

$$\delta_i(t) = |I_i(t) - P_i(t)|. \tag{B.1}$$

At time $t = 0$ the fitness distribution for the finite population is taken to be the infinite population distribution $\vec{I}(0)$. Using equation 6.1 for the transition probabilities we can calculate the probability that $\delta_i(1) > \epsilon$ for some arbitrary component, $0 \leq i \leq N$. The vector $\vec{P}(1)$ is given by a multinomial sample of size M of the expected distribution $\vec{I}(1) = \mathbf{G}(\vec{I}(0))$. Using Chebysev's inequality we find that

$$\Pr[\delta_i(1) > \epsilon] \leq \frac{I_i(1)(1 - I_i(1))}{M\epsilon^2} \leq \frac{1}{4M\epsilon^2}, \tag{B.2}$$

using the inequality $x(1-x) \leq 1/4$.

With the above inequality on the transition probabilities we can prove that for any $\gamma < 1$, any $\beta > 0$, and any finite number of time steps T, there is a population size M such that for populations larger than M with probability at least γ, any component $\vec{P}_i(t)$ of the fitness distributions $\vec{P}(t)$ stays within β of the infinite population trajectory. Specifically, we want to prove that for sufficiently large populations one has

$$\Pr\left[\delta_i(1) \leq \epsilon(1) \text{ and } \delta_i(2) \leq \epsilon(2) \text{ and} \ldots \text{and } \delta_i(T) \leq \epsilon(T)\right] > \gamma, \tag{B.3}$$

where the $\epsilon(t)$ are uniformly smaller than the chosen bound β for all t. Since the process has the Markovian property that the next fitness distribution depends only on the current one, we can factor the left-hand side of B.3 into conditional probabilities:

$$\Pr\left[\delta_i(1) \leq \epsilon(1)\right] \prod_{t=1}^{T-1} \Pr\left[\delta_i(t+1) \leq \epsilon(t+1) | \delta_i(t) \leq \epsilon(t)\right]. \tag{B.4}$$

Thus, we need to bound each of these conditional probabilities. Given a population with fitness distribution $\vec{P} = \vec{I}(t) + \vec{\delta}(t)$, the expected distribution $\vec{P}'$ at the next time step is $\mathbf{G}(\vec{P})$. To first order in $\vec{\delta}(t)$ this is given by

$$\langle \vec{P}(t+1) \rangle = \mathbf{G}(\vec{I}(t) + \vec{\delta}(t)) = \vec{I}(t+1) + \mathbf{DG} \cdot \vec{\delta}(t) + \mathcal{O}(\delta^2), \tag{B.5}$$

where the Jacobian matrix $\mathbf{DG}$ is evaluated with $\vec{I}(t)$:

$$\mathbf{DG}_{ij} = \frac{\tilde{\mathbf{G}}_{ij} - j I_i(t+1)}{f(t)}, \tag{B.6}$$

where we denote the average fitness of the infinite population distribution $\vec{I}(t)$ as $f(t)$. We now need to place an upper bound B on the absolute values of the eigenvalues of this matrix such that we can write

$$\left| \sum_j^N \mathbf{DG}_{ij} \delta_j(t) \right| \leq B \delta_i(t). \tag{B.7}$$

The bound B is nothing other than the norm of the matrix $\mathbf{DG}$, which we can obtain by explicitly substituting a vector $\vec{\delta}$ into equation B.6,

$$\mathbf{DG} \cdot \vec{\delta} = \frac{\tilde{\mathbf{G}} \cdot \vec{\delta} - \vec{I}(t+1) \langle \delta \rangle}{f(t)}, \tag{B.8}$$

where $\langle \delta \rangle = \sum_j^N j \delta_j$. Since $f_i \geq 0$, $I_i(t+1) > 0$, and $\delta_i > 0$ for all i, the two terms in the above expression are of opposite sign and therefore the norm of the above expression is bounded by the norm of the larger of the two terms. That is, we have

$$\left| \sum_j^N \mathbf{DG}_{ij} \delta_j(t) \right| \leq \max\{ \frac{|(\tilde{\mathbf{G}} \cdot \vec{\delta})_i|}{f(t)}, \frac{I_i(t+1)\langle \delta \rangle}{f(t)} \} \leq \frac{N}{f(t)} \delta_i. \tag{B.9}$$

Here we have used the inequalities $\langle \delta \rangle \leq N$, $I_i(t+1) \leq 1$, and $f_N \leq N$, with f_N being the largest eigenvalue of $\tilde{\mathbf{G}}$. We can therefore write

$$|\langle P_i(t+1) \rangle - I_i(t+1)| \leq \frac{N}{f(t)} \delta_i(t). \tag{B.10}$$

Using Chebysev's inequality again on the multinomial transition probabilities we obtain

$$\Pr\left[|P_i(t+1) - \langle P_i(t+1) \rangle| > \epsilon \right] \leq \frac{1}{4M\epsilon^2}. \tag{B.11}$$

Furthermore, we have that

$$|P_i(t+1) - \langle P_i(t+1)\rangle| \leq \epsilon \Rightarrow \delta_i(t+1) \leq \epsilon + \frac{N}{f(t)}\delta_i(t) \ . \tag{B.12}$$

Now, if we define

$$\epsilon(t+1) = \epsilon + \frac{N}{f(t)}\epsilon(t), \tag{B.13}$$

then we find that

$$\Pr\left[\delta_i(t+1) \leq \epsilon(t+1) | \delta_i(t) \leq \epsilon(t)\right] > 1 - \frac{1}{4M\epsilon^2}. \tag{B.14}$$

Looking over a series of time steps, if we define

$$\epsilon(t) = \sum_{n=0}^{t-1} \epsilon \left[\frac{N}{f(t)}\right]^n \tag{B.15}$$

and take $\epsilon(1) \equiv \epsilon$, then we find for the joint probability of equation B.4 that

$$\Pr\left[\delta_i(1) \leq \epsilon(1) \text{ and } \delta_i(2) \leq \epsilon(2) \text{ and} \ldots \text{and } \delta_i(T) \leq \epsilon(T)\right] > \left[1 - \frac{1}{4M\epsilon^2}\right]^T. \tag{B.16}$$

Requiring this probability to be greater than γ we obtain

$$M > \frac{1}{4\epsilon^2(1-\gamma^{\frac{1}{T}})}. \tag{B.17}$$

We also require that $\epsilon(t) < \beta$ for all $t \leq T$. From the definition of $\epsilon(t)$, we see that this requirement implies

$$\epsilon(T) < \beta \Rightarrow \epsilon < \beta \left[\sum_{n=0}^{T-1} \left(\frac{N}{f(t)}\right)^n\right]^{-1} \equiv \frac{\beta}{c[T]}, \tag{B.18}$$

where

$$c[T] = \sum_{n=0}^{T-1} \left(\frac{N}{f(t)}\right)^n. \tag{B.19}$$

Concluding, we see that if we choose M such that

$$M > \frac{c[T]^2}{4\beta^2(1-\gamma^{\frac{1}{T}})}, \tag{B.20}$$

then, for all $t \leq T$, the $\delta_i(t)$ are smaller than β with probability greater than γ. Since the right-hand side in B.20 is finite for any finite T, $\gamma < 1$, and $\beta > 0$ we can always find a population size M such that the above inequality is satisfied. This concludes the proof. Note that if we take the limit $T \to \infty$ there is no finite population M for which the deviations $\delta_i(t)$ remain arbitrarily small for arbitrary time.

References

[1] R. Baserga, editor. *Cell Proliferation, Cancer, and Cancer Therapy*, volume 397 of *Ann. New York Acad. Sci.*, 1982.

[2] R. K. Belew and L. B. Booker, editors. *Proceedings of the Fourth International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, CA, 1991.

[3] A. Bergman and M. W. Feldman. Question marks about the period of punctuation. Technical report, Santa Fe Institute Working paper 96-02-006, 1996.

[4] T.D. Brock. *Biology of microorganisms*, chapter 23, pages 942–943. Prentice Hall, 1997.

[5] T.D. Brock. *Biology of microorganisms*, chapter 23, pages 956–957. Prentice Hall, 1997.

[6] J. P. Crutchfield. Subbasins, portals, and mazes: Transients in high dimensions. *J. Nucl. Phys. B*, 5A:287, 1988.

[7] J. P. Crutchfield. The calculi of emergence: Computation, dynamics, and induction. *Physica D*, 75:11 – 54, 1994.

[8] J. P. Crutchfield and M. Mitchell. The evolution of emergent computation. *Proc. Natl. Acad. Sci. U.S.A.*, 92:100742–10746, 1995.

[9] R. Das, J. P. Crutchfield, M. Mitchell, and J. E. Hanson. Evolving globally synchronized cellular automata. In L. J. Eshelman, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 336–343, San Francisco, CA, 1995. Morgan Kaufmann.

[10] L. D. Davis, editor. *The Handbook of Genetic Algorithms*. Van Nostrand Reinhold, 1991.

[11] M. Eigen. *Naturwissenschaften*, 58:465–522, 1971.

[12] M. Eigen, J. McCaskill, and P. Schuster. The molecular quasispecies. *Adv. Chem. Phys.*, 75:149–263, 1989.

[13] M. Eigen and P. Schuster. *Naturwissenschaften*, 64:541–565, 1977.

[14] M. Eigen and P. Schuster. *The Hypercycle*. Springer, Berlin, 1979.

[15] S.F. Elena, V.S. Cooper, and R.E. Lenski. Punctuated evolution caused by selection of rare beneficial mutations. *Science*, 272:1802–1804, 1996.

[16] L. Eshelman, editor. *Proceedings of the Sixth International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, CA, 1995.

[17] W. Feller. *An Introduction to Probability Theory and its Applications*. Wiley, New York, third, revised edition, 1970.

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

[19] A. D. Fokker. *Ann. Phys. (Leipzig)*, 43(310), 1915.

[20] S. Forrest, editor. *Proceedings of the Fifth International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, CA, 1993.

[21] C. W. Gardiner. *Handbook of Stochastic Methods*. Springer-Verlag, 1985.

[22] J. H. Gillespie. Chapter 4. In M. W. Feldman, editor, *Mathematical Evolutionary Theory*. Princeton University Press, 1989.

[23] S.J. Gould and N. Eldredge. Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology*, 3:115–251, 1977.

[24] R. Haygood. The structure of royal road fitness epochs. *Evolutionary Computation, submitted*, ftp://ftp.itd.ucdavis.edu/pub/people/rch/StrucRoyRdFitEp.ps.gz, 1997.

[25] M. Huynen, P. F. Stadler, and W. Fontana. Smoothness within ruggedness: The role of neutrality in adaptation. *Proc. Natl. Acad. Sci.*, 93:397–401, 1996.

[26] M. Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47:713–719, 1962.

[27] M. Kimura. Diffusion models in population genetics. *J. Appl. Prob.*, 1:177–232, 1964.

[28] M. Kimura. *The neutral theory of molecular evolution*. Cambridge University Press, 1983.

[29] M. Kimura and T. Ohta. The average number of generations until fixation of a mutant gene in a finite population. *Genetics*, 61:763–771, 1969.

[30] A. N. Kolmogorov. *Math. Annal.*, 104:415–418, 1931.

[31] K. Lindgren. Evolutionary phenomena in simple dynamics. In C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, editors, *Artificial Life II*, pages 295–312, Reading, MA, 1992. Addison-Wesley.

[32] J. S. McCaskill. A stochastic theory of macromolecular evolution. *Biological Cybernetics*, 50:63–73, 1984.

[33] M. Mitchell, J. P. Crutchfield, and P. T. Hraber. Evolving cellular automata to perform computations: Mechanisms and impediments. *Physica D*, 75:361 – 391, 1994.

[34] M. Mitchell, J. H. Holland, and S. Forrest. When will a genetic algorithm outperform hill climbing? In J. D. Cowan, G. Tesauro, and J. Alspector (editors), *Advances in Neural Information Processing Systems 6*. San Mateo, CA: Morgan Kaufmann.

[35] A. E. Nix and M. D. Vose. Modeling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence*, 5, 1991.

[36] A. S. Perelson and S. A. Kauffman, editors. *Molecular evolution on rugged landscapes: proteins, RNA and the immune system*, volume IX of *Santa Fe Institute Studies in the Science of Complexity*, 1989.

[37] M. Planck. *Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl.*, 325, 1917.

[38] A. Pruğel-Bennett and J. L. Shapiro. Analysis of genetic algorithms using statistical mechanics. *Physical Review Letters*, 72(9):1305–1309, 1994.

[39] A. Prügel-Bennett and J. L. Shapiro. The dynamics of a genetic algorithm in simple random Ising systems. *Physica D*, 104 (1):75–114, 1997.

[40] M. Rattray and J. L. Shapiro. The dynamics of a genetic algorithm for a simple learning problem. *J. of Phys. A*, 29(23):7451–7473, 1996.

[41] T. S. Ray. An approach to the synthesis of life. In C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, editors, *Artificial Life II*, pages 371–408, Reading, MA, 1992. Addison-Wesley.

[42] J. Swetina and P. Schuster. Self replicating with error, a model for polynucleotide replication. *Biophys. Chem.*, 16:329–340, 1982.

[43] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, 1992.

[44] E. van Nimwegen, J. P. Crutchfield, and M. Mitchell. Finite populations induce metastability in evolutionary search. *Phys. Lett. A*, 229:144–150, 1997.

[45] M. D. Vose and G. E. Liepins. Punctuated equilibria in genetic search. *Complex Systems*, 5, 1991.

[46] S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proc. of the Sixth International Congress of Genetics*, volume 1, pages 356–366, 1932.