Canonical Approximation of Landscapes

By

Peter F. Stadler[a,b,*] and Robert Happel[a]

[a]Institut für Theoretische Chemie der Universität Wien

[b]The Santa Fe Institute, Santa Fe, New Mexico

Completely Revised Version

*Mailing Address: Peter F. Stadler, Institut für Theoretische Chemie,
Universität Wien, Währingerstraße 17, A 1090 Wien, Austria
Phone: (**43 1) 40 480 678
Fax: (**43 1) 40 28 525
E-Mail: Peter.Stadler@tbi.univie.ac.at, stadler@santafe.edu

Abstract

Correlation functions are important characteristics of (fitness) landscapes. We use the Fourier decomposition of landscapes in order to characterize the set of all the possible autocorrelation functions on highly symmetric graphs, as well as the isotropic random fields on such graphs. A canonical approximation procedure is then proposed allowing empirical landscapes to be replaced by statistical models with the same correlation structure. This procedure makes use of elementary landscapes fulfilling an analogue of the Helmholtz equation. We show some applications to the random energy model, Kauffman's Nk models, the Traveling Salesman Problem, and RNA free energy landscapes.

Keywords

Landscapes, Combinatorial Optimization, Fourier Expansion, Isotropy, Correlation Functions

1. Introduction

Since Sewall Wright's seminal paper (Wright 1932) the notion of a *fitness landscape* underlying the dynamics of evolutionary optimization has proved to be one of the most powerful concepts in evolutionary theory. Implicit in this idea is a collection of genotypes arranged in an abstract metric space, with each genotype next to those other genotypes which can be reached by a single mutation, as well as a value assigned to each genotype. Such a construction is by no means restricted to biological evolution; Hamiltonians of disordered systems, such as spin glasses (Binder and Young 1986, Mézard *et al.* 1987), and the cost functions of combinatorial optimization problems (Garey and Johnson 1979) have the same structure. It has been known since Eigen's (1971) pioneering work on the molecular quasispecies that the dynamics of evolutionary adaptation (optimization) on a landscape depends crucially on detailed structure of the landscapes itself. Extensive computer simulations (see, e.g., Fontana and Schuster 1987, Fontana *et al.* 1989) have made it very clear that a complete understanding of the dynamics is impossible without a thorough investigation of the underlying landscape (Eigen *et al.* 1989).

The landscapes of a number of well known combinatorial optimization problems[1], such as the Traveling Salesman Problem `TSP`, the Graph Bipartitioning Problem `GBP`, or the Graph Matching Problem `GMP` have been investigated in some detail (Stadler and Schnabl 1992, Stadler and Happel 1992, Stadler 1992). The distribution of local optima and the statistical characteristics of down-hill walks have been computed for the uncorrelated landscape of the random energy model (Macken *et al.* 1991, Flyvbjerg and Lautrup 1993, Bak *et al.* 1993). Furthermore, two one-parameter families of tunably rugged landscapes have been studied extensively: the Nk model and its variants (Kauffman and Levin 1987, Kauffman *et al.* 1988,

[1]The most common combinatorial optimization problems are described in the appendix.

Weinberger 1991a, Fontana *et al.* 1993a), and the p-spin models (Amitrano *et al.* 1989, Weinberger and Stadler 1993). A detailed survey of a variety of model landscapes derived from folding RNA molecules into their secondary structures has been performed recently (Bonhoeffer *et al.* 1993, Fontana *et al.* 1987, 1989, 1991, 1993ab, Hofacker *et al.* 1994, Huynen and Hogeweg 1994, Schuster 1991, Schuster *et al.* 1994, Stadler and Grüner 1993, Tacker *et al.* 1994).

In this contribution we develop a theory of Fourier Series on the highly symmetric graphs that occur as configuration spaces underlying combinatorial optimization problems and evolutionary dynamics. It will be applied to both random fields (that is, stochastic models of landscapes) and individual landscapes. In section 2 we present technical, in particular graph theoretic preliminaries, and prove a few preparatory results. The mathematical language described here seems to be the appropriate basis for a general theory of (fitness) landscapes.

Section 3 contains a brief overview over *isotropic* random fields. A long the lines of Weinberger (1991b) we give a complete characterization of isotropic random fields on Cayley graphs of commutative groups. Then we briefly discuss the relation between the autocorrelation function of random field and the correlation function of the time series obtained from a random walk on the random field. As a surprising by-product we obtain a relation between eigenvalues and eigenvectors of the collapsed adjacency matrix of a fairly general class of vertex transitive graphs. Furthermore, we extend the class of graphs for which it is known that the Fourier basis of the graph coincides with the eigenvectors of the covariance matrix of isotropic random fields. These results will be used in section 5 for a discussion of the random energy model on vertex transitive graphs and for an analysis of Kauffman's Nk models.

The main result of this paper is theorem 8 in section 4, showing that *elementary* landscapes (which are eigenvectors of the graph Laplacian Δ) have autocorrelation functions which are left eigenvectors of the collapsed adjacency matrix of the

underlying configuration space, or equivalently, which exhibit an exponentially decaying correlation function along random walks on the landscapes. The latter property identifies them with Weinberger's AR(1) landscapes. While we have a complete proof of theorem 8 only for the special case of Cayley graphs of commutative groups there is ample evidence (in the form of additional examples) that it in fact holds for a much larger class of graphs, maybe even for all vertex transitive graphs. The relevance of theorem 8 stems from the fact that the landscapes of the best-studied combinatorial optimization problems, including the symmetric TSP, the graph bipartitioning problem, and the p-spin models, are elementary.

As a first application we obtain a refinement of the classification of landscapes proposed in Weinberger & Stadler (1993). In particular, we show that there are no "smooth" landscapes on sequence spaces (apart from the trivial case of "flat landscapes"), the least rugged case being the Fujijama-type landscapes which already has an exponentially decaying correlation function along random walks.

In section 6 we discuss a few examples of Fourier decompositions of well-known combinatorial optimization problems. These landscapes are either elementary or they are composed of a very small number of modes. We use the knowledge about these modes to explain the behavior of heuristic algorithms on these landscapes. An even more important application of our formalism is the *approximation* of given landscapes by (a superposition of) simpler (elementary) landscapes. In particular, computational studies of the evolutionary dynamics on RNA landscapes are very expensive in terms of computer time even on modern workstations, because of the computational efforts necessary to calculate the RNA structure. A numerical determination of the Fourier spectrum of RNA landscapes will therefore help to reduce the immense time and memory requirements for the studies of evolutionary dynamics (see also the numerical work by Amitrano *et al.* (1989) and Schober *et al.* (1993).) In the best of all worlds we could even hope to obtain analytically treatable models based on simple (spin glass) landscapes. In section 7 we summarize our results and discuss a few directions for future research.

2. Technical Preliminaries

2.1. Landscapes and Random Fields

Two features are common to all landscapes:

(1) A *cost function* or *fitness function* f assigns a value to each *configuration*,

$$f : \mathcal{C} \to \mathbb{R}, \tag{1}$$

e.g, an energy to each spin configuration, a fitness value to each RNA sequence, a length to each tour in a traveling salesman problem.

(2) There is a rule that determines whether two configurations x and y are nearest neighbors. This rule describes the elementary moves in an optimization heuristic and the elementary mutations in molecular evolution.

As a consequence of property (2), the set $\mathcal{C}$ of all configurations can be viewed as a graph with each vertex corresponding to a configuration and an edge between any two nearest neighbors. Therefore, there is a natural metric distance $d(x, y)$ between any two configurations given by the shortest path in the graph $\mathcal{C}$.

Many of the model landscapes mentioned in the introduction contain a stochastic element: a particular instance is generated by assigning a usually large number of parameters at random. This procedure suggests the following

Definition. The set $\{f : \mathcal{C} \to \mathbb{R}\}$ together with a measure $\mu\{f\}$ forms the probability space Ξ, which we will call a *random field* on $\mathcal{C}$ (Besag, 1974).

The measure μ can be recast in the form $P(c_1, c_2, \ldots, c_N)$ which is the probability that for all configurations x_i holds simultaneously $f(x_i) < c_i$, where $c_i \in \mathbb{R}$, and

N is the total number of configurations. Then the *expected value* of a random variable X defined on the random field is given by $\int X \, d\boldsymbol{\mu}$ which takes the form

$$\mathcal{E}[X] = \int_{\mathbb{R}^N} X \, \mathrm{d}P(c_1, \ldots, c_N) \, .$$

We will restrict the use of the term *landscape* to mappings $f : \mathcal{C} \to \mathbb{R}$. Therefore, an element of a random field on $\mathcal{C}$ is a landscape.

In may cases of practical importance, however, we are not given the random field, instead we have only a single instance (i.e., a landscape) and little or no clue from what kind of a random field it might have been drawn. This is the case, e.g., for the landscapes of RNA and protein folding, viral adaptation, and for some technical optimization problems such as the "low autocorrelated string problem" (Golay 1977). In this situation we have to be content with averages over values taken from our single instance. We will use the following notation: $\mathcal{E}[\,.\,]$ denotes averages over random fields, while angular brackets $\langle\ \rangle$ denote averages over configurations in a landscape[2]. For example we have

$$\mathrm{Var}[f(x)] = \mathcal{E}[f(x)^2] - \mathcal{E}[f(x)]^2 \qquad \text{and} \qquad \sigma^2 = \langle f^2 \rangle - \langle f \rangle^2,$$

where the first average is taken over all sample functions of the random field f evaluated at a *fixed position* $x \in \mathcal{C}$, while the second average runs over all vertices x in a *fixed instance* f. Of course the two variances $\mathrm{Var}[f(x)]$ and σ^2 are not the same in general. This can be seen, for instance, form the following (trivial)

Example. Consider a Gaussian random field with $\mathcal{E}[f(x)] = 0$ for all x and $\mathcal{E}[f(x)f(y)] = 1$ for all x and y. A sample function drawn from this random field is almost always constant since the correlation of any two points is 1. The constant value value $\bar{f}$ of this flat landscape, however, is drawn from a Gaussian

[2]This notation should not be confused with $\langle\,,\,\rangle$ which is used for scalar products lateron.

distribution with mean zero and unit variance. In fact: $\mathrm{Var}[f(x)] = 1$ for all x, while the empirical variance measured on a sample function is almost always $\sigma^2 = 0$ since almost all sample landscapes drawn form this random field are completely flat.

If, for a given quantity X, both averages give the same result in a particular model (at least in the limit of infinite system size), one says that X is a *self-averaging* quantity (Binder & Young 1986, Mèzard *et al.* 1987).

2.2. Vertex Transitive Graphs

Let Γ be a graph with $N < \infty$ vertices. A graph automorphism α is a one-to-one mapping of Γ onto itself such that the vertices $\alpha(x)$ and $\alpha(y)$ are connected by an edge if and only if x and y are connected by an edge. The set of all automorphism of a Graph Γ forms a group, the automorphism group $\mathbf{Aut}[\Gamma]$. Let V be the vertex set of Γ. Two vertices of Γ are said to be equivalent if there is an automorphism α such that $y = \alpha(x)$.

Definition. Γ is *vertex transitive* if $\mathbf{Aut}[\Gamma]$ acts transitively on V, i.e., if all vertices are equivalent.

A vertex transitive graph is always regular, i.e., each vertex has the same degree[3] D. In the main part of this paper we will need a more detailed description of the symmetry properties of the configuration spaces $\mathcal{C}$.

[3]The degree of a vertex is the number of its neighbors.

Denote by $\ddot{\mu}$ an orbit of $\mathbf{Aut}[\Gamma]$ on $V \times V$, i.e., a set of vertex pairs of the form

$$\ddot{\mu}(x,y) = \{(u,v) \in V \times V \,|\, \exists \alpha \in \mathbf{Aut}[\Gamma] : \, u = \alpha(x), v = \alpha(y) \,\}.$$

It will turn out to be very useful to consider the subgroup $\mathbf{A}_0 < \mathbf{Aut}[\Gamma]$ of all automorphisms that map a reference vertex 0 onto itself[4]. We will refer to the orbits of $\mathbf{A}_0$ acting on V as the symmetry classes of Γ. Of course, the symmetry classes of Γ and the orbits of the full automorphism group $\mathbf{Aut}[\Gamma]$ acting on $V \times V$ are closely related; we have a one-to-one correspondence via

$$\begin{aligned} \mu &= \{v \in V \,|\, (0,v) \in \ddot{\mu}\}, \\ \ddot{\mu} &= \{\, (u,v) \in V \times V \,|\, \exists \alpha \in \mathbf{Aut}[\Gamma] : \alpha(u) = 0 \,\wedge\, \alpha(v) \in \mu \,\} \,. \end{aligned} \tag{2}$$

The matrix $R^{(\mu)}$ associated with the orbit $\ddot{\mu}$, and hence with the symmetry class μ is defined by

$$R^{(\mu)}_{xy} = \begin{cases} 1 & \text{if} \quad (x,y) \in \ddot{\mu}(x,y) \\ 0 & \text{otherwise} \end{cases}$$

It will play a prominent role in the theory developed in this paper.

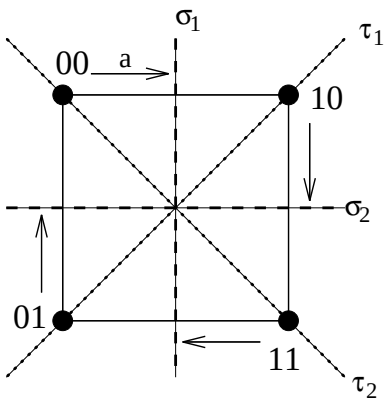


Figure 1: The square graph and its automorphisms. See text for details.

[4]The choice of the reference vertex in Γ is of course arbitrary.

Example. The definition of the symmetry classes μ and $\ddot{\mu}$ is crucial for the understanding of this contribution. As simple example consider the square graph $\mathcal{Q}_2^2$ shown in figure 1. Its automorphism group consists of the following 8 elements: the identity ι, the three rotations a, a^2, and a^3 of the square, and the four reflections τ_1, τ_2, σ_1, and σ_2. The subgroup $\mathbf{A}_0$ which leaves the vertex 00 invariant consists only of the identity ι and the reflections τ_2, which maps 01 to 10 and 10 to 01.

Therefore 01 and 10 belong to the same symmetry class. The vertex set of the square graph decomposes into the following symmetry classes:

$$\mu_0 = \{00\}, \qquad \mu_1 = \{01, 10\}, \qquad \mu_2 = \{11\}.$$

The corresponding classes of vertex pairs are

$$\ddot{\mu}_0 = \{(00,00), (01,01), (10,10), (11,11)\}$$

$$\ddot{\mu}_1 = \{(00,01), (00,10), (01,00), (01,11), (10,00), (10,11), (11,01), (11,10)\}$$

$$\ddot{\mu}_2 = \{(00,11), (01,10), (10,01), (11,00)\}$$

More complicated examples, including examples of graphs where the symmetry classes do not coincide with distance classes are given in section 2.10.

2.3. Cayley Graphs

An important subclass of vertex transitive graphs are the so-called *Cayley graphs* $\Gamma(G, \Phi)$ (see, e.g., Biggs and White 1979). They are constructed as follows:

Definition. Let $(G, \circ)$ be a finite group, and let Φ be a set of generators[5] of G such that the group identity is not contained in Φ, and for each $x \in \Phi$ the inverse

[5] Φ is a set of generators means that each group element $z \in G$ can be represented as a finite product of elements of Φ.

group element x^{-1} is also contained in Φ. Then the Cayley graph $\Gamma(G, \Phi)$ is the graph with vertex set G and an edge connecting two vertices x and y if and only if $xy^{-1} \in \Phi$, i.e., two vertices are connected if and only if there is a $g \in \Phi$ such that $y = gx$.

The set of generators Φ can be interpreted as the set of all possible elementary mutations, or — in the context of an optimization heuristic — as the move set of the algorithm. For Cayley graphs we have $(x, y) \in \ddot{\mu}$ if and only if $xy^{-1} \in \mu$.

An interesting relation between the structure of the group G and the symmetry of the Cayley graph $\Gamma(G, \Phi)$ is the following

Lemma 1. Denote by C_ℓ the conjugacy classes of G. Suppose $C_\ell \cap \Phi = \emptyset$ or $C_\ell \cap \Phi = C_\ell$, i.e., the set of generators consists of one or more conjugacy classes of the group G. Then

(i) The group of all conjugations $\tilde{\psi} : G \to G, x \mapsto \psi x \psi^{-1}$, with $\psi \in G$, is a subgroup of $\mathbf{A}_0$.

(ii) Let μ be a symmetry class of $\Gamma(G, \Phi)$. All conjugacy classes of G fulfil either $C_\ell \cap \mu = \emptyset$ or $C_\ell \cap \mu = C_\ell$, i.e., the conjugacy classes of the group G are contained in symmetry classes of the Cayley graph $\Gamma(G, \Phi)$.

Proof. Clearly $\tilde{\psi}(e) = e$ for all $\psi \in G$, thus the group of a conjugations forms a subgroup of $\mathbf{A}_0$ if we use e as reference vertex in $\Gamma(G, \Phi)$. The conjugation $\tilde{\psi}$ is a graph automorphism if and only of $xy^{-1} \in \Phi$ is equivalent with $\tilde{\psi}(x)[\tilde{\psi}(y)]^{-1} \in \Phi$. The latter expressions can be rewritten as

$$\tilde{\psi}(x)[\tilde{\psi}(y)]^{-1} = \psi x \psi^{-1} \psi y^{-1} \psi^{-1} = \tilde{\psi}(xy^{-1}).$$

The assumption on Φ implies immediately that conjugation is an automorphism of the Cayley graph. Therefore each conjugation class is contained in one of the symmetry classes of $\Gamma(G, \Phi)$. ∎

2.4. Distance Transitive Graphs

A graph is *distance transitive* if for any two pairs of vertices (x, y) and (u, v) with $d(x, y) = d(u, v)$ there is an automorphism α such that $\alpha(x) = u$ and $\alpha(y) = v$, i.e., if the sets $\{(x, y)|d(x, y) = d\}$ are orbits of $\mathbf{Aut}[\Gamma]$ on $V \times V$. (See, e.g., Buckley and Harrary 1990). Given an arbitrary reference point 0, the set of all vertices with common distance d will be referred to as a distance class.

The Cayley graph property and distance transitivity are independent properties: The Petersen graph, figure 4a in section 2.10, is distance transitive but it is not the Cayley graph of any group (see, e.g., Biggs and White 1979). Conversely, the trigonal prism, figure 4b in section 2.10, is clearly a Cayley graph. It is not distance transitive because a pair of vertices within a triangular face and a pair of vertices connecting the two triangular faces are obviously not equivalent.

2.5. The Collapsed Adjacency Matrix

A graph Γ is defined by its *adjacency matrix* A, with $A_{xy} = 1$ if the vertices x and y are connected by an edge, and $A_{xy} = 0$ otherwise.

Following Bollobás (1979) we may then partition the vertex set by the following procedure: Choose an arbitrary vertex as reference point, which in the following will be labeled 0. Then construct pairwise disjoint sets of vertices $\mu_0 = \{0\}, \mu_1, \mu_2, \ldots, \mu_M$ such that

$$\hat{A}_{\mu\nu} := \sum_{x \in \mu} A_{xy} \qquad \text{is independent of } y \in \nu. \tag{3}$$

This partitioning is induced by the symmetry classes of Γ: the vertices in each of the sets μ are permutated by automorphisms in $\mathbf{A}_0$, i.e., by those automorphisms that leave 0 invariant. In the worst case each set contains a single vertex.

For highly symmetric graphs the *collapsed adjacency matrix* $\hat{A} = (\hat{A}_{\mu\nu})$ can be much smaller than the adjacency matrix A. In case of distance transitive graphs, for instance, the distance classes form a partition fulfilling equ.(1), and the collapsed adjacency matrix $\hat{A}$ coincides with the intersection matrix as defined by Biggs (1974) (see also Cvetković *et al.* 1980). We will use d instead of Greek letters for indexing the distance classes of distance transitive graphs.

2.6. Adjacency Algebra of a Graph

The set of all polynomials in A forms the *adjacency algebra* $\mathcal{A}(\Gamma)$ of the graph Γ under the usual operations of addition and multiplication of matrices. Regarded as a complex vector space, this algebra has a finite dimension which is given by the number of distinct eigenvalues of the adjacency matrix A.

Definition. A *path* of length ℓ is a sequence of vertices $x_0, x_1, \ldots, x_\ell$ such that $d(x_{i-1}, x_i) = 1$ for $1 \leq i \leq \ell$.

A well known result, see, e.g., Biggs (1994), is the following

Lemma 2. The entry x, y in the s-th power of the adjacency matrix, A^s, equals the number of paths of length s from x to y.

Definition. A *simple random walk* (Spitzer 1976) on Γ starting at 0 is a Markov process with transition matrix

$$T = \frac{1}{D} A \tag{4}$$

and initial condition $p_0(x) = \delta_{x,0}$.
If Γ is vertex transitive, let us define the probability $\varphi_{s\mu}$ that a simple random walk of s steps ends in a vertex contained in the symmetry class μ.

Explicit expressions for $\varphi_{s\mu}$ can be found in Stadler (1994a), see also section 4.2.

Corollary. Suppose Γ is vertex transitive. Then

$$A^s = \sum_{\mu} \frac{D^s}{|\mu|} \varphi_{s\mu} R^{(\mu)}. \tag{5}$$

Proof. Of course the same number of paths ends in each vertex of a given symmetry class. There is a total of D^s paths of length s, of which $D^s \varphi_{s\mu}$ wind up in symmetry class μ, and therefore $D^s \varphi_{s\mu}/|\mu|$ end in a particular vertex in μ. ∎

The matrices $R^{(\mu)}$ form the basis of a vector space of dimension M. The coefficients

$$\Theta_{s\mu} = \frac{D^s \varphi_{s\mu}}{|\mu|} \tag{6}$$

form a $(M+1) \times (M+1)$ matrix which can be invertible only if the number of distinct eigenvectors of A equals the number of symmetry classes, i.e., if all eigenvalues of the collapsed adjacency matrix are simple.

Definition. We will say that a vertex transitive graph has the Θ-property, if all matrices $R^{(\mu)}$ are elements of the adjacency algebra $\mathcal{A}(\Gamma)$, i.e., if they can be written as polynomials of A.

It is well known that all distance transitive graphs have the Θ-property (Damerell, 1973). The examples at the end of this section suggest the following

Conjecture 1. Let Γ be vertex transitive and suppose the collapsed adjacency matrix has only simple eigenvalues. Then Γ has the Θ-property.

2.7. Spectra of Graphs

The set of eigenvalues of the adjacency matrix is called the spectrum of the graph Γ. Let e be an eigenvector of A with eigenvalue λ. Note that in our terminology $e(x)$ is a landscape. Then v defined by

$$v(\mu) = \sum_{x \in \mu} e(x) \tag{7}$$

is a *right* eigenvector of $\hat{A}$ with the same eigenvalue. If v is a right eigenvector, then ρ, defined by

$$\rho(\mu) = \frac{1}{|\mu|} v(\mu) \tag{8}$$

is a *left* eigenvector of $\hat{A}$. We use $|\mu|$ to denote the number of configurations in symmetry class μ. As shown in Stadler (1994a)

Lemma 3. Let $\hat{A}$ be the collapsed adjacency matrix of a vertex transitive graph Γ. Then there is always a basis $\{\rho_i\}$ of left eigenvectors of $\hat{A}$ with the following properties:

(i) The eigenvectors are normalized such that $\rho(0) = v(0) = 1$.

(ii) The left eigenvectors are orthogonal with respect to the scalar product

$$\langle a; b \rangle = \sum_{\mu} a(\mu) b(\mu) |\mu|. \tag{9}$$

This basis is unique if all eigenvalues of $\hat{A}$ are simple.

Proof. The main idea is to use that A is symmetric and then to use equ.(2). The details can be found in Stadler (1994a). ∎

For Cayley graphs of commutative groups (CGCGs) the eigenvectors and eigenvalues have a particularly simple form. Let $x = (x_1, \ldots, x_m)$ denote the component

wise representation of the group element x which is obtained from the decompositions of the commutative group G into a direct product of m cyclic groups of prime orders N_k. The group action is represented by component-wise addition modulo N_k:

$$x \circ y = (x_1 + y_1 \bmod N_1, x_2 + y_2 \bmod N_2, \ldots, x_m + y_m \bmod N_m).$$

The eigenvectors of a Cgcg are then of the form (Lovács 1975)

$$e_g(x) = \exp\left(2\pi i \sum_k \frac{x_k g_k}{N_k}\right), \tag{10}$$

and the corresponding eigenvalues are given by

$$\lambda_g = \sum_{x \in \Phi} e_g(x).$$

For convenience we recall a few basic properties of the functions $e_x(y)$.

(i) $e_g(x) = e_x(g)$.

(ii) $e_g(x \circ y) = e_g(x)e_g(y)$ and $e_{g \circ h}(x) = e_g(x)e_h(x)$.

(iii) $e_g(x)e_g^*(x) = 1$, where the asterisk denotes complex conjugation.

(iv) $\sum_{x \in G} e_g(x)e_h^*(x) = N\delta_{gh}$ and $\sum_{l \in G} e_l(x)e_l^*(y) = N\delta_{xy}$, where N is the total number of configurations.

Remark. The landscapes $e_g(x)$ are of course the characters of the group G. Except for property (ii), which will be crucial for some results in the following, the above statements are immediate consequences of Schur's Lemma, and are therefore true for the characters of arbitrary groups (see, e.g., Serre 1977).

2.8. Incidence Matrix and Graph Laplacian

In this section we will introduce the notion of the Laplacian matrix of a graph and briefly discuss its most prominent properties. The incidence matrix defined below is an extremely useful tool in algebraic graph theory (Biggs 1994). It will not be used explicitly in this paper; it is mentioned here only as a motivation for the definition of the graph Laplacian.

Definition. Let Γ be an arbitrary graph with vertex set V and edge set $E\Gamma$. For each edge $h = \{v, w\}$ we choose one of the two vertices as the "positive end" and the other one as the "negative end" of the edge. The choice of the orientation is completely arbitrary. The matrix[6]

$$\nabla^+_{ij} = \begin{cases} +1 & \text{vertex } v_i \quad \text{is the positive end of edge } e_j \\ -1 & \text{vertex } v_i \quad \text{is the negative end of edge } e_j \\ 0 & \text{otherwise} \end{cases}$$

is called the *incidence matrix* of Γ.

The choice of the symbol ∇ is intentional. In fact, let $f : V \to \mathbb{R}$ be an arbitrary landscape. Then $\nabla f : E\Gamma \to \mathbb{R}$ is given by

$$(\nabla f)(h) = f(v) - f(w) \qquad \text{where } h = \{v, w\}$$

and v is the positive end of the edge h. This is as close to a differential operator as one can get on a graph.

Definition. Let D be the diagonal matrix of vertex degrees, i.e., D_{xx} is the number of edges incident into x. Then the matrix

$$\Delta = A - D \tag{11}$$

[6]The supercript $^+$ indicates transposition.

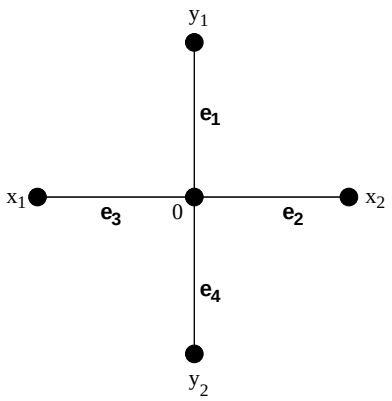


Figure 2: The graph Laplacian Δ is a generalization of the discrete approximation of the familiar Laplacian differential operator in $\mathbb{R}^n$.
This approximation is commonly performed by replacing the continuous space by a square lattice. Rescaling the space coordinates we can assume that the lattices points have integer coordinates. The first derivatives, evaluated at the mid-points of the edges, e_i, are computed from

$$\frac{\partial f}{\partial x}(e_2) = f(x_2) - f(0) \qquad \frac{\partial f}{\partial x}(e_3) = f(0) - f(x_1),$$

and analogous expressions for $\partial f/\partial y$. Consequently, the second derivates evaluated in 0 become

$$\frac{\partial^2 f}{\partial x^2}(0) = (f(x_2) - f(0)) - (f(0) - f(x_1)) = f(x_1) + f(x_2) - 2f(0),$$

and an analogous expression for $\partial^2 f/\partial x^2$. The discrete approximation of the usual Laplacian hence coincides with the graph Laplacian of the square lattice.

is called the *(graph) Laplacian* of Γ. For regular graphs this becomes $\Delta = A - DE$, where D is now the constant vertex degree of Γ and E is the identity matrix.

Denote by $N(x) = \{y \in V \mid d(x,y) = 1\}$ the set of neighbors of a vertex $x \in \Gamma$. For an arbitrary function $f : V\Gamma \to \mathbb{R}$ we have

$$(\Delta f)(x) = \sum_{y \in N(x)} \big(f(y) - f(x) \big). \tag{12}$$

The graph Laplacian shares its most important properties with the familiar differential operator $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ in $\mathbb{R}^n$, see also figure 2.

Lemma 4. The graph Laplacian has the following properties

(i) Δ is symmetric.

(ii) Δ is non-positive definite.

(iii) Δ is singular; the eigenvector $(1, \ldots, 1)$ belongs to the eigenvalue $\mu_0 = 0$. If Γ is connected (as we will always assume), then μ_0 has multiplicity 1.

(iv) $\Delta = -\nabla^+\nabla$, that is, it corresponds to "second derivatives" on the graphs.

(v) for all landscapes f and g holds *Green's formula* in the following form

$$\sum_{x\in V} f(x)(\Delta g)(x) = \sum_{x\in V} g(x)(\Delta f)(x) = -\sum_{h\in E\Gamma} (\nabla f)(h)(\nabla g)(h).$$

Proof. (i) is obvious, (ii) and (iii) are well known, see, e.g., Biggs (1994) or Cvetković *et al.* 1980), and (iv) is Proposition 4.8 of Biggs (1994). Green's formula (v) is checked by explicit calculation:

$$\begin{aligned}\sum_{h\in E\Gamma} (\nabla f)(h)\,(\nabla g)(h) &= \sum_{x\in V}\sum_{y\in V}\sum_{h\in E\Gamma} f(x)\nabla_{hx}\nabla_{hy}g(y)\\ &= \sum_{x\in V}\sum_{y\in V} f(x)\left(\sum_{h\in E\Gamma}\nabla^+_{xh}\nabla_{hy}\right)g(y)\\ &= -\sum_{x\in V}\sum_{y\in V} f(x)\Delta_{xy}g(y) = -\sum_{x\in V} f(x)(\Delta g)(x),\end{aligned}$$

An analogous calculation shows $\sum_h (\nabla f)(h)(\nabla g)(h) = \sum_x g(x)(\Delta f)(x)$. ∎

As an example for its applications we remark that the graph Laplacian is central to the theory of electrical networks: Let $\eta(x)$ be the current flowing into the network at vertex x. Then there is a potential $\Phi : V \to \mathbb{R}$ satisfying $\Delta\Phi = \eta$, and the vector ζ of currents along the edges is given by $\zeta = \nabla\,\Phi$. A recent book on potential theory on discrete lattices is Soardi (1994).

Finally we note the following (almost trivial) connection between the spectrum of a graph and the Graph Laplacian:

Lemma 5. Let Γ be a regular graph with adjacency matrix A and vertex degree D. Then A and the graph Laplacian have the same eigenvectors and the eigenvalues of the graph Laplacian are given by $\tilde{\Lambda}_k = \Lambda_k - D$, where Λ_k denotes the corresponding eigenvalue of the graph Γ.

Proof. Follows inmediately from $\Delta = A - DE$ for regular graphs. ∎

Notation. An eigenvalue of the adjacency matrix A will be denoted by Λ_k, the corresponding eigenvalues of the graph Laplacian is $\tilde{\Lambda}_k$, and the corresponding eigenvalue of the transition matrix T is $\lambda_k = \Lambda_k / D$.

2.9. Fourier Expansion of Landscapes

A series expansion in terms of a complete and orthonormal system of eigenfunctions of the Laplace operator is commonly termed Fourier expansion. We will adopt the same terminology here following Weinberger's (1991) paper. Thus let f be landscape on Γ and let $\{\theta_s\}$ denote a complete orthonormal set of eigenvectors of the graph Laplacian Δ. Then we call

$$f(x) = \sum_{y \in \Gamma} b_y \theta_y(x) \tag{13}$$

the *Fourier expansion* of the landscape.

It will turn out to be convenient to label the eigenvectors θ_y by the vertices of the underlying graph Γ. This is possible because the eigensystem of the finite symmetric operator Δ is complete. In general, this labeling is arbitrary. Note, however, that there is a canonical labeling for the eigenvectors e_g of the adjacency matrix of a Cggcg. Since Δ is symmetric, the Fourier basis $\{\theta_y\}$ can always be

chosen real valued; it is convenient, however, to allow for complex conjugate pairs of eigenvectors. In particular, we will use the following identity for landscapes on a CGCG:

$$f(x) = \sum_{l \in G} a_l e_l(x) = \sum_{l \in G} a_l^* e_l^*(x). \tag{14}$$

Since we deal with a finite vector space with a scalar product (which we will use the notation $\langle \, . \, , \, . \, \rangle$ in the paper), spanned by eigenvectors $\{\theta_y\}$ of the graph Laplacian, the familiar properties of Fourier series, such as Parseval's equation

$$\|f\|^2 = \langle f, f \rangle = \sum_{y \in \Gamma} \left(\frac{\langle f, \theta_y \rangle}{\langle \theta_y, \theta_y \rangle} \right)^2 \ , \tag{15}$$

and the mean square approximation theorem hold for all landscapes on all connected graphs. For the convenience of the reader we recall

Proposition. [*Mean Square Approximation Theorem*]
Consider a landscape f on Γ with Fourier expansion

$$f(x) = \sum_{y \in V} a_y \theta_y(x)$$

Let X be a subset of V, let

$$g(x) = \sum_{y \in X} b_y \theta_y(x)$$

be an approximation to f. Then the squared approximation error $\|f - g\|^2 = \langle (f-g), (f-g) \rangle$ is minimized by choosing $b_y = a_y \equiv \langle f, \theta_y \rangle$ for all $y \in X$.

The Fourier basis of a CGCG is the set of functions $\{e_g(x) | g \in G\}$, defined in section 2.7. These functions have the well-known form of the Fourier basis on $\mathbb{C}^n$.

2.10. Examples of Configuration Spaces

2.10.1. Sequence Spaces

Consider the set of all sequences of length n, which are composed of letters taken from some finite alphabet of size α. The canonical distance is given by the number of positions in which two sequences x and y differ; this metric distance measure is the Hamming distance (Hamming 1950). The graph obtained from connecting two sequences if and only if their Hamming distance is $d(x, y) = 1$ is called the *Hamming graph* or *sequence space* $\mathcal{Q}_\alpha^n$. The sequence space $\mathcal{Q}_\alpha^n$ is the n-fold product[7] of the complete graph with α vertices with itself.

Hamming graphs constitute the most common class of configuration spaces. Their automorphism group is well known (Dress and Rumschitzki 1988):

$$\mathbf{Aut}[\mathcal{Q}_\alpha^n] = S_\alpha \wr S_n = S_n \ltimes (S_\alpha)^n \,, \tag{16}$$

where S_m denotes the symmetric group on m elements.

Lemma 6. A sequence space $\mathcal{Q}_\alpha^n$ is a Cayley graph of some commutate group, and it is distance transitive.

Proof. It is easily checked that the product of two Cayley graphs $(G_1, \Phi_1) \times (G_2, \Phi_2)$ is the Cayley graph $(G_1 \times G_2, \Phi)$, where $G_1 \times G_2$ is the direct product of the two groups and $\Phi = (\Phi_1 \times \{e_2\}) \cup (\Phi_2 \times \{e_1\})$, e_1 and e_2 being the group identities. A complete graph is a Cayley graphs of an arbitrary group G with the set of generators $\Phi = G \setminus \{e\}$. Hence the sequence spaces are Cayley graphs, and

[7] The (Cartesian) product $\Gamma_1 \times \Gamma_2$ of two graphs has vertex set $V(\Gamma_1 \times \Gamma_2) = V(\Gamma_1) \times V(\Gamma_2)$. Two nodes (x_1, x_2) and (y_1, y_2) are connected by an edge if either (i) $x_1 = y_1$ and x_2, y_2 are adjacent in Γ_2, or (ii) $x_2 = y_2$ and x_1, y_1 are adjacent in Γ_1.

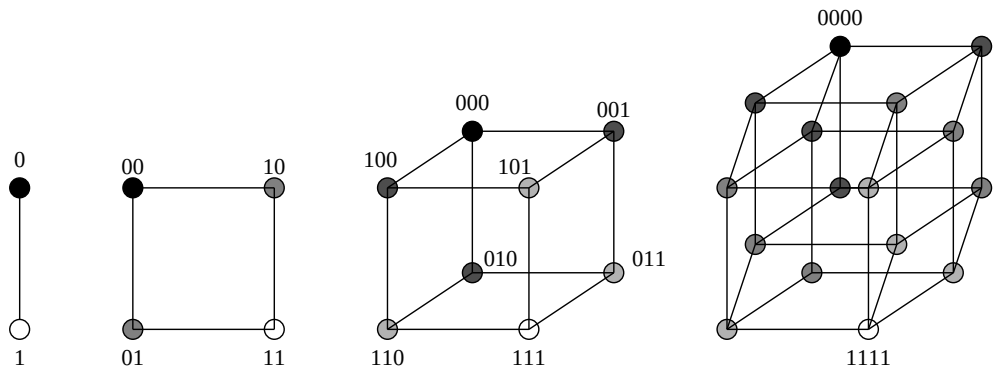


Figure 3: Hypercubes $\mathcal{Q}_2^n$ for $n = 1$ through $n = 4$. Vertices belonging to the same symmetry class μ are marked by a common symbol. Vertices within a symmetry class have common distance from the reference vertex $\bullet$. The hypercubes are distance transitive, since their symmetry classes are defined by the distance from the reference vertex.

G can be chose to be commutative.

Distance transitivity is well known result, see, e.g., Biggs (1974). ∎

Lemma 7. The eigenvalues and left eigenvectors of the collapsed adjacency matrix $\hat{A}$ of a sequence space $\mathcal{Q}_\alpha^n$ are given by

$$\begin{aligned} \Lambda_p &= n(\alpha - 1) - p\alpha, \\ \rho_p(d) &= \frac{1}{(\alpha-1)^p \binom{n}{p}} \mathbf{K}_p^{(\alpha,n)}(d), \end{aligned} \tag{17}$$

where $\mathbf{K}_p^{(\alpha,n)}$ are the *Krawtchouk polynomials*

$$\mathbf{K}_p^{(\alpha,n)}(d) = \sum_{j=0}^{p} (-1)^j \binom{d}{j} \binom{n-d}{p-j} (\alpha-1)^{p-j}. \tag{18}$$

The multiplicity of Λ_p w.r.t. the full adjacency matrix A is $m(\Lambda_p) = (\alpha-1)^p \binom{n}{p}$; all eigenvalues are simple w.r.t. the collapsed adjacency matrix $\hat{A}$.

Proof. See, e.g., Dunkl (1976), Koornwinder (1982), and Rumschitzky (1987). ∎

Lemma 8. The left eigenvectors ρ_p of $\hat{A}$ fulfill the orthogonality relation

$$\langle \rho_k ; \rho_l \rangle = \frac{\kappa^n}{(\kappa-1)^k \binom{n}{k}} \delta_{kl}$$

with respect to the scalar product

$$\langle a; b\rangle = \sum_{d=0}^{n} (\kappa - 1)^d \binom{n}{d} a(d)\, b(d),$$

and the symmetry relation

$$\rho_p(d) = \rho_d(p)\,. \tag{19}$$

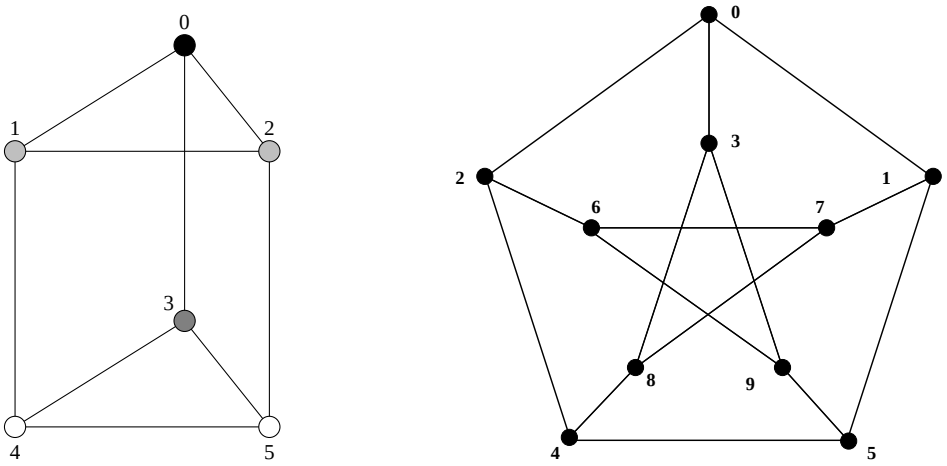
Proof. The above expressions are immediate consequences of the basic properties of Krawtchouk polynomials. For details on Krawtchouk polynomials see e.g., Brouwer *et al.* 1989, Dunkl 1976, 1979, Koornwinder 1982, Krawtchouk 1929, MacWilliams and Sloane 1991, vanLint 1982. ∎

Remark. When considered as a continuous function on $[0, n]$, the normalized Krawtchouk polynomial $\rho_p(d)$ has exactly p zeros in $(0, n)$. If x_i are the zeros of ρ_p, $x_0 = 0$ and $x_{n+1} = n$, then there is exactly one zero of ρ_{p+1} in each open interval (x_i, x_{i+1}), $i = 0, \ldots, n$, see, e.g., Szegö (1959).

2.10.2. Generalized Hamming Graphs

The name *Generalized Hamming Graph* is sometimes used for direct products of sequence spaces. They correspond to sequences where the alphabet is not the same for all positions. Such graphs have turned out to be useful when considering the inverse folding problem for RNA secondary structures (Reidys *et al.* 1994): An RNA secondary structure can be partioned into the unpaired regions where each position can be realized by one of the 4 nucleotides **G**, **C**, **A**, and **U**, and the paired regions, where each base pair can be realized by one of the 6 types of base pairs (**GC**, **CG**, **AU**, **UA**, **GU**, and **GU**). The set of sequences compatible with the given structure can then be considered as the graph

$$\mathcal{C} \cong \mathcal{Q}_4^{n_u} \times \mathcal{Q}_6^{n_p},$$



a **b**

Figure 4: **a)** The trigonal prism graph is the smallest vertex transitive graph which is not edge transitive. Its vertex set decomposed into four symmetry classes which are the reference vertex $\mathbf{0} = \{0\}$ itself, its neighbors within the same triangle $\mathbf{1a} = \{1, 2\}$, its neighbor in the other triangle $\mathbf{1b} = \{3\}$, and the remaining vertices $\mathbf{2} = \{4, 5\}$ which have distance two from the reference vertex.
b) The Petersen graph is the smallest distance transitive graph which is not a Cayley graph of any group. It has three classes of vertices: the reference vertex $\mathbf{0} = \{0\}$, its neighbors $\mathbf{1} = \{1, 2, 3\}$, and all other vertices $\mathbf{2} = \{4, 5, 6, 7, 8, 9\}$.

where n_u and n_p are the numbers of the unpaired positions and of the base pairs, respectively.

The simplest case of a Hamming Graph which is not a sequence space is the trigonal prism $\mathcal{Q}_2 \times \mathcal{Q}_3$, see figure 5a. Its collapsed adjacency matrix is

$$\hat{A} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$

The eigenvalues of $\hat{A}$, the corresponding *left* eigenvectors, and the multiplicities of these eigenvalues in A are

$$\begin{array}{llll} \Lambda_0 = & 3 & \rho_0 = & (1, 1, 1, 1) & m(\Lambda_0) = 1 \\ \Lambda_1 = & 1 & \rho_1 = & (1, 1, -1, -1) & m(\Lambda_1) = 1 \\ \Lambda_2 = & 0 & \rho_2 = & (1, -\frac{1}{2}, 1, -\frac{1}{2}) & m(\Lambda_2) = 2 \\ \Lambda_3 = & -2 & \rho_3 = & (1, -\frac{1}{2}, -1, \frac{1}{2}) & m(\Lambda_3) = 2 \end{array}$$

A simple calculation shows that the matrices $R^{(\mu)}$ are contained in the adjacency algebra; we have explicitly:

$$\begin{aligned} R^{(\mathbf{0})} &= E \\ R^{(\mathbf{1a})} &= \frac{1}{5}(-2A^3 + 3A^2 + 14A - 5E) \\ R^{(\mathbf{1b})} &= \frac{1}{5}(2A^3 - 3A^2 - 9A + 5E) \\ R^{(\mathbf{2})} &= \frac{1}{5}(A^3 + A^2 - 7A - 5E) \end{aligned}$$

2.10.3. The Petersen Graph

The Petersen Graph, Figure 4b, is the smallest graph which is distance transitive but not a Cayley graph (see, e.g., Biggs and White 1979). The symmetry classes are the three distance classes $\mathbf{0} = \{0\}$, $\mathbf{1} = \{1, 2, 3\}$, and $\mathbf{2} = \{4, 5, 6, 7, 8, 9\}$, with $|\mathbf{0}| = 1$, $|\mathbf{1}| = 3$, and $|\mathbf{2}| = 6$. The collapsed adjacency matrix is

$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

The eigenvalues of $\hat{A}$, the corresponding *left* eigenvectors, and the multiplicities of these eigenvalues in A are

$$\begin{array}{llll} \Lambda_0 = & 3 & \rho_0 = & (1, 1, 1) & m(\Lambda_0) = 1 \\ \Lambda_2 = & 1 & \rho_1 = & (1, \frac{1}{3}, -\frac{1}{3}) & m(\Lambda_1) = 5 \\ \Lambda_3 = & -2 & \rho_2 = & (1, -\frac{2}{3}, \frac{1}{6}) & m(\Lambda_2) = 4 \end{array}$$

Note that the multiplicities $m(\Lambda_i)$ of the eigenvectors of A are 1, 5, and 4, while the the symmetry classes contain 1, 3, 6 vertices, respectively.

Since the Petersen graph is distance transitive we can write the matrices $R^{(\mu)}$ as polynomials in A, see Biggs (1994), Lemma 20.6. We find explicitly:

$$\begin{aligned} R^{(0)} &= E \\ R^{(1)} &= A \\ R^{(2)} &= A^2 - 3E \end{aligned}$$

2.10.4. Cayley Graphs of the Symmetric Group

Let S_n denote the symmetric group on n elements, and let $\mathcal{T}$ be the set of all transpositions (i, j). It is well known that the transpositions generate the symmetric group. The corresponding Cayley-graph $\Gamma(S_n, \mathcal{T})$ is hence a possible choice for the configuration space of a n-city travelling salesman problem. Little is known in general on these graphs. We will use them here in order to explain the definitions in the above sections on a non-trivial example. $\Gamma(S_n, \mathcal{T})$ has $N = n!$ vertices and is regular with degree $n(n-1)/2$. Furthermore they are bipartite since transpositions change the sign of a permutation.

The cases $n = 1$ and $n = 2$ are trivial: $\Gamma(S_1, \mathcal{T})$ is an isolated vertex, and $\Gamma(S_2, \mathcal{T})$ is the connected graph with $N = 2$ vertices.

The first case of interest is $\Gamma(S_3, \mathcal{T})$, see figure 5. It has $N = 6$ vertices, and its diameter is 2. It coincides with the *complete bipartite graph* $K_{3,3}$, which is known to be distance transitive. Therefore the symmetry classes are the distance classes. $K_{3,3}$ may serve as an example for the fact that Cayley graphs of non-commutative groups can coincide with Cayley graphs of commutative groups. Consider the commutative group with $N = 6$ elements, namely $C_2 \times C_3$, with the notation $C_2 = \{e, a\}$ and $C_3 = \{e, b, \bar{b}\}$. It is easy to check that $\Phi = \{a, ab, a\bar{b}\}$ generates $C_2 \times C_3$, and that

$$\Gamma(C_2 \times C_3, \Phi) = K_{3,3} = \Phi(S_3, \mathcal{T}).$$

We emphasize at this point that the results presented in this paper depend only on the structure of the configuration space and not on the particular group representation used to construct this graph. For instance we may regard a landscape on $\Gamma(S_3, \mathcal{T})$ as a landscape on $\Gamma(C_2 \times C_3, \Phi)$, i.e., as a landscape on a Cgcg, even though the physical model might force us to use the symmetric group.

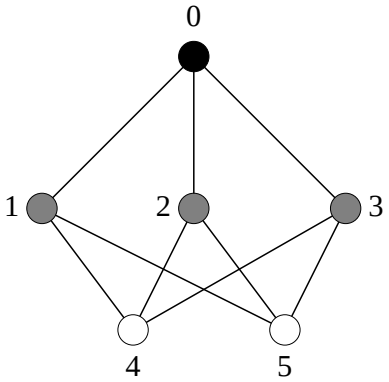


Figure 5: The Cayley graph $\Gamma(S_3, \mathcal{T})$ coincides with the distance transitive graph $K_{3,3}$, which is also obtained as $\Gamma(C_2 \times C_3, \{a, ab, a\bar{b}\})$, where $C_2 = \{e, a\}$ and $C_3 = \{e, b, \bar{b}\}$ are the cyclic groups with 2 and 3 elements, respectively.

The adjacency matrix and the collapsed adjacency matrix of $\Gamma(S_3, \mathcal{T})$ are

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} \qquad \hat{A} = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 3 \\ 2 & 0 & 2 \end{pmatrix}$$

The eigenvalues, the corresponding left eigenvectors of $\hat{A}$, and the multiplicities of the eigenvalues in A are

$$\begin{array}{llll} \Lambda_0 = & 3 & \rho_0 = & (1,1,1) \qquad m(\Lambda_0) = 1 \\ \Lambda_1 = & 0 & \rho_1 = & (1,0,-\frac{1}{2}) \qquad m(\Lambda_1) = 4 \\ \Lambda_2 = & -3 & \rho_2 = & (1,-1,1) \qquad m(\Lambda_2) = 1 \end{array}$$

The graph $\Gamma(S_4, \mathcal{T})$ has $N = 24$ vertices; it is shown in figure 6, and its adjacency matrix is given in table 1.

A close inspection of the graph shows that there are 5 symmetry classes. The collapsed adjacency matrix $\hat{A}$ of $\Gamma(S_4, \mathcal{T})$ is

$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 6 & 0 & 2 & 3 & 0 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 4 & 0 & 0 & 4 \\ 0 & 0 & 4 & 3 & 0 \end{pmatrix},$$

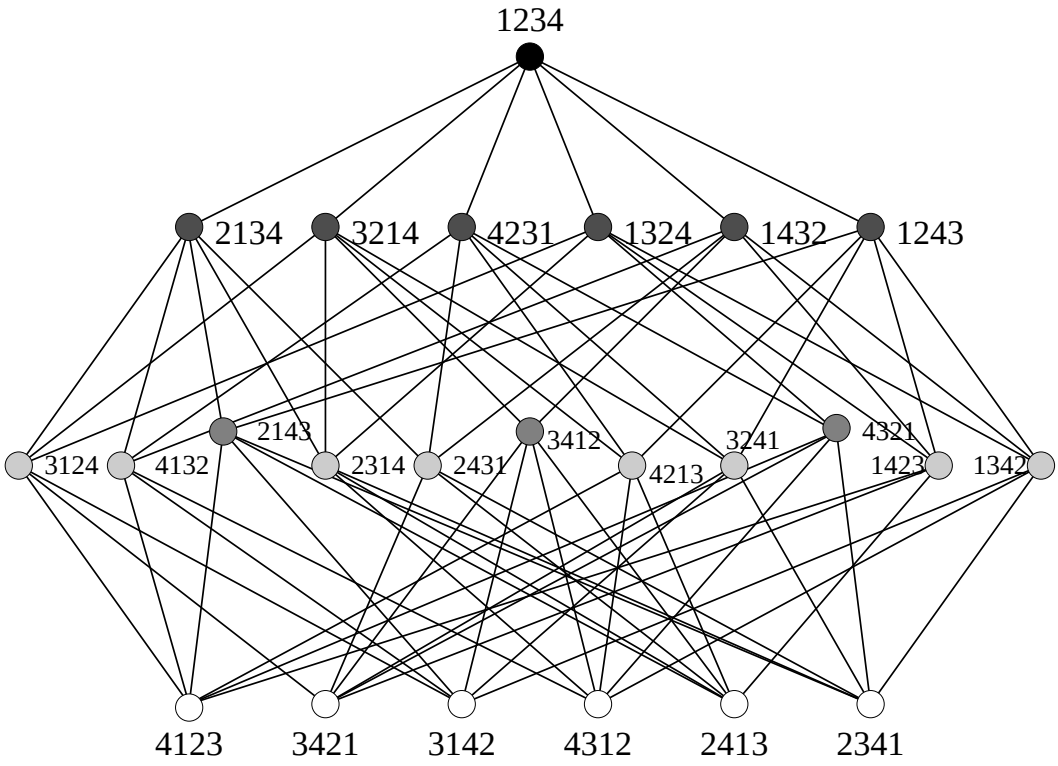


Figure 6: The Cayley graph $\Gamma(S_4, \mathcal{T})$.

which has the following eigenvalues and left eigenvectors:

$$\begin{array}{lll} \Lambda_0 = & 6 & \rho_0 = \quad (1,1,1,1,1) \\ \Lambda_1 = & 2 & \rho_1 = \quad (1,\frac{1}{3},-\frac{1}{3},0,-\frac{1}{3}) \\ \Lambda_2 = & 0 & \rho_2 = \quad (1,0,1,-\frac{1}{2},0) \\ \Lambda_3 = & -2 & \rho_3 = \quad (1,-\frac{1}{3},-\frac{1}{3},0,\frac{1}{3}) \\ \Lambda_4 = & -6 & \rho_4 = \quad (1,-1,1,1,-1) \end{array}$$

A lengthy calculation shows that the matrices $R^{(\mu)}$ can be expressed in terms of

Table 1. Adjacency Matrix for $\Gamma(S_4, T)$.

	1234	2134	3214	4231	1324	1432	1243	2143	3412	4321	3124	4132	2314	2431	4213	3241	1423	1342	4123	4321	3142	4312	2413	2341	
1234		1	1	1	1	1	1																		(.)(.)(.)(.)
2134	1							1			1	1	1	1											
3212	1								1		1		1		1	1									
4231	1									1		1		1	1	1									(..)(.)(.)
1324	1									1	1		1				1	1							
1432	1								1			1		1			1	1							
1243	1							1							1	1	1	1							
2143		1					1												1		1		1	1	
3412			1			1														1	1	1	1		(..)(..)
4321				1	1														1	1		1		1	
3124		1	1		1														1	1	1				
4132		1		1		1													1		1	1			
2314		1	1		1																	1	1	1	
2431		1		1		1														1			1	1	(...)(.)
4213			1	1			1												1			1	1		
3241			1	1			1													1	1			1	
1423					1	1	1												1	1			1		
1342					1	1	1														1	1		1	
4123								1		1	1	1			1		1								
3421									1	1	1			1		1	1								
3142								1	1		1	1				1		1							
4312									1	1		1	1		1			1							(....)
2413								1	1				1	1	1		1								
2341								1		1			1	1		1		1							

The last column shows the cycle decomposition of the permutations. Permutations are grouped into the conjugate classes of S_4. A close inspection of the graph shows that the symmetry classes of the Cayley graph $\Gamma(S_4, \mathcal{T})$ correspond in fact to the conjugate classes. We will index them by

$$\mu_0 = [1^4], \quad \mu_1 = [1^2 2], \quad \mu_2 = [2^2], \quad \mu_3 = [13], \quad \mu_4 = [4].$$

A as follows

$$\begin{aligned} R^{[1^4]} &= E \\ R^{[1^2 2]} &= A \\ R^{[2^2]} &= \frac{1}{32}A^4 - \frac{9}{4}A^2 + 3E \\ R^{[31]} &= -\frac{1}{48}A^4 + \frac{13}{12}A^2 - 4E \\ R^{[4]} &= \frac{1}{16}A^3 - \frac{5}{4}A \end{aligned}$$

Let $\eta = [k_1^{l_1} k_2^{l_2} \dots k_r^{l_r}] = (\eta_1, \eta_2, \dots, \eta_m)$ be a partition of n, i.e., $\sum_{i=1}^{r} l_i k_i = \sum_{j=1}^{m} \eta_j = m$, with $k_i > k_j$ and $\eta_i \geq \eta_j$ whenever $i < j$. It is well know that the conjugacy classes of the S_n correspond to the partitions of n via the cycle-decomposition of the permutations (see, e.g., Jacobson 1985, p. 48). Lemma 2 in section 2.6 implies now that the conjugacy classes of the S_n are contained in the symmetry classes of the graph $\Gamma(S_n, \mathcal{T})$ since the transpositions $\mathcal{T}$ form a single conjugacy class $[2^1 1^{n-2}] = (2, 1, 1, \dots, 1)$ with $n-1$ cycles.

The eigenvalues of $\Gamma(S_n, \mathcal{T})$ have been obtained from the characters of the symmetric group. One finds

$$\Lambda_\eta = \frac{1}{2}\left[n + \sum_{j=1}^{m} \eta_j (1 - 2j\eta_j) \right] ,$$

see, e.g., Diaconis and Shashahani (1981), Ingram (1950), and MacDonald (1979).

2.11. Karhunen-Loève Decomposition of Random Fields

Recall that a random field on Γ is a probability space $\Xi = (\{f : \mathcal{C} \to \mathbb{R}\}, \boldsymbol{\mu})$. An important characteristic of Ξ is its covariance matrix $C = (C_{xy})_{x,y \in V}$, which is

defined element-wise by

$$C_{xy} = \mathcal{E}[f(x)f(y)] - \mathcal{E}[f(x)]\mathcal{E}[f(y)]. \tag{20}$$

Denote by $\{\phi_s\}$ a complete set of orthonormal eigenvectors (which exists by symmetry of C). Although $\{\phi_s\}$ can be chosen real, we will admit complex vectors as well. Then the random field f on Γ can be represented in mean square sense as

$$f(x) \doteq \sum_{y\in\Gamma} a_y \phi_y(x), \quad \text{i.e.,} \quad \mathcal{E}\left[\left(f(x) - \sum_{y\in\Gamma} a_y\phi_y(x)\right)^2\right] = 0\,. \tag{21}$$

This series representation is known as the *Karhunen-Loève* decomposition of the random field Ξ. The Karhunen-Loève expansion coincides for finite sets with the well known *principal component analysis* introduced by Hotelling (1933), see also Rao (1973). Its most important property is

Lemma 9. [Hotelling, 1933] The coefficients a_y are uncorrelated, i.e., $\mathcal{E}[a_x a_y^*] = \mathcal{E}[a_x]\mathcal{E}[a_y^*]$ if $x \neq y$, with respect to the measure $\boldsymbol{\mu}$.

3. Isotropic Random Fields

3.1. Characterization of Isotropic Random Fields

Definition. A random field on a graph Γ is *isotropic* if and only if

(i) $\mathcal{E}[f(x)] = a_0$ for all configurations $x \in \Gamma$.

(ii) Suppose the pairs of configurations (x, y), (u, v) belong to the same symmetry class $\ddot{\mu}$.

Then $\mathcal{E}[f(x)f(y)] = \mathcal{E}[f(u)f(v)]$.

Lemma 10. Let Γ be a vertex transitive graph. Then a random field on Γ is isotropic if and only if

(i) $\mathcal{E}[f(x)] = a_0$ for all $x \in V$, and

(ii) there exist a constant $C_{00} = \mathrm{Var}[f(x)]$ and a real valued function ρ of the symmetry classes with $\rho(0) = 1$ such that

$$C = C_{00} \sum_{\mu} \rho(\mu) R^{(\mu)}. \tag{22}$$

Proof. This is a simple rewriting of the above definition, observing that all entries in the covariance matrix belonging to a fixed symmetry class μ are the same and correspond to the non-zero entries of $R^{(\mu)}$. ∎

The function $\rho(\mu)$ is called the *autocorrelation function* of the random walk. It will be discussed in section 3.3.

Theorem 1. [Weinberger 1991b, Stadler 1994b] Let f be an isotropic random field on a Cayley graph $\Gamma(G, \Phi)$ with a commutative group G, or on a vertex transitive

graph Γ which has the Θ-property. Then the adjacency matrix A of Γ and the covariance matrix C of the random field f commute, i.e., Fourier decomposition and Karhunen-Loève decomposition coincide.

Proof. The case of CGCG is dealt with in Weinberger (1991b).
The case of distance transitive graphs is proved in Stadler (1994b). If Γ is vertex transitive with the Θ-property then $R^{(\mu)}$ is a polynomial in A, and hence C is a polynomial in A. Consequently we have $CA = AC$; thus Fourier expansion and Karhunen-Loève expansion coincide. ∎

As a consequence we immediately conclude that the Fourier coefficients are uncorrelated for an isotropic random field on any vertex transitive graph with the Θ-property. A more detailed result can be obtained for isotropic landscapes on a CGCG.

Remark. The notion of isotropy for random fields is analogous to the defintion of stationarity for stochastic processes. Following the conventions of Karlin and Taylor (1975) our notion of isotropy should be called "covariance isotropic", "weakly isotropic", or "wide sense isotropic". Weinberger's (1991c) definition corresponds to "strictly isotropic". For a Gaussian random field the notions of (weak) isotropy and strict isotropy coincide.

Theorem 2. A random field on a CGCG is isotropic if and only if
(i) $\mathcal{E}[a_l] = 0$ for all $l > 0$; the coefficient a_0 is arbitrary.
(ii) The coefficients are pairwise uncorrelated, $\mathcal{E}[a_k a_l] = \mathcal{E}[|a_l|^2]\delta_{kl}$.
(iii) $\mathcal{E}[|a_l|^2] = V(\mu)$ depends only on the symmetry class μ to which l belongs.

Proof. ($\Longrightarrow$) Uncorrelatedness of the Fourier coefficients follows immediately from theorem 1 and Hotelling's results (1933). An explicit proof can be found in

Weinberger (1991b). Properties (i) and (ii) can be obtained after some computations from Weinberger's (1991b) proof. For the convenience of the reader we give a complete proof here. (i) Using the Karhunen-Loève decomposition we have

$$\frac{1}{N}\sum_x \mathcal{E}[f(x)]e_k^*(x) = \frac{1}{N}\sum_x \bar{f}e_k^*(x) = \bar{f}\delta_{0k}$$
$$\frac{1}{N}\sum_x \mathcal{E}[f(x)]e_k^*(x) = \frac{1}{N}\sum_l\sum_x \mathcal{E}[a_l]e_l(x)e_k^*(x) = \mathcal{E}[a_k],$$

and hence $\mathcal{E}[a_l] = 0$ for $l \neq 0$ as claimed.

(iii) can be shown analogously: We have

$$\begin{aligned}\frac{1}{N^2}\sum_{x,y}\mathcal{E}[f(x)f(y)]e_p^*(x)e_q(y) &= \frac{1}{N^2}\sum_{x,z}U(z)e_p^*(x)e_q(x)e_q(z) = \\ &= \frac{1}{N}\sum_z\sum_\mu U(\mu)\sum_{z\in\mu}e_q(z) = \\ &= \delta_{pq}\frac{1}{N}\sum_\mu U(\mu)v_\kappa(\mu) = \delta_{pq}W(\kappa), \text{ where } p\in\kappa.\end{aligned}$$

$$\begin{aligned}\frac{1}{N^2}\sum_{x,y}\mathcal{E}[f(x)f(y)]e_p^*(x)e_q(y) &= \frac{1}{N^2}\sum_{k,l}\mathcal{E}[a_la_k^*]\sum_x e_l(x)e_p^*(x)\sum_y e_k^*(y)e_q(y) = \\ &= \mathcal{E}[a_pa_q^*],\end{aligned}$$

i.e., $\mathcal{E}[a_pa_q^*] = \delta_{pq}W(\kappa)$, where $p\in\kappa$, showing that (ii) and (iii) are true.

($\Longleftarrow$) (i) One finds explicitly $\mathcal{E}[f(x)] = \sum_l \mathcal{E}[a_l]e_l(x) = \mathcal{E}[a_0]e_0(x) = \mathcal{E}[a_0]$ for all x and this $\mathcal{E}[f(x)]$ is independent of x.

From (ii) and (iii) we find

$$\begin{aligned}\mathcal{E}[f(x)f(y)] &= \sum_{k,l}\mathcal{E}[a_la_k^*]e_l(x)e_k^*(y) = \sum_l \mathcal{E}[|a_l^2|]e_l(x)e_l^*(x)e_l^*(z)) = \\ &= \sum_l \mathcal{E}[|a_l^2|]e_l(z) = \sum_\mu V(\mu)\sum_{l\in\mu}e_l(z) = \sum_\mu V(\mu)v_\kappa(\mu) = W(\kappa),\end{aligned}$$

where κ is the symmetry class to which $z = xy^{-1}$ belongs. Hence the covariances depend in fact only on the symmetry class $\ddot{\kappa}$ to which (x, y) belongs. ∎

Note that theorem 2 depends on the fact the we can relate the eigenfunctions $e_p(x)$ directy with symmetry classes. The examples in the previous section show,

however, that this is not possible for more general graphs. The relation of Fourier coefficients and isotropy on more general graphs certainly deserves attention in forthcoming investigations.

3.2. Landscapes on the Hypercube and the p-Spin Models

The vertices of a Boolean hypercube are the sequences of length n taken from an alphabet of size $\alpha = 2$. Without loosing generality we may use the alphabet $\{+1, -1\}$. A configurations is then a string σ of "spins" $\sigma_k \in \{+1, -1\}$. An alternative encoding uses a binary string x, with $x_i \in \{0, 1\}$. The following result is well known in the literature:

Lemma 11. Any landscape f on the Boolean Hypercube can be written as

$$f(\sigma) = J_0 + \sum_{p=1}^{n} \sum_{i_1<i_2<\ldots<i_p} J_{i_1 i_2 \ldots i_p} \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p}, \tag{23}$$

where the $J_{i_1 i_2 \ldots i_p}$ are constants.

Proof. Using the correspondences $\sigma_k = 2x_k - 1$ and $\varepsilon_q(\sigma) = e_q(x)$ one finds explicitly

$$\varepsilon_q(\sigma) = \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p} \quad \text{where} \quad q_k = 1 \quad \text{if and only if} \quad k \in \{i_1, i_2, \ldots, i_p\}, \tag{24}$$

and hence (13) is simply the Fourier expansion of landscape, see also Weinberger (1991a). ∎

The p-spin Hamiltonian

$$\mathcal{H}_p(\sigma) = \sum_{i_1<i_2<\ldots<i_p} J_{i_1 i_2 \ldots i_p} \sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_p} \tag{25}$$

was introduced by Derrida (1980a) in order to bridge the gap between the SK model (Sherrington & Kirkpatrick 1975), which is the special case $p = 2$ and the random energy model (Derrida, 1980ab, 1981, Gardner and Derrida 1989). The coefficients $J_{i_1 i_2 \ldots i_p}$ are i.i.d. Gaussian random variables with mean 0 and unit variance.[8] The autocorrelation function of the p-spin Hamiltonian $\mathcal{H}_p(\sigma)$ is simply $\rho = \rho_p$ for all p (Weinberger & Stadler 1993); see also equ.(6).

Note that we regard a p-spin model as random field on the Boolean Hypercube. For the next result we have to state precisely what we mean by a "sum" of random fields.

Definition. Let $\Xi_i = (\{f : V \to \mathbb{R}\}, \boldsymbol{\mu}_i)$ be random fields on Γ. The *superposition* of two random fields on Γ with fixed coefficients $a, b \in \mathbb{R}$ is the probability space

$$a\Xi_1 + b\Xi_2 \stackrel{\text{def}}{=\!=} (\{f : V \to \mathbb{R}\}, \boldsymbol{\mu}^*) \quad \text{with} \quad \boldsymbol{\mu}^*(f) = \int_g \boldsymbol{\mu}_1(\frac{1}{a}g)\boldsymbol{\mu}(\frac{1}{b}(f-g)).$$

The effect of this definition is that elements of the probability spaces Ξ_i are added as independent random variables.

Theorem 3. A Gaussian isotropic random field on the Boolean hypercube is a superposition of p-spin models with independent coefficients.

Proof. An isotropic random field has Fourier coefficients as given in Theorem 2. Since for Gaussian distribution functions uncorrelatedness implies independence, the coefficients are i.i.d. Gaussian random variables for each distance class. Let $S_p^2 = \mathcal{E}[J_{i_1 i_2 \ldots i_p}^2]$ be the variance of the coefficients with p indices. Then any Gaussian isotropic random field is of the form $\Xi = \sum_p S_p \mathcal{H}_p(\sigma)$ as a consequence of theorem 2. ∎

[8] In Derrida's original formulation the variance was chosen to be $S_p^2 = p!/n^p$, such that the variance of the Hamiltonian itself is independent of n.

3.3. Autocorrelation Functions of Isotropic Random Fields

As usual let Γ be a vertex transitive graph. Consider a simple random walk $\{x_0, x_1, \ldots\}$ on Γ. Weinberger (1990) has suggested to use the time series obtained by evaluating the random field at each step in order to gather information on the random field itself. The correlation function along such a time series is defined by

$$r(s) = \frac{\mathcal{E}[f(x_{t+s})f(x_t)] - \mathcal{E}[f(x_{t+s})f(x_t)]}{\sqrt{\mathrm{Var}[f(x_{t+s})]\mathrm{Var}[f(x_t)]}}. \tag{26}$$

For isotropic random fields this reduces to

$$r(s) = \frac{\mathcal{E}[f(x_s)f(x_0)] - \mathcal{E}[f(x_0)]^2}{\mathrm{Var}[f(x_0)]} = \sum_{\mu} \varphi_{s\mu}\rho(\mu) \tag{27}$$

In the light of this equation it is not surprising that there is a simple relation between the autocorrelation functions $\rho(\mu)$ and $r(s)$.

Theorem 4. Let Γ be vertex transitive graph and consider an isotropic random field on Γ with correlation functions $\rho(\mu)$ and $r(s)$, respectively. Then:

(i) $r(s) = \lambda^s$ if and only if $\rho\hat{A} = \Lambda\rho$.

(ii) Let $\{\rho_p\}$ be the orthonormal system of left eigenvector of $\hat{A}$ described in section 2.7. We have $r(s) = \sum_p b_p\lambda_p$ if and only if $\rho(\mu) = \sum_p b_p\rho_p$. Of course, such a decomposition always exists.

Proof. This theorem is the content of Stadler (1994a). ∎

An immediate consequence is the following surprising relation between eigenvalues and eigenvectors of a large class of graphs.

Corollary. Let Γ be vertex transitive and assume that Γ has a symmetry class $\dot{\mathbf{1}} = \{(x,y)|d(x,y) = 1\}$ consisting of all pairs of neighbors. Then $\rho_p(\mathbf{1}) = \lambda_p$.

Proof. First note that $r(1) = \rho(\mathbf{1})$ for all landscapes on such graphs. Now consider a landscape which has autocorrelation function ρ_p. Then we have as a consequence of the theorem above $r(1) = \lambda_p$. ∎

4. Empirical Landscapes

4.1. Empirical Correlation Functions

The empirical autocorrelation function of a landscape on a vertex transitive graph is defined as (Fontana *et al.* 1991, 1993a, 1993b, Schuster & Stadler 1994)

$$\hat{\rho}(\mu) = \frac{\langle f(x)f(y)\rangle_{(x,y)\in\ddot{\mu}} - \langle f\rangle^2}{\langle f^2\rangle - \langle f\rangle^2} \tag{28}$$

Of course, the autocorrelation function is invariant under the transformation $f \to f - \langle f\rangle$. Therefore we can assume without loosing generality $\langle f\rangle = 0$ for the reminder of this chapter. Under this assumption we may rewrite equ.(16) in the form

$$\hat{\rho}(\mu) = \frac{1}{|\mu|}\frac{\langle f, R^{(\mu)}f\rangle}{\langle f, f\rangle}\,, \tag{29}$$

where $\langle\,.\,,\,.\,\rangle$ denotes as usual the standard scalar product on Euclidean vector spaces.

In the following we characterize the set of all possible empirical autocorrelation functions for sufficiently symmetric configuration spaces.

Theorem 5. Let $\{\rho_\mu\}$ be an orthogonal basis of left eigenvectors of the collapsed adjacency matrix of a Cgcg Γ. Then $\hat{\rho}$ is an autocorrelation function of a landscape on Γ if and only if

$$\hat{\rho} = \sum_\mu \Omega_\mu \rho_\mu, \quad \Omega_0 = 0, \quad \Omega_\mu \geq 0, \quad \sum_\mu \Omega_\mu = 1. \tag{30}$$

Proof. The first term in the numerator can be written as

$$\langle f(x)f(y)\rangle_{(x,y)\in\ddot{\mu}} = \frac{1}{N}\frac{1}{|\mu|}\sum_{x\in G}\sum_{(x,y)\in\ddot{\mu}} f(x)f(y) = \frac{1}{N}\frac{1}{|\mu|}\sum_{x\in G}\sum_{z\in\mu} f(x)f(zx).$$

Substituting the Fourier decomposition of f yields

$$\begin{aligned}\langle f(x)f(y)\rangle_{(x,y)\in\ddot{\mu}} &= \frac{1}{N}\frac{1}{|\mu|}\sum_{g\in G}\sum_{h\in G}\sum_{z\in\mu}\sum_{x\in G} a_g e_g(x) a_h e_h(xz) = \\ &= \sum_{g\in G}\sum_{h\in G} a_g^* a_h \frac{1}{N}\sum_{x\in G} e_g^*(x)e_h(x)\,\frac{1}{|\mu|}\sum_{z\in\mu'} e_h(z) = \\ &= \sum_{g\in G}\sum_{h\in G} a_g^* a_h \delta_{hg}\rho_\kappa(\mu) = \sum_{g\in G}|a_g|^2\rho_\kappa(\mu) = \sum_\kappa\left[\sum_{g\in\kappa}|a_g|^2\right]\rho_\kappa(\mu).\end{aligned}$$

It is straightforward to express the denominator in terms of the Fourier coefficients a_g:

$$\frac{1}{N}\sum_{x\in G} f(x)^2 = \frac{1}{N}\sum_{x\in G}\sum_{g\in G}|a_g|^2 e_g(x)e_g^*(x) = \frac{1}{N}\sum_{x\in G}\sum_{g\in G}|a_g|^2 = \sum_{g\in G}|a_g|^2$$

The autocorrelation function is invariant under the addition of an arbitrary constant to the landscape f. Hence we can assume that $\langle f\rangle = 0$, which in turn implies that $\kappa = 0$ does not contribute to the autocorrelation function. Consequently we can decompose the empirical autocorrelation function of any landscape on Γ as

$$\hat{\rho}(\mu) = \sum_{\kappa\neq 0}\left[\frac{\sum_{g\in\kappa}|a_g|^2}{\sum_{g\in G}|a_g|^2}\right]\rho_\kappa(\mu) =: \sum_\kappa \Omega_\kappa\rho_\kappa(\mu) \tag{31}$$

Obviously the coefficients are all non-negative and their sum is 1. ∎

Corollary. All empirical autocorrelation functions fulfill

$$\sum_\mu \hat{\rho}(\mu)|\mu| = 0. \tag{32}$$

Proof. We have $\rho_0 \equiv 1$, hence the above expression is $\langle\rho_0;\hat{\rho}\rangle = \Omega_0$, which vanishes by theorem 4. ∎

4.2. Simple Random Walks on Landscapes

Before we proceed to landscapes on more general graph let us consider the empirical correlation function along a simple random walk. The geometric relaxation of such a random walk is conveniently described by the probabilities $\varphi_{s\kappa}$, defined in section 2.6, that a random walks of length s ends in symmetry class κ. An explicit representation of $\varphi_{s\kappa}$ for vertex transitive graphs is (Stadler 1994a)

$$\varphi_{s\kappa} = \sum_{\mu} \frac{|\kappa|\, \rho_\mu(\kappa)}{\langle \rho_\mu ; \rho_\mu \rangle} \lambda_\mu^s. \tag{33}$$

On a sequence space with alphabet size α we obtain explicitly

$$\varphi_{sd} = \sum_{p=0}^{n} \mathbf{K}_p^{(\alpha,n)}(d) \frac{(\alpha-1)^p \binom{n}{p}}{\alpha^n} \left(1 - \frac{\alpha}{\alpha-1}\frac{p}{n}\right)^s. \tag{34}$$

As in section 3.3, a random walk $\{x_0, x_1, \ldots\}$ induces a time series $\{f(x_0), f(x_1), \ldots\}$ on a landscape f.

The autocorrelation function of this time series is defined by

$$\hat{r}(s) = \frac{\langle f(x_t) f(x_{t+s}) \rangle_t - \langle f \rangle^2}{\langle f^2 \rangle - \langle f \rangle^2}. \tag{35}$$

Without loosing generality we assume as above that $\langle f \rangle = 0$. Clearly the denominator is just the variance of the landscape, $\sigma^2 = \langle f^2 \rangle$. The landscape autocorrelation and the random walk autocorrelation are related by

$$\hat{r}(s) = \sum_{\kappa} \varphi_{s\kappa} \hat{\rho}(\kappa), \tag{36}$$

see, e.g., Weinberger (1990). The relation between the correlation functions $\hat{\rho}$ and $\hat{r}$ is completely analogous to the relation of ρ and r in section 3. In particular, we have

Theorem 6. Let Γ be a CGCG with an adjacency matrix with left eigenvalues λ_μ and corresponding eigenvectors ρ_μ. Then $\hat{r}(s)$ is an autocorrelation function of a simple random walk of a landscape defined on Γ if and only if

$$\hat{r}(s) = \sum_\mu \Omega_\mu \lambda_\mu^s, \qquad \Omega_0 = 0, \qquad \Omega_\mu \geq 0, \qquad \sum_\mu \Omega_\mu = 1. \tag{37}$$

Proof. The covariance term is

$$\begin{aligned}
&\left\langle \lim_{\tau\to\infty} \frac{1}{\tau} \sum_{t=1}^{\tau} f(x_t) f(x_{t+s}) \right\rangle_{\text{all possible walks}} \\
&\quad = \frac{1}{N} \sum_{x\in\Gamma} f(x) \sum_{y\in\Gamma} f(y) \text{Prob}\{y = x_{t+s} | x = x_t\} \\
&\quad = \frac{1}{N} \sum_{x\in\Gamma} f(x) \sum_{y\in\Gamma} [T^s]_{xy} f(y)
\end{aligned}$$

where $T = (1/D)A$ is the transition matrix of the simple random walk on Γ. Using the Fourier decomposition of the landscape on a CGCG we find hence

$$\begin{aligned}
\hat{r}(s) &= \frac{1}{\sigma^2} \frac{1}{N} \sum_{x\in G} \sum_{m\in G} a_m^* e_m^*(x) \sum_{y\in G} [T^s]_{xy} \sum_{l\in G} a_l e_l(y) \\
&= \frac{1}{\sigma^2} \sum_{m\in G} \sum_{l\in G} a_m^* a_l \frac{1}{N} \sum_{x\in G} \lambda_l^s e_m^*(x) e_l(x) \\
&= \frac{1}{\sigma^2} \sum_{l\in G} |a_l|^2 \lambda_l^s
\end{aligned}$$

As in the proof of theorem 4 the contributions of a_0 cancel. Hence, using that λ_l depends only on the symmetry class μ to which l belongs yields

$$\hat{r}(s) = \sum_\mu \left[\frac{\sum_{l\in\mu} |a_l|^2}{\sum_{l\in G} |a_l|^2} \right] \lambda_\mu^s = \sum_{\mu\neq 0} \Omega_\mu \lambda_\mu^s. \tag{38}$$

Clearly the coefficients are the same as for the decomposition of the autocorrelation function $\hat{\rho}$ of the landscape itself. ∎

The eigenvalues of the transition matrix T are explicitly known for sequence spaces:

$$\lambda_d = \Lambda_d / D = 1 - \frac{d}{n} \frac{\alpha}{\alpha - 1}. \tag{39}$$

4.3. General Vertex Transitive Graphs

Let us denote the set of indices of eigenvectors θ_p belonging to the same eigenvalue $\tilde{\Lambda}_i$ of Δ by $L(i)$.

Definition. Let Γ be vertex transitive with the Θ-property, and let $\{\theta_y\}$ be a Fourier basis. Then define

$$\omega_i(\mu) = \frac{1}{|\mu|} \frac{\langle \theta_y, R^{(\mu)} \theta_y \rangle}{\langle \theta_y, \theta_y \rangle}, \tag{40}$$

where $y \in L(i)$.

Theorem 7. Let f be a landscape on a vertex transitive graph Γ which has the Θ-property. Then its autocorrelation function is of the form

$$\hat{\rho}(\mu) = \sum_i \left[\frac{\sum_{q \in L(i)} a_q^2}{\sum_{q \neq 0} a_q^2} \right] \omega_i(\mu) \tag{41}$$

Proof. Since Γ has the Θ-property we know that θ_q is an eigenvector of $R^{(\mu)}$ with an eigenvector $r_i^{(\mu)}$ which is a polynomial of the eigenvalue Λ_i of A corresponding to the eigenvector θ_q. Inserting the Fourier expansion into equation (17) yields then

$$\hat{\rho}(\mu) = \frac{1}{|\mu|} \frac{\langle f, R^{(\mu)} f \rangle}{\langle f, f \rangle} = \frac{1}{|\mu|} \frac{\sum_q a_q^2 \langle \theta_q, R^{(\mu)} \theta_q \rangle}{\sum_q a_q^2 \langle \theta_q, \theta_q \rangle} \tag{42}$$

Rearranging the terms and using the definition of ω_i yields the claim of the theorem. ∎

Theorem 5 shows immediately that $\omega_i \equiv \rho_i$ for all i is true for Cayley graphs of commutative groups. Furthermore we find this result for all example graphs discussed in section 2.10. This, and the very form of the above decomposition, strongly suggests the following

Conjecture 2. Let Γ be a vertex transitive graph with the Θ-property. Then $\omega_i \equiv \rho_i$.

We have not been able to find a counter example to this conjecture. If conjecture 2 is true, then the above theorems 5 and 6 hold at least for all vertex transitive graphs with the Θ-property.

4.4. Elementary Landscapes

Definition. Let Γ be a graph, Δ the Laplacian of Γ and f a landscape on Γ. We say f is *elementary* if

$$\Delta f + K'(f - \bar{f}) = 0, \tag{43}$$

where $\bar{f} = \langle f \rangle = 1/N \cdot \sum_{x \in \Gamma} f(x)$ is the average value of f, and K' is a constant.

Grover (1992) observed that the landscapes of a number of classical combinatorial optimization problems are of this form, see Table 2. In order to keep the notation consistent with Grover's work for regular graphs we introduce $K = K'/D$ where D is as usual the vertex degree. Equation (43) is the discrete analogue of the Helmholtz equation, or reduced wave equation, (Miller, 1984). An equivalent interpretation of (43) is that f is an eigenfunction of the Laplacian operator. It is not surprising that property (43) is closely related to the properties of the empirical autocorrelation function of the landscapes:

Theorem 8. Let Γ be a Cgcg with vertex degree D. Then the following statements are equivalent:

(i) f is an elementary landscape with parameter K.

(ii) The empirical autocorrelation function $\hat{\rho}$ of f is a left eigenvector of the collapsed adjacency matrix $\hat{A}$ with eigenvalue $\Lambda = D\lambda = D(1-K) = D-K'$.

(iii) The empirical autocorrelation function $\hat{r}(s)$ of f measured along a simple random walk on Γ is $\hat{r}(s) = \lambda^s$.

Proof. (i $\Longleftrightarrow$ ii) First recall that $\Delta = A - DE$. Thus an elementary landscape has a Fourier expansion of the form $f = \sum_j a_j e_j$ with $Ae_j = D\lambda e_j$, and from the proof of theorem 5 we recall that the autocorrelation function of f can be written as

$$\hat{\rho}(\mu) = \sum_g |a_g|^2 \frac{1}{|\mu|} \sum_{z \in \mu} e_g(z) = \sum_\kappa \sum_{g \in \kappa} |a_g|^2 \rho_\kappa(\mu)$$

in terms of the eigenvectors of A and $\hat{A}$, respectively. Hence we conclude that $\hat{\rho}$ is a eigenvector of $\hat{A}$ if and only if all e_i with non-zero coefficients belong to the same eigenvalue of A. Consequently f itself is an eigenvector of A if and only if this is the case.

(ii $\Longleftrightarrow$ iii) has been shown in Stadler (1994a). ∎

Theorem 8 forms the main result of this paper. While we have a complete proof only for the special case of Cayley graphs of commutative groups, there is ample evidence that it holds for a much broader class of configuration spaces. In particular it holds true whenever $\rho_i = \omega_i$ for all eigenvalues i, see section 4.3. A related result linking elementary landscapes to properties of their correlation functions on a larger class of graphs is given below:

Let us call $r(1)$ the *empirical nearest neighbor correlation* of the landscape. If Γ is symmetric, that is, if all pairs of neighbors belong to the same orbit, then $\hat{r}(1) = \hat{\rho}(\mathbf{1})$, see the corollary to theorem 4 in section 3.3.

Lemma 12. Let f be an elementary landscape on a symmetric graph Γ. Then

$$\hat{r}(1) = 1 - K, \tag{45}$$

Table 2. Elementary Landscapes.

Problem	Move Set	K	$1-\lambda$
NAES	Hamming	$\frac{4}{n}$	
p-spin	Hamming	$\frac{2p}{n}$	
WP	Hamming	$\frac{4}{n}$	
GC	Hamming	$\frac{2\alpha}{(\alpha-1)n}$	
GBP	Exchange	$\frac{8}{n}-\frac{8}{n^2}$	$\frac{8}{n}-\frac{8}{n^2}$
symmetric TSP	Transposition	$\frac{4}{n}$	$\sim\frac{4}{n}$
	Inversions	$\frac{2}{n-1}$	$\sim\frac{2}{n}$
GMP	Transposition	$\frac{4}{n}$	$\sim\frac{4}{n}$

For Hamming graphs (i.e., generalized sequence spaces) the relation $K = 1 - \lambda$ is established by theorem 8. The configuration space of the GBP is the Johnson scheme $J(n, n/2)$. The definitions of the optimization problems compiled here is discussed in the appendix.
The values K for NAES (Non-All-Equal-Satisfyability), WP (Weight Partition), GC (Graph Coloring), GBP (Graph Bipartitioning), and TSP (Traveling Salesman Problem) are taken from Grover (1992). The value of K for the GMP (Graph Matching Problem) is derived in appendix B. The values of λ for the GBP and the GMP problem are taken from Stadler and Happel (1992) and Stadler (1992), respectively; the remaining values are from Fontana *et al.* (1993a).

$-DK$ being the eigenvalue of the graph Laplacian corresponding to the landscape.

Proof. First we observe that

$$r(1) = \frac{1}{D}\frac{\langle f, Af\rangle}{\langle f, f\rangle} = 1 + \frac{1}{D}\frac{\langle f, \Delta f\rangle}{\langle f, f\rangle}. \tag{46}$$

If f is elementary, and $-DK$ is the eigenvalue of the graph Laplacian belonging to f, then (45) reduces to (46). ∎

The structure of an elementary landscape is closely related to the constant K', i.e., the corresponding eigenvector of the graph Laplacian. The solutions of the *Laplace equation* $\Delta f = 0$ form the *harmonic functions* on Γ. It is well known that there are no non-trivial harmonic functions on finite connected graphs (see, e.g., Soardi 1994). The harmonic functions correspond to the flat landscapes.

Let f be an arbitrary landscape on Γ. Let

$$V_+ = \{x \in V | f(x) \geq 0\} \qquad \text{and} \qquad V_- = \{x \in V | f(x) \leq 0\}$$

be the set of all vertices on which f is non-negative or non-positive, respectively. Let Γ_+ and Γ_- be the corresponding induced subgraphs of Γ. The connected components of Γ_+ and Γ_- are called the *nodal domains* of f. Clearly the geometry of the nodal domains is a very important characteristic of the landscape f.

The second largest eigenvalue $\tilde{\Lambda}_1$ of the graph Laplacian, which is negative for all connected graphs, is often called the *algebraic connectivity* of Γ. Theorem 2.5.7 in Brualdi and Ryser (1991) states that if f is an eigenvector of Δ with eigenvalue $\tilde{\Lambda}_1$, then both Γ_+ and Γ_- are connected, i.e., there are exactly two nodal domains. Following a suggestion by Kauffman we term this type of landscapes "Fujijama" landscapes as they consist of only a single "mountain" Γ_+.

For Riemannian manifolds even more is known. *Courant's Nodal Domain Theorem* states that if all eigenvalues of the Laplacian on a Riemannian manifold are ordered as $0 = \tilde{\Lambda}_1 \geq \tilde{\Lambda}_2 \geq \ldots$, then the number of nodal domains of an eigenfunction θ_k belonging to $\tilde{\Lambda}_k$ is at most k; see, e.g., the book by Chavel (1984). We expect that an analogous theorem holds on graphs. Results along these lines will be presented elsewhere.

Closely related to these observations is theorem 6 of Grover (1992). He showed that the local optima have a characteristic distribution on elementary landscapes: Let $z_{\min}$ and $z_{\max}$ be a local minimum and a local maximum, respectively. Then

$$f(z_{\min}) \leq \bar{f} \leq f(z_{\max}), \tag{44}$$

i.e., if f is an eigenvector of Δ then all local maxima are in Γ_+ and all local minima are in Γ_-. Grover also gives bounds on the length of adaptive walks for elementary

landscapes. We remark in this context that Palmer (1991) has suggested to use the distribution of local optima (instead of correlation measures) for defining and quantifying "rugged" landscapes. The eigenstructure of the graph Laplacian offers a promising formalism for relating the correlation structure of a landscape to the geometry of its local optima.

4.5. Classification of Landscapes

The characterization of autocorrelation functions (theorem 5 and theorem 6) has direct implications on the *classification* of autocorrelation function. We can now relate the classes proposed by Weinberger & Stadler (1993) with the spectral decomposition of the autocorrelation function.

We assume that we have a family $\{f_n\}$ of landscapes (or random fields) on a distance transitive graphs Γ_n. The index $n = 1, 2, \ldots$ is a "natural" measure of the size of the configuration space. For sequence spaces, n will be simply the chain length. Let $\rho_{[n]}(d)$ be the corresponding autocorrelation function. We define the scaled autocorrelation function $\zeta(h)$ for all rational numbers h by

$$\zeta(h) = \lim_{n\to\infty} \rho_{[n]}(nh).$$

Note that ζ does not necessarily exist! An example is the family of p-spin models with $p = n$. $\rho_{[n]}(n/2) = \pm 1$ depending on whether $n \equiv 0 \bmod 4$ or $n \equiv 2 \bmod 4$.

Lemma 13. Suppose ζ exists for all h in a neighborhood of 0.
(i) If ζ is continuous in 0, then $\lim_{n\to\infty} \rho_{[n]}(1) = 1$.
(ii) If ζ is of the form $\zeta(h) = 1 - c|h|^\gamma + o(|h|^\gamma)$, then $\lim_{n\to\infty} n^\gamma[1 - \rho_{[n]}(1)] = c$ with $c > 0$.

Proof. (i) Continuity implies $\lim_{h\to 0}\zeta(h) = \lim_{n\to\infty}\zeta(1/n) = \lim_{n\to\infty}\rho_{[n]}(1) = 1$.
(ii) We have $\zeta(1/n) = 1 - c\cdot n^{-\gamma} + o(n^{-\gamma})$. Hence $n^{\gamma}(1-\rho_n(1)) = c - o(1) \to c$. The constant c must be positive since $\rho_{[n]}(1) < 1$. ∎

On the sequence spaces we have explicitly

$$\zeta(\frac{1}{n}) = 1 - \frac{1}{n}\frac{\alpha}{\alpha-1}\sum_{p=0}\Omega_p\cdot p =: 1 - \frac{1}{n}\frac{\alpha}{\alpha-1}\bar{\pi}\,. \tag{47}$$

The parameter $\bar{\pi}$ can be interpreted as the "average mode" of the landscape. Hence the form of ζ depends on the scaling of $\bar{\pi}$. First we note that $1 \le \bar{\pi} \le n$ holds for all landscapes on sequences spaces. If $\bar{\pi} = \mathcal{O}(n)$ then $\zeta(\frac{1}{n})$ does not converge to 1 and hence ζ is not continuous. The initial slope of ζ can be estimated as $\partial\zeta/\partial h|_{0+} = -\frac{\alpha}{\alpha-1}\bar{\pi}$, if it exists at all. Hence we have an asymptotically finite slope only if $\bar{\pi} = \mathcal{O}(1)$.

Theorem 9. The smoothest landscape on a sequence space has autocorrelation function $\rho_1(d) = 1 - \frac{\alpha}{\alpha-1}\frac{d}{n}$.

Proof. The fact that the eigenvectors ρ_p are orthogonal polynomials implies immediately that ρ_1 is the slowest decaying mode. ∎

Table 3. Tentative Classification of Landscapes.

	Class I	rugged "Fractals"	Class II	smooth "Fractals"	Class III
ζ	not continuous	$\zeta = 1 - c\lvert h\rvert^{\gamma}$ $0<\gamma<1$	$\zeta = 1 - c\lvert h\rvert$	$\zeta = 1 - c\lvert h\rvert^{\gamma}$ $1<\gamma<2$	$\zeta = 1 - c\lvert h\rvert^{2}$
$\bar{\pi}$	$\bar{\pi} = \mathcal{O}(n)$	$\bar{\pi} \sim o(n)$, $\bar{\pi}\to\infty$	$\bar{\pi} = \mathcal{O}(1)$	do not exist on $\mathcal{Q}_\alpha^n$	
	Class A	Class B	Class C		

The classification scheme in the top line is taken from Weinberger and Stadler (1993), the scheme in the last line is explained in the text.

This implies that there are three different classes of landscapes, see also table 3:

A $\bar{\pi} = \mathcal{O}(n)$ implies $\lim_{n\to\infty} \zeta(\frac{1}{n}) < 1$ and hence the asymptotic autocorrelation function is discontinuous at the origin.

B $\bar{\pi} \to \infty$ but $\bar{\pi} = o(n)$ allows ζ to be continuous with infinite initial slope.

C $\bar{\pi} \sim \mathcal{O}(1)$ implies an autocorrelation function which decays continuously with finite slope.

We do *not* claim that there is one-to-one correspondence between the two classification schemes, since we have no general conditions for the existence of ζ.

5. Model Random Fields

5.1. The Random Energy Model

The results derived in the previous section allow us to calculate a "spectrum" $\Omega(\ell) = \Omega_\ell$ of the autocorrelation functions: $\rho(d) = \sum_{\ell=0}^{n} \Omega_\ell \rho_\ell(d)$.

We consider here a random field on a vertex transitive graph, where the values $f(x)$ are i.i.d. random variables (see, e.g., Gardiner 1990). This Random Energy Model (REM) introduced by Derrida (1980ab, 1981) has been applied to a variety of problems, in particular to the maturation of the immune system (Kauffman *et al.* 1988, Macken *et al.* 1991).

Lemma 14. Let Γ be vertex transitive. The REM is isotropic and has autocorrelation function $\rho(\mu) = 0$ for all $\mu \neq 0$.

Proof. Since $f(x)$ are i.i.d. random variables we have $\mathcal{E}[f(x)] = \bar{f}$ and $\mathcal{E}[f^2(x)] = S^2$. Independence implies uncorrelatedness, hence $\mathcal{E}[f(x)f(y)] - \bar{f}^2 = 0$ for $x \neq y$. Hence the autocorrelation function vanishes for all symmetry classes μ except for the symmetry class consisting only of the references vertex 0. ∎

As a consequence a Gaussian REM on a Boolean hypercube – that is, Derrida's original model – is a superposition of p-spin models. Derrida argued that the REM is the limit $p \to \infty$ of the p-spin models. The following theorem makes the sense in which this is true more precise.

Theorem 10. Let Γ be vertex transitive. Then the spectrum of the REM is

$$\Omega(\ell) = \frac{1}{\langle \rho_\ell ; \rho_\ell \rangle} \tag{48}$$

On a Boolean hypercube $\mathbf{S}_2^n$ we have explicitly: $\Omega(p) = \frac{1}{2^n}\binom{n}{p}$.

Proof. From equ.(9) we find

$$\Omega(\ell) = \frac{1}{\langle \rho_\ell ; \rho_\ell \rangle} \sum_\mu \rho_\ell(\mu)\delta_{0\mu}|\mu| = \frac{|0|\rho_\ell(0)}{\langle \rho_\ell ; \rho_\ell \rangle} = \frac{1}{\langle \rho_\ell ; \rho_\ell \rangle}.$$

On sequence spaces we have explicitly $\frac{1}{\langle \rho_\ell ; \rho_\ell \rangle} = \alpha^{-n}(\alpha-1)^\ell \binom{n}{\ell}$. ∎

Hence the REM on the Boolean hypercube can be represented as a superposition of a "continuum" of p-spin models with $p \approx n/2$. Note that the average mode $\bar{\pi}$, as defined in the previous section, becomes $\bar{\pi} = n/2$.

5.2. Nk Models

The Nk model assigns a real valued "fitness" to the bit string $\mathbf{b}$ by first assigning a real valued "fitness contribution", f_i , to the i^{th} bit, b_i, in $\mathbf{b}$. Each such assignment depends, not just on i and the value of b_i, but also on $0 \leq k < n$ other bits, which we call its "neighbors". The fitness contribution of each site is a random function, $f_i(\mathbf{b}_i)$, of the substring, $\mathbf{b}_i$, formed by the i^{th} bit and its k neighbors. $f_i(\mathbf{b}_i)$ is assigned by selecting an independent random variable from some distribution, such as the uniform or Gaussian distributions, for each of the 2^{k+1} possible values of $\mathbf{b}_i$, thus generating a "fitness table" for the i^{th} site. There is a different, independently generated table for each of the n sites. Then, given any string of n bits, the total fitness of the string, f, is defined as the average of the fitness contributions of each site; that is,

$$f(\mathbf{b}) = \frac{1}{n}\sum_{i=1}^{n} f_i(\mathbf{b}_i). \tag{49}$$

The Nk model comes a different "flavors" depending on the choice of neighbors influencing a particular site fitness. The simplest — but not the only — way of choosing neighbors, at least for even k, is to use the k sites adjacent to site i; that is, the bits at sites $i - k/2$ through $i + k/2$. As in the original formulation of the model, we introduce periodic boundary conditions to assign neighbors to sites i with $i \leq k/2$ and $i \geq n - k/2$. In other words, we assume that the sites are arranged in a circle, such that site n is next to site 1. (Periodic boundary conditions are chosen because they minimize chain length dependent end effects and because we are interested in bulk properties only). This choice of neighbors gives rise to a class of short range spin glasses. We will call this arrangement the "adjacent neighborhood" model, AN. Alternatively, we could assign the neighbors by randomly selecting, for each site i, k other sites. This assignment of neighbors makes the model similar to a long range, dilute spin glass. It will be referred to as "random neighborhood" model, RN. In a third variation we drop the requirement that f_i depends on bit i. This version will be called "purely random" model, PR. The Nk model reduced to the REM for $k + 1 = n$, see e.g., Kauffman (1993).

Lemma 15. The autocorrelation functions of the the Nk model with random, purely random, and adjacent neighborhood, respectively are given by

$$\begin{aligned}
\rho_{RN}(d) &= (1 - \frac{d}{n})(1 - \frac{k}{n-1})^d \\
\rho_{PR}(d) &= (1 - \frac{k+1}{n})^d \\
\rho_{AN}(d) &= 1 - \frac{k+1}{n}d + \frac{1}{\binom{n}{d}} \sum_{j=1}^{\min(k,n+1-d)} (k-j+1)\binom{n-j-1}{d-2}.
\end{aligned} \tag{50}$$

Proof. For a proof see, e.g., the appendix in Fontana *et al.* 1993a. ∎

Theorem 11. The autocorrelation function of the adjacent neighbor Nk model, ρ_{AN}, is a polynomial of degree $k + 1$. Its autocorrelation function is therefore of the form

$$\rho_{AN}(d) = \sum_{p=0}^{k+1} \Omega_p \rho_p(d). \tag{51}$$

Proof. We start by rewriting ρ_{AN} without using binomial coefficients:

$$\begin{aligned}\rho_{AN}(d) &= 1 - \frac{k+1}{n}d + \sum_{j=1}^{\min(k,n+1-d)} (k+1-j)\frac{(n-j-1)!\,d!\,(n-d)!}{(n-j-d+1)!\,(d-2)!\,n!} \\ &= 1 - \frac{k+1}{n}d + \frac{d(d-1)}{n(n-1)}\sum_{j=1}^{k}(k+1-j)\prod_{\ell=0}^{j-2}\frac{n-d-\ell}{n-2-\ell}\end{aligned}$$

where an empty product is 1 by definition. This is obviously a polynomial in d. We have for $k \geq 1$

$$\text{Degree} = \max\left\{0;\ 1;\ 2 + \max_{1\leq j\leq k}[j-1]\right\} = k+1,$$

since d^{k-1} appears with non-zero coefficient in the product. For $k = 0$ we have explicitly $\rho(d) = 1 - d/n$, i.e., the degree is 1. Since the ρ_j are polynomials of degree j we can write ρ_{AN} as a superposition of the eigenfunctions ρ_j, $j \leq k+1$.

∎

The coefficients $\Omega(\ell)$ of the spectrum can be obtained explicitly for sequence spaces, if the correlation function $\rho(d)$ is known.

$$\begin{aligned}\Omega(\ell) = \frac{\langle \rho_\ell ; \rho \rangle}{\langle \rho_\ell ; \rho_\ell \rangle} &= \frac{(\alpha-1)^\ell \binom{n}{\ell}}{\alpha^n}\sum_{d=0}^{n}(\alpha-1)^d \binom{n}{d}\rho_\ell(d)\rho(d) = \\ &= \frac{n!}{\alpha^n}\sum_{i=0}^{n}\rho(i)\sum_{j=0}^{p}\frac{(-1)^j(\kappa-1)^{p+i-j}}{j!(i-j)!(p-j)!(n+j-i-p)!}\end{aligned} \tag{52}$$

For Nk models this expression has been evaluated numerically, results are shown in figure 1. In particular, in case of the AN-version the data suggest the following

Conjecture 3. The coefficients $\Omega(p)$ of the AN-Nk model are independent of the system size n and fulfill the symmetry relation

$$\Omega(p) = \Omega(k+1-p) \qquad \text{for all} \quad 0 \leq p \leq k+1. \tag{53}$$

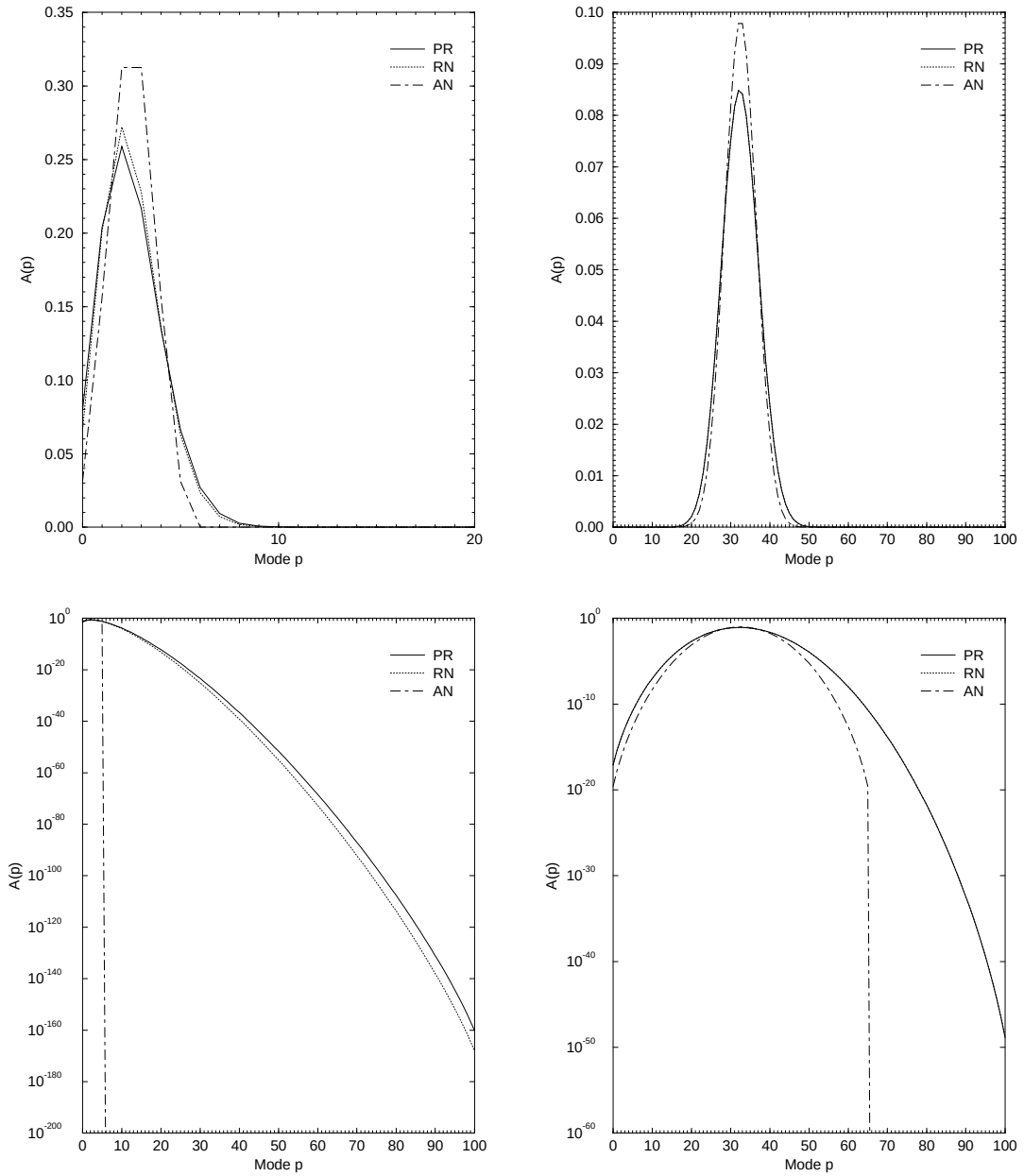


Figure 7: Spectral decompositions of Nk models on Boolean hypercubes with $n = 100$. Numerical data strongly suggest that for the AN model the symmetry relation $\Omega(p) = \Omega(k+1-p)$ holds for all n. The logarithmic plots (bottom) emphasize that the two random neighborhood Nk models involve all modes $0 \leq p \leq n$ for all k while the AN models are superpositions of the modes $0 \leq p \leq k+1 \leq n$.

Evidence. Evaluation of equ.(52) for a large number of special cases in the range $n \leq 200$ and $0 \leq k < n$ using `Mathematica` are fully consistent with the conjecture.

Weinberger and Stadler (1993) concluded bases on a comparison of correlation lengths, that Nk models roughly resemble p-spin models with $p = (k+1)/2$. This view is supported by the spectra of the Nk models showing that the dominating modes are those centered around $\bar{\pi} = (k+1)/2$ for all three versions of the Nk model, see Figure 7. The convergence of Nk models and p-spin models is similar to the case of the p-spin models and the REM. If k scales linearly with n, i.e., if $(k+1)/n \to \chi > 0$, then we expect that the spectrum in terms of the scaled variables $\psi = p/n$ will be concentrated at $\bar{\psi} = \chi/2$ in the limit $n \to \infty$. The corresponding numerical calculations are not shown here.

The major difference between the different Nk models is that for the AN model only the first $k+1$ coefficients play a role, while for the other two models all modes of the Fourier expansion contribute. An interesting observation in this context is due to Weinberger (1991c): He considered the Nk decision problem *"Is the global optimum of a given Nk model larger than some specified value M?"*, and showed that the AN version can be solved in $\mathcal{O}(2^{k+1}n)$, i.e., in polynomial time for constant k, while the RN and PR versions are NP-complete for $k \geq 3$.

5.3. Superposition of Nk Models

In theorem 3 we have seen that all isotropic random fields on the Boolean hypercube can be constructed by superpositions of p-spin models. In the following we

will show that Nk-models cannot be used for this purpose. Let $\{\mathcal{N}_k(x)\}$ denote a set of Nk models with parameter k, such that

$$\mathcal{E}[\mathcal{N}_i(x)\mathcal{N}_j(y)] - \mathcal{E}[\mathcal{N}_i(x)]\mathcal{E}[\mathcal{N}_j(y)] = 0 \tag{54}$$

whenever $i \neq j$, i.e., we suppose that the entries in the fitness tables are uncorrelated also between the random fields. Given the construction of the Nk model this is the natural assumption. Then we have

Theorem 12. Denote by $\Omega(p)$ the spectrum of an isotropic random field of the form $\Xi = \sum_{k=0}^{k_{\max}} \beta_k \mathcal{N}_k$, i.e., Ξ is a superposition of Nk models. Then $\Omega(p) > 0$ for all $p \leq k_{\max} + 1$. Furthermore $\Omega(p) > 0$ for all $p \leq n$ if at least one of the Nk models $\mathcal{N}_k$ is an RN or a PR version of the Nk model.

Proof. We have

$$\begin{aligned}
&\mathcal{E}[f(x)f(y)] - \mathcal{E}[f(x)]\mathcal{E}[f(y)] = \sum_{i,j} \beta_i\beta_j \mathcal{E}[\mathcal{N}_i(x)\mathcal{N}_j(y)] - \mathcal{E}[\mathcal{N}_i(x)]\mathcal{E}[\mathcal{N}_j(y)] = \\
&= \sum_{k=0}^{k_{\max}} \beta_k^2 \left\{\mathcal{E}[\mathcal{N}_k(x)\mathcal{N}_k(y)] - \mathcal{E}[\mathcal{N}_k(x)]\mathcal{E}[\mathcal{N}_k(y)]\right\} = \\
&= \sum_{k=0}^{k_{\max}} \beta_k^2 \Big[\mathcal{E}[\sum_p \Omega_k(p)\mathcal{H}_p(x) \sum_q \Omega_k(q)\mathcal{H}_q(y)] - \\
&\qquad\qquad - \mathcal{E}[\sum_p \Omega_k(p)\mathcal{H}_p(x)]\mathcal{E}[\sum_q \Omega_k(p)\mathcal{H}_q(x)]\Big] = \\
&= \sum_{k=0}^{k_{\max}} \beta_k^2 \sum_p \sum_p \Omega_k(p)\mathcal{E}[\mathcal{H}_p(x)\mathcal{H}_p(y)] = \sum_{k=0}^{k_{\max}} \beta_k^2 \sum_p \sum_p \Omega_k(p)\rho_p(x,y)\sigma_p^2 = \\
&= \sum_{p=0}^{n} \left[\sigma_p^2 \sum_{k=0}^{k_{\max}} \Omega_k(p)\beta_k^2\right] \rho_p(d(x,y)) := \sum_{p=0}^{n} B_p \rho_p(d(x,y))
\end{aligned}$$

The coefficients B_k are strictly positive as long as $\Omega_k(p) > 0$ for some $k \leq k_{\max}+1$, and this is true for all $p \leq k_{\max} + 1$ in the AN case, and for all p in the other two versions of the Nk-model. ∎

Corollary. A p-spin model cannot be represented as a superposition of Nk models for all $p \geq 1$. The empirical correlation function of the 1-spin model and the Nk-model with $k = 0$ coincide. The empirical correlation function of a p-spin model with $p \geq 2$ cannot be modeled by a superposition of Nk models.

Proof. The first claim follows immediately from the theorem. Since the zero-component ρ_0 is projected out by the definition of the empirical correlations function, we have $\hat{\rho}_{\mathrm{Nk:k=0}} = \hat{\rho}_{p=1} = \rho_1$. Since ρ and $\hat{\rho}$ coincide for p-spin models and differ only by the zero-component for Nk-models, the second claim is again an immediate consequence of theorem 9. ∎

An obvious consequence is then the following

Corollary. Almost all instances of Nk models with $k > 0$ are not elementary.

6. Model Landscapes

6.1. Traveling Salesman Problems

Let $W \in \mathbb{R}^{n,n}$ be a an arbitrary matrix, and let $\pi \in S_n$ be permutations of the cities. Then

$$f(\pi) = \sum_j w_{\pi(j)\pi(j-1)} \tag{55}$$

is the cost functions of the TSP with distance matrix W. We define

$$f^*(\pi) = f(\pi^*) = \sum_j w_{\pi(j)\pi(j+1)}; \tag{56}$$

this is the cost of the tour π walked through in reverse order, π^* being the "reverse order" permutation of π. Recalling that an arbitrary matrix W can be uniquely decomposed into its symmetric component $W^\sigma = (W + W^+)/2$ and its antisymmetric component $W^\alpha = (W - W^+)/2$ we introduce furthermore

$$\begin{aligned} f^\sigma(\pi) &= \sum_j w^\sigma_{\pi(j)\pi(j-1)} = \frac{f(\pi) + f^*(\pi)}{2}, \\ f^\alpha(\pi) &= \sum_j w^\alpha_{\pi(j)\pi(j-1)} = \frac{f(\pi) - f^*(\pi)}{2}. \end{aligned} \tag{57}$$

Note that f^σ and f^α can be viewed as cost functions of TSPs with "distance matrices" W^σ and W^α, respectively.

Theorem 13. Both f^σ and f^α are elementary landscapes on the Cayley graphs of the symmetric group with the transpositions and the inversions as generators, respectively. In particular we have for transpositions

$$\Delta f^\sigma + 2(n-1)(f^\sigma - \bar{f}) = 0 \qquad\qquad \Delta f^\alpha + 2n f^\alpha = 0, \tag{58}$$

and for inversions (2opt moves, Lin & Kernighan 1965) we find

$$\Delta f^{\sigma} + n(f^{\sigma} - \bar{f}) = 0 \qquad\qquad \Delta f^{\alpha} + \frac{n(n+1)}{2} f^{\alpha} = 0. \tag{59}$$

Proof. Without loosing generality we can consider the neighborhood of the identity permutation $\imath$, since the numbering of the cities is arbitrary.

(i) In the case of transpositions we start with Grover's (1992) formula

$$\begin{aligned} f((i,j)) - f(\imath) = \quad & w_{j,i-1} + w_{i+1,j} + w_{i,j-1} + w_{j+1,i} \\ & - w_{i,i-1} - w_{i+1,i} - w_{j,j-1} - w_{j+1,j} \\ & + (\delta_{j,i+1} + \delta_{j,i-1})(w_{ij} + w_{ji}). \end{aligned}$$

Summing over all $i \neq j$ yields

$$2\Delta f(\imath) = 4X - 4n f(\imath) + 2f(\imath) + 2f^{*}(\imath);$$

and analogously we obtain

$$2\Delta f^{*}(\imath) = 4X - 4n f^{*}(\imath) + 2f^{*}(\imath) + 2f(\imath).$$

Here X is the sum over all non-diagonal entries of W. The proposition follows immediately.

(ii) In case of inversions we find

$$f([k,l]) - f(\imath) = w_{l,k-1} + w_{l+1,k} - w_{k,k-1} - w_{l+1,l} + \sum_{j=k}^{l-1} (w_{j,j+1} - w_{j+1,j}).$$

Summing over all $k \neq l$ yields

$$2\Delta f(\imath) = 2X - 2n f(\imath) + \frac{n(n-1)}{2}[f^{*}(\imath) - f(\imath)].$$

A completely analogous result is obtained for $\Delta f^{*}(\imath)$, and a short calculation then completes the proof. ∎

Corollary. The landscape of a TSP with transpositions or inversions is elementary if and only if W is either symmetric or antisymmetric.

Asymmetric TSPs hence provide an example of fairly simple composite landscapes. They consist of two modes corresponding to the symmetric and the antisymmetric part of the distance matrix W, respectively. It is also interesting to note in this context that canonical transpositions $(i, i+1)$ do not lead to an elementary landscape.

Consequently, the nearest neighbor correlations of the symmetric and antisymmetric components of a TSP with transpositions are $r(1) = 1 - 4/n$ and $r(1) = 1 - 4/(n-1)$, respectively, i.e., very similar. In fact, numerical estimates (Stadler & Schnabl 1992) had been consistent with $r(1) \sim 1 - 4/n$ for large n in both cases. In the case of inversions we have a symmetric mode with nearest neighbor correlation $r(1) = 1 - 2/(n-1)$ and an antisymmetric contribution with a vanishingly small contribution $r(1) = -2/(n-1) \sim 0$.

It is interesting to correlate these values of $r(1)$ with known facts about the performance of heuristic optimization algorithm, in particular with the simulated annealing. It has been observed by serveral authors that simulated annealing on symmetric TSP is much more effective when reversals instead of transpositions are used as move set, see, e.g., the books by Aarts and Korst (1990) or Otten and vanGinneken (1989). Furthermore, Miller and Pekny (1991) have observed that reversals are a remarkably bad move set for asymmetric TSPs. These observations are in accordance with the conjecture that landscape with smoother correlation functions have fewer local optima and are thus easier to optimize on (Stadler and Schnabl 1992). In particular the difference between symmetric and asymmetric TSPs when reversals are used is easily explained in these terms: while for the symmetric TSP the landscape is as smooth as possible, it is completely rugged for the antisymmetric case.

Numerical data by Stadler & Schnabl (1992) indicate that the correlation functions $\hat{r}(s)$ of both the symmetric and the antisymmetric components should be of the form $r(1)^s$, we do not, however, have a proof for this conjecture for all n. To this end we would have to show that theorem 8 holds also for the Cayley graphs of the symmetric group. This would follow, for instance, if conjecture 2 were true. We are confident that the theory developed in section 4 for the Cayley graphs of commutative groups works in fact for all vertex transitive graphs. This view is also supported by recent numerical calculations (Krakhofer, personal communication).

6.2. Graph Matching Problem

In this subsection we show that another well-studied optimization problem leads to an elementary landscape on Cayley graphs of the symmetric group. For a definition of the Graph Matching Problems see the appendix.

Theorem 14. The landscape of the graph matching problem `GMP` with transposition metric is elementary with $DK = 2(n-1)$, i.e., $K = 4/n$.

Proof. Without loosing generality we may evaluate the cost function of the identity permutation $\imath$ since the labeling of the vertices is arbitrary. Using the labels k of the matched pairs as indices we have to following five types of neighbors, and the corresponding changes of the cost function for each of them

[11'] $[W_{2j-1,2i} + W_{2i-1,2j}] - [W_{2i-1,2i} + W_{2j-1,2j}]$,

[12'] $[W_{2j,2i} + W_{2j-1,2i-1}] - [W_{2i-1,2i} + W_{2j-1,2j}]$,

[21'] $[W_{2i-1,2j-1} + W_{2i,2j}] - [W_{2i-1,2i} + W_{2j-1,2j}]$,

[22'] $[W_{2i-1,2j} + W_{2j-1,2i}] - [W_{2i-1,2i} + W_{2j-1,2j}]$, and

[11] $W_{2i,2i-1} - W_{2i-1,2i}$.

Summing the first four terms over all $i \neq j$ (i.e., counting each neighbor twice) yields

$$[ug] + [ug] - 2f(\imath) - 2(n/2-1)f(\imath) \quad + \quad [gg] + [uu] - d - 2(n/2-1)f(\imath) \quad +$$
$$[uu] + [gg] - d - 2(n/2-1)f(\imath) \quad + \quad [ug] + [ug] - 2f(\imath) - 2(n/2-1)f(\imath),$$

where $[ug]$, etc., denotes the sum over all W_{ij} with odd i and even j, and d is the sum over all diagonal entries W_{kk}. Using that W is symmetric, i.e., $[ug] = [gu]$ this sum my be written as

$$2\Delta f(\imath) = 2\sum_{i \neq j} W_{ij} - 4(n-1)f(\imath),$$

since the transposition $(2i-1, 2i)$ gives a contribution of 0. Thus we have

$$\Delta f(\imath) = 2(n-1)\,[\bar{f} - f(\imath)], \tag{60}$$

i.e., $DK = 2(n-1)$, and f is elementary on the Cayley graph $\Gamma(S_n, \mathcal{T})$. ∎

It is interesting to note that theorem 14 is false if W is not symmetric, in complete accordance with the asymmetric `TSP`.

6.3. Graph Bipartitioning Problem

The `GBP` is also defined in appendix A. As a close relative of the Sherrington-Kirkpatrick model it has received considerable attention. The most natural move set is the exchange of a vertex in partition A with a vertex in partition B. The resulting graph, the Johnson graph $J(n, n/2)$, see, e.g., Cameron and vanLint

(1991), is distance transitive (Dunkl 1979). It is not known (to the authors) whether $J(n, n/2)$ can be constructed as a Cayley graph, and if so, if the group can be chosen to be commutative.

The random field associated with the GBP is constructed by assigning the weights $H_{ij} = H_{ji}$ independently and identically distributed. Stadler and Happel (1993) have shown by explicit calculations that the correlation functions of this random field are

$$\begin{aligned} \rho(d) &= 1 - 8\frac{d}{n}\left(1 - \frac{1}{n-2}\right) + 16\left(\frac{d}{n}\right)^2\left(1 - \frac{1}{n-2}\right) \\ r(s) &= \left(1 - \frac{8}{n} + \frac{8}{n^2}\right)^s, \end{aligned} \tag{61}$$

irrespective of the particular distribution from which the H_{ij}'s are drawn. From theorem 4 we conclude that $\rho(d)$ is an eigenvector with eigenvalue $r(1)$ of the Johnson graph $J(n, n/2)$.

Grover (1992) found that each instance of the GBP is an elementary landscape with $K = 8/n - 8/n^2$, see also table 1. Hence we find in fact the expected result $K = 1 - \lambda$. We take this as one further hint that the theorems proved in the previous sections for Cgcg are in fact true for a much larger class of configuration spaces.

6.4. The Low Autocorrelated Binary String Problem

The LABSP (Golay 1977, Bernasconi 1987) consists of finding binary strings σ over the alphabet $\{-1, +1\}$ with low aperiodic off-peak autocorrelation $R_k(\sigma) = \sum_{i=1}^{N-k} \sigma_i \sigma_{i+k}$ for all lags k. These strings have technical applications such as the

synchronization in digital communication systems and the modulation of radar pulses. The quality of a string σ is measured by the fitness function

$$f(\sigma) = \sum_{k=1}^{n-1} R_k(\sigma)^2. \tag{62}$$

In most of the literature on the LABSP the *merit factor* $F(\sigma) = n^2/(2f(\sigma))$ is used. This transformation gives rise to a non-Gaussian distribution of merit factors, see Bernasconi (1987) for details.

Lemma 16. The landscape f of the LABSP can be written as

$$f(\sigma) = a_0 + \sum_{k=1}^{\lceil \frac{n}{2} \rceil - 1} \sum_{i=1}^{n-2k} 2\varepsilon_{i,i+k}(\sigma) + \sum_{k=1}^{n-1} \sum_{i=1}^{n-1} \sum_{j \neq i-k,i,i+k} \varepsilon_{i,i+k,j,j+k}(\sigma), \tag{63}$$

and thus the empirical autocorrelation functions of the LABS are

$$\begin{aligned} \hat{\rho}(d) &= [1 - \mathcal{O}(1/n)]\rho_4(d) + \mathcal{O}(1/n)\rho_2(d), \\ \hat{\varrho}(s) &= [1 - \mathcal{O}(1/n)] \left(1 - \frac{8d}{n}\right)^s + \mathcal{O}(1/n) \left(1 - \frac{4d}{n}\right)^s. \end{aligned} \tag{64}$$

Proof. Recall that the functions $\varepsilon_{i_1,i_2,\ldots,i_p} = \sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_p}$, where all indices are different, are eigenvalues of the Laplacian of the hypercube belonging to the eigenvalues Λ_p, sect. 3.2. Provided the above representation of the landscapes $f(\sigma)$ is true we conclude immediately from theorems 5 and 6 that the correlation functions have the required form since $\Omega_2 = \mathcal{O}(n^2)$ and $\Omega_4 = \mathcal{O}(n^3)$. It is convenient to allow arbitrary indices, setting $\sigma_q = 0$ for all $q < 1$ or $q > n$. With this convention we have

$$\begin{aligned} f(\sigma) &= \sum_{k=1}^{n-1} \sum_{i=1}^{n-k} \sum_{j=1}^{n-k} \sigma_i\sigma_{i+k}\sigma_j\sigma_{j+k} \\ &= \sum_{k=1}^{n-1} \sum_{i=1}^{n-k} \left[\sum_{j \neq i-k,i,i+k} \sigma_i\sigma_{i+k}\sigma_j\sigma_{j+k} + \sigma_{i-k}\sigma_i^2\sigma_{i+k} + \sigma_i^2\sigma_{i+k}^2 + \sigma_i\sigma_{i+k}^2\sigma_{i+2k} \right] \\ &= \frac{n(n-1)}{2} + \sum_{k=1}^{n-1} \sum_{i=1}^{n-k} [\sigma_{i-k}\sigma_{i+k} + \sigma_i\sigma_{i+2k}] + \sum_{k=1}^{n-1} \sum_{i=1}^{n-k} \sum_{j \neq i-k,i,i+k} \sigma_i\sigma_{i+k}\sigma_j\sigma_{j+k} \end{aligned}$$

Using the definition of $\varepsilon_{\ldots}(\sigma)$ and eliminating all terms containing a σ_q with $q < 1$ or $q > n$ completes the proof. ∎

The landscape of the LABPS is thus not elementary, it consists of a superposition of two modes, namely $p = 2$ and $p = 4$. The smoother $p = 2$ contribution becomes negligible for large n, so that f behaves for long strings almost like an elementary $p = 4$ landscape. This fact explains why the LABPS has been found to be much harder for simulated annealing than, say, the SK spin glass (Bernasconi 1987). Stadler (1992) has computed the empirical autocorrelation function $\hat{r}(s)$ numerically based on the merit factor F. The numerical estimate for the correlation length

$$\ell \stackrel{\text{def}}{=\!=} -\frac{1}{\ln r(1)} \approx 0.1230 \cdot n - 0.983$$

is in excellent agreement with the asymptotic value $\ell = n/8 + \mathcal{O}(1)$ implied by lemma 16.

6.5. RNA Free Energy Landscapes

RNA sequences are strings of length n over the alphabet {**G**,**C**,**A**,**U**}. The canonical distance measure between two sequences of equal length n is the Hamming distance.

The physico-chemical properties of RNA molecules, and, in particular, their free energies of folding, ΔG, depend crucially on the spatial structure of the molecule. In the case of RNA the process of folding the one-dimensional primary structure (sequence) into the three-dimensional tertiary structure can be decomposed into two steps:

1) Folding of the sequence into a planar secondary structure by formation of complementary Watson-Crick base pairs **G≡C** and **A=U** and the weaker **G−U** pairs.
2) Formation of the 3D tertiary structure from the planar pattern.

The reason for this decomposition is that the intramolecular forces stabilizing the secondary structures — base pairing and base pair stacking — are much stronger than those accounting for the particular spatial arrangement. Thus the free energy for the formation of the entire spatial structure can be estimated by the free energy for the formation of the secondary structure. The dominant role of the secondary structure is also well documented in nature by the conservation of secondary structure elements in evolution (Cech 1988). There are efficient algorithms for the prediction of the secondary structure (Zucker and Sankoff 1984, McCaskill 1990, Hofacker *et al.* 1994).

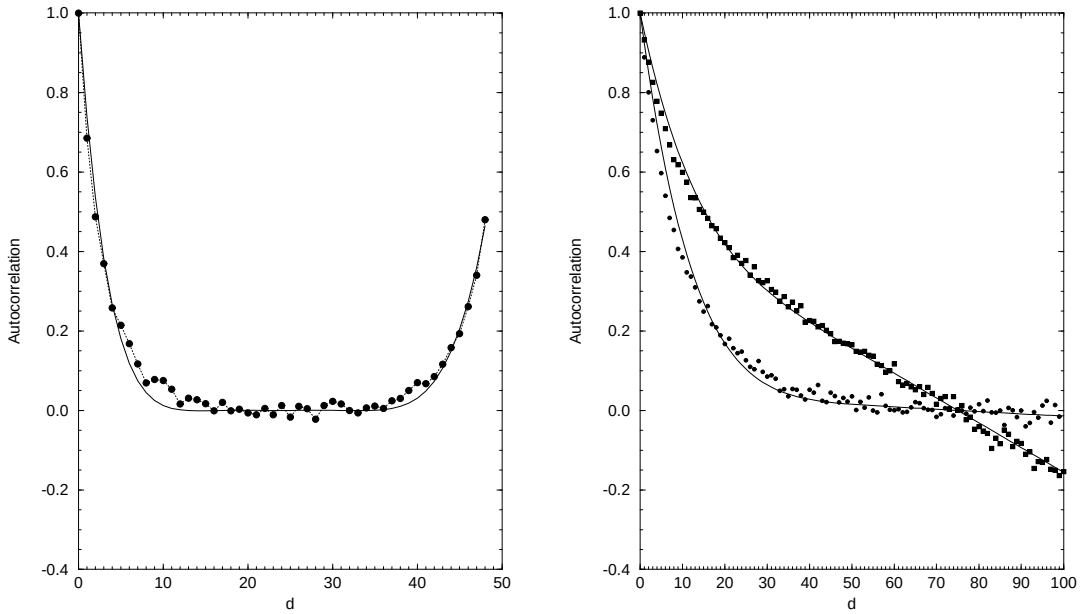


Figure 8: a) **GC** alphabet, chain length 50. Fitted curve is $0.73\rho_6 + 0.27\rho_7$. b) **GCAU** alphabet, □, and **GCXK** alphabet, •. The chain length is $n = 100$. The fitted curves are $\Omega_1\rho_1 + \Omega_6\rho_6$.

For our study we used the autocorrelation functions of RNA free energy landscapes discussed in Fontana *et al.* 1993a. An example, together with a fit by Krawtchouk polynomials is shown in figure 2. We find a pronounced difference between the natural **GCAU** alphabet and an artificial alphabet **GCXK** with two base pairs, **GC** and **XK**, of equal strength. While the latter is represented very well by the $\ell = 6$ mode, we obtain a superposition of the modes $\ell = 1$ and $\ell = 6$ in case of the natural alphabet. This does not come as a surprise, however. As shown in Stadler and Grüner (1993) there is an essentially linear increase of the free energy with increasing **GC** content which accounts for the $\ell = 1$ mode. (Recall that $\ell = 1$ corresponds to the additive fitness model, or 'Fujijama' landscape). Numerical data are given in table 4.

Table 4. Fits to RNA Free Energy Landscapes.

n	a_6			a_1		a_7
	GCAU	**GCXK**	**GC**	**GCAU**	**GCXK**	**GC**
20	0.68	0.86		0.32	0.14	
30	0.65	0.88		0.35	0.12	
50	0.57	0.93	0.73	0.43	0.07	0.27
70	0.57	0.93		0.43	0.07	
100	0.53	0.96		0.47	0.04	

The spectrum of the RNA data cannot be computed from equation (52) due to the noise in data. One has to resort to a least square fitting procedure that takes the non-negativity of the coefficients explicitly into account and uses only a small number of non-zero coefficients. We do not expect the RNA free energy landscapes to be *exactly* superpositions of only two modes, we rather expect a broad peak of the spectrum concentrated at about $\ell = 5 \ldots 8$. Using more than two modes increases the quality of the fits only slightly. Furthermore the fits with more than two coefficients do not compare well when independent data sets are used. We conclude hence that our numerical estimates of the correlation functions $\hat{\rho}(\mu)$ are not accurate enough (approximate noise level $\pm 5\%$) for obtaining a more detailed decomposition of the RNA free energy landscapes.

7. Discussion

We have developed a mathematical framework for a detailed analysis of fitness landscapes and random field models of fitness landscapes. The basic ingredient of this theory are the Fourier series implied by the geometry of the configuration space. The main results are the following:

Isotropic random fields are characterized in terms of their Fourier coefficients. All landscapes on the Boolean hypercube can be represented as a superposition of p-spin Hamiltonians, and isotropic Gaussian landscapes correspond exactly to a superposition of Derrida's version of p-spin models: the coupling coefficients are uncorrelated and the variance of the coefficients depends only on the order p of the coupling.

The Fourier expansion of (fitness) landscapes can be used to find the set of all possible correlation functions; it turns out that these are determined by the geometry of the configuration spaces underlying the landscapes. More precisely, the set of all autocorrelation functions forms a simplex spanned by suitably normalized left eigenvectors of the collapsed adjacency matrix of the configuration space. The corresponding eigenvalue determines the decay of the correlation function along a simple random walk on the landscape.

We have shown that elementary landscapes — that is landscapes fulfilling the analogue of the Helmholtz equation: $\Delta f + K f = 0$, or in other words, which are eigenfunctions of the graph Laplacian — are precisely the landscapes whose correlation functions are eigenvectors of the collapsed adjacency matrix, at least if the configuration space is a Cayley graph of a commutative group. Elementary landscapes exhibit a characteristic distribution of local optima on the configuration space which depends crucially on the corresponding eigenvalue of the Laplacian.

The formalism derived in this contribution suggests to approximate a given landscape by a superposition of statistical models that are elementary landscapes (which are hopefully easier to understand). The coefficients of such a superposition can be obtained by observing that the correlation function is a superposition of the correlation functions of the elementary landscapes under consideration. The latter can be computed analytically for a number of problems, since they are the left eigenvectors of the collapsed adjacency matrix of the underlying configuration space. Instead of the correlation function we can then use the "spectrum", that is the coefficients of this decomposition, as a means of describing the landscape. On a Boolean hypercube, for instance, this spectrum can be interpreted directly as the relative importance of p-ary (spin) interactions.

We have studied the spectra of a number of random fields and landscapes. In case of an uncorrelated (random energy) landscape the spectrum has a peak at the average distance in configuration space, and a width of the order $\mathcal{O}(\sqrt{n})$. Nk models also lead to a broad peak in the spectrum, the location of which depends on the k parameter of the model. TSPs decompose into two elementary landscapes corresponding to the symmetric and the antisymmetric part of the distance matrix, respectively. In the latter case the nearest neighbor correlations depend on the choice of the move set. The data are consistent with the conjecture that landscapes with higher correlation have fewer local optima and are hence easier to optimize on.

As a further example we have studied RNA free energy landscapes. It turns out that the natural **GCAU** landscape can be fitted reasonably well by a superposition of a 1-spin and a 6-spin model with equal amplitude, while the **GC** and **GCXK** landscapes can be approximated by a 6-spin model. We do not expect that RNA landscapes are elementary, we rather expect a distribution of p-spin models centered around $p = 6$. The numerical correlation data are to noisy, however,

to permit a higher resolution, and we have to be content with the approximate location of the peak. The order of the dominating modes allows also to draw conclusions on the approximate size of the building blocks underlying the landscape: in case of RNA free energies the data indicate that on average a nucleotide interacts strongly with about 5 other position, leading to a building block size of 6. This is a perfectly reasonable number given the energy model (Zuker and Sankoff 1984) for RNA secondary structure formation: a major contribution of the energy comes from stacking of base pairs, hence we expect at least the pairing partners and their neighbors to be important.

A number of questions remain only partially solved. The most important of them is whether the results summarized above can be generalized to less symmetric types of configuration spaces. Analytical results for a variety of special cases as well as numerical calculations suggest that this is in fact possible, the main implications of this paper should hold for all vertex transitive configuration spaces. The details of the relation between local optima and correlation structure are far from being understood. A first step in this direction would be a proof (or a counterexample) for a discrete version of Courant's nodal domain theorem. A related problem concerns the number of local optima within a single nodal domain.

The relation between the Fourier expansion of a landscape and the computational complexity of the optimization problem on the landscape is of great importance for devising practical optimization heuristics. Disappointingly, there seems to be **no** simple relation. Most of the elementary landscapes discussed in this paper, such as the `TSP`, are $\mathcal{NP}$-complete, but have only a fairly small number of Fourier coefficients, namely those corresponding to a single eigenvalue of the configuration space. On the other hand, the AN-version of the Nk model, which has only modes up to $k+1$-spin interactions is solvable in polynomial time, while other variants containing all possible modes are $\mathcal{NP}$-complete (Weinberger 1991c).

Even a complete theory of elementary landscapes is only a part of the full story: What can be said about the superposition of elementary landscapes even if their structures are known in detail? How much of the theory outlined in this contribution carries over to configuration space with less symmetry than the sequence space such as the permutation spaces underlying the `TSP` or the `GMP`? Is a comparably theory possible for irregular configuration space such as spaces of finite trees? And finally, can we devise a comparable formalism for the combinatory maps of sequence-structure relations (Fontana *et al.* 1993a), where the image of a configuration is not a real-valued fitness but an element of an abstract metric space.

Acknowledgments

Stimulating discussions with P. Anderson, A.W.M. Dress, W. Fontana, R. Palmer, and G.P. Wagner are gratefully acknowledged. Christian Reidys from Jena was a great help in preparing the revised version of the manuscript. P.F.S. is grateful to the German *Max Planck Gesellschaft* for supporting his stay at the Santa Fe Institute in 1993.

References

- Aarts E.H.L. and J. Korst. *Simulated Annealing and Boltzman Machines.* J. Wiley & Sons, New York, 1990.
- Amitrano C., L. Peliti, and M. Saber. Population Dynamics in Spin Glass Model of Chemical Evolution. *J. Mol. Evol.* **29**, 513-525 (1989).
- Bak P., H. Flyvbjerg, and B. Lautrup. Coevolution in a rugged fitness landscape. *Phys. Rev. A[15]* **46**, 6724-6730 (1992).
- Bernasconi J. Low autocorrelation binary sequences: Statistical mechanics and configuration space analysis. *J. Physique*, **48**, 559-567 (1987).
- Besag J. Spatial Interactions and the Statistical Analysis of Lattice Systems. *Amer. Math. Monthly* **81** 192-236 (1974).
- Biggs N. *Algebraic Graph Theory.* [1st edition] Cambridge University Press, Cambridge, UK.
- Biggs N. *Algebraic Graph Theory.* [2nd edition] Cambridge University Press, Cambridge, UK.
- Biggs N.L. and A.T. White *Permutation Groups and Combinatorial Structures.* Cambridge Univ. Press, Cambridge, 1979.
- Binder K. and A.P. Young. Spin Glasses: Experimental Facts, Theoretical Concepts, and Open Questions. *Rev. Mod. Phys.* **58**, 801-976 (1986).
- Bollobás B. *Graph Theory – An Introductory Course.* Springer-Verlag, New York, 1979.
- Bonhoeffer S., J.S. McCaskill, P.F. Stadler and P. Schuster. RNA Multi-Structure Landscapes. A Study Based on Temperature Dependent Partition Functions. *Eur. Biophys. J.* **22**, 13-24 (1993).
- Brouwer A.E., A.M. Cohen, and A. Neumaier. *Distance-regular Graphs.* Springer Verlag, Berlin, New York, 1989.
- Brualdi R.A. and H.J. Ryser. *Combinatorial Matrix Theory.* Cambridge Univ. Press, Cambridge UK, 1991.

- Buckley F. and F. Harary. *Distance in Graphs.* Addison-Wesley, Reading, MA, 1990.
- Cameron P.J. and J.H. vanLint. *Designs, Graphs, Codes, and their Links.* London Math. Soc. Student Texts Vol. 22, Cambridge University Press, Cambridge UK, (1991).
- Cech T.R. Conserved sequences and structures of groupI introns: building an active site for RNA catalysis - a review. *Gene* **73**, 259-271 (1988).
- Chavel I. *Eigenvalues in Riemannian Geometry.* Academic Press, Orlando Fl., 1984.
- Cvetković D.M., M. Doob, and H. Sachs. *Spectra of Graphs – Theory and Applications*, Academic Press, New York, 1980.
- Damerell R.M. On Moore Graphs. *Proc. Cambridge Phil. Soc.* **74**, 227-236 (1973).
- Derrida B. Random Energy Model: Limit of a Family of Disordered Models. *Phys. Rev. Lett.* **45**, 79-82 (1980a).
- Derrida B. The Random Energy Model. *Phys. Rep.* **67**, 29-35 (1980).
- Derrida B. Random Energy Model: An exactly solvable model of disordered systems.This basis is unique if all eigenvalues of $\hat{A}$ are simple. *Phys. Rev. B* **24**, 2613-2626 (1981).
- Diaconis P., and M. Shahshahani. Generating a Random Permutation with Random Transpositions. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **57**, 159–179 (1981).
- Dress A.W.M., and D.S. Rumschitzki. Evolution on Sequence Space and Tensor Products of Representation Spaces. *Acta Appl. Math.* **11**, 103-115 (1988).
- Dunkl C.F. Orthogonal Functions on Some Permutation Groups. In: Ray-Chaudhuri D.K. (ed.) *Proceeding of Symposia in Pure Mathematic* **34**. American Mathematical Society, New York 1979.

- Dunkl C.F., A Krawtchouk polynomial addition theorem and wreath products of symmetric groups. *Indiana Univ. Math. J.* **25** 335-358 (1976).
- Eigen M. Selforganization of Matter and the Evolution of Biological Macromolecules. *Naturwissenschaften*, **10**, 465-523 (1971).
- Eigen M., J.S. McCaskill, and P. Schuster. The Molecular Quasispecies. *Adv. Phys. Chem.* **75**, 149–263 (1989).
- Flyvbjerg H. and B. Lautrup. Evolution in a rugged fitness landscape. *Phys. Rev. A[15]* **46**, 6714-6723 (1992).
- Fontana W. and P. Schuster. A Computer Model of Evolutionary Optimization. *Biophys. Chem.* **26**, 123-147 (1987).
- Fontana W., W. Schnabl, and P. Schuster. Physical aspects of evolutionary optimization and adaptation. *Phys. Rev. A* **40**, 3301-3321 (1989).
- Fontana W., T. Griesmacher, W. Schnabl, P.F. Stadler, and P. Schuster. Statistics of Landscapes Based on Free Energies, Replication and Degradation Rate Constants of RNA Secondary Structures. *Mh. Chem.* **122** 795-819 (1991).
- Fontana W, 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 Maps *Phys. Rev. E* **47**, 2083–2099 (1993a).
- Fontana W., D.A.M. Konings, P.F. Stadler, P. Schuster. Statistics of RNA Secondary Structures. *Biopolymers* **33**, 1389-1404 (1993b).
- Fu Y. and P.W. Anderson. Application of statistical mechanics to $\mathcal{NP}$- complete optimization problems in combinatorial optimization. *J. Phys. A* **19** 1605-1620 (1986).
- Gardiner C.W. *Handbook of Stochastic Methods.* (2nd Edition) Springer-Verlag, Berlin, 1990.
- Gardner E. and B. Derrida. The probability distribution of the partition function of the random energy model. *J. Phys. A* **22**, 1975-1982 (1989).
- Garey M.R. and D.S. Johnson. *Computers and Intractability. A Guide to the Theory of NP-Completeness.* Freeman, San Francisco, 1979.

- Golay M.J.E. Sieves for low-autocorrelation binary sequences. *IEEE Trans. Inform. Theory* **IT-23**, 43-51 (1977).
- Grover L.K. Local Search and the Local Structure of NP-Complete Problems. *Oper. Res. Lett.* **12**, 235-243 (1992).
- Hamming R.W. Error Detecting and Error Correcting Codes *Bell Syst. Tech. J.* **29** 147-160 (1950).
- Hofacker I.L., W. Fontana, P.F. Stadler, L.S. Bonhoeffer, M. Tacker, and P. Schuster. Fast Folding and Comparison of RNA Secondary Structures (The Vienna RNA Package). *Mh. Chem.* **125**, 167-188 (1994).
- Hotelling H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psych.* **24**, 417-441 and 498-520 (1933).
- Huynen M.A., and P. Hogeweg. Pattern Generation in Molecular Evolution. Exploitation of the Variation in RNA Landscapes. *J. Mol. Evol.* **39**, 71-79 (1994).
- R.E. Ingram. Some Characters of the Symmetric Group. *Proc. Amer. Math. Soc.* **1**, 358–369 (1950).
- Jacobson N. *Basic Algebra I.* (2nd edition), Freeman & Co., New York (1985).
- Kauffman S.A. and S. Levin. Towards a General Theory of Adaptive Walks on Rugged Landscapes. *J. Theor. Biol.* **128**, 11-45 (1987).
- Kauffman S.A. *The Origins of Order.* Oxford University Press, New York, Oxford, 1993.
- Kauffman S.A., E.D. Weinberger, and S.A. Perelson. Maturation of the Immune Response Via Adaptive Walks on Affinity Landscapes. In: *Theoretical Immunology, Part I*, ed.: Alan Perelson. Addison-Wesley, Redwood City, 1988.
- Koornwinder T.H., Krawtchouk Polynomials. A Unification of Two Different Group Theoretic Interpretations. *SIAM J. Math. Anal.* **13** 1011-1023 (1982).
- Krawtchouk M. Sur une gènèralisation des polynomes d'Hermite. *Comptes Rendus* **189**, 620-622 (1929).

- Lawler E.L., J.K. Lenstra A.H.G. Rinnoy Kan, and D.B. Shmoys. *The Traveling Salesman Problem. A Guided Tour of Combinatorial Optimization.* John Wiley & Sons, New York, 1985.
- Lin, S. and B.W. Kernighan. An Effective Heuristic Algorithm for the Traveling Salesman Problem. *Oper.Res.* **21**, 498-516 (1965).
- Lovász L. Spectra of graphs with transitive groups. *Periodica Math. Hung.* **6**, 191-195 (1975).
- Macken C.A., P.S. Hagan, and A.S. Perelson. Evolutionary Walks on Rugged Landscapes. *SIAM J. Appl. Math.* **51**, 799-827 (1991).
- MacDonald I.G. *Symmetric Functions and Hall Polynomials.* Oxford Univ. Press, Oxford UK, 1979.
- MacWilliams F.J., N.J.A. Sloane. *The Theory of Error-Correcting Codes.* North-Holland, Amsterdam, New York, Oxford, Tokyo 1991.
- McCaskill J.S. The Equilibrium Partition Function and Base Pair Binding Probabilities for RNA Secondary Structure. *Biopolymers* **29**, 1105-1119 (1990).
- Mézard M., G. Parisi, and M.A. Virasoro. *Spin Glass Theory and Beyond.* World Scientific, Singapore, 1987.
- Miller Jr., W. *Symmetry and Separation of Variable.* (Enceyclopedia of Mathematics and its Applications, Vol. 4) Cambridge Univ. Press, Cambridge, 1984.
- Miller D.L. and J.F. Pekny. Exact Solution of Large Asymmetric Traveling Salesman Problems. *Science* **251**, 754-761 (1991).
- Otten R.H.J.M. and L.P.P.P. vanGinneken. *The Annealing Algorith.* Kluwer Acad. Publ., Boston, 1989.
- Palmer R. Optimization on rugged landscapes. In: A.S. Perelson and S.A. Kauffman (eds.), *Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System*, Addison Wesley, Redwood City, CA, 1991, pp.3-25.
- Rao C.R. *Linear Statistical Interference and Its Applications.* 2nd ed., Wiley, New York (1973).

- Rumschitzky D. Spectral Properties of Eigen's evolution matrices. *J. Math. Biol.* **24**, 667-680 (1987).
- Schober A., M. Thürk, and M. Eigen. Optimization by hierarchical mutant production. *Biol. Cybern.* **69**, 493-502 (1993)
- Schuster P. Complex Optimization in an Artificial RNA World. In: D. Farmer, C. Langton S. Rasmussen and C. Taylor: *Artificial Life II*, SFI Studies in the Science of Complexity XII Addison-Wesley, Reading, Mass. 1991.
- Schuster P., Fontana W., Stadler P.F., and Hofacker I.L. From Sequences to Shapes and Back: A Case Study in RNA Secondary Structures. *Proc. Roy. Soc. B* **255**, 279-284 (1994).
- Serre, J.-P. *Linear Representations of Finite Groups.* Springer-Verlag, New York, Berlin, Heidelberg, 1977.
- Sherrington D., and S. Kirkpatrick. Solvable Model of a Spin Glass. *Phys. Rev. Lett.* **35**, 402-416 (1975).
- Soardi P.M. *Potential Theory on Infinite Networks.* Springer-Verlag, Berlin (1994).
- Spitzer F. *Principles of Random Walks.* Springer-Verlag, New York. (1976)
- Stadler P.F. and W. Schnabl. The Landscape of the Traveling Salesman Problem. *Phys. Lett. A* **161**, 337-344 (1992).
- Stadler P.F. and R. Happel. Correlation Structure of the Graph Bipartitioning Problem. *J. Phys. A: Math. Gen.* **25**, 3103-3110 (1992).
- Stadler P.F. Correlation in Landscapes of Combinatorial Optimization Problems *Europhys. Lett.* **20** 479-482 (1992).
- Stadler P.F. and W. Grüner. Anisotropy in Fitness Landscapes. *J. Theor. Biol.* **165**, 373-388 (1993).
- Stadler P.F. Random Walks and Orthogonal Functions Associated with Highly Symmetric Graphs. *Disc. Math.* in press 1994a, and *SFI preprint 93-06-31*
- Stadler P.F. Linear Operators on Correlated Landscapes. *J. Physique I (Paris)* **4**, 681-696 (1994b).

- Szegö G. *Orthogonal Polynomials.* Colloquium Publications Vol. 23. American Mathematical Society, New York; revised edition 1959.
- Tacker M., W. Fontana, P.F. Stadler, and P. Schuster. Statistics of RNA Melting Kinetics. *Eur. Biophys. J.* **23**, 29-38 (1994).
- vanLint J.H. *Introduction to Coding Theory.* Springer-Verlag, New York, 1982.
- Warren R.H. Special Cases of the Travelling Salesman Problem *Appl. Math. & Comp.* **60**, 171-177 (1994).
- Weinberger E.D. Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference. *Biol. Cybern.* **63**, 325-336 (1990).
- Weinberger E.D. Local Properties of the *Nk*-Model, a Tunably Rugged Landscape. *Phys. Rev. A* **44**, 6399-6813 (1991a).
- Weinberger E.D. Fourier and Taylor Series on Landscapes. *Biol. Cybern.* **65**, 321-330 (1991b).
- Weinberger E.D. NP Completeness of Kauffman's N-k Model, a Tunably Rugged Fitness Landscape.

  Preprint, MPI Biophysical Chemistry, Göttingen, Germany (1991c).
- Weinberger E.D. and P.F. Stadler. Why Some Fitness Landscapes are Fractal. *J. Theor. Biol.* **163** 255-276 (1993).
- Wright S.A. The role of mutation, inbreeding, cross-breeding, and selection in evolution. *Proc. 6th Int. Congress on Genetics*, Vol. I, pp. 356-366 (1932).
- Zuker M. and D. Sankoff. RNA secondary structures and their prediction. *Bull. Math. Biol.* **46**, 591-621 (1984).

Appendix A: Combinatorial Optimization Problems

Graph Bipartitioning (GBP). G is a graph with an even number n of vertices and H is a symmetric matrix of edge weights. A configuration is a partition of the vertex set into two subsets A and B of equal size. The cost function is

$$f([A;B]) = \sum_{i \in A} \sum_{j \in B} H_{ij},$$

i.e., the total weight of edges connecting the two subsets. The usual move set for this problem is exchanging a vertex in A with a vertex in B, see e.g., Fu and Anderson (1986)

Travelling Salesman (TSP). An instance of a TSP (Lawler *et al.* 1985) is defined by a set of n cities and matrix D of costs for connecting them. A tour is a permutation π of cities, and its cost is

$$f(\pi) = \sum_{i=1}^{n} D_{\pi(i),\pi(i-1)}$$

where the indices are taken modulo n. Different versions of the TSP are defined by the properties of D (arbitrary, or symmetric, or with entries additionally obeying the triangle inequality, etc.; see Warren, 1994). A variety of move sets is conceivable on the set of all permutations. The most common ones are the sets of all transpositions (i, j), the set of all canonical transpositions $(i, i+1)$, and the set of all reversals (also called inversions or 2opt-moves, Lin and Kernighan 1965).

Not-All-Equal-Satisfiability (NAES). Consider a vector of n binary variable. A literal is a variable or its complement. A clause is a set of three literals that does not contain both a variable and its complement. A clause is said to be satisfied if at least on literal is 0 and at least one literal is 1. An instance of NAES is given by a set of c clauses, and the cost function is the number of non-satisfied clauses. The

move set is defined by flipping the value of a single variable, thus the configuration space is the Boolean Hypercube.

Graph Coloring (GC). An instance of a graph coloring problem consists of a graph $G(V, E)$ and a number α of colors. A configuration x is a α-coloring of the vertex set V, i.e., an assignment of one color $x(p)$ to each vertex p of the graph. The cost function is the number edges $(p, q) \in E$ such that $x(p) = x(q)$:

$$f(x) = \sum_{(p,q)\in E} \delta_{x(p)x(q)}.$$

A move is the replacement of one color by another one at single vertex. The configuration spaces are thus the general Hamming graphs, i.e., sequence spaces over the alphabet of the α colors.

Weight Partition (WP). Given a string of n "spins" $x = (x_i) \in \{-1, +1\}^n$ and corresponding weights w_i, the cost function is given by

$$f(x) = \left(\sum_{i=1}^{n} w_i x_i \right)^2 .$$

The move set is given by flipping a single spin, hence the configuration space is again a hypercube.

Graph Matching Problem (GMP). Given a graph X with n vertices and a symmetric matrix W of edge weights, the task is to partition the graph into $n/2$ pairs of vertices such that the sum of the edge weights corresponding to these pairs is optimal. A convenient encoding of the problem is the following. Let $\pi \in S_n$ be a permutation of the vertices. We assume that the vertices are arranged such that $[\pi(2k-1), \pi(2k)]$ form a pair. The cost function is hence

$$f(\pi) = \sum_{k=1}^{n/2} W_{\pi(2k-1),\pi(2k)}.$$

Again, the configuration space is the symmetric group, and hence the set of all transpositions forms a canonical move set.

Appendix B: Notation

$\langle\,.\,\rangle$	average over all points in a landscape
$\langle\,.\,,\,.\,\rangle$	standart scalar product in an Euclidean vector space
$\langle\,.\,;\,.\,\rangle$	scalar product of correlation functions defined in equ.(9).
$[\,.\,]$	cycle decomposition of a permutation
$\vert\,.\,\vert$	number of vertices in a symmetry class; absolute value of a number
$\Vert\,.\,\Vert$	norm of a vector
0	reference vertex
1	symmetry class consisting of all neighbors of 0
α	number of letters in an alphabet
a_x	Fourier coefficients
A	adjacency matrix of a graph
$\hat{A}$	collapsed adjacency matrix
$\mathcal{A}$	adjacency algebra of a graph
$\mathbf{A}_0$	group of all automorphisms fixing the reference vertex
Aut	automorphism group
C	covariance matrix
Δ	graph Laplacian
d	distance on a graph
$d(\,.\,,\,.\,)$	a metric
D	vertex degree of a regular graph Γ
e	identity of a group
$e(\,.\,)$	an eigenvector of A
$e_x(\,.\,)$	an eigenvector of a Cgcg
E	identity matrix
$E\Gamma$	edge set of Γ
$\mathcal{E}[\,.\,]$	expected value on a random field

$\varepsilon_{i_1 \ldots i_p}$	eigenvector of the Laplacian of the Boolean hypercube in terms of spin variables
f	a landscape
Γ	a graph
$\Gamma(G, \Phi)$	a Cayley graph
G	a group
$\mathcal{H}_p$	Hamiltonian of a p-spin model
ι	identity permutation
$J_{i_1 \ldots i_p}$	interaction coefficients of a p-spin model
κ	a symmetry class of vertices, i.e., an orbit of $\mathbf{A}_0$
$\ddot{\kappa}$	an orbit of Aut[Γ] on the set of all vertex pairs of Γ
K	Parameter of an elementary landscape
$\mathbf{K}_p^{(n,\alpha)}$	Krawtchouk polynomial over $\mathcal{Q}_\alpha^n$ of order p.
λ	eigenvalue of $\frac{1}{D}A$
Λ	eigenvalue of A
$\tilde{\Lambda}$	eigenvalue of Δ
μ	a symmetry class of vertices, i.e., an orbit of $\mathbf{A}_0$
$\ddot{\mu}$	an orbit of Aut[Γ] on the set of all vertex pairs of Γ
$\boldsymbol{\mu}$	measure of a random field
n	system size, chain lenght of string
N	number of vertices of a the configuration space
$N(x)$	number of neighbors of a vertex x
$\mathcal{N}_k$	random field of a Nk model
∇	gradient operator on graphs
∇^+	incidence matrix of a graph
ω_i	special functions defined in equ.(40)
$\Omega_p, \Omega(p)$	coefficient in the expansion of the autocorrelation function, Fourier spectrum

$\phi(\,.\,)$	eigenvectors of the covariance matrix
$\varphi_{s\mu}$	Probability for a simple random walk with s steps to end in distance class μ
Φ	a set of generators of a group
p	order of interaction
π	a permutation
$\mathcal{Q}$	a Hamming graph
$\rho(\,.\,)$	autocorrelation function of random field in terms of symmetry classes
$\hat{\rho}(\,.\,)$	(empirical) autocorrelation function of landscape
$\rho_p(\,.\,)$	left eigenvector of $\hat{A}$
$r(s)$	autocorrelation function along a simple random walk on a random field
$\hat{r}(s)$	(empirical) autocorrelation function along a random walk on a landscape
$R^{(\mu)}$	matrix associated with the orbit $\ddot{\mu}$
$\mathbb{R}$	real numbers
S_n	symmetric group over n elements
s	lag
σ^2	variance over all configurations in a landscape
$\theta(\,.\,)$	eigenvectors of a graph Laplacian
T	Transition matrix of a simple random walk $T = (1/D)A$
$\mathcal{T}$	set of all transpositions
$\upsilon_p(\,.\,)$	right eigenvector of $\hat{A}$
V	vertex set of a graph Γ
$\mathrm{Var}[\,.\,]$	variance on a random field
W	distance matrix in a TSP or matching problem
x	vertex of a graph, group element
Ξ	a random field
y	vertex of a graph, group element
z	vertex of a graph, group element

Table of Contents

1. Introduction 2

2. Technical Preliminaries 5

2.1. Landscapes and Random Fields 5

2.2. Vertex Transitive Graphs 7

2.3. Cayley Graphs 9

2.4. Distance Transitive Graphs 10

2.5. The Collapsed Adjacency Matrix 11

2.6. Adjacency Algebra of a Graph 12

2.7. Spectra of Graphs 13

2.8. Incidence Matrix and Graph Laplacian 15

2.9. Fourier Expansion of Landscapes 19

2.10. Examples of Configuration Spaces 20

2.10.1. Sequence Spaces 20

2.10.2. Generalized Hamming Graphs 23

2.10.3. The Petersen Graph 25

2.10.4. Cayley Graphs of the Symmetric Group 25

2.11. Karhunen-Loève Decomposition of Random Fields 30

3. Isotropic Random Fields 32

3.1. Characterization of Isotropic Random Fields 32

3.2. Landscapes on the Hypercube and the p-Spin Models 35

3.3. Autocorrelation Functions of Isotropic Random Fields 36

4. Empirical Landscapes 38

4.1. Empirical Correlation Functions 38

4.2. Simple Random Walks on Landscapes 39

4.3. General Vertex Transitive Graphs 42

4.4. Elementary Landscapes 43

4.5. Classification of Landscapes 47

5. Model Random Fields 50

5.1. The Random Energy Model 50

5.2. Nk Models 51

5.3. Superposition of Nk Models 55

6. Model Landscapes 58

6.1. Traveling Salesman Problems 58

6.2. Graph Matching Problem 61

6.3. Graph Bipartitioning Problem 62

6.4. The Low Autocorrelated Binary String Problem 63

6.5. RNA Free Energy Landscapes 65

7. Discussion 68

References 72

Appendix A: Combinatorial Optimization Problems 79

Appendix B: Notation 81