Evolvability of Complex Characters:
Population Dependent Fourier Decomposition of Fitness Landscapes over Recombination Spaces

PETER F. STADLER
Institut für Theoretische Chemie, Universität Wien
Währingerstraße 17, A-1090 Wien, Austria
Phone: **43 1 4277 52737 Fax: **43 1 4277 52793
and
The Santa Fe Institute
1399 Hyde Park Road, Santa Fe, NM 87501, USA
E-mail: studla@tbi.univie.ac.at or stadler@santafe.edu

RUDI SEITZ
Dept. of Electrical Engineering & Computer Science
Massachusetts Institute of Technology MIT 38-401
Cambridge, MA 02139-4307
E-mail: rseitz@mit.edu

GÜNTER P. WAGNER
Dept. of Ecology and Evolutionary Biology
Center for Computational Ecology
Yale University, New Haven, CT 06520, USA
E-mail: gpwag@peaplant.biology.yale.edu

The effect of recombination on genotypes can be represented in the form of P-structures, i.e. a map from the set of pairs of genotypes to the power set of genotypes. The interpretation is that the P-structure maps the pair of parental genotypes to the set of recombinant genotypes which result from the recombination of the parental genotypes. A recombination fitness landscape is then a function from the genotypes in a P-structure to the real numbers. In previous papers we have shown that the eigenfunctions of (a matrix associated with) the P-structure provide a base for the Fourier decomposition of arbitrary recombination landscapes.
Here we generalize this framework to include the effect of genotype frequencies, assuming linkage equilibrium. We find that the autocorrelation on the eigenfunctions of the population weighted P-structure is independent of the population composition. As a consequence we can directly compare the performance of mutation and recombination operators by comparing the auto-correlations on the finite set of elementary landscapes. This comparison predicts that point mutation is a superior search strategy on landscapes with low and moderate order of interaction $p < n/3$ (n is the number of loci). For more complex landscapes one point recombination is superior to both mutation and uniform recombination, but only if the distance among the interacting loci (defining length) is minimal.

Furthermore we find that the autocorrelation on any landscape is increasing as the distribution of genotypes becomes more extreme, i.e. if the population occupies a location close to the boundary of the frequency simplex. Landscapes are smoother the more biased the distribution of genotype frequencies is. We suggest that this result explains the paradox that there is little epistatic interaction for quantitative traits detected in natural populations if one uses variance decomposition methods while there is evidence for strong interactions in molecular mapping studies for quantitative trait loci.

Key words: Fitness Landscapes, Recombination, P-Structures, Autocorrelation, Epistatic Interactions

1. Introduction

Evolutionary change is caused by the spontaneously generated genetic variation and its subsequent fixation by drift and/or selection. Consequently, the main focus of evolutionary theory has been to understand the genetic structure and dynamics of populations, see e.g. [41]. This objective of neo-Darwinian theory influenced the mathematical face of evolutionary theory. It mainly consists of mathematical statistics, necessary to understand the genetic structure of populations and the stochastic forces acting on it, and dynamical systems theory (see e.g. [26]), which is used to model the deterministic forces, such as natural selection, acting on the population. In recent years, however, alternative approaches have gained increasing prominence in evolutionary theory. This development has been stimulated to some extent by the application of evolutionary models to design so-called *evolutionary algorithms* as for instance Genetic Algorithms, Evolution Strategies, and Genetic Programming, as well as the theory of Complex Adaptive Systems [27, 34, 17]. In these areas mathematical problems arose that are beyond the reach of traditional evolutionary models. Here we will deal with the question how the adaptive landscape approach can be made useful for understanding the performance of recombination as a search strategy in complex adaptations.

Given the bi-phasic nature of the evolution process, in which the genetic variation is generated independently from the natural selection acting on it, the generic structure of an evolutionary model can be written as

$$x' = S\left(x, \mathsf{w}\right) \circ T\left(x, \mathsf{t}\right) \ , \tag{1}$$

where x is e.g. the vector of haplotype frequencies, and $S(x, \mathsf{w})$ is a term describing the selection forces acting on x. The parameters w form the so-called fitness function, since they can be regarded as a mapping from the set of types into the real numbers. The second term, $T(x,\mathrm{t})$ describes the transmission processes by determining the probability of transforming one type into another one by mutation and/or recombination [2]. Hence, evolution models can be seen as dynamical systems of genotype frequencies which live on an algebraic structure [38] determined by the genetic processes like mutation and recombination.

If one considers just a limited number of types or some highly symmetrical cases the behavior of the system can be understood by the selection forces plus perturbations caused by the transmission processes (see [9] for a review of mutation-selection theory). The dynamical systems approach, however, turned out to be less informative if applied to systems containing a large number of types and/or highly irregular fitness functions. These situations occur for instance in practical application of mutation/recombination selection algorithms or if one wants to understand the evolution of complex adaptations. In such cases the algebraic structure on which the dynamical system is realized becomes more important for an understanding of the the system's behavior than the dynamical equations themselves. The reason is that the global behavior of a genetic algorithm, for instance, is mostly determined by the accessibility of superior genotypes by mutation and recombination from those that are already realized in the population. Accessibility of a genotype is largely determined by the mutation/recombination steps necessary to reach the genotype and whether the intermediary genotypes have higher fitness or not.

Metaphorically, the dynamics of a mutation-selection process can therefore be seen as a walk on a landscape, where uphill moves are preferred. The ability to evolve certain genotypes of organisms is influenced by the topological features of the fitness landscape. This realization led to what is now called *landscape theory.* The idea of landscape models has several roots. In evolutionary theory it can be traced back to Wright's ideas about adaptive landscapes (see [44, pp. 304-317], and became important in theories of molecular evolution and the origin of life [14, 15, 18, 20, 31, 35, 43, 47, 48] and in evolutionary computer science [32, 33]. Similar developments exist in physics, where free energy landscapes of disordered systems such as spin glasses are considered [39], and in search theory [51]. The main challenge to landscape theory is to determine which features of the fitness landscape determine the evolvability of the systems on the landscape.

A landscape results from the combination of three elements. A set of genotypes V, a set of genetic operators like mutation or recombination which act on the genotypes to transform them into other genotypes, and a fitness function. Genetic operators induce a topological structure on V which is called the configuration space. A fitness function imposed on a configuration space is then a landscape in which terms such as "peak", "valley", "ridge", etc. can be defined. A simple example of a configuration space is the set of strings of characters, like the nucleotides A, G, C, and T, together with point mutations. In this case, the configuration space can be represented as a graph, where edges connect all those genotypes that can be converted into each other by replacing one character by another. The resulting graph is known as Hamming graph or point mutation space. The ease with which the point mutation space can be constructed comes from the fact that the mutation operator has only one argument and the set of genotypes accessible by a single mutation defines a topological neighborhood. Furthermore the resulting topology can be represented as a simple graph.

In biology and genetic algorithm theory, the landscape concept can only be useful if it can include recombination. There are two difficulties in achieving this. One is that recombination is an operator with two arguments, and the other is that the probability of a particular recombination event depends on the frequency of matings between the two types.

The most immediate consequence of the fact that recombination acts on two arguments is that the recombination induced configuration space can not be represented as a simple graph with the set of genotypes representing the set of vertices. These leaves two alternatives: One can change the nature of the vertex set and have pairs of types as vertices. Then one obtains again a (di-)graph, since each elementary recombination event creates up to two different strings. This approach was pioneered by Culberson [13] and Jones [33]. The alternative is to leave the vertices to represent individual genotypes and to make the edges more complex. In Gitchoff and Wagner [21] it was shown that recombination spaces can be represented as hypergraphs[†], where the hyper-edges are the sets of all recombinants that can arise from the recombination of two types. With this approach it is was easy to show that string recombination spaces and point mutations spaces are homomorphic. Hypergraphs are still not completely satisfactory, since they do not indicate which pair of types produces which set of recombinants, i.e., which hyper-edge arises from which mating. This led us to invent *P-structures*, which are mappings of pairs of types to the hyper-edges of the hypergraph [54, 59]. Using this concept it was possible to derive a Fourier decomposition theory for recombination landscapes, generalizing an approach that was found useful in the case of mutation landscapes. The theory of P-structures, however, so far has neglected the fact that the probability of recombination events depends on the frequency of the recombination partners in a population. In other words, the structure of the configuration space that is induced by a realistic recombination operator will be population dependent. This problem will be addressed here.

[†]A *hypergraph* [5] consists of a vertex set V and a collection $\mathcal{E}$ of (not necessarily) distinct subsets of V which are called (hyper)edges.

2. Cross-over, Recombination, and P-Structures

Throughout this contribution we consider a genome consisting of n loci. For each locus k, there are α_k alleles. The set of all the $\prod_k \alpha_k$ possible genotype will be denoted by V. For each locus k, we label the alleles using a letter from the alphabet $\mathcal{A}_k = \{0, \ldots, \alpha_k - 1\}$. Thus $V = \prod_k \mathcal{A}_k$. A particular genotype (or sequence) $x \in V$ can be regarded as a vector with components $x_k \in \mathcal{A}_k$.

Definition 1. *A* cross-over operator *is a map* $\chi : V \times V \to V \times V$ *with the following property: Suppose* $\chi(y,z) = (u,v)$. *Then for each* k *either* $y_k = u_k \wedge z_k = v_k$ *or* $z_k = u_k \wedge y_k = v_k$. *By abuse of notation we write* $x \in \chi(y,z)$ *if* $x = u$ *or* $x = v$, *i.e., if* x *is an* offspring *of* (y,z).

As an immediate consequence we see that $\chi(x,x) = (x,x)$. We follow here the spirit of [33] when we regard a crossover operator as producing pairs of sequences rather than a single sequence from a pair of "ancestors".

We write $\boldsymbol{\chi} = \{k | x_k = u_k \wedge y_k = v_k\}$ and $\bar{\boldsymbol{\chi}} = \{k | y_k = u_k \wedge x_k = v_k\}$ for the two subsets of loci (sequence positions) that are *separated* by the cross-over operator χ. There is of course a one-to-one correspondence between a cross-over operator $\chi : V \times V \to V \times V$ and the associated set $\boldsymbol{\chi} \subseteq \{1, \ldots, n\}$: the set lists exactly those loci that are inherited from the first parent x by the first child u. Analogously, $\bar{\boldsymbol{\chi}}$ is the list of loci that the first child u inherits from the second parent y. Of course, $\bar{\boldsymbol{\chi}} = \{1, \ldots, n\} \setminus \boldsymbol{\chi}$.

The offsprings x and $\bar{x}$ of the two parents y and z have therefore the following representations:

$$x_k = \begin{cases} y_k & \text{if} \quad k \in \boldsymbol{\chi} \\ z_k & \text{if} \quad k \in \bar{\boldsymbol{\chi}} \end{cases} \qquad \bar{x}_k = \begin{cases} y_k & \text{if} \quad k \in \bar{\boldsymbol{\chi}} \\ z_k & \text{if} \quad k \in \boldsymbol{\chi} \end{cases} \tag{2}$$

A *recombination operator* in the sense of much of the GA literature is a family $\mathcal{F}$ of cross-over operators that act on $V \times V$ with probability $\pi(\chi)$.

Definition 2. *Let* V *be a finite set with power set* $\mathcal{P}(V)$. *A P-structure*[‡] *is a pair* $(V, \mathcal{R})$ *where* $\mathcal{R} : V \times V \to \mathcal{P}(V)$. *We say that the P-structure is symmetric if* $\mathcal{R}(x,y) = \mathcal{R}(y,x)$ *for all* $x, y \in V$. *In a weighted P-structure we attach a positive weight* $\mathbf{H}_{x,(y,z)}$ *to each triple* (x,y,z) *for which* $x \in \mathcal{R}(y,z)$ *and we set* $\mathbf{H}_{x,(y,z)} = 0$ *if* $x \notin \mathcal{R}(y,z)$. *We call* $\mathbf{H}$ *the incidence matrix of the P-structure.*

[‡]The term "P-structure" has been chosen in [54] because its image set is the **p**ower set of V.

In particular, there is a weighted P-structure associated in a natural way with each cross-over operator χ:

$$\begin{aligned}\mathcal{R}^{\chi}(y,z) &= \{x \in V | x \in \chi(y,z)\} \\ \mathbf{H}^{\chi}_{x,(y,z)} &= \begin{cases} 2 & \text{if} \quad x = y = z \\ 1 & \text{if} \quad x \in \chi(y,z) \qquad \text{and } y \neq z \\ 0 & \qquad \text{otherwise}\,. \end{cases}\end{aligned} \tag{3}$$

We observe that $\mathbf{H}^{\chi}_{x,(y,z)} > 0$ if and only if x is an offspring of (y,z). The doubled weight in the "diagonal", $\mathbf{H}^{\chi}_{x,(x,x)} = 2$, is mostly a technical convenience: It implies immediately $\sum_x \mathbf{H}_{x,(y,z)} = 2$, since any crossover operator produces exactly 2 offsprings from a pair of parents. If $y = z$, we simply count the offspring $y = z$ twice. The following technical detail turns out the be one of the most useful tools in this contribution:

Lemma 1. *The incidence matrix associated with the cross-over operator χ has the position-wise representation*

$$\mathbf{H}^{\chi}_{x,(y,z)} = \prod_{k\in\boldsymbol{\chi}} \delta_{x_k,z_k} \prod_{k\in\bar{\boldsymbol{\chi}}} \delta_{x_k,y_k} + \prod_{k\in\boldsymbol{\chi}} \delta_{x_k,y_k} \prod_{k\in\bar{\boldsymbol{\chi}}} \delta_{x_k,z_k}\,. \tag{4}$$

Proof. Clearly equ.(4) implies $\mathbf{H}^{\chi}_{x,(x,x)} = 2$. If $x \neq y$, then at most one of the two products is non-zero. This is the case if and only if x coincides with y on the subset $\boldsymbol{\chi}$ of loci while it matches z on $\bar{\boldsymbol{\chi}}$, or *vice versa*. □

Definition 3. *The weighted P-structure associated with a recombination operator $\mathcal{F}$ is*

$$\begin{aligned}\mathbf{H} &= \sum_{\chi\in\mathcal{F}} \pi(\chi)\mathbf{H}^{\chi} \\ \mathcal{R}(y,z) &= \bigcup_{\chi\in\mathcal{F}} \mathcal{R}^{\chi}(y,z) = \big\{x \in V \big| \exists \chi \in \mathcal{F} \,:\, x \in \chi(y,z)\big\}\end{aligned} \tag{5}$$

The interpretation of this definition is straight forward: $\mathbf{H}_{x,(y,z)}$ is the chance that x is an offspring of the parents y and z under $\mathcal{F}$-recombination [59].

The *recombination hypergraph* imag$\mathcal{R}$ has vertex set V and hyper-edges $\mathcal{R}(y,z)$, $y,z \in V$. Gitchoff and Wagner [21] introduced the following axioms to describe the action of recombination in terms of what we call here P-structures:

Definition 4. *A P-structure $(V,\mathcal{R})$ is a recombination structure if for all $x,y,z \in V$ holds:*

(i) $\mathcal{R}(x,x) = \{x\}$.

(ii) $\mathcal{R}(x,y) = \mathcal{R}(y,x)$.
(iii) $\{x,y\} \subseteq \mathcal{R}(x,y)$.
(iv) *For all* $z \in \mathcal{R}(x,y)$ *holds* $|\mathcal{R}(x,z)| \le |\mathcal{R}(x,y)|$.

In [54, Lemma C2] we showed that any recombination operator $\mathcal{F}$ on $\prod_k \mathcal{A}_k$ forms a recombination structure if and only if $\imath \in \mathcal{F}$, where $\iota : V\times V \to V\times V$ is the identity map: $\iota(y,z) = (y,z)$ for all $y,z \in V$. The characteristic set belonging to ι is $\iota = \{1,\dots,n\}$, $\bar{\iota} = \emptyset$. Condition (iii) is hence equivalent to requesting that offsprings can be produced without any crossover at all. We will therefore sometimes drop this requirement.

Before we proceed, we introduce the two most important recombination operators:

[∞] Uniform recombination contains all 2^n possible crossover operators. In this case it is natural to include the identity $\imath$.

[1] 1-point recombination contains all cross-over operators χ for which the characteristic set is of the form $\chi = \{1,\dots,k\}$. In this case we obtain more elegant results if the identity is excluded.

3. Population Weighted Cross-over

A *cross-over walk* [28, 29] on V is the Markov process based on the following rule: The "father" y is mated with a randomly chosen "mother" z. The offsprings are the "son" x and the "daughter" $\bar{x}$. The "son" x becomes the "father" of the next mating. We regard the sequence of "fathers" as a random walk on V. It is straightforward to derive the transition matrix of this Markov process:

$$\mathbf{S}_{xy} = \frac{1}{2}\sum_{z\in V} \mathbf{H}_{x,(y,z)}\wp(z) \tag{6}$$

The factor $1/2$ stems from the fact that the offspring x is the "son" and not the "daughter" of the parents y and z with probability $1/2$. By $\wp(z)$ we denote the probability that z is the "mother" of the mating, i.e., $\wp(z)$ is the frequency of genotype z in the *population* $\vec{\wp}$ under random mating. In previous studies [54, 59] we have always assumed that $\wp(z) = 1/|V|$, i.e., a uniform population. We use the notion $\mathbf{S}$ or $\mathbf{S}^{\mathcal{F}}$ for the case of a uniform populations. For the general case we write

$$\mathbf{S}^{\mathcal{F},\vec{\wp}}_{xy} = \sum_{\chi\in\mathcal{F}} \pi(\chi)\frac{1}{2}\sum_{z\in V} \mathbf{H}^{\chi}_{x,(y,z)}\wp(z) = \sum_{\chi\in\mathcal{F}} \pi(\chi)\mathbf{S}^{\chi,\vec{\wp}}_{xy} \tag{7}$$

where $\mathbf{S}^{\chi,\vec{\wp}}$ is the *population weighted* transition operator associated with the the crossover operator χ. It is easy to check by direct computation that $\mathbf{S}^{\chi,\vec{\wp}}$ is a stochastic matrix for any choice of the cross-over operator χ and the population vector $\vec{\wp}$. As an immediate consequence, $\mathbf{S}^{\mathcal{F},\vec{\wp}}$ is also a stochastic matrix for any family $\mathcal{F}$ of recombination operators.

Random walks with transition matrix $\mathbf{S}$ were used in [28, 29] in order to generate the "times-series" for a numerical correlation analysis of Nk landscapes under recombination. The fitness values $\{f(x_0), f(x_1), \dots\}$ collected along such a walk forms a "time-series" that generalizes Weinberger's [60, 61] procedure for mutation spaces [54, 59]. Decompositions of fitness landscapes with respect to recombination walk operators are reported in [30] for models derived from RNA folding and evolving CAs.

In this contribution we shall restrict ourselves to population vectors on the *Wright manifold*

$$\mathcal{W} = \left\{ \vec{\wp} \middle| \wp(z) = \prod_k p_k(z_k) \right\} \tag{8}$$

where $p_k(a)$ denotes the frequency of allele a at locus k. The starting point of our investigation is the following explicit representation of $\mathbf{S}^{\chi,\vec{\wp}}$ for $\vec{\wp} \in \mathcal{W}$:

$$2\mathbf{S}^{\chi,\vec{\wp}}_{xy} = \prod_{k\in\boldsymbol{\chi}} p_k(x_k) \prod_{k\in\bar{\boldsymbol{\chi}}} \delta_{x_k,y_k} + \prod_{k\in\bar{\boldsymbol{\chi}}} p_k(x_k) \prod_{k\in\boldsymbol{\chi}} \delta_{x_k,y_k} \tag{8}$$

which is easily verified.

Theorem 1. *Let $\vec{\wp} \in \mathcal{W}$. Then $\vec{\wp} = \mathbf{S}^{\mathcal{F},\vec{\wp}}\vec{\wp}$ for any family $\mathcal{F}$ of cross-over operators.*

Proof. We proof that $\mathbf{S}^{\chi,\vec{\wp}}\vec{\wp} = \vec{\wp}$ holds for all cross-over operators χ. The theorem then follows from $\mathbf{S}^{\mathcal{F},\vec{\wp}} = \sum_{\chi\in\mathcal{F}} \pi(\chi)\mathbf{S}^{\chi,\vec{\wp}}$.
We have explicitly

$$\begin{aligned} 2\sum_y \mathbf{S}^{\chi,\vec{\wp}}_{xy}\wp(y) &= \prod_{k\in\boldsymbol{\chi}} p_k(x_k) \underbrace{\Bigl(\sum_{y_k} p_k(y_k)\Bigr)}_{1} \prod_{k\in\bar{\boldsymbol{\chi}}} \underbrace{\Bigl(\sum_{y_k} \delta_{x_k,y_k} p_k(y_k)\Bigr)}_{p_k(x_k)} \\ &+ \prod_{k\in\bar{\boldsymbol{\chi}}} p_k(x_k) \underbrace{\Bigl(\sum_{y_k} p_k(y_k)\Bigr)}_{1} \prod_{k\in\boldsymbol{\chi}} \underbrace{\Bigl(\sum_{y_k} \delta_{x_k,y_k} p_k(y_k)\Bigr)}_{p_k(x_k)} \\ &= \prod_{k\in\boldsymbol{\chi}} p_k(x_k) \prod_{k\in\bar{\boldsymbol{\chi}}} p_k(x_k) + \prod_{k\in\bar{\boldsymbol{\chi}}} p_k(x_k) \prod_{k\in\boldsymbol{\chi}} p_k(x_k) \quad = 2\wp(x)\,. \end{aligned}$$

Thus the population vector $\vec{\wp}$ is the Perron-Frobenius eigenvector (stationary distribution) of all the transition matrices $\mathbf{S}^{\chi,\vec{\wp}}$. □

This result confirms the population genetic fact that linkage equilibrium is maintained under recombination. This fact was first proven for two alleles and arbitrary number of loci by Robbins [46] and for multiple alleles by Bennett [4].

COROLLARY 1. *If $\vec{\wp}$ is a uniform population, i.e., $p_k(x_k) = 1/\alpha_k$, then $\mathbf{S}^{\chi,\vec{\wp}}$ is symmetric for any cross-over operator χ.*

This result was obtained in [54] as a consequence of the symmetries of string recombination P-structures.

LEMMA 2. *$\mathbf{S}^{\mathcal{F},\vec{\wp}}_{xy}\wp(y) = \mathbf{S}^{\mathcal{F},\vec{\wp}}_{yx}\wp(x)$ for all $x, y \in V$ and $\vec{\wp} \in \mathcal{W}$.*

Proof. Direct substitution yields: $2\mathbf{S}^{\chi,\vec{\wp}}_{xy}\wp(y) =$

$$\prod_{k\in\boldsymbol{\chi}} p_k(x_k)p_k(y_k) \prod_{k\in\bar{\boldsymbol{\chi}}} \delta_{x_k,y_k}p_k(y_k) + \prod_{k\in\bar{\boldsymbol{\chi}}} p_k(x_k)p_k(y) \prod_{k\in\boldsymbol{\chi}} \delta_{x_k,y_k}p_k(y_k) =$$

$$\prod_{k\in\boldsymbol{\chi}} p_k(y_k)p_k(x_k) \prod_{k\in\bar{\boldsymbol{\chi}}} \delta_{x_k,y_k}p_k(x_k) + \prod_{k\in\bar{\boldsymbol{\chi}}} p_k(y_k)p_k(x_k) \prod_{k\in\boldsymbol{\chi}} \delta_{x_k,y_k}p_k(x_k)$$

$$= 2\mathbf{S}^{\chi,\vec{\wp}}_{yx}\wp(x).$$ □

The fact that $\mathbf{S}^{\mathcal{F},\vec{\wp}}$ is reversible (and thus self-adjoint) for any family of recombination operators considerably will simplify the analysis of the autocorrelation functions associated with $\mathbf{S}^{\mathcal{F},\vec{\wp}}$-walks. In the following section we consider the spectral properties of $\mathbf{S}^{\mathcal{F},\vec{\wp}}$ in some detail.

4. GENERALIZED WALSH FUNCTIONS

The following notation will be convenient for constructing eigenfunctions of the transition operators $\mathbf{S}^{\chi,\vec{\wp}}$: For each $x_k \in \mathcal{A}_k = \{0, \ldots, \alpha_k - 1\}$ we set

$$\hat{x}_k = \exp(2\pi\imath\, x_k/\alpha_k) \tag{9}$$

Note that $\hat{x}_k$ is a root of unity. Furthermore, for $I \in \prod_k \mathcal{A}_k$ set $\tilde{I} = \{k | I_k \neq 0\}$. For each "index" I we define the *generalized Walsh function*

$$\varepsilon_I : \prod_k \mathcal{A}_k \to \mathbb{C} : \varepsilon_I(x) = \prod_k \hat{x}_k^{I_k} = \prod_{k\in\tilde{I}} \hat{x}_k^{I_k} \tag{10}$$

We remark that $\{\varepsilon_I | I \in V\}$ is the standard Fourier basis of the Abelian group $\prod Z_{\alpha_k}$. The functions ε_I are of course the characters of $\prod Z_{\alpha_k}$ [37, 49]. These functions are also eigenvectors of the adjacency matrix of the Hamming graphs $\prod \mathcal{Q}_{\alpha_k}$, see e.g., [51]. Note that the formal association of the index sets I with the vertices in V is a mere book-keeping device that does not readily lend itself to a meaningful interpretation.

In the following we repeatedly make use of the identity

$$\sum_{x_k=0}^{\alpha_k-1} \hat{x}_k^{I_k} = \begin{cases} 0 & \text{if} \quad I_k \neq 0 \\ \alpha_k & \text{if} \quad I_k = 0 \end{cases} \tag{11}$$

For the population-weighted case we set

$$\psi_I^{\vec{\wp}}(x) = \prod_{k \in \tilde{I}} \frac{1}{p_k(x_k)} \hat{x}_k^{I_k} \tag{12}$$

The main technical result of this paper is the following, rather surprising

Theorem 2. *Let $\vec{\wp} \in \mathcal{W}$. Then the population weighted generalized Walsh function $\psi_I^{\vec{\wp}}$ is a left eigenvector of $\mathbf{S}^{\chi,\vec{\wp}}$ with eigenvalue*

$$\lambda_I^{\chi} = \begin{cases} 1 & \textit{if} \qquad \tilde{I} = \emptyset \\ 1/2 & \textit{if} \quad \emptyset \neq \tilde{I} \subseteq \boldsymbol{\chi} \ \textit{or}\ \emptyset \neq \tilde{I} \subseteq \bar{\boldsymbol{\chi}} \\ 0 & \qquad \textit{otherwise} \end{cases} \tag{13}$$

Proof. We have $2 \sum_{x \in V} \psi_I^{\vec{\wp}}(x) \mathbf{S}_{xy}^{\chi,\vec{\wp}} =$

$$\prod_{k \in \boldsymbol{\chi} \cap \tilde{I}} \Big(\sum_{x_k} \frac{p_k(x_k)}{p_k(x_k)} \hat{x}_k^{I_k} \Big) \prod_{k \in \boldsymbol{\chi} \setminus \tilde{I}} \underbrace{\Big(\sum_{x_k} p_k(x_k) \Big)}_{1} \prod_{k \in \bar{\boldsymbol{\chi}} \cap \tilde{I}} \Big(\sum_{x_k} \frac{\delta_{x_k y_k}}{p_k(x_k)} \hat{x}_k^{I_k} \Big) \prod_{k \in \bar{\boldsymbol{\chi}} \setminus \tilde{I}} \underbrace{\Big(\sum_{x_k} \delta_{x_k y_k} \Big)}_{1} +$$

$$\prod_{k \in \bar{\boldsymbol{\chi}} \cap \tilde{I}} \Big(\sum_{x_k} \frac{p_k(x_k)}{p_k(x_k)} \hat{x}_k^{I_k} \Big) \prod_{k \in \bar{\boldsymbol{\chi}} \setminus \tilde{I}} \underbrace{\Big(\sum_{x_k} p_k(x_k) \Big)}_{1} \prod_{k \in \boldsymbol{\chi} \cap \tilde{I}} \Big(\sum_{x_k} \frac{\delta_{x_k y_k}}{p_k(x_k)} \hat{x}_k^{I_k} \Big) \prod_{k \in \boldsymbol{\chi} \setminus \tilde{I}} \underbrace{\Big(\sum_{x_k} \delta_{x_k y_k} \Big)}_{1}$$

$$= \prod_{k \in \boldsymbol{\chi} \cap \tilde{I}} \Big(\underbrace{\sum_{x_k} \hat{x}_k^{I_k}}_{\substack{0 \text{ unless} \\ \boldsymbol{\chi} \cap \tilde{I} = \emptyset}} \Big) \prod_{k \in \bar{\boldsymbol{\chi}} \cap \tilde{I}} \frac{1}{p_k(y_k)} \hat{y}_k^{I_k} + \prod_{k \in \bar{\boldsymbol{\chi}} \cap \tilde{I}} \Big(\underbrace{\sum_{x_k} \hat{x}_k^{I_k}}_{\substack{0 \text{ unless} \\ \bar{\boldsymbol{\chi}} \cap \tilde{I} = \emptyset}} \Big) \prod_{k \in \boldsymbol{\chi} \cap \tilde{I}} \frac{1}{p_k(y_k)} \hat{y}_k^{I_k}$$

We have to consider three cases now:

(i) If $\tilde{I} = \emptyset$, then $\boldsymbol{\chi} \cap \tilde{I} = \bar{\boldsymbol{\chi}} \cap \tilde{I} = \emptyset$ and we have $2 \sum_{x \in V} \psi_{\emptyset}^{\vec{\wp}}(x) \mathbf{S}_{xy}^{\chi,\vec{\wp}} = 2$. Noting that $\psi_{\emptyset}^{\vec{\wp}}(x) = 1$ for all x, we observe that $\psi_{\emptyset}^{\vec{\wp}}(x)$ is an eigenvector $\mathbf{S}_{xy}^{\chi,\vec{\wp}}$ with eigenvalue $\lambda_I^{\chi} = 1$.

(ii) Suppose $\tilde{I} \cap \boldsymbol{\chi} \neq \emptyset$ and $\tilde{I} \cap \bar{\boldsymbol{\chi}} \neq \emptyset$. Then the products over terms of the form $\sum_{x_k} \hat{x}_k^{I_k} = 0$ are non-empty and we find that $\psi_I^{\vec{\wp}}$ is an eigenvector of $\mathbf{S}_{xy}^{\chi,\vec{\wp}}$ with eigenvalue $\lambda_I^{\chi} = 0$.

(iii) If $\tilde{I} \neq \emptyset$ but $\tilde{I} \cap \boldsymbol{\chi} = \emptyset$, then $\tilde{I} \subseteq \bar{\boldsymbol{\chi}}$ and hence

$$2 \sum_{x \in V} \psi_I^{\vec{\wp}}(x) \mathbf{S}_{xy}^{\chi,\vec{\wp}} = \prod_{k \in \tilde{I}} \frac{1}{p_k(y_k)} \hat{y}_k^{I_k} = \psi_I^{\vec{\wp}}(y)$$

Thus $\psi_I^{\vec{\wp}}$ is an eigenvector of $\mathbf{S}_{xy}^{\chi,\vec{\wp}}$ with eigenvalue $\lambda_I^{\chi} = 1/2$. The same results is obtained for $\emptyset \neq \tilde{I} \cap \bar{\boldsymbol{\chi}} = \emptyset$.

Hence $\psi_I^{\vec{\wp}}$ is an eigenvector of $\mathbf{S}^{\chi,\vec{\wp}}$ in each case. For non-constant landscapes the eigenvalue is 0 or 1/2 depending on whether or not χ separates the interacting loci I. □

Corollary 2. *The Walsh functions ε_I are eigenfunctions of $\mathbf{S}^{\chi}$ in the case of a uniform population for all χ and hence also of $\mathbf{S}^{\mathcal{F}}$ for any string recombination operator $\mathcal{F}$.*

Definition 5. *The* population weighted scalar product *of two landscapes f and g on V is*

$$\langle f, g \rangle_{\wp} := \sum_{x \in V} p(x) f(x) g(x)^* \,, \tag{14}$$

where the star denotes complex conjugation.

Lemma 3. *Suppose $\tilde{I} \neq \tilde{J}$. Then $\langle \psi_I^{\vec{\wp}}, \psi_J^{\vec{\wp}} \rangle_{\wp} = 0$.*

Proof. We have

$$\begin{aligned}\langle \psi_I^{\vec{\wp}}, \psi_J^{\vec{\wp}} \rangle_{\wp} = & \prod_{k \notin (\tilde{I} \cup \tilde{J})} \underbrace{\left(\sum_{x_k} p_k(x_k) \right)}_{1} \prod_{k \in \tilde{I} \setminus \tilde{J}} \underbrace{\left(\sum_{x_k} \frac{p_k(x_k)}{p_k(x_k)} \hat{x}_k^{I_k} \right)}_{0} \\ & \times \prod_{k \in \tilde{J} \setminus \tilde{I}} \underbrace{\left(\sum_{x_k} \frac{p_k(x_k)}{p_k(x_k)} \hat{x}_k^{-J_k} \right)}_{0} \prod_{k \in (\tilde{J} \cap \tilde{I})} \left(\sum_{x_k} \frac{p_k(x_k)}{p_k^2(x_k)} \hat{x}_k^{I_k - J_k} \right)\end{aligned}$$

We have used here that the only non-real terms arise in powers of $\hat{x}_k$, and $(\hat{x}_k^{I_k})^* = \hat{x}_k^{-I_k}$. Clearly, this product vanishes unless $\tilde{I} \setminus \tilde{J} = \tilde{J} \setminus \tilde{I} = \emptyset$. □

Note that for two alleles at each locus, we have $\langle \psi_I^{\vec{\wp}}, \psi_J^{\vec{\wp}} \rangle_{\wp} = 0$ whenever $I \neq J$. This is **not** true in general for larger sets of alleles.

On the other hand, $\mathbf{S}^{\mathcal{F}}$ is self-adjoint w.r.t. the scalar product $\langle\, . \,; \, . \,\rangle_{\wp}$. Thus, there is a basis of eigenfunctions that are orthonormal w.r.t. this scalar product. It is obtained from the $\{\psi_I^{\vec{\wp}}\}$ by separately orthogonalizing and normalizing those basis functions that belong to a given index set $\tilde{I}$.

Since the eigenvalue to which $\psi_I^{\vec{\wp}}$ belongs, depends only on $\tilde{I}$ but not on I itself, we can ignore this complication in the following. It will be sufficient to know that such an ONB exists.

As a first application, we explicitly derive the eigenvalues belonging to ψ_I for the two most prominent families of cross-over operators: uniform recombination and 1-point recombination. In the context of 1-point recombination the *defining length* $\ell(\tilde{I}) = \max(\tilde{I}) - \min(\tilde{I}) + 1$ of the interaction scheme I in the eigenvector $\psi_I^{\vec{\wp}}$ plays a crucial role.

COROLLARY 3. *For the family of 1-point cross-over operators we obtain*

$$\lambda_I^{[1]} = \frac{1}{2}\frac{n-\ell(\tilde{I})}{n-1} \quad \text{if } \tilde{I} \neq \emptyset \qquad \text{and} \qquad \lambda_{\emptyset}^{[1]} = 1\,. \tag{15}$$

Hence a landscape f is elementary under 1-point cross-over if and only if its Fourier series representation contains only terms with a single defining length ℓ.

Proof. χ_r does not disrupt I if either $r \le \min\{k \in \tilde{I}\}$ or if $r > \max\{k \in \tilde{I}\}$. Thus there are $\ell(\tilde{I}) - 1$ crossover operators in $\mathcal{F}_{[1]}$ that disrupt the interaction pattern $\tilde{I}$ and $n - \ell(\tilde{I})$ ones that don't. The desired expression follows now directly from theorem 2. □

COROLLARY 4. *For the family of uniform cross-over operators we obtain*

$$\lambda_I^{[\infty]} = 2^{-|\tilde{I}|}\,. \tag{16}$$

Hence a landscape is elementary under uniform cross-over if and only if its Fourier series representation contains only terms with a single interaction order $p = |\tilde{I}|$.

Proof. The set of all uniform crossover operators can be identified with the set of all 01-strings of length n such that bit r is 1 if $r \in \boldsymbol{\chi}$ and 0 otherwise. The $\lambda_I^{\chi} = 1/2$ if and only if the string representing χ is either 1 or 0 on all $s \in I$. On the remaining bits this string is arbitrary, hence we have $2 \times 2^{n-|\tilde{I}|}$ choices of χ that do not separate I, including the two strings that code for the identity and for the exchange of all positions between the two parents. □

Theorem 2 implies that the main effect of a non-uniform population structure leads to a deformation of the eigenspaces of $\mathbf{S}^{\mathcal{F}}$. This effect is conveniently quantified by the "angles"

$$\begin{aligned}\langle \psi_I^{\vec{\wp}}; \varepsilon_J \rangle_{\wp} = & \prod_{k \notin (\tilde{I} \cup \tilde{J})} \underbrace{\left(\sum_{x_k} p_k(x_k) \right)}_{1} \prod_{k \in \tilde{I} \cap \tilde{J}} \underbrace{\left(\sum_{x_k} \hat{x}_k^{I_k} \hat{x}_k^{-J_k} \right)}_{\substack{0 \text{ if } I_k \neq J_k \\ \alpha_k \text{ if } I_k = J_k}} \\ & \times \prod_{k \in \tilde{I} \setminus \tilde{J}} \underbrace{\left(\sum_{x_k} \hat{x}_k^{I_k} \right)}_{0} \prod_{k \in \tilde{J} \setminus \tilde{I}} \underbrace{\left(\sum_{x_k} p_k(x_k) \hat{x}_k^{-J_k} \right)}_{G_k}\end{aligned} \tag{17}$$

A close inspection of this expression shows that $\langle \psi_I; \varepsilon_J \rangle_{\wp} = 0$ unless $\tilde{I} \subseteq \tilde{J}$ and $I_k = J_k$ for all $k \in \tilde{I} \cap \tilde{J}$. By abuse of notation we shall denote this situation by $I \subseteq J$ and call I a *sub-pattern* of the *interaction pattern* J. The quantity $G_K(\vec{\wp}) = \prod_{k \in K} G_k$ measures the deviations from uniformity of the population structure $\vec{\wp}$ restricted to the subset K of loci.

Thus, if we are given a Walsh expansion of a fitness function to which only a certain collection $\mathcal{J}$ of interaction patterns I contribute, $f(x) = \sum_{I \in \mathcal{J}} a_I \varepsilon_I(x)$, then in non-uniform population only the sub-patterns $J \subseteq I \in \mathcal{J}$ contribute to the population-weighted Fourier decomposition.

This observation has an important interpretation in terms of the perceived ruggedness of a landscape, as we shall see at the end of the following section.

5. Autocorrelation Functions

In this section we consider the autocorrelation function $r(t)$ of a time-reversible random walk on the finite set V. Let $\mathbf{S}$ be the transition matrix of such a Markov process and denote its stationary distribution (Perron-Frobenius eigenvector) by $\vec{\wp}$.

The statistics of such a random walk are captured taking expectation of the form

$$\mathcal{E}[g] := \sum_{x \in V} g(x) \wp(x) \tag{18}$$

It will be convenient to define $\tilde{f}(x) = f(x) - \mathcal{E}[f]$. The correlation function is then defined by

$$r(t) = \frac{1}{\mathcal{E}[\tilde{f}^2]} \sum_{y \in V} \tilde{f}(x) \mathbf{S}_{xy}^t \tilde{f}(y) \wp(y) = \frac{\langle \tilde{f}, \mathbf{S}^t \tilde{f} \rangle_{\wp}}{\langle \tilde{f}, \tilde{f} \rangle_{\wp}} = \mathcal{R}_{\mathbf{S}^t}[g] \tag{19}$$

where the *Rayleigh quotient* of a matrix $\mathbf{M} \in \mathbb{R}^{V \times V}$ is given by

$$\mathcal{R}_{\mathbf{M}}[g] := \frac{\langle g, \mathbf{M} g \rangle_{\wp}}{\langle g, g \rangle_{\wp}} \tag{20}$$

Definition 6. *We say a landscape $f : V \to \mathbb{R}$ is elementary w.r.t. the transition operator* $\mathbf{S}$*, if $\tilde{f}$ is an eigenfunction of* $\mathbf{S}$.

The notion of elementary landscapes was introduced in a series of papers about the geometry of fitness landscapes for mutation operators [51, 25, 52, 53] and recombination spaces [54, 59]. In these papers a Laplacian operator was used instead of the transition operator of a random walk. It is not difficult to verify, however, that the two definitions agree as long as the transition operator is bistochastic. In the more general cases considered here, the above definition seems to be more transparent.

Lemma 4. *A non-constant landscape f is elementary if and only if $r(t) = \lambda^t$ for some $\lambda < 1$.*

Proof. Follows directly from $\tilde{f}\mathbf{S} = \lambda\tilde{f}$ or $\mathbf{S}\tilde{f} = \lambda\tilde{f}$. □

If $\mathbf{S}$ is time-reversible, i.e., if $\mathbf{S}_{xy}\wp(y) = \mathbf{S}_{yx}\wp(x)$ for holds for the stationary distribution $\wp = \mathbf{S}\wp$, then $\mathbf{S}$ is self-adjoint and hence admits a basis of eigenvectors $\{\psi_k\}$ that are orthonormal w.r.t. the scalar product $\langle\, .\, ;\, .\, \rangle_\wp$ [6]. In this case we call an expansion of the form $f = \sum_k a_k\psi_k$ a *Fourier series* expansion of the landscape. The Fourier coefficients are given by

$$a_j = \langle f; \psi_j \rangle_\wp \tag{21}$$

For the auto-correlation function we have the following result which follows by direct computation:

Theorem 3. *Let f be an arbitrary landscape and suppose the transition operator* $\mathbf{S}$ *is self-adjoint with stationary distribution $\wp$. Then*

$$r(t) = \sum_{q>1} B_q \lambda_q^t\,, \tag{22}$$

where $1 = \lambda_1 > \lambda_2 > \cdots > \lambda_M$ are the distinct eigenvectors of f and the coefficients are given by

$$B_q = \frac{1}{\langle \tilde{f}; \tilde{f} \rangle_\wp} \sum_{j \in H(q)} |a_j|^2 \tag{23}$$

in terms of the Fourier coefficients a_j that belong to a common eigenvector λ_q of $\mathbf{S}$*. Here $H(q) = \{k | \mathbf{S}\psi_k = \lambda_q\psi_k\}$ denotes the set of indices of the eigenfunctions belonging to the eigenvalue λ_q.*

An analogous result for landscapes on regular graphs is proved in [50, 52].

In the previous section, equ.(17), we convinced ourselves that deviations from a uniform population distribution imply that the population-dependent

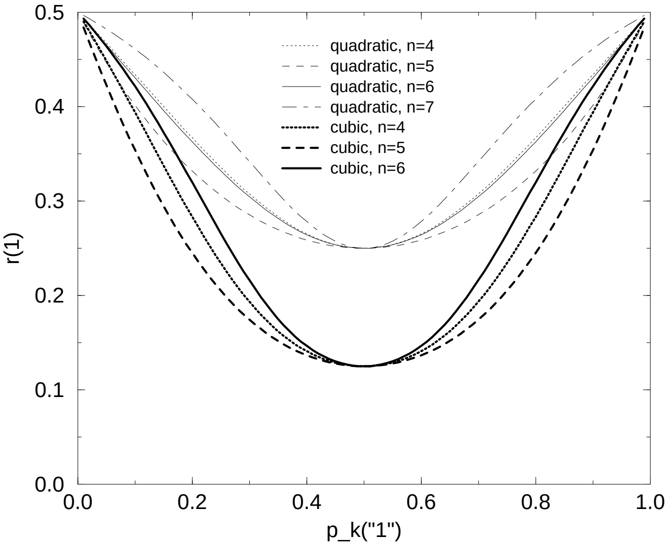

Figure 1. Dependence of the nearest neighbor correlation $r(1)$ on the population structure for quadratic and cubic landscapes. We vary $p_k('1') = p$ from almost 0 to almost 1. The pair-correlation approaches 0.5 for the extreme cases, indicating an effectively additive fitness function in this regime.

Fourier decomposition will involve not only the basis functions $\psi_I^{\vec{\wp}}$ corresponding to ε_I in the uniform case, but also contributions belonging to the sub-patterns $J \subset I$. Intuitively, this can be understood by visualizing that certain alleles in $I \setminus J$ are fixed with a non-zero probability (which is determined by the differences in allele frequencies at these loci in question), and hence effectively only the alleles at the other loci vary. The influence of these sub-patterns is determined by the "angles" $\langle \psi_J^{\vec{\wp}}; \varepsilon_I \rangle_{\wp}$. This quantity vanishes for a uniform population distribution. The last term in equ.(17) furthermore indicates that the influence of smaller sub-patterns increases with the differences G_k in the allele frequencies at a given locus. A numerical example for quadratic and cubic spin glasses (two alleles at each locus) is displayed in figure 1.

As a consequence the landscape becomes "less rugged" when seen from the point of view of a non-uniform population, at least if one considers uniform cross-over or 1-point crossover. The reason for this is, that clearly both the defining length ℓ and the order of an interaction "scheme" decreases when one goes from I to a subset of I.

This result sheds light on a phenomenon known from quantitative genetics. If one considers a pair of loci with AxA epistatic interactions the amount of additive variance is minimal when the two alleles at the two loci are of equal frequency. In contrast, the additive genetic variance increases towards to edges of the simplex as the allele frequencies get more extreme [10]. One can see the amount of additive variance as a measure of the "smoothness" of the genotype-phenotype function, since it determines the linear component of the regression of offspring phenotype on parent phenotype. Our result shows that the contribution of higher order interactions becomes added to lower order contributions if the frequencies of some of the chromosomes becomes small. The reason is that higher order interaction only make a contribution if all the genotypes representing the interaction are present in the population. Hence, with decreasing frequency of some of the genotypes, the effective degree of epistasis is decreasing and the level of additive and lower order effects is increasing.

6. Mutation Operators

Before we turn to the behavior of mutation operators in non-uniform populations we review their basic algebraic properties.

An elementary mutation event $y \to x$ is defined by the position k in which the strings x and y differ, and by the "kind" of mutation. In the case of a binary alphabet, there is only one "kind" of mutation, namely replacing 0 by 1 and *vice versa*. For more than 2 alleles, however, different mutants are possible. Mathematical tractability demands to specify the kind of mutation as the difference $x_k - y_k \mod \alpha_k = d \neq 0$. In particular, this will imply that the mutation rates do not depend on the individual alleles but only on the loci k and on the difference d in arbitrary circular "ordering" of the alphabet $\mathcal{A}_k$. While the k-dependence is an important effect in nature, it seems that the d-dependence is rather artificial.

We denote an elementary mutation by (k, d). The associated transition matrix is

$$\mathbf{M}_{xy}^{(k,d)} = \begin{cases} 1 & \text{if} \quad x_k - y_k \mod \alpha_k = d \\ & \quad \text{and } x_j = y_j, \forall\, j \neq k \\ 0 & \quad \text{otherwise} \end{cases} = \mu_{x_k,y_k}^{(k,d)} \prod_{j \neq k} \delta_{x_j,y_j} \,. \tag{24}$$

Here $\mu_{p,q}^{(k,d)} = 1$ if $p = q + d \mod \alpha_k$ and 0 otherwise. More complicated mutation operators can be built from these. For instance the product $\mathbf{M}^{(k,d_1)}\mathbf{M}^{(l,d_2)}$ introduces the simultaneous mutation $x_k = y_k + d_1 \mod \alpha_k$ and mutation $x_l = y_l + d_2 \mod \alpha_l$. A simple computation verifies $\mathbf{M}^{(k,d_1)}\mathbf{M}^{(k,d_2)} = \mathbf{M}^{(k,d_1+d_2)}$. Obviously, therefore, the elementary mutation operator commute. More complicated mutation operators are readily constructed as sums and products of elementary mutation operators. For instance, the point mutation operator

$$\tilde{\mathbf{M}} = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{\alpha_k - 1} \sum_{d=1}^{\alpha_k - 1} \mathbf{M}^{(k,d)} \tag{25}$$

introduces one mutation of arbitrary kind in an arbitrary position. Note that the matrix $n(\alpha-1)\tilde{\mathbf{M}}$ is the adjacency matrix of the Hamming graph $\mathcal{Q}^n_\alpha$, and $\tilde{\mathbf{M}}$ is the transition matrix of the simple random walk on $\mathcal{Q}^n_\alpha$ that was used in [60] to collect "time-series" from a landscape. The following theorem generalizes results from [52]:

Theorem 4. *The generalized Walsh functions ε_I are eigenfunctions of the elementary mutations operators:*

$$\mathbf{M}^{(k,d)}\varepsilon_I = \zeta_k^{d\cdot I_k}\varepsilon_I \tag{26}$$

where $\zeta_k = \exp(2\pi\imath/\alpha_k)$.

Proof. We have to evaluate

$$\begin{aligned}\sum_y \mathbf{M}^{(k,d)}_{xy}\varepsilon_I(y) &= \sum_{y_k=0}^{\alpha-1}\mu^{(k,d)}_{x_k,y_k}\hat{y}_k^{I_k}\prod_{j\neq k}\sum_{y_j=0}^{\alpha_k-1}\delta_{x_j,y_j}\hat{y}_j^{I_j}\\ &= \hat{x}_k^{I_k-d}\prod_{j\neq k}\hat{y}_j^{I_j} = \hat{\zeta}_k^{-dI_k}x_k^{I_k}\prod_{j\neq k}\hat{x}_j^{I_j}\end{aligned}$$

Substituting the definition of $\varepsilon_I(x)$ completes the proof. □

Corollary 5. *The generalized Walsh functions ε_I are eigenfunctions of all mutation operators.*

Proof. Follows by direct computation. □

Since the d-dependence has little physical background it seems reasonable to consider only mutation operators that are polynomials of the point-mutation operator

$$\mathbf{M}^{(k)} = \frac{1}{\alpha_k-1}\sum_{d=1}^{\alpha_k-1}\mathbf{M}^{(k,d)} \tag{27}$$

Corollary 6. *The eigenvalue $\lambda^{(k)}_I$ of the point mutation matrix $\mathbf{M}^{(k)}$ is*

$$\lambda^{(k)}_I = \begin{cases} 1 & \text{if} \quad k\notin\tilde{I}\\ -\frac{1}{\alpha_k-1} & \text{if} \quad k\in\tilde{I}\end{cases} \tag{28}$$

Proof. $\mathbf{M}^{(k)}\varepsilon_I = \dfrac{1}{\alpha_k-1}\displaystyle\sum_{d=1}^{\alpha_k-1}\zeta_k^{-dI_k}\varepsilon_I = \dfrac{1}{\alpha_k-1}\left(\sum_{d=0}^{\alpha_k-1}(\zeta_k^{-dI_k}-1)\right)\varepsilon_I$
$= \dfrac{\delta_{I_k,0}\alpha_k-1}{\alpha_k-1}\varepsilon_I.$ □

Corollary 7. *Each recombination matrix* $\mathbf{S}$ *(with a uniform population) commutes with each mutation matrix* $\mathbf{M}$*. Furthermore, the generalized Walsh function* ε_I *is an eigenvector of an arbitrary mutation-recombination operators for uniform populations.*

Proof. Two symmetric matrices commute if and only if the have a common basis of eigenvectors. □

This results will allow us to compare the correlation structure of a fitness function as seen by different operators in a straight forward way.

Let us now turn to the mutation operators in the case of an arbitrary population structure. Since mutation operators change the population structure (they are bi-stochastic, hence $(1,\dots,1)$ is the Perron Frobenius eigenvector, and hence the population exponentially converges to the uniform distribution), we only consider nearest-neighbor correlations. It is easy to check that ψ_I is not an eigenvector of a mutation operator unless the population is uniform.

It seems natural to compare ruggedness under mutation and recombination, respectively, by introducing a "population-weighted" autocorrelation function. The population dependence enters by drawing the starting point of the walk from the population $\vec{\wp}$ instead of uniformly. Hence we are interested in

$$r_{\mathbf{M}}(t) = \frac{1}{\langle \tilde{f}; \tilde{f}\rangle_{\wp}} \sum_x \tilde{f}(x) \sum_y (\mathbf{M}^t)_{xy} \tilde{f}(y) \wp(y) = \frac{\langle \tilde{f}; \mathbf{M}^t \tilde{f}\rangle_{\wp}}{\langle \tilde{f}; \tilde{f}\rangle_{\wp}} \tag{29}$$

It follows immediately that $r_{\mathbf{M}}(t) = \lambda_I^{\mathbf{M}}$ for any Walsh function ε_I, since $\varepsilon_I \mathbf{M} = \lambda_I^{\mathbf{M}} \varepsilon_I$. Thus, if f is elementary w.r.t. to the mutation operator $\mathbf{M}$, then the correlation function $r_{\mathbf{M}}$ is independent of the population structure. Recall that the landscapes become smoother for non-uniform population structures for 1-point and uniform recombination.

Converse, we are interested in the influence of the population structure on the mutational ruggedness of landscapes that are elementary w.r.t. to a recombination operator. More precisely, we fix a population structure $\vec{\wp} \in \mathcal{W}$ and consider a landscape that is elementary for the population weighted recombination operator $\mathbf{S}^{\mathcal{F},\vec{\wp}}$. Then we may ask how the mutational pair-correlation differs between the population-weighted elementary landscape and its counterpart with the same interaction pattern but uniform population structure. Since the expressions in the following are rather unwieldly, we shall restrict ourselves to two alleles at each locus and nearest neighbor correlations only. The crucial quantities are the matrix elements

$\langle\psi_L; \mathbf{M}^{(k)}\psi_I\rangle_\wp$. A rather tedious calculation yields:

$$\langle\psi_L; \tilde{\mathbf{M}}^{(k)}\psi_I\rangle_\wp = \langle\psi_I;\psi_I\rangle_\wp \begin{cases} 1 & \text{if} \quad I = L \wedge k \notin I \\ -1 & \text{if} \quad I = L \wedge k \in I \\ q_k & \text{if} \quad L = I \cup \{k\} \wedge k \neq I \\ 0 & \text{if} \quad \text{otherwise} \end{cases} \tag{30}$$

where $q_k = \sum_{x_k} \hat{x}_k / p_k(x_k)$ is a measure for the deviation of the allele distribution at locus k from uniformity. Note that $q_k = 0$ if and only if $p_k(1) = p_k(0) = 1/2$.

Setting $f = \sum_I a_I \psi_I$ we find for the population weighted mutational nearest neighbor correlation:

$$\begin{aligned} r_{\tilde{\mathbf{M}},\vec{\wp}}(1) = \frac{\langle\tilde{f}; \tilde{\mathbf{M}}\tilde{f}\rangle_\wp}{\langle\tilde{f};\tilde{f}\rangle_\wp} &= \frac{\sum_{I\neq\emptyset} a_I^2 \left(1 - \frac{2|I|}{n}\right)\langle\psi_I;\psi_I\rangle_\wp}{\sum_{I\neq\emptyset} a_I^2 \langle\psi_I;\psi_I\rangle_\wp} \\ &+ \frac{\sum_{I\neq\emptyset} \frac{1}{n}\sum_{k\neq I} a_I a_{I\cup\{k\}} q_k \langle\psi_I;\psi_I\rangle_\wp}{\sum_{I\neq\emptyset} a_I^2\langle\psi_I;\psi_I\rangle_\wp} \end{aligned} \tag{31}$$

In order to compare the nearest neighbor correlations between different operators for the same population structure, we consider the case of a landscape of the form $f = \sum_I a_I \psi_I^{\vec{\wp}}$ with non-zero Fourier coefficients only if $|I| = q$. In this case the products $a_I a_{I\cup\{k\}}$ in equ.(31) are always 0 and we are left with the mutational nearest neighbor correlation

$$r_{\tilde{\mathbf{M}},\vec{\wp}}(1) = 1 - \frac{2q}{n} \tag{32}$$

which is the same result as for mutation on $\tilde{\mathbf{M}}$-elementary landscapes in the uniform population case. Analogously, we obtain $r_{\mathbf{S}[\infty],\vec{\wp}}(1) = 2^{-q}$ for uniform crossover. For 1-point crossover the correlation is maximal if the landscape consists of compact blocks only, i.e., if there are non-zero coefficients only for the eigenfunctions belonging to $|I| = \ell(I) = q$.

Conversely, we consider a landscape of the form $f = \sum_I a_I \varepsilon_I$ with $a_I = 0$ unless $|I| = q$. Equ.(17) now implies that we have non-zero population weighted coefficients a_I only for modes with $|I| \leq q$. Thus the q_k-dependent correction terms above have non-zero coefficients a_J only when $J \cup \{k\} \leq q$ i.e., for the sub-modes $J \subset I$. Thus the net effect of population weighting is to increase the nearest neighbor correlations by adding contributions of the smoother sub-modes.

7. Discussion

This and our previous papers [21, 54, 59] on the theory of recombination landscapes are motivated by the success of the Fourier Decomposition Approach (FDA) to the analysis of landscapes of single argument operators, most importantly mutation operators [52, 30]. The goal is to make this approach also available to the analysis of recombination processes or any other two argument operators in general. The power of the FDA derives from the following facts: (1) The Fourier basis for a landscape is given by the eigenfunctions of the operator itself and encapsulates its intrinsic symmetries. (2) any fitness landscape can be represented as a superposition of a small set of elementary landscapes. (3) There is a close relationship between the eigenvalue of an elementary landscape, auto-correlation measures of the landscape, and operator performance [52, 54]. (4) Correlation measures of search operators on an arbitrary landscape f are weighted averages of the same correlation measure for the elementary landscapes from which f is built up.

For many purposes it is sufficient therefore, to consider the performance of an operator on a small finite set of elementary landscapes instead of the uncountable set of all landscapes. In addition, the set of elementary landscapes provides us with an useful classification of landscapes, since the various elementary landscapes differ in the number interacting loci and hence they differ with respect to the degree of epistasis among the sites. The degree of epistasis influences the difficulty to find a fitness optimum. In other words, the Fourier analysis of landscapes is closely linked to the eigen-systems of an operator and together they provide a powerful tool to summarize information about the expected performance of an operator on generic fitness landscapes.

The mathematical approach utilized in the FDA to mutation operators, however, depends on a number of formal properties of the mutation operator which are not realized in the case of recombination. Since mutation only acts on a single genotype, the actions of a mutation operator can approximately be represented as a graph, where the vertices are the genotypes and the edges represent the effects of single mutation events. This representation is even exact if each mutation has the same probability. As a consequence the FDA of mutation landscapes is a derivative of the spectral theory of graphs [51]. Furthermore since mutations act independent of the rest of the population, the exploration of genotype space by mutation can be represented as a random walk independent of the population distribution. These conditions are not realized for recombination operators and thus the mathematical apparatus developed for FDA of mutation operators cannot be applied directly to the analysis of recombination operators.

Recombination differs from mutation in two important respects. First, a recombination event requires the interaction of two genotypes. As a consequence, the actions of a recombination operator can not be represented by a simple graph where the vertices are single genotypes. We have addressed this difficulty in our previous papers by introducing a hypergraph representation of recombination effects [21] and later so-called P-structures, which are a maps from pairs of vertices onto the edges of the recombination hypergraph [54, 59]. Furthermore, since recombination acts on two genotypes, the probability to reach a genotype x from a given one y depends on the availability (frequency) of recombination partners z in the population, i.e., the recombination operator and its elementary landscapes depend on the frequency distribution of genotypes in the population. This fact, which has been neglected in our work, is the focus of the present contribution.

We treat the influence of population structure on the performance of recombination operators with two restrictions: (1) we only consider homologous recombination events among strings of genes or "sites", and (2) we limit the analysis to populations in linkage equilibrium. The set of all population structures in linkage equilibrium forms the Wright manifold. We approach the problem of population structure in recombination dynamics by generalizing the recombination operators introduced in [54] by incorporating population dependent weights and show that each element in the Wright manifold is a fixed point under all string recombination operators for arbitrary number of loci and alleles per locus (see Theorem 1).

Our mathematical results provide a complete description of the spectral properties of population weighted recombination operators and a wealth of information of the structure of their eigenspaces. In the following we will interpret these results by addressing two questions: (1) How does the nearest neighbor correlation for recombination operators compare with that of the point mutation operator for landscapes of different ruggedness? (2) Given a particular landscape, how does population structure influence the nearest neighbor correlation for a given recombination operator?

The comparison of string recombination and point mutation operators is easiest if we first focus on populations with uniform allele frequency distribution (below called *uniform population*), i.e., the case where all alleles at each locus have the same frequency and thus the genetic variation is maximal. Then all string recombination operators and the point mutation operators have the same eigenfunctions. This was already noted in [59] for binary strings. Here, the result is generalized to arbitrary numbers of alleles. This means that any landscape that is elementary for mutation is also elementary for string recombination. The eigenvalues, however, and by implication the nearest neighbor correlations on these elementary landscapes depend on the kind of the operator. To asses the relative effect of mutation and recombination in different kinds of landscapes one only needs to compare the nearest neighbor correlations of the different operators on the same elementary landscapes. This provides information on how the number and location of interacting loci on the genome influences the relative performance of mutation and recombination operators as predicted by the nearest neighbor correlation. For illustration purposes this has been done in Figure 2 for point mutations, uniform recombination and one point recombination on binary strings of length $n = 10$.

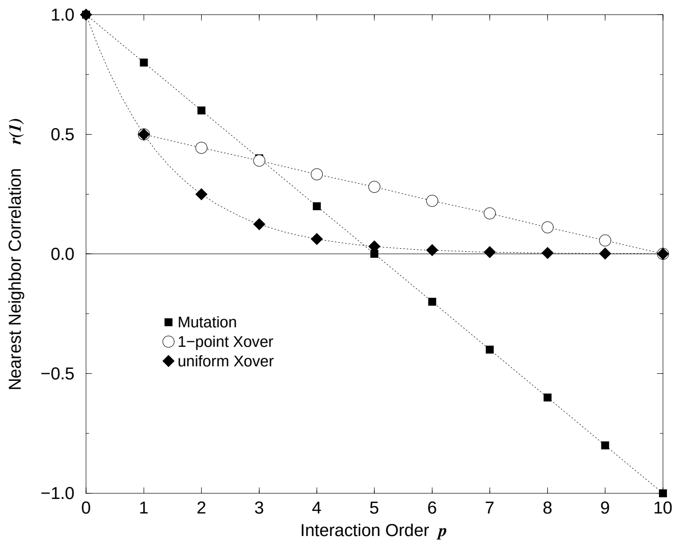

Figure 2. Nearest neighbor correlations $r(1)$ for landscapes with different interaction orders p on the hypercube $\mathcal{Q}_2^{10}$. The values for 1-point recombination refer to the smoothest possible cases, that is, to those landscapes for which defining length $\ell(\tilde{I})$ and interaction order coincide.

For point mutation and uniform recombination the nearest neighbor correlation $r(1)$ on an elementary landscape only depends on the number of interacting loci, the so-called interaction order p. This is not true in general for one point recombination (see below).

For point mutation $r(1)$ is a linear function of the interaction order which crosses the line of $r(1) = 0$ at the point where the interaction order equals half the number of sites in the string. For free recombination, $r(1)$ starts from $1/2$ for additive landscapes independent of the number of loci, and exponentially approaches zero with increasing interaction order. This means that for smooth and moderately complex fitness landscapes, i.e., interaction order less than $n/2$, the nearest neighbor correlation is higher for the mutation operator than for uniform recombination. For high interaction order, uniform recombination has a higher $r(1)$ than mutation, but it is approximately zero, while $r(1)$ is negative for mutation. We interpret this to mean that neither of these operators will perform well on landscapes dominated by high interaction orders.

The inclusion in this comparison of one point recombination is complicated by the fact that $r(1)$ for one point recombination depends on the defining length ℓ, i.e., the distance on the chromosome of the two most distant interacting sites, rather than the number of interacting sites. Since $r(1)$ is strictly decreasing with defining length, the maximal correlation for a given interaction order is reached when the defining length is equal to the interaction order, $p = \ell(\tilde{I})$. The observation provides a starting point for a discussion of the *building block hypothesis* or the *schema theorem* [27, 1, 55] in the context of our theory, see [54]. The graph plotted in Figure 2 for one point crossover is the maximal $r(1)$ for one point crossover given a certain interaction order. The minimal correlation, on the other hand, is zero for all landscapes of the interaction order larger or equal to 2. For very smooth landscapes, i.e., for interaction orders 1 and 2, mutation still outperforms recombination, i.e., the $r(1)$ for mutation is larger then $\max r(1)$ for one point crossover. For moderately high and high interaction orders (i.e. moderately complex to complex fitness functions), however, the value of $\max r(1)$ becomes higher for one-point crossover than for mutation. Note, however, that the minimal $r(1)$ for one point crossover is always zero, i.e., even lower than the correlation for uniform recombination. Hence, it is predicted for complex fitness functions that one point crossover will be superior to both mutation and uniform recombination, provided the interacting sites are closely spaced on the chromosome, i.e., if the defining length ℓ is close to minimal possible value p.

As a thumb nail sketch the results can be summarized to mean that for simple fitness functions mutation is a superior search strategy. For more complex fitness functions, i.e. higher interaction orders, one-point recombination is superior to mutation as well as uniform recombination if the interacting genes are clustered on the chromosomes and hence the defining length is minimal. In the worst case, when the defining length is large, however, one point crossover is worse than mutation and uniform recombination. Hence if the distribution of interacting loci is random it is saver to use uniform recombination, while mutation alone will always be better in simple and moderately complex situations. With many interacting genes only one point crossover with minimal or small defining length gives any hope for successful adaptation. These comparisons so far are done assuming that the population is maximally polymorphic. Below we will consider whether population structure affects this characterization.

The main result from the analysis of population weighted recombination operator is that the population structure only affects the eigenfunctions of the recombination operator but not its spectrum which still only depends on the interaction order or the defining length respectively. This means that the nearest neighbor correlations on elementary landscapes are invariant with respect to different population structures. What counts as an elementary landscape, however, is changing with allele frequencies. This means that the diagram in Figure 2 does not change if different population structures are considered. All that is different is the interpretation of the points on the x-axis. In this case the points represent families of population dependent elementary landscapes.

The *representation problem* is the hypothesis that the evolution of complex adaptations by random variation and selection is only possible if the genetic representation of the phenotype matches the functional constraints on the phenotypic characters [58]. This idea has been developed under various names in the area of evolutionary algorithms [17] and, independently, in the theory of biological evolution, most notably by Riedl [45] and Frazzetta [19], as well as others. The present theory puts some quantitative constraints to this idea. From inspecting Figure 2, it is obvious that there is no representation problem as long as the interaction order is much less than half the number of loci. In fact, for low and moderately complex fitness functions, mutation is by far the superior search strategy. The representation problem kicks in as the interaction order p is large, more precisely, if

$$p > \frac{n^2 - n}{3n - 1}, \tag{33}$$

which reduces to $p > n/3$ for $n \gg 1$.

With interaction order larger than this limit one-point crossover can do better than mutation provided that the distribution of the interacting loci on the chromosomes is clustered, i.e., the defining length is minimal or close to minimal. A good representation is given (i.e. one that is predicted to allow adaptation of organisms with high interaction order among their genes) if the interacting genes are clustered on the chromosomes. Taken literally, this limit suggests that mutation is a fairly good search strategy up to quite high interaction orders for realistically large numbers of loci. Hence, the question whether complex organisms face a representation problem can be restated as: is the interaction order among genes contributing to a functional system higher than one third of the number of genes affecting the character? If yes, then an appropriate distribution of loci on the chromosome might lead to higher evolvability. If not, then mutation is predicted to do always better.

The conclusion that very high interaction orders are required to produce serious problems in evolvability confirms earlier work on quantitative genetic models in which it has been shown that reasonable smooth fitness functions on additive genetic traits lead to either no effect on the rate of evolution [8, 56, 57] or only to moderate effects [3].

Another obvious set of implications of the present results concern the question of the evolutionary advantage of recombination. One of the hypotheses states that recombination might increase the rate of adaptive evolution by combining independently derived mutations into the same genotype [40, 16, 11, 42]. Population genetic models already indicated that the conditions under which this is correct are quite restrictive from a population dynamical point of view [11, 12]. The present results further support this conclusion, by pointing out that mutation consistently has higher nearest neighbor correlations in the search on smooth and moderately complex fitness landscapes than either uniform recombination or one point crossover. Hence, it is unlikely that recombination would consistently provide significant advantages in terms of evolvability if there are only moderate levels of interactions and the distribution of interacting genes is random (i.e. if the defining length can be large).

Now we want to turn to the question how the performance of a given operator is influenced by the population structure on a given fitness landscape. As shown at the end of section 5, deviations from the uniform distribution of allele frequencies leads to a decrease of the *effective interaction order*. This is also seen in Figure 1 which shows that the nearest neighbor correlation is lowest in the center of the simplex where there is the maximum of genetic variation and it increases as the allele frequencies get more extreme. Here this fact has been proven for arbitrary fitness landscapes and string recombination operators. The intuitive interpretation of this result is that the lower the genetic variation in a population, the smaller the area of the fitness landscape explored or covered by the variation in the population. In the extreme of a monomorphic population, only a single point of the landscapes is covered. If the population assumes extreme allele frequencies, the locally explored area of the fitness landscape looks more smooth, just as the curvature experienced in a small window of a curve is less than with a large window.

This result is related to a number of phenomena known from quantitative genetics. Surprisingly it was found that a bottle neck, during which inbreeding and drift leads to the loss of genetic variation, can lead to an increase of the heritability of a quantitative trait [22, 23, 24] due to increased additive genetic variance of the character [10]. This effect has been shown in simulations as well as in experimental populations of mice (Chevrud, unpublished results), or house flies [7]. The cause of this phenomenon is that epistatic effects get lost as genes get fixed or approach near fixation. As a consequence part of the epistatic effects get converted into additive effects [24]. Similarly, it has been found that in outbred populations there is little evidence for epistasis if one uses standard variance decomposition methods from quantitative genetics [62]. In contrast, crosses of inbred lines used to find QTL's by genetic mapping show considerable epistasis for some characters. The solution to this apparent paradox is that close to the boundaries of the simplex, the phenotype landscape is more smooth and thus epistasis has a smaller effect than in the center of the simplex [36]. Hence, outbred populations close to mutation-selection equilibrium might show less evidence of gene interaction since they are closer to the boundaries of the simplex, while the recombinant lines from crosses of inbred lines have equal allele frequencies and thus show more evidence of gene interaction. This effect is also explained by or result that the effective order of interaction decreases as the allele frequencies are approaching the boundaries of the simplex (see Figure 1).

Finally we want to point out that our results show that point mutation and string recombination operators commute; this means that the effect of a combined action of mutation and recombination is independent of the order in which they are applied. This is important to note since most higher organisms use both, mutation as well as recombination for their evolution. The commutativity shows that the two classes of operators are in a deep sense compatible. There is no bias induced the search procedure by the order in which mutation and recombination are applied. This fact may lay to rest the question whether mutation and recombination "experience" fundamentally different fitness landscapes as proposed in the One-Operator-One-Landscape Hypothesis [33]. Some evidence to that effect already emerged in our previous papers [54, 59] but remained questionable in these papers since the influence of population structure was not accounted for. This gap has now been closed.

Acknowledgement. This is publication number CCE# 56 of the *Center of Computational Ecology* at Yale. GPW thanks for the financial support by the Yale Institute for Biosperic Studies to CCE.

References

[1] L. Altenberg. The schema theorem and the Price's theorem. In D. Whitley and M. D. Vose, editors, *Foundations of Genetic Algorithms 3*, pages 23–49, Cambridge, MA, 1995. MIT Press.

[2] L. Altenberg and M. W. Feldman. Selection, generalized transmission, and the evolution of modifier genes. I. The reduction principle. *Genetics*, 117:559–572, 1987.

[3] M. Baatz and G. P. Wagner. Adaptive inertia caused by hidden pleiotropic effects. *Theor. Pop. Biol.*, 51:49–66, 1997.

[4] J. H. Bennett. On the theory of random mating. *Ann. Eugen.*, 18: 311–317, 1954.

[5] C. Berge. *Hypergraphs*. Elsevier, Amsterdam NL, 1989.

[6] R. N. Bhattacharya and E. C. Waymire. *Stochastic Processes with Applications*. Wiley, New York, 1990.

[7] E. H. Bryant, S. A. McCommas, and L. M. Combs. The effect of an experimental bottleneck upon genetic variation in the house fly. *Genetics*, 114:1191–1211, 1986.

[8] R. Bürger. Constraints for the evolution of functionally coupled characters: A nonlinear ananlysis of a phenotpyic model. *Evolution*, 40: 182–193, 1986.

[9] R. Bürger. Mathematical properties of mutation-selection models. *Genetika*, 102/103:279–298, 1998.

[10] J. M. Cheverud and E. J. Routman. Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution*, 50:1042–1051, 1996.

[11] J. F. Crow and M. Kimura. Evolution in sexual and asexual populations. *Am. Nat.*, 99:439–450, 1965.

[12] J. F. Crow and M. Kimura. *An Introduction to Population Genetics Theory.* Harper and Row, New York, 1970.

[13] J. C. Culberson. Mutation-crossover isomorphism and the construction of discriminating functions. *Evol. Comp.*, 2:279–311, 1995.

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

[15] M. Eigen and P. Schuster. *The Hypercycle.* Springer-Verlag, New York, Berlin, 1979.

[16] R. A. Fisher. *The genetical theory of natural selection.* Clarendon Press, Oxford, 1930.

[17] D. B. Fogel. *Evolutionary Computation.* IEEE Press, New York, 1995.

[18] W. Fontana, P. F. Stadler, E. G. Bornberg-Bauer, T. Griesmacher, I. L. Hofacker, M. Tacker, P. Tarazona, E. D. Weinberger, and P. Schuster. RNA folding and combinatory landscapes. *Phys. Rev. E*, 47:2083 – 2099, 1993.

[19] T. H. Frazzetta. *Complex Adaptations in Evolving Populations.* Sinauer Ass. Inc., Sunderland, MA, 1975.

[20] J. H. Gillespie. *The Causes of Molecular Evolution.* Oxford University Press, New York, Oxford, 1991.

[21] P. Gitchoff and G. P. Wagner. Recombination induced hypergraphs: A new approach to mutation-recombination isomorphism. *Complexity*, 2: 47–43, 1996.

[22] C. J. Goodnight. On the effect of founder events on epistatic genetic variance. *Evolution*, 41:80–91, 1987.

[23] C. J. Goodnight. Epistasis and the effect of founder events on the additive genetic variance. *Evolution*, 42:441–454, 1988.

[24] C. J. Goodnight. Epistasis and the increase in additive genetic variance: implications for phase 1 of Wright's shifting-balance process. *Evolution*, 49:502–511, 1995.

[25] R. Happel and P. F. Stadler. Canonical approximation of fitness landscapes. *Complexity*, 2:53–58, 1996.

[26] J. Hofbauer and K. Sigmund. *Dynamical Systems and the Theory of Evolution.* Cambridge University Press, Cambridge U.K., 1988.

[27] J. H. Holland. *Adaptation in Natural and Artificial Systems.* MIT Press, Cambridge, MA, 1993.

[28] W. Hordijk. A measure of landscapes. *Evolutionary Computation*, 4 (4):335–360, 1996.

[29] W. Hordijk. Correlation analysis of the synchronizing-CA landscape. *Physica D*, 107:255–264, 1997.

[30] W. Hordijk and P. F. Stadler. Amplitude spectra of fitness landscapes. *J. Complex Systems*, 1:39–66, 1998.

[31] M. A. Huynen, P. F. Stadler, and W. Fontana. Smoothness within ruggedness: the role of neutrality in adaptation. *Proc. Natl. Acad. Sci. (USA)*, 93:397–401, 1996.

[32] T. Jones. *Evolutionary Algorithms, Fitness Landscapes, and Search.* PhD thesis, Univ. of New Mexico, Albuquerque, NM, 1995.

[33] T. Jones. One operator, one landscape. Technical Report #95-02-025, Santa Fe Institute, 1995.

[34] S. A. Kauffman. *The Origin of Order.* Oxford University Press, New York, Oxford, 1993.

[35] S. A. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.*, 128:11–45, 1987.

[36] M. Laubichler. *Identifying Units of Selection: Conceptual and Methodological Issues.* PhD thesis, Yale University, New Haven, CT, 1997.

[37] L. Lovász. Spectra of graphs with transitive groups. *Periodica Math.Hung.*, 6:191–195, 1975.

[38] Y. I. Lyubich. *Mathematical structures in population genetics.* Springer-Verlag, Berlin, 1992.

[39] M. Mézard, G. Parisi, and M. Virasoro. *Spin Glass Theory and Beyond.* World Scientific, Singapore, 1987.

[40] T. H. Morgan. *Heredity and Sex.* Columbia University Press, New York, 1913.

[41] T. Nagylaki. *Introduction to Theoretical Population Genetics.* Springer-Verlag, Berlin, 1992.

[42] S. P. Otto and N. H. Barton. The evolution of recombination: removing the limits to natural selection. *Genetics*, 147:879–906, 1997.

[43] A. S. Perelson and S. A. Kauffman, editors. *Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System*, volume 9 of *Santa Fe Institute Studies*, Reading, MA, 1991. Addison-Wesley.

[44] W. B. Provine. *Sewall Wright and Evolutionary Biology.* Univ. Chicago Press, Chicago, London, 1986.

[45] R. Riedl. A systems-analytical approach to macroevolutionary phenomena. *Quart. Rev. Biol.*, 52:351–370, 1977.

[46] R. B. Robbins. Some applications of mathematics to breeding problems. III. *Genetics*, 3:375–389, 1918.

[47] P. Schuster and P. F. Stadler. Landscapes: Complex optimization problems and biopolymer structures. *Computers & Chem.*, 18:295–314, 1994.

[48] P. Schuster, P. F. Stadler, and A. Renner. RNA structures and folding: From conventional to new issues in structure predictions. *Curr. Opinions Structural Biol.*, 7:229–235, 1997.

[49] J.-P. Serre. *Linear Representations of Finite Groups.* Springer-Verlag, New York, Heidelberg, Berlin, 1977.

[50] P. F. Stadler. Random walks and orthogonal functions associated with highly symmetric graphs. *Discr. Math.*, 145:229–238, 1995.

[51] P. F. Stadler. Towards a theory of landscapes. In R. Lopéz-Peña, R. Capovilla, R. García-Pelayo, H. Waelbroeck, and F. Zertuche, editors, *Complex Systems and Binary Networks*, pages 77–163, Berlin, New York, 1995. Springer Verlag.

[52] P. F. Stadler. Landscapes and their correlation functions. *J. Math. Chem.*, 20:1–45, 1996.

[53] P. F. Stadler and R. Happel. Random field models for fitness landscapes. *J. Math. Biol.*, 1999. in press, SFI preprint 95-07-069.

[54] P. F. Stadler and G. P. Wagner. The algebraic theory of recombination spaces. *Evol. Comp.*, 5:241–275, 1998.

[55] C. R. Stephens and H. Waelbroeck. Effective degrees of freedom in genetic algorithms. *Phys. Rev. E*, 57:3251–3264, 1998.

[56] G. P. Wagner. The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *J. Evol. Biol.*, 1:45–66, 1988.

[57] G. P. Wagner. The significance of developmental constraints for phenotypic evolution by natural selection. In G. deJong, editor, *Population Genetics and Evolution*, pages 222–229. Springer-Verlag, Berlin, Heidelberg, 1988.

[58] G. P. Wagner and L. Altenberg. Complex adaptations and the evolution of evolvability. *Evolution*, 50:967–976, 1995.

[59] G. P. Wagner and P. F. Stadler. Complex adaptations and the structure of recombination spaces. In C. Nehaniv and M. Ito, editors, *Algebraic Engineering*, pages 151–170, Singapore, 1998. World Scientific. Proceedings of the Conference on Semi-Groups and Algebraic Engineering, University of Aizu, Japan.

[60] E. D. Weinberger. Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biol. Cybern.*, 63:325–336, 1990.

[61] E. D. Weinberger. Fourier and Taylor series on fitness landscapes. *Biological Cybernetics*, 65:321–330, 1991.

[62] M. C. Whitlock, P. C. Phillips, F. B. Moore, and S. J. Tonsor. Multiple fitness peaks and epistasis. *Ann. Rev. Ecol. System.*, 26:601–629, 1995.