Fitness Landscapes

Peter F. Stadler

[1] Institut für Theoretische Chemie und Molkulare Strukturbiologie, Universität Wien, Währingerstrasse 17, A-1090 Wien, Austria
[2] The Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501

Abstract. Fitness landscapes are a valuable concept in evolutionary biology, combinatorial optimization, and the physics of disordered systems. A fitness landscape is a mapping from a configuration space that is equipped with some notion of adjacency, nearness, distance or accessibility, into the real numbers. Landscape theory has emerged as an attempt to devise suitable mathematical structures for describing the "static" properties of landscapes as well as their influence on the dynamics of adaptation. This chapter gives a brief overview on recent developments in this area, focusing on "geometrical" properties of landscapes.

1 Introduction

The concept of a *fitness landscape* originated in theoretical biology more than seventy years ago [1]. It can be thought of as a kind of "potential function" underlying the dynamics of evolutionary optimization. Implicit in this idea is both a *fitness function* f that assigns a fitness value to every possible genotype (or organism), and the arrangement of the set of genotypes in some kind of abstract space that describes how easily or frequently one genotype is reached from another one.

The same abstract construction arises in a natural way in the physics of disordered systems. Spin-glasses, for example, can be cast into the same form [2,3]. Each spin configuration is assigned an energy by virtue of the Hamiltonian that specifies the model; the dynamic properties invoke a collection of transitions between configurations. In biophysics energy landscapes govern the folding of biopolymers, including proteins [4–6] and nucleic acids [7,8]. Conceptually, there is a close connection with the *potential energy surfaces* of theoretical chemistry [9,10]: As a consequence of the validity of the Born-Oppenheimer approximation, the PES provides the potential energy $U(\boldsymbol{R})$ of a molecule with n atoms as a function of its nuclear geometry $\boldsymbol{R} \in \mathbb{R}^{3n}$. Electoral Landscapes are used to explain party platform behavior in spatial voting models [11,12].

In combinatorial optimization the fitness function is usually referred to as the *cost function*, and a *move-set* allows to inter-convert the elements of the *search space* [13]. The application of evolutionary models to combinatorial optimization problems has lead to the design of so-called *evolutionary algorithms* such as Genetic Algorithms, Evolution Strategies, and Genetic Programming [14–18].

The intuitive notion of *ruggedness* is closely related to the difficulty of optimizing (or adapting) on a given landscape. It depends obviously on both the fitness function and the geometry of the search space, which is induced by the search process. On the other hand, simulations of adaptation of biologically realistic landscapes derived from RNA folding [19] have shown that *neutrality*, that is, the occurrence of adjacent configurations with the same fitness, can play a dominating role in evolutionary dynamics as well.

One of the main topic in "landscape theory", and the focus of this contribution, is therefore a detailed understanding of the geometric features of landscapes: Mountain massives, valleys, basins, peaks, plains and ridges in multidimensional combinatorial objects may look quite different from our 3D experience and oftentimes require a mathematical description in terms of algebraic combinatorics rather than calculus.

Landscapes can also be studied from a "dynamical" point of view, focusing on the features of a dynamical system, for instance an evolving population, that uses the landscape as its substrate. The challenge for a *theory of landscapes* is therefore to link these two points of views, for instance by determining how geometric properties influence the dynamical behavior.

Given that landscapes arise naturally in many different fields, it is not surprising that the concept of a *fitness landscape* has emerged as a unifying theme in the literature on complex systems [17,20–22]. In formal terms, a landscape consists of three ingredients

1. A set X of configurations,
2. a notion $\mathcal{X}$ of neighborhood, nearness, distance, or accessibility on X, and
3. a fitness function $f : X \to \mathbb{R}$,

The set X together with the "structure" $\mathcal{X}$ forms the configuration space. The definition of $\mathcal{X}$ is purposefully left vague at this point as we will elaborate on the structure $\mathcal{X}$ in section 2. A common source of confusion is the fact that biologists like to maximize fitness on their landscapes, while physicists minimize energy on theirs. Obviously, replacing f by $-f$ maps one picture into the other.

Despite its wide range of applicability, the usefulness of fitness landscapes is limited to certain situations. Let us consider a general evolutionary process as an example for the limitations of the landscape concept. Since genetic variation is generated independently from the natural selection acting on it, the generic structure of an evolutionary model in discrete time can be written as

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

where x is e.g. the vector of haplotype frequencies [23,24]. As usual, $\circ$ denotes the Schur (Hadamard, component-wise) product of vectors. The transmission term $T(x, \mathsf{t})$ describes the probability of transforming one type into another one by mutation and/or recombination [25] and hence determines the structure $\mathcal{X}$ on the set X of all vectors of haplotype frequencies. In genetics this structure can be understood in terms of certain classes of algebraic structures [26] that depend on the details of transmission mechanism represented by the paramters t. The term $S(x, \mathsf{w})$ describes the selection forces acting on x. The parameters w form the *fitness function*, since they can be regarded as a mapping from the set of types into the real numbers. Whether or not the dynamics of equ.(1) is consistent with or determined by a *fitness landscape*, depends on the particularities of of the selection term. Setting $S(x, \mathsf{w}) = x \circ F(x, \mathsf{w})$, or in component-wise notation, $S_k(x, \mathsf{w}) = x_k F_k(x, \mathsf{w})$, we have selection proportional to a *growth-rate function* F_k for each type [27]. We suggest that one should speak about *fitness landscapes* only when $F_k(x, \mathsf{w}) \approx f(k)$, where $f(k)$ is a constant that is characteristic for the type k, since otherwise the fitness (growth rate) of type k depends on and changes with the frequencies of all other types. Models of co-evolution are sometimes viewed as "coupled dancing landscapes" [17] where a species A changes the landscape of species B, and B changes the landscape of A, at the same time scale at which both species adapt to their respective landscapes. We prefer here to limit the term fitness landscape to situations in which (1) fitness is characteristic of a type and (2) constant in time, at least approximately.

2 Configuration Spaces

There appear to be three distinct approaches to organizing the configuration space.

1. In computer science one typically specifies a "move set" or "genetic operator" that inter-converts one or more configurations into a new one.
2. Sometimes transition probabilities are specified that describe how frequently a system attempts to move from one configuration to another one.
3. A rigorous mathematical analysis often starts with specifying a metric distance or a topology on X.

2.1 Move Sets

In its most abstract form a *move set* assigns to a k-tuple $(x_1, \dots, x_k) \in X^k$ of "parents" a list $N(x_1, \dots, x_k) \subseteq X$ of "children". We will restrict our attention here to the two most commonly used move-set types, mutation and recombination.

A mutation operator simply assigns a set $N(x)$ of "accessible neighbors" or "elementary mutants" to each configuration x. This allows us to interpret X as a (possibly directed) graph with vertex set X and $N(x)$ the (out)neighbors of $x \in X$. Most commonly the move sets are constructed such that $y \in N(x)$ if and only if $x \in N(y)$, in which case the graph is symmetric, or, equivalently, undirected.

In the case of strings (i.e., sequences of characters taken from a fixed alphabet $\mathcal{A}$) typical moves consist of the replacement of a character at a single position by another one. The resulting graphs are the Hamming graphs. In particular, for a two letter alphabet such as *spin-up* ($\uparrow$), and *spin-down* ($\downarrow$), one obtains hypercubes, see Figure 1.

In some cases the configurations are naturally interpreted as algebraic objects. For instance, the tours of a Traveling Salesman Problem are permutations of a list of cities. Configurations hence are group elements and moves become

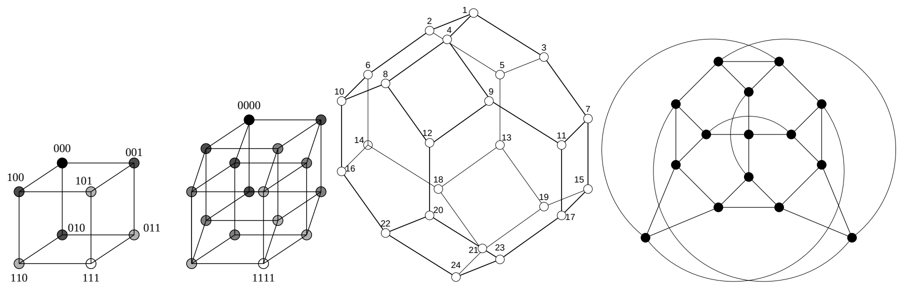

Fig. 1. Some Configuration space graphs: hypercubes $\mathcal{Q}_2^3$ and $\mathcal{Q}_2^4$, the permutohedron graph $\Gamma[\mathsf{S}_5, \mathcal{K}]$, the line graph of the Petersen graph $L[P]$ which equals the Robinson graph for $n = 5$ taxa.

generators of the group. Let G be a permutation group acting on a finite set X. Furthermore, let $\Omega \subset \mathsf{G}$ be a set of generators of G such that (i) $\imath \notin \mathsf{G}$ and (ii) $x \in \Omega \implies x^{-1} \in \Omega$. A graph $G = G(\mathsf{G}, \Omega)$ with vertex set $V = \mathsf{G}$ and edges $\{x, y\} \in E$ if and only if $xy^{-1} \in \Omega$ is called a *Cayley graph* of the group G. In case of the symmetric group S_n suitable generator sets are e.g. transpositions, reversals, or the "canonical" transpositions of two *subsequent* cities which, for $n = 4$, lead to the permutohedron graph shown in Fig. 1. It is not hard to show that Hamming graphs are also Cayley graphs of a suitable group, see e.g. [28,29].

In biology evolutionary relationships between species or individual genes are customarily represented as phylogenetic trees. The vertices of a phylogenetic tree represent taxonomic units, the graph's topology delineates the genealogical relationships between them, and the branch lengths reflect the time of divergence. Many methods exist for the construction of phylogenetic trees, the more sophisticated among them seeking those trees in which the taxonomic units evolve with the least evolutionary change [30] (most parsimonious trees) or trees of maximum likelihood given a stochastic model of sequence evolution [31]. The search for the optimal tree is hence recast as a combinatorial optimization problem on the set of all phylogenetic trees with a given number of leaves (taxa). The basic variants of these tree reconstruction problems are all NP-complete [32,33], hence search heuristics are used in practice which employ a variety of editing operations on phylogenetic trees. So-called "nearest neighbor interchange" (swapping of subtrees separated by an inner edge of the tree), for instance, leads to a rather well-studied family graph sometimes referred to as Robinson graphs [34–36].

All move sets discussed above are symmetric and regular, i.e., any two configurations have the same number of neighbors. Of course, this is not always the case. Biological sequences, for instance, not only undergo point mutations but also insertions and deletions leading to highly irregular graphs. Other mutation operators of interest in this context include gene duplications and genomic rearrangements [37,38]. A graph is faithfully represented by its adjacency matrix $\mathbf{A}$ which has the entries $\mathbf{A}_{xy} = 1$ if $x \in N(y)$ and $\mathbf{A}_{xy} = 0$ otherwise. Obviously, $\mathbf{A}$ is symmetric if and only if the graph is undirected.

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. This 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 [39] and Jones [40]. The alternative is to leave the vertices to represent individual genotypes and to make the edges more complex. In Gitchoff and Wagner [41] it was shown that recombination spaces can be represented as hypergraphs (which consist of a vertex set X and a collection $\mathcal{E}$ of (not necessarily) distinct subsets of X called (hyper)edges), where the hyperedges 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 hyperedges of the hypergraph [42,43].

Let us first consider homologous recombination on 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 genotypes 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 $x_k \in \mathcal{A}_k$. A particular cross-over operator χ is determined by the list $\boldsymbol{\chi}$ of loci that the child inherits from the first parent. Thus the loci in $\overline{\boldsymbol{\chi}} = \{1, \ldots, n\} \setminus \boldsymbol{\chi}$ come from the second parent. More formally, given $\boldsymbol{\chi}$, the offspring $x = \chi(y, z)$ of the two parents y and z has the component-wise representation

$$x_k = \begin{cases} y_k \text{ if } k \in \boldsymbol{\chi} \\ z_k \text{ if } k \in \overline{\boldsymbol{\chi}} \end{cases} \tag{2}$$

It will be convenient in the following to express equ.(2) by means of an "incidence operator"

$$\mathbf{H}^{\chi}_{x,(y,z)} = \begin{cases} 2 \text{ if } x = y = z \\ 1 \text{ if } \quad y \neq z \text{ and } \quad x = \chi(y, z) \\ 0 \quad\quad \text{otherwise}\,. \end{cases} \tag{3}$$

Here we shall restrict ourselves to recombination on strings. Crossover operators for permutation, such as Traveling Salesman tours, are reviewed for instance in [44].

A *recombination operator* in the sense of most of the GA literature is then a family $\mathcal{F}$ of cross-over operators that act on $X \times X$ with probability $\pi(\chi)$. The incidence "matrix" associated with a recombination operator is simply

$$\mathbf{H}^{\mathcal{F}} = \sum_{\chi \in \mathcal{F}} \mathbf{H}^{\chi} \tag{4}$$

The two most important recombination operators are

[∞] 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 $\boldsymbol{\chi} = \{1, \ldots, k\}$.

Homologous recombination (of strings) under very general conditions leads to very regular configuration spaces. In particular, one can show that the automorphism group of $\mathbf{H}^{\mathcal{F}}$ is generously transitive [42]. (A permutation group G on a set X is generously transitive if for each pair $x, y \in X$ there is a group element $\mathsf{g} \in \mathsf{G}$ such that $\mathsf{g}(x) = y$ and $\mathsf{g}(y) = x$, see e.g. [45].) This picture, however, changes radically, if unequal crossover is considered, where the number of genes on a chromosome can change [46].

2.2 Transition Matrices

Regarding X as a set of "states" we may alternatively specify transition probabilities $\mathbf{T}_{xy}$ for moving from y to x. The Markov process with transition matrix $\mathbf{T}$ organizes the configuration space in this case. Typically one requires $\mathbf{T}$ to be ergodic (i.e., every state can be reached from every other state) and reversible, i.e., to satisfy

(E) $\mathbf{T}$ is irreducible, or, equivalently,
there is a unique stationary distribution p on X such that $\mathbf{T}p = p$. Furthermore $p(x) > 0$ for all $x \in X$.

(R) $\mathbf{T}_{xy}p(y) = \mathbf{T}_{yx}p(x)$. This condition is also known as "detailed balance".

In other words, $\mathbf{T}$ is self-adjoined w.r.t. to the scalar product

$$\langle f, g \rangle_p = \sum_x p(x) f(x) g(x)^* \tag{5}$$

where the star denotes complex conjugation.

A most useful observation is that the matrix $\mathbf{S}$ defined by

$$\mathbf{S}_{xy} = p(x)^{-1/2} \mathbf{T}_{xy} p(y)^{1/2} \tag{6}$$

is symmetric and similar to $\mathbf{T}$. Hence given a non-symmetric transition matrix $\mathbf{T}$ and a landscape f we may transform the model to new coordinates with the symmetric operator $\mathbf{S}$ and the transformed landscape

$$f^{\sigma}(x) = p(x)^{-1/2} f(x) \tag{7}$$

This allows the application of the spectral landscape theory discussed in section 4 also to the non-symmetric case.

The move sets discussed in the previous section can be translated into the Markov chain setting in a natural way. With each (directed or undirected) graph there is an associated Markov process on its vertex set [47] defined by the transition matrix

$$\mathbf{T} = \mathbf{A}\mathbf{D}^{-1} \tag{8}$$

where $\mathbf{D}$ is the so-called degree matrix, which is diagonal and $\mathbf{D}_{xx} = |N(x)|$ is the number of neighbors of x, and $\mathbf{A}$ is the adjacency matrix introduced in the previous section. This Markov process describes a *random walk* on X which has been suggested as a means to sample information about a landscape by Ed Weinberger [48,49]. We remark that in the case of undirected and symmetric directed graphs the stationary distribution is given by

$$p(x) = \frac{\mathbf{D}_{xx}}{2E} \tag{9}$$

where E is the number of undirected edges, which, for symmetric directed graphs, is of course $E = \sum_x |N(x)|/2$.

A *cross-over walk* [50,51] on X is the Markov process based on the following rule: The "father" y is mated with a randomly chosen "mother" z. The offspring is "son" x which becomes the "father" of the next mating. We regard the sequence of "fathers" as a random walk on X. It is straightforward [52] to derive the transition matrix of this Markov process for homologous recombination from the incidence "matrix" $\mathbf{H}^{\mathcal{F}}$. One obtains

$$\mathbf{S}^{\mathcal{F},\wp}_{xy} = \sum_{\chi\in\mathcal{F}} \pi(\chi)\frac{1}{2}\sum_{z\in X} \mathbf{H}^{\chi}_{x,(y,z)}\wp(z) = \sum_{\chi\in\mathcal{F}} \pi(\chi)\mathbf{S}^{\chi,\wp}_{xy} \tag{10}$$

where $\wp(z)$ denotes the frequency distribution of the genotypes in the equilibrium population.

2.3 Configuration Space Topologies

We shall see in the following section that finite ("discrete", or combinatorial) landscapes are treated quite differently from their manifold ("continuous") counterparts. The reason is that functions on $\mathbb{R}^n$, or more generally Riemannian manifolds can be analyzed in terms of differential operators such as gradients, while finite sets are usually discussed in terms of graph-theoretical properties. It seems desirable therefore, to find a basic framework that allows to deal with landscapes on arbitrary configurations spaces. A suitably general language is provided by the theory of *pretopological spaces.*

A pretopological space consists of an arbitrary set X and a collection $\mathcal{N}(x)$ of neighborhoods for every point $x \in X$, such that
(P1) $N \in \mathcal{N}(x)$ implies $x \in N$;
(P2) $N \in \mathcal{N}(x)$ and $N \subseteq N'$ implies $N' \in \mathcal{N}(x)$
(P2) $N, N' \in \mathcal{N}(x)$ implies $N \cap N' \in \mathcal{N}(x)$ Pretopologies are more general then the much more familiar topological spaces. In fact, $(X, \mathcal{N})$ is a topological space if and only if
(T) For each $N \in \mathcal{N}(x)$ there is an $N' \in \mathcal{N}(x)$ such that $N \in \mathcal{N}(y)$ for all $y \in N'$.

Directed graphs are exactly the finite pretopological spaces. Their neighborhood systems consists of all sets N' containing x and all vertices adjacent to x, i.e., $N(x) \cup \{x\} \subseteq N'$. Notions such as minima, maxima, or continuity of a function, connectedness, convergence, limits, etc. can be defined on pretopological spaces [53–56]. Their usefulness in the context of genotype-phenotype maps and fitness landscapes will be discussed in forthcoming manuscripts [57,58].

3 Basic Properties of Landscapes

3.1 Local Optima

Combinatorial optimization is concerned with finding "optimal" i.e., minimal or maximal values of the cost function. Local optima thus play an important role since they might be obstacles on the way to the optimal solution. In the theory of disordered systems, local minima of the energy function are usually called metastable states. For the sake of definiteness we shall consider local minima in the following. Analogous expressions for local maxima can be obtained by replacing f with $-f$. Let us start with a formal

Definition. A configuration $\hat{x} \in X$ is a *local minimum* if there is a neighborhood $N \in \mathcal{N}(\hat{x})$ such that $f(\hat{x}) \leq f(y)$ for all $y \in N$.

Clearly, this definition makes sense on arbitrary pretopological spaces and it coincides with the usual definition in the graph case, which requires $f(\hat{x}) \leq f(y)$ for all $y \in N(\hat{x})$. A minimum $\hat{x}$ is global, of course, if $f(\hat{x}) \leq f(y)$ for all $y \in X$. Note that landscapes need not have local or even global minima unless they are defined on a compact configuration space.

The number of local optima is a measure for the "ruggedness" of landscape. Richard Palmer [59], for instance, suggested to call a landscape f *rugged* if the number M_f of local optima scales exponentially with some measure of "system size" such as the number of cities in a TSP or the number of spins in spin glass. Unfortunately, there is in general no simply way of computing M_f without exhaustively generating the landscape. Alternatively, one can of course estimate M_f by checking whether a randomly generated $x \in X$ is a local minimum. Numerical data of this kind are reported e.g. in [60–62]. Methods from statistical mechanics can be used, however, to obtain the scaling of the expected value $\mathbb{E}[\mathsf{M}]$ with the system size for a variety of disordered systems, see e.g. [63–70].

3.2 Basins

To each local minimum $\hat{x}$ there is an associated *basin* $\mathcal{B}(\hat{x})$. On manifolds it can be defined as the set of all $y \in X$ such that $\hat{x}$ is the ω-limit of the gradient dynamics $\dot{z} = -\nabla f(z)$ with initial condition y. In the graph case one can use the steepest descent algorithm instead: Starting with $z_0 = y$ we choose at each step the neighbor $z_{k+1} \in N(z_k)$, $f(z_{k+1}) < f(z_k)$ with the smallest fitness value and repeat the procedure until it terminates when $z_{k+1} = \hat{x}$ is a local minimum. The notion of a basin hence may become ambiguous when there is "local neutrality" in $N(x)$, i.e., if there are $x \in X$ and $y, y' \in N(x)$ with $f(y) = f(y')$. It is an open question how the basin should be defined in full generality, or what kind of structure on X must be required in order to properly define basins. It is not surprising that the distribution of basin sizes is crucial for the performance of simple optimization heuristics [71]. So far there does not appear to be a good method for estimating basin sizes beyond exhaustive enumeration or random sampling, however.

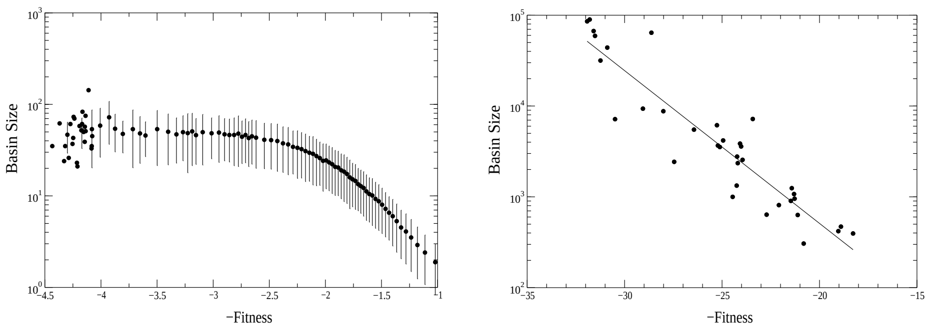

Fig. 2. Size distribution of basins of attraction for the SK-Model (quadratic Ising spin glass with i.i.d. Gaussian coefficients $a_{i_1 i_2}$ in equ.(20) and all other coefficients 0, $n = 20$), and a random assignment of Gaussian random numbers with mean 0 and variance 1 to the vertices of the 20-dimensional hypercube $\mathcal{Q}_2^{20}$. Error bars show the standard deviation of the distribution of basin sizes in a fixed fitness interval for the random Gaussian landscape, which has $\mathsf{M}_f = 49935$ local minima. The SK model on the r.h.s has only $\mathsf{M}_f = 70$ local minima and therefore much larger basins. The important observation is, however, that the basin size scales exponentially with fitness (energy) in this case.

An important aspect is the correlation between basin size and fitness of the minimum: In general, deeper minima have larger basins. Figure 2 shows that this is not only true for well-behaved landscapes such as the Sherrington-Kirkpatrick spin glass (r.h.s.), but also for random landscapes (l.h.s.). The difference is, however, that basin sizes appear to scale exponentially with fitness in well-behaved landscapes, while they approach a constant in essentially random landscapes.

3.3 Gradient Walks and Adaptive Walks

A simple measure for the size of a basin $\mathcal{B}(\hat{x})$ is the average length L of the steepest descent walks from $y \in \mathcal{B}(\hat{x})$ to $\hat{x}$. The average length L of a gradient walk has been investigated as a ruggedness measure in a few models, including random landscapes, Kauffman's NK landscapes [72], fitness landscapes derived from RNA folding [73].

An *adaptive walk* accepts a neighbor $x_{k+1} \in N(x_k)$ provided $f(x_{k+1}) < f(x_k)$ instead of looking for the steepest descent. Gillespie [74] suggested to use adaptive walks as models of evolutionary adaptation. They have been studied extensively in NK models [72,75,76], in particular in the context of the maturation of the immune response [77–79], in RNA folding landscapes [73] and in a model of early vascular land plants [80]. The lengths distribution of adaptive walks appears to be linked to the size distribution of the basins; the details of this connection, however, remain to be elucidated.

3.4 Barriers

The basins of local minima are separated by saddle points and fitness barriers. Let $\hat{x}$ and $\hat{y}$ be two local minima and let $\mathbf{p}$ be a path in X from $\hat{x}$ to $\hat{y}$. Then the fitness barrier separating $\hat{x}$ from $\hat{y}$ is

$$f[\hat{x},\hat{y}] = \min\left\{ \max\left[f(z)\middle| z\in\mathbf{p}\right] \,\middle|\, \mathbf{p} : \text{path from } \hat{x} \text{ to } \hat{y} \right\}, \tag{11}$$

A point $\hat{z} \in X$ satisfying the minimax condition in equ.(11) is a *saddle point* of the landscape. It should be noted that this definition is meaningful both in the graph case and on $\mathbb{R}^n$. However, in the context of potential energy surfaces one typically defines a saddle point as a critical point $\nabla f = 0$ that is neither a minimum nor a maximum. The saddle-point energies $f[\hat{x},\hat{y}]$ form an ultrametric distance measure on the set of local minima, see e.g. [81–83]. This hierarchical structure can be represented by the *barrier tree* of the landscape, Figure 3. Its leaves are the local minima and its internal nodes correspond to saddle points.

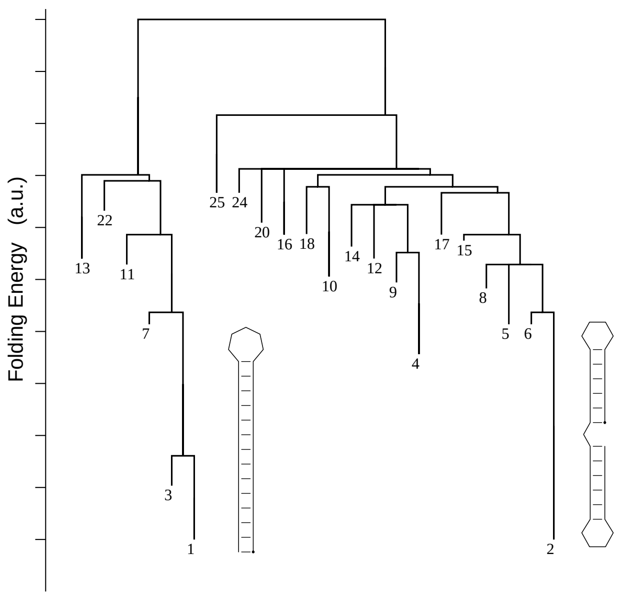

Fig. 3. Example of a barrier tree for the folding energy landscape of a bi-stable RNA molecule. The secondary structures of the two lowest energy states are indicated. For details see [84].

The *barrier* enclosing a local minimum is the height of the lowest saddle point that gives access to a more favorable minimum. In symbols:

$$B(\hat{x}) = \min\left\{ f[\hat{x},\hat{y}] - f(\hat{x}) \middle| \hat{y} : f(\hat{y}) < f(\hat{x}) \right\} \tag{12}$$

If $B(s) = 0$ then the local minimum s is degenerate. It is easy to check that eq.(12) is equivalent to the definition of the depth of a local minimum in [85]. For metastable states it agrees with the more general definition of the depth of a "cycle" in the literature on inhomogeneous Markov chains [86–88].

3.5 Depth

The information contained in the energy barriers is conveniently summarized by two global parameters that e.g. determine the convergence behavior of Simulated Annealing and related algorithms. Let Ω_f be the set of all global minima of f. Now consider the following two quantities

$$\begin{aligned} \mathsf{D} &= \max\left\{ B(s) \middle| s \notin \Omega_f \right\} \\ \psi &= \max\left\{ \frac{B(s)}{f(s) - f(\min)} \middle| s \notin \Omega_f \right\} \end{aligned} \tag{13}$$

Both parameters are easily obtained from the barrier tree. The *depth* D and *difficulty* ψ [85,87–90] play a crucial role in theory of Simulated Annealing. For instance, Simulated Annealing converges almost surely to a ground state if and only if the cooling schedule T_k satisfies $\sum_{k\geq 0} \exp(-\mathsf{D}/T_k) = \infty$ [89]. The difficulty parameter is directly related to the optimal speed of convergence of Simulated Annealing.

3.6 Correlation

Correlation measures are by the far the most accessible indicators of ruggedness. Weinberger [48] considers the autocorrelation function $r(s)$ of the "time series" of fitness values $f(x^{(t)})$ sampled along a random walk $\{x^{(0)}, \ldots, x^{(t)}, \ldots\}$ on X with transition matrix T and initial conditions distributed as p. In [91,92] distance dependent correlation functions $\rho(d)$ are considered, where d is a metric on X. The walk correlation function $r(s)$ of a landscape can be obtained without reference to the stochastic sampling process as [29]

$$r(s) = \langle f, \mathbf{T}^s f \rangle \tag{14}$$

The relationship between the walk correlation function $r(s)$ and the distance correlation function $\rho(d)$ is described in detail in [93] for highly symmetric transition operators. In many applications the correlation length

$$\ell = \sum_{s=0}^{\infty} r(s) \tag{15}$$

is used as a convenient measure of ruggedness, see e.g. [73,94].

The correlation length ℓ, the length L of gradient walks, and the expected number M of local optima appear to be closely related in "typical" landscapes. The notion of a "typical" landscape is made precise in [95]. Denoting by $X(x_0, \ell)$ the set of configurations that can be reached in at most ℓ applications of $\mathbf{T}$ from x_0 on a graph Γ, the "correlation length conjecture" [96] states that there should be roughly $\mathsf{M} \approx |X|/|X(x_0, \ell)|$ local optima. This estimate is based on the assumption that the correlation length determines the diameter of the large mountains and valleys, and that due to the high dimensional nature of typical search spaces each mountain typically contains only a small number of local optima. The correlation length conjecture has been tested on a variety of combinatorial optimization problems and appears to be a very good approximation [62,97].

4 Spectral Landscape Theory

Spectral approaches to fitness landscapes start with one of the symmetric non-negative operators on X discussed in section 2.2 above. The basic idea is to interpret the adjacency matrix $\mathbf{A}$ of a symmetric graph, or the operator $\mathbf{S}$, as a representation of the configuration space and to discuss the fitness function in terms of the regularities of $\mathbf{S}$. From an algebraic point of view it appears to be more natural to start with a discrete *Laplace operator*

$$-\boldsymbol{\Delta} = \mathbf{D_S} - \mathbf{S} \qquad \text{with } (\mathbf{D_S})_{xx} = \sum_{y\in X} \mathbf{S}_{xy} \tag{16}$$

since it has number of desirable mathematical properties:
$-\boldsymbol{\Delta}$ is symmetric and has non-positive off-diagonal entries.
$-\boldsymbol{\Delta}$ has 0 as an eigenvalue with eigenvector $\mathit{1} = (1, \dots, 1)$. The eigenvalue 0 is unique if and only if the graph associated with the off-diagonal entries is irreducible.
$-\boldsymbol{\Delta}$ is non-negative definite. The graph Laplacian arises naturally as the discretization of the Laplacian differential operator for instance in finite element computations. For recent surveys on graph Laplacians see [98–101].

Let $\{\varphi_k\}$ be an orthonormal basis of eigenvectors of $-\boldsymbol{\Delta}$. Of course, we can interpret φ_k as a fitness function on X, hence we use the "function" notation $\varphi_k(x)$ for the coordinate of φ_k indexed by x. It appears natural to expand a fitness function f into a *Fourier series*

$$f(x) = \sum_k a_k \varphi_k(x) \tag{17}$$

On so-called quasi-abelian Cayley graphs, that is, Cayley graphs for which the generator set is a union of conjugacy classes of the underlying group, this graph-theoretical Fourier series and the group theoretical Fourier transformation [102] coincide (apart from a different conventional normalization), see [103].

Since $-\boldsymbol{\Delta}$ is symmetric one can of course choose the basis functions φ_k to be real valued. In many instances it is much more convenient, however, to allow for complex valued eigenfunctions. For instance, the basis functions for the Hamming graph $\mathcal{Q}_\alpha^n$ on the α-letter alphabet $\mathcal{A} = \{0, \dots, \alpha - 1\}$ can be written in the form

$$\varphi_a(x) = \alpha^{-n/2} \prod_{k=1}^{n} \exp\left(2\pi\iota \frac{a_k x_k}{\alpha}\right) \tag{18}$$

for each index $a \in \mathcal{A}^n$. A real-valued basis for this case is described e.g. in [28]. The corresponding Laplacian eigenvalue is $\Lambda_a = \alpha \mathrm{ord}(a)$, where

$$\mathrm{ord}(a) = |\{k | a_k \neq 0\}| \tag{19}$$

can be interpreted as the *interaction order* of the eigenfunction φ_a. This notion becomes more intuitive by considering Ising spin models. The most general spin glass Hamiltonian is

$$f(x) = a_0 + \sum_{p=1}^{n} \sum_{i_1 < i_2 < \dots < i_p} a_{i_1 i_2 \dots i_p} x_{i_1} x_{i_2} \dots x_{i_p} \tag{20}$$

with Ising spins $x_j = \pm 1$. In other words $f(x)$ is a superposition of p-spin models, where $p = \mathrm{ord}(i_1, \dots, i_p)$ is the interaction order. In fact, the Fourier basis on the hypercube *are* the (normalized) Walsh functions

$$\varphi_I(x) = 2^{-n/2} \prod_{i \in I} x_i = 2^{-n/2} x_{i_1} x_{i_2} \dots x_{i_p} \tag{21}$$

with the index set $I = \{i_1, i_2, \dots, i_p\}$, and hence $p = \mathrm{ord}(I) = |I|$. The standard way of specifying a spin glass model therefore is its Fourier expansion (17). In the following we shall assume without loosing generality that the landscape is normalized such that

$$\begin{aligned} \overline{f} &= \sum_{x \in X} p(x) f(x) = 0 \\ \mathrm{Var}[f] &= \sum_{x \in X} p(x)(f(x) - \overline{f})^2 = \langle f, f \rangle_p - \overline{f}^2 = 1 \end{aligned} \tag{22}$$

where $p(x)$ is the stationary distribution of the transition operator in questions, see section 2.2. Thus we may assume $a_0 = 0$ in equ.(20).

Walsh functions, equ.(21), are used extensively in the analysis of Genetic Algorithms [42,105–109]. It is shown in [43,52] that the Walsh functions are also eigenvectors of the crossover transition matrices $\mathbf{S}^\chi$ defined in equ.(10) with uniform population distribution $\wp(z) = 1/|X|$. The corresponding eigenvectors are

$$\lambda_a^\chi = \begin{cases} 1 & \text{if } \tilde{a} = \emptyset \\ 1/2 & \text{if } \tilde{a} \neq \emptyset \quad \text{and } \tilde{a} \subseteq \boldsymbol{\chi} \text{ or } \tilde{a} \subseteq \overline{\boldsymbol{\chi}} \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

Table 1. Some Elementary Landscapes.
For a detailed discussion see section 4 of [29]. The *order* $\mathrm{ord}(\Lambda)$ of an eigenspace is its position in the spectrum of $-\boldsymbol{\Delta}$, not counting multiplicities and starting to count with 0 for the "flat landscape" with $\Lambda = 0$. For strings the order is given by equ.(19).

Problem	Graph	Degree	Λ	$\mathrm{ord}(\Lambda)$
p-spin glass	$\mathcal{Q}_2^n$	n	$2p$	p
NAES$^{(1)}$	$\mathcal{Q}_2^n$	n	4	2
Weight Partitioning	$\mathcal{Q}_2^n$	n	4	2
Graph α-Coloring	$\mathcal{Q}_2^\alpha$	$(\alpha-1)n$	2α	2
XY-spin glass	$\mathcal{Q}_\alpha^n$	$(\alpha-1)n$	2α	2
for $\alpha > 2$:	$\mathcal{C}_\alpha^n$	2	$8\sin^2(\pi/\alpha)$	2
Linear Assignment	$\Gamma(\mathcal{S}_n,\mathcal{T})$	$n(n-1)/2$	n	1
TSP symmetric	$\Gamma(\mathcal{S}_n,\mathcal{T})$	$n(n-1)/2$	$2(n-1)$	2
	$\Gamma(\mathcal{S}_n,\mathcal{J})$	$n(n-1)/2$	n	2
	$\Gamma(\mathcal{A}_n,\mathcal{C}_3)$	$n(n-1)(n-2)/6$	$(n-1)(n-2)$	?
antisymmetric	$\Gamma(\mathcal{S}_n,\mathcal{T})$	$n(n-1)/2$	$2n$	3
	$\Gamma(\mathcal{S}_n,\mathcal{J})$	$n(n-1)/2$	$n(n+1)/2$	$\mathcal{O}(n)$
Graph Matching	$\Gamma(\mathcal{S}_n,\mathcal{T})$	$n(n-1)/2$	$2(n-1)$	2
Graph Bipartitioning	$J(n,n/2)$	$n^2/4$	$2(n-1)$	2

$^{(1)}$Not-All-Equal-Satisfiability, see [104].

where $\tilde{a} = \{k|a_k \neq 0\}$ is the set of non-zero indices for the corresponding eigenfunction φ_a. This fact allows a direct comparison of the landscapes formed by the same fitness function for a variety of crossover and mutation operators. The bottom line of such an analysis is that fitness functions with low interaction order look smoother with mutation, while recombination appears to be favorable for high interaction orders. For the details we refer to [42,43,52].

The usefulness of the spectral approach is by no means limited to Walsh functions. Its general applicability is established by the following observations: (1) The landscapes of many of the most studied combinatorial optimization problems are *elementary*, i.e., their normalized fitness functions, equ.(22), are eigenvectors of the graph Laplacian, when X is organized according to the most natural move sets [29,110,111]. Examples include the Traveling Salesman Problem, Graph Bipartitioning, certain Satisfiability problems, Graph Coloring with a fixed number of colors, see Table 1. Furthermore, most of the examples belong to the 2nd non-zero eigenvalue of the Laplacian, see the last column of Table 1. Not all of the "classical" landscape are elementary, of course, but many of the

non-elementary ones have non-zero projections to only a few eigenspaces with small eigenvalues; an example is the Quadratic Assignment Problem [103,112] or the asymmetric TSP [29].
(2) Eigenvectors of graph Laplacians have local minima and maxima that are well-separated on the fitness scale. Lov Grover [110] showed that for any local minimum $\hat{x}$ and any local maximum $\hat{y}$ one has

$$f(\hat{x}) \leq \overline{f} \leq f(\hat{y}) \tag{24}$$

(3) Eigenvectors of graph Laplacians satisfy a discrete version of Courant's nodal domain theorem [113]. A nodal domain of $f : X \to \mathbb{R}$ is a maximal connected subset of X such that f does not change sign. Suppose the eigenvalues of $-\boldsymbol{\Delta}$ arranged in ascending order $0 = \Lambda_1 < \Lambda_2 \leq \cdots \leq \Lambda_{|X|}$, counting multiplicities. Note that the "ground state" Λ_1 corresponds to the *flat* landscape. On Hamming graphs we have therefore $\mathrm{ord}(\Lambda_k) = k - 1$.
The nodal theorem now states that if $-\Delta f = \Lambda_k f$, then f has at most k nodal domains, [114–118]. Thus landscapes may have more nodal domains, and hence more "mountain massives" if they belong to larger eigenvalues of the Laplacian. We remark that the nodal domain theorem holds for the class of so-called discrete Schrödinger operators which includes the symmetric transition matrices discussed in section 2.2.

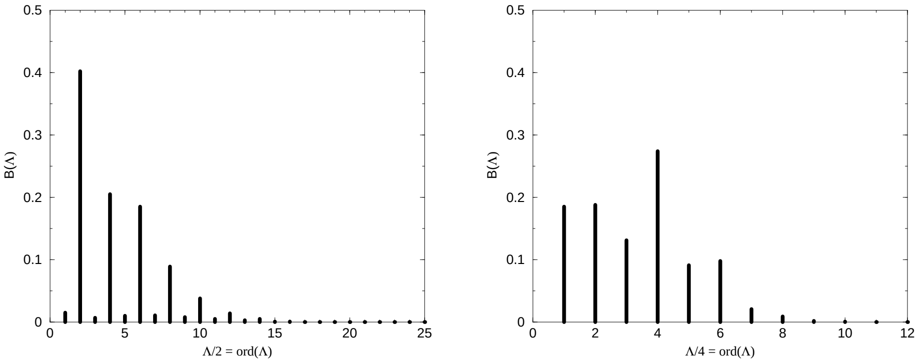

Fig. 4. Amplitude Spectra of RNA folding energy landscapes for **GC** sequences of length $n = 25$ and **GCAU**-sequences of length $n = 12$. The amplitude spectra are computed from explicit FFTs of the landscape as described in [103]. Note that the **GC** landscape has large amplitudes $B(\Lambda)$ only for even interaction orders, while the **GCAU**-landscape also contains a substantial linear, $\mathrm{ord}(a) = 1$, component. For a discussion of the biophysical reasons see [119,120].

Given an arbitrary landscape, we can measure the importance of a particular eigenspace of $-\Delta$ by means of the *amplitude spectrum*

$$B(\Lambda) = \sum_{k: -\Delta\varphi_k = \Lambda\varphi_k} |a_k|^2 \tag{25}$$

using $\sum_k |a_k|^2 = 1$ for normalized landscapes. Thus we have $B(\Lambda) \geq 0$ and $\sum_{\Lambda \neq 0} B(\Lambda) = 1$. For regular graphs, or bi-stochastic transition matrices, the Laplacian (16) and $\mathbf{T}$ have the same eigenvectors. In this case we can express random walk correlation functions and correlation lengths in terms of the amplitude spectrum

$$r(s) = \sum_{\Lambda \neq 0} B(\Lambda)\,(1 - \Lambda/d)^s \qquad \text{and} \qquad \ell = d \sum_{\Lambda \neq 0} \frac{B(\Lambda)}{\Lambda} \tag{26}$$

where $d = (\mathbf{D_S})_{xx}$ for all $x \in X$. The amplitude spectrum, or an aggregate parameter such as the average eigenvalue $\Lambda^* = \sum_\Lambda B(\Lambda)\Lambda$, thus may serve as an alternative measure of ruggedness. It is interesting to note that Davidor's "epistasis variance" [121] corresponds to $\sum_{\Lambda > \Lambda_2} B(\Lambda)$, while $B(\Lambda_2)$ measures the linear (additive) part of a landscape defined on a set of strings. Elementary landscapes belonging to Λ_3, or equivalently, $\mathrm{ord}(\Lambda) = 2$ on Hamming graphs, thus belong to the simplest class of landscapes with epistasis. In biology, epistasis is the interaction between genes such that the contribution of a gene to the fitness depends on the value of other genes in the chromosome [122,123].

5 Concluding Remarks

The present contribution is by no means an exhaustive survey of fitness landscapes. There is a great number of topics that have not been discussed here. Most importantly, we have excluded a thorough discussion of ensembles of landscapes and all the topics associated with them such as statistical mechanics methods, see e.g. [3], the notion of isotropic ensembles of landscapes [95], a stochastic treatment of neutrality [124], or the random graph approach to neutral landscapes [125–129].

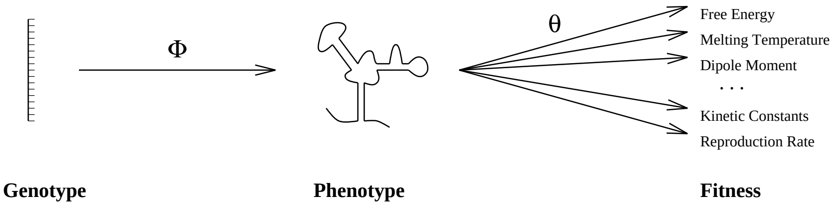

Fig. 5. Biologically realistic fitness landscapes are compositions of the genotype-phenotype map Φ and the evaluation θ of the phenotype by the environment.

The second topic that we have not touched yet is the structure of *biologically* realistic fitness landscapes, which typically can be viewed as the composition of a genotype-phenotype map and the fitness evaluation of the phenotype, Figure 5. Such landscapes inherit many of their properties, including ruggedness and neutrality, essentially from the genotype-phenotype map Φ [130]. Genotype-Phenotype maps have been studied extensively for RNA molecules. In this model the RNA sequence serves as genotype, while the secondary structure approximates the phenotype [19,131,132].

Another important topic concerns the connection of Genetic Algorithm and landscape structure. Schemata, i.e., hyperplanes in $\mathcal{Q}_a^n$ play an important role here [105–109,133–135]. For a discussion of the Schema Theorem and the Building Block Hypothesis we refer to the literature [16,136–140]. The fitness function $f : X \to \mathbb{R}$ can be extended in a natural way to subsets of X by setting

$$f(A) = \frac{1}{|A|} \sum_{x \in A} f(x) \tag{27}$$

If A is a schema, then $f(A)$ is the schema-fitness. A variety of landscape classes can be defined in terms of schema fitnesses, most notably *deceptiveness* of landscape [141,142]. Some of these are compared in [143,144]. The impact of properties such as ruggedness, neutrality, deceptiveness, isotropy, etc. on the performance of particular optimization strategies is the subject of ongoing research [112,145–147].

Acknowledgements

Thanks to Ivo Hofacker and Bärbel Stadler for useful discussion. Partial support by the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, Proj. No. P14094-MAT is gratefully acknowledged.

References

1. S. Wright: "The roles of mutation, inbreeding, crossbreeeding and selection in evolution", in *International Proceedings of the Sixth International Congress on Genetics*, ed. by D.F. Jones, Vol. 1 (1932), pp. 356–366
2. K. Binder, A.P. Young: Rev. Mod. Phys. **58**, 801–976 (1986)
3. M. Mézard, G. Parisi, M. Virasoro: *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987)
4. H.S. Chan, K.A. Dill: J. Chem. Phys. **95**, 3775–3787 (1991)
5. K.A. Dill, S. Bromberg, K. Yue, K.M. Fiebig, D.P. Yeo, P.D. Thomas, H.S. Chan: Prot. Sci. **4**, 561–602 (1995)
6. J.N. Onuchic, H. Nymeyer, A.E. Garcia, J. Chahine, N.D. Socci: Adv. Protein Chem. **53**, 87–152 (2000)
7. C. Flamm, I.L. Hofacker, P.F. Stadler: Adv. Complex Syst. **2**, 65–90 (1999)
8. C. Flamm, W. Fontana, I. Hofacker, P. Schuster: RNA **6**, 325–338 (2000)
9. P.G. Mezey: *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987)
10. D. Heidrich, W. Kliesch, W. Quapp: *Properties of Chemically Interesting Potential Energy Surfaces*, Vol. 56 of Lecture Notes in Chemistry (Springer-Verlag, Berlin, 1991)
11. K. Kollman, J.H. Miller, S.E. Page: Amer. Pol. Sci. Rev. **86**, 929–937 (1992)
12. B.M. Stadler: Adv. Complex Syst. **2**, 101–116 (1999)
13. M. Garey, D. Johnson: *Computers and Intractability. A Guide to the Theory of $\mathcal{NP}$ Completeness* (Freeman, San Francisco, 1979)

14. I. Rechenberg: *Evolutionstrategie* (Frommann-Holzboog, Stuttgart, 1973)
15. J.R. Koza: *Genetic Programming: On the Programming of Computers by Means of Natural Selection* (MIT Press, Cambridge, MA, 1992)
16. J.H. Holland: *Adaptation in Natural and Artificial Systems* (MIT Press, Cambridge, MA, 1993)
17. S.A. Kauffman: *The Origin of Order* (Oxford University Press, New York, Oxford, 1993)
18. D.B. Fogel: *Evolutionary Computation* (IEEE Press, New York, 1995)
19. P. Schuster, W. Fontana, P.F. Stadler, I.L. Hofacker: Proc. Roy. Soc. Lond. B **255**, 279–284 (1994)
20. A.S. Perelson, S.A. Kauffman (Eds.): *Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System* (Addison-Wesley, Reading, MA, 1991) (1991)
21. H. Frauenfelder, A.R. Bishop, A. Garcia, A. Perelson, P. Schuster, D. Sherrington, P.J. Swart (Eds.): *Landscape Paradigms in Physics and Biology: Concepts, Structures, and Dynamics* (Elsevier, Amsterdam, 1997), special Issue of *Phyica D* vol. 107(2-4)
22. J.P. Crutchfield, P. Schuster (Eds.): *Evolutionary Dynamics — Exploring the Interplay of Selection, Neutrality, Accident, and Function* (Oxford Univ. Press, Oxford UK, 2000), to appear
23. J. Hofbauer, K. Sigmund: *Dynamical Systems and the Theory of Evolution* (Cambridge University Press, Cambridge U.K., 1988)
24. R. Bürger: *The Mathematical Theory of Selection, Recombination, and Mutation* (John Wiley & Sons, Chichester UK, 2000)
25. L. Altenberg, M.W. Feldman: Genetics **117**, 559–572 (1987)
26. Y.I. Lyubich: *Mathematical structures in population genetics* (Springer-Verlag, Berlin, 1992)
27. J.W. Weibull: *Evolutionary Game Dynamics* (MIT Press, Cambridge MA, 1996)
28. A. Dress, D. Rumschitzki: Acta Appl. Math. **11**, 103–111 (1988)
29. P.F. Stadler: J. Math. Chem. **20**, 1–45 (1996)
30. J.S. Farris: "The logical basis of phylogenetic analysis", in *Advances in Cladistics*, ed. by N.I. Platnick, V.A. Funk (Columbia University Press, New York, 1983), pp. 1–36
31. J. Felsenstein: J. Mol. Evol. **17**, 368–376 (1981)
32. L.R. Foulds, R.L. Graham: Adv. Appl. Math. **3**, 43–49 (1982)
33. W.H.E. Day, D.S. Johnson, D. Sankoff: Math. Biosci. **81**, 33–42 (1986)
34. D.F. Robinson: J. Combin. Theory B **11**, 105–119 (1971)
35. M. Li, J. Tromp, L. Zhang: J. Theor. Biol. **182**, 463–467 (1996)
36. O. Bastert, D. Rockmore, P.F. Stadler, G. Tinhofer: (2000), "Landscapes on spaces of trees", submitted, SFI preprint 01-01-006
37. J. Kececioglu, D. Sankoff: Algorithmica **13**, 180–210 (1995)
38. D. Sankoff, G. Sundaram, J. Kececioglu: Internatl. J. Foundations Computer Sci. **7**, 1–9 (1996)
39. J.C. Culberson: Evol. Comp. **2**, 279–311 (1995)
40. T. Jones: (1995), "One operator, one landscape", Tech. Rep. #95-02-025, Santa Fe Institute
41. P. Gitchoff, G.P. Wagner: Complexity **2**, 47–43 (1996)
42. P.F. Stadler, G.P. Wagner: Evol. Comp. **5**, 241–275 (1998)
43. G.P. Wagner, P.F. Stadler: "Complex adaptations and the structure of recombination spaces", in *Algebraic Engineering*, ed. by C. Nehaniv, M. Ito (World Scientific, Singapore, 1999), pp. 96–115, proceedings of the Conference on Semi-Groups and Algebraic Engineering, University of Aizu, Japan

44. P. Larrañaga, C.M.H. Kuijpers, R.H. Murga, I. Inza, S. Dizdarevic: Articial Intelligence Review **13**, 129–170 (1999)
45. A. Brouwer, A. Cohen, A. Neumaier: *Distance-regular Graphs* (Springer Verlag, Berlin, New York, 1989)
46. M. Shpak, G.P. Wagner: Artificial Life **6**, 25–43 (2000)
47. L. Lovasz: "Random walks on graphs: A survey", in *Combinatorics, Paul Erdős is Eighty*Vol. 2Budapest (1996), pp. 353–398, keszthely (Hungary) 1993
48. E.D. Weinberger: Biol. Cybern. **63**, 325–336 (1990)
49. E.D. Weinberger: Biol. Cybern. **65**, 321–330 (1991)
50. W. Hordijk: Evolutionary Computation **4**(4), 335–360 (1996)
51. W. Hordijk: Physica D **107**, 255–264 (1997)
52. P.F. Stadler, R. Seitz, G.P. Wagner: Bull. Math. Biol. **62**, 399–428 (2000)
53. H.R. Fischer: Math. Annalen **137**, 269–303 (1959)
54. D.C. Kent: Fund. Math. **54**, 125–133 (1964)
55. E. Čech: *Topological Spaces* (Wiley, London, 1966)
56. D.C. Kent: Fund. Math. **62**, 95–100 (1968)
57. B.M.R. Stadler, P.F. Stadler, W. Fontana, G.P. Wagner: J. Theor. Biol. (2001). Submitted, SFI preprint 00-12-070
58. B.M.R. Stadler, P.F. Stadler, M. Shpak, G.P. Wagner: Artificial Life (2001). Submitted, SFI preprint 01-02-011
59. R. Palmer: "Optimization on rugged landscapes", in *Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System*, ed. by A.S. Perelson, S.A. Kauffman (Addison Wesley, Redwood City, CA, 1991), pp. 3–25
60. P.F. Stadler, W. Schnabl: Phys. Lett. A **161**, 337–344 (1992)
61. P.F. Stadler, B. Krakhofer: Rev. Mex. Fis. **42**, 355–363 (1996)
62. R. García-Pelayo, P.F. Stadler: Physica D **107**, 240–254 (1997)
63. D.J. Thouless, P.W. Anderson, R.G. Palmer: Phil. Mag. **35**, 593–601 (1977)
64. F. Tanaka, S.F. Edwards: J. Phys. F **10**, 2769–2778 (1980)
65. A.J. Bray, M.A. Moore: J. Phys. C **14**, 1313–1327 (1981)
66. D.J. Gross, M. Mèzard: Nucl. Phys. B **240**, 431–452 (1984)
67. B. Derrida, E. Gardner: J. Physique **47**, 959–965 (1986)
68. H. Rieger: Phys. Rev. B **46**, 14 655–14 661 (1992)
69. V.M. de Oliveira, J.F. Fontanari, P.F. Stadler: J. Phys. A: Math. Gen. **32**, 8793–8802 (1999)
70. F.F. Ferreira, J.F. Fontanari, P.F. Stadler: J. Phys. A: Math. Gen. **33**, 8635–8647 (2000)
71. J. Garnier, L. Kallel: SIAM J. Discr. Math. (2000). Submitted
72. E.D. Weinberger: Phys. Rev. A **44**, 6399–6413 (1991)
73. W. Fontana, P.F. Stadler, E.G. Bornberg-Bauer, T. Griesmacher, I.L. Hofacker, M. Tacker, P. Tarazona, E.D. Weinberger, P. Schuster: Phys. Rev. E **47**, 2083 – 2099 (1993)
74. J.H. Gillespie: Evolution **38**, 1116–1129 (1984)
75. S.A. Kauffman, S. Levin: J. Theor. Biol. **128**, 11–45 (1987)
76. H. Flyvbjerg, B. Lautrup: Phys. Rev. A **46**, 6714–6723 (1992)
77. C.A. Macken, A.S. Perelson: Proc. Natl. Acad. Sci. USA **86**, 6191–6195 (1989)
78. C.A. Macken, P.S. Hagan, A.S. Perelson: SIAM J. Appl. Math. **51**, 799–827 (1991)
79. A.S. Perelson, C.A. Macken: Proc. Natl. Acad. Sci. USA **92**, 9657–9661 (1995)
80. K.J. Niklas: Amer. J. Botany **84**, 16–25 (1997)
81. R. Rammal, G. Toulouse, M.A. Virasoro: Rev. Mod. Phys. **58**, 765–788 (1986)

82. A.M. Vertechi, M.A. Virasoro: J. Phys. France **50**, 2325–2332 (1989)
83. S.R. Morgan, P.G. Higgs: J. Phys. A **31**, 3153–3170 (1998)
84. C. Flamm, I.L. Hofacker, S. Maurer-Stroh, P.F. Stadler, M. Zehl: RNA **7**, 254–265 (2001)
85. W. Kern: Discr. Appl. Math. **43**, 115–129 (1993)
86. R. Azencott: *Simulated Annealing* (John Wiley & Sons, New York, 1992)
87. O. Catoni: Ann. Probab. **20**, 1109–1146 (1992)
88. O. Catoni: "Simulated annealing algorithms and Markov chains with rate transitions", in *Seminaire de Probabilites XXXIII*, ed. by J. Azema, M. Emery, M. Ledoux, M. Yor (Springer, Berlin/Heidelberg, 1999), Vol. 709 of Lecture Notes in Mathematics, pp. 69–119
89. B. Hajek: Math. Operations Res. **13**, 311–329 (1988)
90. J. Ryan: Discr. Appl. Math. **56**, 75–82 (1995)
91. G.B. Sorkin: (1988), "Combinatorial optimization, simulated annealing, and fractals", Tech. Rep. RC13674 (No.61253), IBM Research Report
92. M. Eigen, J. McCaskill, P. Schuster: Adv. Chem. Phys. **75**, 149 – 263 (1989)
93. P.F. Stadler: Discr. Math. **145**, 229–238 (1995)
94. P. Schuster, P.F. Stadler: Computers & Chem. **18**, 295–314 (1994)
95. P.F. Stadler, R. Happel: J. Math. Biol. **38**, 435–478 (1999)
96. P.F. Stadler, W. Schnabl: Phys. Letters A **161**, 337–344 (1992)
97. B. Krakhofer, P.F. Stadler: Europhys. Lett. **34**, 85–90 (1996)
98. B. Mohar: "The Laplacian spectrum of graphs", in *Graph Theory, Combinatorics, and Applications*, ed. by Y. Alavi, G. Chartrand, O. Ollermann, A. Schwenk (John Wiley and Sons, Inc., New York, 1991), pp. 871–898
99. R. Merris: Lin. Alg. Appl. **39**, 19–31 (1995)
100. F.R.K. Chung: *Spectral Graph Theory*, Vol. 92 of CBMS (American Mathematical Society, Providence RI, 1997)
101. B. Mohar: "Some applications of Laplace eigenvalues of graphs", in *Graph Symmetry: Algebraic Methods and Applications*, ed. by G. Hahn, G. Sabidussi (Kluwer, Dordrecht, 1997), Vol. 497 of NATO ASI Series C, pp. 227–275
102. D. Rockmore: "Some applications of generalized FFTs", in *Groups and Computation II*, ed. by L. Finkelstein, W. Kantor (American Mathmatical Society, Providence, RI, 1995), Vol. 28 of DIMACS, pp. 329–370
103. D. Rockmore, P. Kostelec, W. Hordijk, P.F. Stadler: Appl. Comput. Harmonic Anal. (2000). In press Santa Fe Institute preprint 99-10-068
104. T.J. Schaefer: "The complexity of satisfiability problems", in *Proceedings of the 10th Annual ACM Symposium on Theory of Computing*, ed. by N.N. (Association for Computing Machinery, New York, 1978), pp. 216–226
105. D.E. Goldberg: Complex Systems **3**, 129–152 (1989)
106. D.E. Goldberg: Complex Systems **3**, 153–176 (1989)
107. S.E. Page, D.E. Richardson: Complex Systems **6**, 125–136 (1992)
108. M.D. Vose, A.H. Wright: Evol. Comp. **6**, 253–274 (1998)
109. M.D. Vose, A.H. Wright: Evol. Comp. **6**, 275–289 (1998)
110. L.K. Grover: Oper. Res. Lett. **12**, 235–243 (1992)
111. B. Codenotti, L. Margara: (1992), "Local properties of some np-complete problems", Tech. Rep. TR 92-021, International Computer Science Institute, Berkeley, CA
112. E. Angel, V. Zissimopoulos: Discr. Appl. Math. **99**, 261–277 (2000)
113. I. Chavel: *Eigenvalues in Riemannian Geometry* (Academic Press, Orlando Fl., 1984)

114. Y.C. de Verdière: Rendiconti di Matematica **13**, 433–460 (1993)
115. J. Friedman: Duke Math. J. **69**(3), 487–525 (1993)
116. H. van der Holst: (1996), "Topological and spectral graph characterizations", Ph.D. thesis, Universiteit van Amsterdam
117. A.M. Duval, V. Reiner: Lin. Alg. Appl. **294**, 259–268 (1999)
118. E.B. Davies, G.M.L. Gladwell, J. Leydold, P.F. Stadler: Lin. Alg. Appl. (2001). In press, see also: math.SP/0009120
119. R. Happel, P.F. Stadler: Complexity **2**, 53–58 (1996)
120. W. Hordijk, P.F. Stadler: J. Complex Systems **1**, 39–66 (1998)
121. Y. Davidor: Complex Systems **4**, 369–383 (1990)
122. M. Whitlock, P.C. Phillips, F.B.G. Moore, S. Tonsor: Ann. Review Ecol. Systematics **26**, 601–629 (1995)
123. J.B. Wolf, E.D. Brodie III, M.J. Wade (Eds.): *Epistasis and the Evolutionary Process* (Oxford Univ. Press, Oxford, UK, 2000)
124. C.M. Reidys, P.F. Stadler: Appl. Math. & Comput. **117**, 321–350 (2001)
125. C.M. Reidys: Adv. Appl. Math. **19**, 360–377 (1997)
126. S. Gavrilets, J. Gravner: J. Theor. Biol. **184**, 51–64 (1997)
127. C.M. Reidys, P.F. Stadler, P. Schuster: Bull. Math. Biol. **59**, 339–397 (1997)
128. S. Gavrilets, H. Li, M.D. Vose: Proc. Roy. Soc. London B **265**, 1483–1489 (1998)
129. S. Kopp, C.M. Reidys: Adv. Complex Syst. **2**, 283–301 (1999)
130. P.F. Stadler: J. Mol. Struct. (THEOCHEM) **463**, 7–19 (1999)
131. W. Fontana, D.A.M. Konings, P.F. Stadler, P. Schuster: Biopolymers **33**, 1389–1404 (1993)
132. W. Fontana, P. Schuster: Science **280**, 1451–1455 (1998)
133. D.E. Goldberg, M. Rudnik: Complex Systems **5**, 265–278 (1991)
134. G.E. Liepins, M.D. Vose: Complex Systems **5**, 45–61 (1991)
135. P. Field: Complex Systems **9**, 11–28 (1995)
136. L. Altenberg: "The schema theorem and the Price's theorem", in *Foundations of Genetic Algorithms 3*, ed. by L.D. Whitley, M.D. Vose (Morgan Kauffman, San Francisco CA, 1995), pp. 23–49
137. A.D. Bethke: (1991), "Genetic algorithms and function optimizers", Ph.D. thesis, University of Michigan
138. J.H. Holland: "Genetic algorithms and classifier systems: foundations and future directions", in *Proceedings of the 2nd International Conference on Genetic Algorithms* (1987), pp. 82–89
139. S. Forrest, M. Mitchell: "Relative building block fitness and the building block hypothesis", in *Foundations of Genetic Algorithms 2*, ed. by L.D. Whitley (Morgan Kaufmann, San Mateo, CA, 1993), pp. 109–126
140. C.R. Stephens, H. Waelbroeck: Phys. Rev. E **57**, 3251–3264 (1998)
141. L.D. Whitley: "Fundamental principles of deception in genetic search", in *Foundations of Genetic Algorithms*, ed. by G. Rawlins (Morgan Kaufmann, San Mateo, CA, 1991), pp. 221–241
142. K. Deb, D.E. Goldberg: "Analyzing deception in trap functions", in *Foundations of Genetic Algorithms 2*, ed. by L.D. Whitley (Morgan Kaufmann, San Mateo, CA, 1993), pp. 93–108
143. P.F. Stadler: "Spectral landscape theory", in *Evolutionary Dynamics—Exploring the Interplay of Selection, Neutrality, Accident, and Function*, ed. by J.P. Crutchfield, P. Schuster (Oxford University Press, New York, 2001), in press
144. B. Naudts, L. Kallel: IEEE Trans. Evol. Comp. (2000). To appear
145. M. Mitchell: *An Introduction to Genetic Algorithms* (MIT Press, Cambridge MA, 1996)

146. D.F. T. Baeck, Z. Michalewicz (Eds.): *Handbook of Evolutionary Computation* (Oxford University Press, New York, 1997)
147. E. Angel, V. Zissimopoulos: Theor. Computer Sci. **191**, 229–243 (1998)