Approximate Scaling Properties of RNA Free Energy Landscapes

By

SUBBIAH BASKARAN[a,b,c], PETER F. STADLER[b,d], AND PETER SCHUSTER[b,c,d,*]

[a]ES76, Marshall Space Flight Center, Huntsville, AL35812, U.S.A.
[b]Institut für Theoretische Chemie, Universität Wien, Vienna, Austria
[c]Institut für Molekulare Biotechnologie e.V., Jena, Germany
[d]Santa Fe Institute, Santa Fe, NM 87501, U.S.A.

*Correspondence to:
Institut für Theoretische Chemie, Universität Wien
Währingerstraße 17, A-1090 Vienna, Austria
Phone: **43 1 40 480 / 669
Fax: **43 1 40 480 / 660
Email: pks@tbi.univie.ac.at

Abstract

RNA free energy landscapes are analyzed by means of "time-series" that are obtained from random walks restricted to excursion sets. The power spectra, the scaling of the jump size distribution, and the scaling of the curve length measured with different yard stick lengths are used to describe the structure of these "time-series". Although they are stationary by construction, we find that their *local* behavior is consistent with both AR(1) and self-affine processes. Random walks confined to excursion sets (i.e., with the restriction that the fitness value exceeds a certain threshold at each step) exhibit essentially the same statistics as free random walks.
We find that an AR(1) time series is in general approximately self-affine on time scales up to approximately the correlation length. We present an empirical relation between the correlation parameter ρ of the AR(1) model and the exponents characterizing self-affinity.

Key Words

RNA Folding — Excursion Sets — Fractal Landscape — AR(1) Process — $1/f$ noise

1. Introduction

Evolutionary optimization as well as combinatorial optimization take place on *landscapes* resulting from mapping (micro)configurations to scalar quantities like fitness values, energies, or costs (Schuster & Stadler, 1994). In most cases one lacks a detailed understanding of the structure of fitness landscapes that underlies a particular instance of biological evolution. One resorts thus to using model landscapes. Well studied examples include combinatorial optimization problems such as the Traveling Salesman Problem, the Graph Matching Problem, or the Graph Bipartitioning Problem, various spin glass models, among them the Sherrington-Kirkpatrick models, and Kauffman's Nk-models (Kauffman, 1993), The ruggedness of a landscape, often measured by means of a correlation function, is of crucial importance for the dynamics of the evolution process (Eigen *et al.*, 1989; Bonhoeffer & Stadler, 1993). Detailed studies of the correlation structure of model landscapes can be found, for instance, in the following references (Stadler & Schnabl, 1992; Stadler & Happel, 1992; Stadler, 1992; Weinberger, 1991a; Weinberger, 1991b; Weinberger & Stadler, 1993).

Exclusively in the case of RNA landscapes do we have a sound biophysical model for the fitness function. Models based on RNA secondary structure prediction algorithms have been analyzed in great details in a series of papers (Fontana *et al.*, 1991; Fontana *et al.*, 1993; Bonhoeffer *et al.*, 1993; Tacker *et al.*, 1994; Schuster *et al.*, 1994). Evolutionary dynamics on such landscapes was the topic of extensive research as well (Fontana & Schuster, 1987; Fontana *et al.*, 1989; Huynen *et al.*, 1995). A detailed understanding of these landscapes is a necessary prerequisite for building simpler models based on spin glass or Nk model landscapes that are significantly less costly in computer simulations and that lend themselves much easier to analytical treatment.

Weinberger (Weinberger, 1990) suggested to characterize a landscape by means of a "time series" obtained by sampling the fitness values along a random walk in sequence space. While this method is rather indirect, it yields a data set that can be analyzed by the standard methods of time series analysis (Hordijk, 1995). In this contribution we shall investigate the "fractal-like" features of landscapes in terms of the approximate self-affinity of these "time-series".

A great variety of systems, physical and biological, exhibit $1/\omega$ power spectra, commonly called 1/f-noise or "flicker" noise. Some examples are resistivity fluctuation in conducting materials (Weissman, 1988), luminosity fluctuations of stars and galaxies (Nolan *et al.*, 1981), flow fluctuations of highway traffic (Musha & Higuchi, 1976) and of deep ocean waters (Taft *et al.*, 1974), frequency variations of quartz oscillator (Attkinson *et al.*, 1963), the loudness fluctuations in music and speech (Voss & Clarke, 1975). In biological systems $1/f$ noise has been reported for nerve membranes (Verveen & Derkson, 1968), for the DNA sequences of the non-coding introns (Voss, 1992; Li & Kaneko, 1992) as well as of coding regions (Buldyrev *et al.*, 1995). In this paper we will show that the "time-series" sampled along a random walk on a RNA free energy landscapes also leads to $1/f$ noise.

This contribution is organized as follows: In section 2 we review some notions that are basic to the theory of fitness landscapes. In particular, we introduce a variety of correlation measures and highlight their relations with each other. In particular we consider the class of landscapes that lead to exponential correlation functions of the "time series" obtained from simple random walks. In section 3 we briefly consider self-affine time-series and show that AR(1) processes mimic self-affinity on time scales up to their correlation length. These finding are applied to free energy landscapes of RNA in section 4. In particular, we shall see that the mountainous parts of the landscapes do not differ significantly from the average fitness regime, at least as long as the excursion sets do not fragment into tiny pieces. Section 5 concludes our discussion. The relaxation time of a simple random walk on a sequence space is computed in the appendix.

2. Landscapes

2.1. Rugged Landscapes

Definition. A landscape is a map $f : \mathcal{C} \to \mathbb{R}$, where $\mathcal{C} = (X, d)$ is a finite metric space with metric $d : X \times X \to \mathbb{R}$.

In most applications of landscapes in biology, physics, or combinatorial optimization the *configuration space* (X, d) can be represented as a graph Γ. Then two configurations x and y are neighbors in Γ if $d(x, y) = 1$. The metric d is often obtained from an *editing procedure* that allows to interconvert two configurations $x, y \in X$ by means of a finite sequence of operations. $d(x, y)$ is commonly defined as the number of operations in the shortest sequence that changes x into y or vice versa. In a biological context the "elementary operations" are in general mutations. We will restrict ourselves there to the case where X is a set of sequences of common length n which are constructed from some alphabet with α letters. In this case d is the so-called Hamming distance (Hamming, 1950), and the graph Γ is known a the *sequence space* $\mathcal{Q}^n_\alpha$, or Boolean hypercube in the special case $\alpha = 2$. For a recent review see (Schuster & Stadler, 1994; Stadler, 1995b).

2.2. Correlation Functions

A very important characteristic of a landscape is its *ruggedness*. Rugged landscapes are characterized by a large number of local optima (Palmer, 1991), the fact that uphill walks are short and easily trapped in local optima, and by short correlation lengths (Kauffman, 1993). There is ample evidence that heuristic optimization procedures work less efficiently the more rugged a landscape is (Stadler

& Schnabl, 1992; Schuster & Stadler, 1994). It will be convenient to define for a given landscape f:

$$\overline{f} = \frac{1}{|X|}\sum_{x\in X} f(x) \qquad \sigma_f^2 = \frac{1}{|X|}\sum_{x\in X}\left(f(x)-\overline{f}\right)^2 . \tag{2.1}$$

It has been suggested by various authors (Eigen *et al.*, 1989; Fontana *et al.*, 1991; Sorkin, 1988; Weinberger, 1990) to measure "ruggedness" by some sort of correlation measure. We shall use the following definition, which was first proposed in ref. (Eigen *et al.*, 1989):

$$\rho(d) = \frac{1}{\sigma_f^2}\cdot\frac{1}{|\mathfrak{D}_d|}\sum_{(x,y)\in\mathfrak{D}_d}\left(f(x)-\overline{f}\right)\left(f(y)-\overline{f}\right) \tag{2.2}$$

Here $\mathfrak{D}_d$ denotes the set of all pair of vertices that have mutual distance d in the graph Γ. For a sequence space we have for instance

$$|\mathfrak{D}_d| = \alpha^n(\alpha-1)^d\binom{n}{d} . \tag{2.3}$$

This definition is useful if Γ is a distance regular graph (Brouwer *et al.*, 1989). A more general mathematical framework is developed in (Stadler, 1995c; Happel & Stadler, 1995; Stadler & Happel, 1995).

Weinberger (Weinberger, 1990; Weinberger, 1991a; Weinberger, 1991b) suggested to investigate the properties of landscapes by sampling the values along a simple random walk in the configuration space $\mathcal{C}$:

$$\begin{array}{ccccccccccccc} x_0 & \rightarrow & x_1 & \rightarrow & x_2 & \rightarrow & \dots & \rightarrow & x_k & \rightarrow & \dots \\ | & & | & & | & & & & | & & \\ f_0 & \rightarrow & f_1 & \rightarrow & f_2 & \rightarrow & \dots & \rightarrow & f_k & \rightarrow & \dots \end{array} \tag{2.4}$$

where x_i and x_{i+1} are neighbors in $\mathcal{C}$. At each step one of the neighbors of x_i in Γ is chosen with uniform probability, i.e., the series $\{x_i\}$ is a simple random walk on $\mathcal{C}$ (Spitzer, 1971). By evaluating the configurations along the walk $\{x_i\}$ we obtain a random walk on the landscape, i.e., the "time series" $\{f_i = f(x_i)\}$. This series is stationary by construction.

The autocorrelation function of a stationary time series is defined by

$$r(s) = \frac{\langle f_t f_{t+s} \rangle - \langle f_t^2 \rangle}{\sigma^2} \tag{2.5}$$

where σ^2 is the variance, which coincides with σ_f^2 defined above, and the angular brackets indicate the expectation value taken over all random walks $\{x_i\}$ and all times t. Provided the graph Γ is D-regular, i.e., each configuration x has exactly D neighbors, we may write the transition matrix of the random walks as $\mathbf{T} = (1/D)\mathbf{A}$. The entry $\mathbf{A}_{xy}$ of the *adjacency matrix* is 1 or 0, depending on whether the configurations x and y are neighbors or not. It is shown in (Stadler, 1995c) that the correlation function $r(s)$ has the following algebraic representation:

$$r(s) = \frac{1}{\sigma^2} \left[\langle f, \mathbf{T}^s f \rangle - \overline{f}^2 \right] . \tag{2.6}$$

Note that $\langle\, . \, , . \,\rangle$ denotes here a scalar product, not an expectation value! Another useful (Fontana *et al.*, 1993) representation is

$$r(s) = 1 - \frac{\langle\, (f_{t+s} - f_t)^2 \,\rangle}{\sigma^2} ; \tag{2.7}$$

The average squared difference $\langle\, (f_{t+s} - f_t)^2 \,\rangle$ was used as a correlation measure in Sorkin's pioneering paper (Sorkin, 1988).

The autocorrelation function $\rho(d)$ of the landscape itself and the autocorrelation $r(s)$ of the "time-series" of the landscape are related via

$$r(s) = \sum_d \varphi_{sd} \rho(d) \tag{2.4}$$

where φ_{sd} is the probability that a simple random walk of length s ends at distance $d(x_0, x_s) = d$. Explicit expressions for φ_{sd} can be found in ref. (Fontana *et al.*, 1991; Stadler, 1995a); we shall not make use of them in this contribution.

2.3. Elementary Landscapes

For a quite large number of model landscapes it has been found that the correlation function $r(s)$ is exactly a decaying exponential (Stadler & Happel, 1992; Weinberger & Stadler, 1993), numerically indistinguishable from a decaying exponential (Stadler & Schnabl, 1992; Stadler, 1992), or at least very close to a decaying exponential (Weinberger, 1990; Weinberger, 1991a). It has been argued that a nearly exponential autocorrelation function $r(s)$ would be generic for landscapes with a Gaussian distribution of fitness values (Weinberger, 1990). This argument is wrong, however.

It is not hard to check that $r(s)$ is exponential whenever f is of the form $f(x) = \overline{f} + \varphi(x)$, where φ is an *eigenvector* of the adjacency matrix $\mathbf{A}$ with eigenvalue Λ. Indeed, under these conditions one finds $r(s) = (\Lambda/D)^s$. In a more general context is useful to assume that φ is an eigenvector of the so-called graph Laplacian (Mohar, 1991); for regular graphs we have $\Delta = \mathbf{A} - D\mathbf{E}$, where $\mathbf{E}$ is the identity matrix, i.e., the eigenvectors of $\mathbf{A}$ are the same as the eigenvectors of the Laplacian Δ. Landscapes of this type have been termed *elementary.* Lov Grover (Grover, 1992) found that a number of well known model landscapes are elementary, for instance the landscape of the Traveling Salesman Problem. In (Stadler, 1995c) it is also shown that $r(s)$ is exponential if *and only if* the landscape is elementary. Note that the possible eigenvalues Λ are uniquely determined by the adjacency matrix $\mathbf{A}$, i.e., by the geometry of the configuration space. As a consequence there is only a finite small number of possible values for the parameter $\varrho \stackrel{\text{def}}{=} \Lambda/D$ of the exponential decay, i.e., it is *not* possible to construct a landscape f with autocorrelation function $r(s) = \varrho^s$ with an arbitrarily prescribed parameter ϱ — in contrast to the case of merely constructing a time series. In other worlds, only a very special set of time series is generated by random walks on landscapes.

2.4. Power Spectra

Instead of a correlation function one can use power spectrum

$$S(\omega) \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{2\pi N} \left[\left(\sum_{t=1}^{N} f_t \cos(\omega t) \right)^2 + \left(\sum_{t=1}^{N} f_t \sin(\omega t) \right)^2 \right], \tag{2.9}$$

of the time series $\{f_t\}$ as a means of characterizing the landscape. Here N is the number of points sampled from the time series $\{f_t\}$. Power spectrum and autocorrelation function of a stationary process are related by the Wiener-Khinchin theorem (see, e.g., (Yaglom, 1986)):

$$\begin{aligned} r(s) &= \frac{2}{\sigma^2} \int_0^{\pi} S(\omega) \cos(\omega s) d\omega \\ S(\omega) &= \frac{\sigma^2}{2\pi} \left(1 + 2 \sum_{s=1}^{\infty} r(s) \cos(\omega s) \right) \end{aligned} \tag{2.10}$$

A negative slope of $S(\omega)$ implies some degree of correlation in f_t. A steeper slope implies a higher degree of correlation. A signal $\{f_t\}$ is called $1/f$ noise if a log-log plot of the power spectrum versus frequency can be approximated by straight line with slope close to -1 *in the frequency range of interest.* More generally, one speaks of $1/f^a$ noise if the slope is $-a$. We shall return to this type of time series in section 3.

The most common definition of a correlation length in physics is simply the integral of the autocorrelation function. In the discrete case it is convenient to use

$$\hat{\ell} \stackrel{\text{def}}{=} \frac{1}{2} + \sum_{s=1}^{\infty} r(s)\,. \tag{2.11}$$

Comparing this definition with the Wiener-Khinchin theorem, equ.(2.10), yields the simple relation

$$S(0) = \frac{\sigma^2}{\pi}\, \hat{\ell} \tag{2.12}$$

which can be used as an alternative way of estimating the correlation length of a time series.

2.5. Excursion Sets

The parts of the landscape in which the values are close to the global maximum or minimum are particular interest. One might ask, for instance, how the "good" solutions are distributed in sequence space? Are they clustered around a globally optimal solution, or are configurations with close-to-optimal values scattered all over the configuration space? A suitable mathematical framework for this type of questions is set by the notion of *excursion sets* (Adler, 1981). In this subsection we collect a few definitions and their immediate corrolaries which will be useful for the discussion of the RNA free energy landscapes in section 4.

Definition. Let $f : X \to \mathbb{R}$ be an arbitrary landscape.

(i) A configuration x is a *local optimum* if for all neighbors y of x holds $f(x) \geq f(y)$. Two configurations x and y are called *neutral* if $f(x) = f(y)$.

(ii) The set $A_E = \{x \in X | f(x) \geq E\}$ is called the *excursion set* of f at level E. A connected component of A_E is called a *cycle* (Freidlin & Wentzell, 1984).

(iii) A connected subgraph $B \subset A$ is called *neutral network in X* if all elements are neutral, and if all neutral neighbors of any $x \in B$ are elements of B as well.

For sufficiently small E we have of course $A_E = X$, the entire configuration space. On the other hand, if E is larger than the global optimum of f, then A_E is empty. Clearly, $E > E'$ implies $A_E \subseteq A_{E'}$, hence excursion sets introduce a hierarchical structure on the landscape. In general, A_E will not be connected, i.e., it will decompose into more than one cycle (connected component). Bounds on the number of cycles can be obtained for elementary landscapes and the special value $E = \overline{f}$, for details see (Stadler, 1995c). Cycles play a prominent role in the analysis of simulated annealing techniques on combinatory landscapes, see (Azencott, 1992) for a recent review.

Excursion sets, local optima, and neutral networks are closely related. We list here only a few simple geometric relationships: (i) Suppose C_E is a cycle and B is a neutral network, then C_E and B are either disjoint or B is subset of C_E. (ii)

Each cycle C_E contains at least one local optimum. (iii) A neutral network B is a cycle if and only if it consists entirely of local optima. Each cycle C_E contains a cycle of this type. (iv) A neutral network which is a cycle contains no other cycles except for itself. (v) If a cycle consists of only one configuration then this configuration is a local optimum.

The notion of excursion sets suggests two percolation problems: (i) At which level E does A_E cease to be a single cycle? (ii) At which level E does A_E decomposes into many small cycles, as opposed to consisting of a single giant component containing almost all vertices of A_E and a number of very small islands? Both problems have not been treated so far, although they seem to be of utmost importance for the understanding of adaptation on combinatory landscapes. In this contribution we shall be content with investigating the structure of landscape at fitness levels for which the cycles are still large in general.

2.6. Random Walks on Excursion Sets

Instead of performing the random walk on the entire configuration space $\mathcal{C}$ one may confine it to an excursion set $A_E \subset \mathcal{C}$. The random walks is then automatically constrained to a connected component of A_E, i.e., to a cycle. We used the following procedure to generate a walk within a cycle C_E. The process starts in a vertex a_0 known to be in the desired excursion set. These initial points are generated by screening a large number of random configurations. (Alternatively one might use configurations obtained from some simple optimization heuristics as starting points for higher excursion levels. This would, however, bias the the sampling, since configurations in large "mountains" would be favored.) Then an attempt is made to move to a neighboring vertex. If it is contained in the same cycle, i.e., if its fitness is above the threshold level E then it is accepted, otherwise the attempt is rejected. The "time-series" is formed by the accepted moves only. This procedure generates a time series provided C_E contains more than one configuration. In fact, we are only interested in large cycles.

It is clear that confining random walks to cycles means that they sample predominantly in the vicinity of local optima. One can hope, therefore, that the resulting time-series provide information about the most interesting regions of the fitness landscape — the region of high fitness. The major drawback is that, by equ.(2.8), the time series contains a superposition of two effects, namely the correlation of fitness values on the landscape and the geometrical relaxation of the walk in C_E. The correlation of a walk in C_E is

$$r^E(s) = \sum_d \varphi^E_{sd} \rho^E(d) \approx \rho^E(\langle d(s) \rangle_E), \quad \text{with } \langle d(s) \rangle_E = \sum_d \varphi^E_{sd} d. \tag{2.13}$$

Here $\rho^E(d)$ is the correlation of the restriction of the landscape to the excursion set A_E and $\langle d(s) \rangle_E$ describes the geometric relaxation of a random walk in a cycle C_E. Since the topology of C_E is not known it is very difficult to retrieve more than qualitative information on the structure of the mountainous parts of the landscape.

3. Self-Affine Time Series and Fractal Landscapes

3.1. Self-Affine Time Series

Definition. A time series $\{F_t\}$ is *self-affine* (or *fractal*) if

$$s^{-H}(F_{t+s} - F_t) \stackrel{\mathrm{d}}{=} (F_{t+1} - F_t), \tag{3.1}$$

where s is the number of steps between the two measurements. The notation $\stackrel{\mathrm{d}}{=}$ indicates equality in the sense of distributions. The parameter H fulfils $0 \leq H \leq 1$.

An example is fractional Brownian motion, see, e.g., (Mandelbrot, 1982). The power spectrum of a time series with a distribution fulfilling (3.1) follows a power law (Mandelbrot & vanNess, 1968) of the form

$$S(\omega) = \omega^{-a} \quad \text{with} \quad a = 1 + 2H \tag{3.2}$$

In case of fractional Brownian motion in continuous time, the parameter H and the Haussdorf dimension D_H of the resulting curve are related by $D_H = H + 1$. We remark that a time series fulfilling (3.1) strictly for all s cannot be stationary.

Instead of using the power spectrum one can use more direct methods for characterizing a self-affine time series. Probably the most immediate approach is to consider the *jump size*

$$J(s) = \langle |F_{t+s} - F_t| \rangle \tag{3.3}$$

As an immediate consequence of (3.1) we have $J(s) \sim s^H$. H can be obtained by means of a least square fit from a log-log plot, see, e.g., (Osborne & Provenzale, 1989). A closely related technique has been proposed by Sorkin (Sorkin, 1988). Multiplying (3.1) by itself and taking the expectation yields

$$s^{-2H} \langle (F_{t+s} - F_t)^2 \rangle = \langle (F_{t+1} - F_t)^2 \rangle \tag{3.4}$$

and one obtains the slope $2H$ from a log-log plot of the mean square differences versus the lag s.

Another approach to self-similarity focuses on the curve length as function of the yard-stick length used for the measurement. The method outlined below was proposed in ref. (Higuchi, 1988) as an improvement of the procedure given by Burlaga and Klein (Burlaga & Klein, 1986). It provides numerically stable scaling exponents even for a small number of data points. We divide the time series $\{F_t\}$ into k partial series

$$\mathcal{F}^m(s) = \{F_m, F_{m+s}, F_{m+2s}, \dots, F_{m+\lfloor\frac{N-m}{s}\rfloor s}\},$$

and define the length of $\mathcal{F}^m(s)$ as

$$L_m(s) = \frac{N-1}{s\lfloor\frac{N-m}{s}\rfloor} \sum_i |F_{m+is} - F_{m-(i-1)s}|,$$

The curve length $L(s)$ measured with step size s is then the average value taken over all the partial series:

$$L(s) = \frac{1}{s} \sum_{m=1}^{s} L_m(s). \tag{3.5}$$

If the time-series is self-affine, then the curve length follows a power law of the form $L(s) \sim s^{-D}$. The correction factor in the definition of $L_m(s)$ approaches 1 for large data sets, and hence we find as an immediate consequence of equ.(3.1) that

$$L(s) \to \frac{N}{s} s^H \langle\, |F_{t+1} - F_t|\,\rangle \;\sim\; s^{H-1}, \tag{3.6}$$

and therefore $D = 1 - H$. The parameters a, H, and D of a self-affine time series are related by means of the equations

$$a = 1 + 2H = 3 - 2D. \tag{3.7}$$

Independent estimates of a, H, and D can thus be used to determine to what extent a given time series is consistent with the assumption of self-affinity.

3.2. Fractal Landscapes

It is obvious that a time series obtained from a random walk on a landscape cannot be strictly self-affine since it must be (at least approximately) stationary. Hence (3.1) is an approximation that holds only for $s \ll n$, where n is the maximal distance in configuration space.

Dividing Sorkin's equ.(3.4) by twice the (finite) variance σ^2 of the landscape and substituting equ.(2.7) we find

$$s^{-2H}(1 - r(s)) = 1 - r(1). \tag{3.8}$$

Solving for the autocorrelation function yields $r(s) = 1 - cs^{2H}$. The parameter c can be obtained as follows: Since a single step along the random walk always leads to distance 1 we have $r(1) = \rho(1) \overset{\text{def}}{=\!=} \varrho$, the nearest neighbor correlation of the landscape. Thus $\varrho = 1 - c$, and we finally obtain an autocorrelation function of form

$$r(s) = 1 - (1 - \varrho)s^{2H}\,. \tag{3.9}$$

Equ.(3.9) holds of course only for s small compared to the maximum distance in the landscape. It has been used for a classification of rugged landscapes (Weinberger & Stadler, 1993; Stadler, 1995b) in terms of the parameter

$$H = \frac{1}{2}\frac{\partial \ln(1 - r(s))}{\partial \ln s} \tag{3.10}$$

for small s.

3.3. "AR(1)-Landscapes" are Locally Fractal

An AR(1) (or Ornstein-Uhlenbeck) process is defined by the following recurrence relation, see, e.g., (Papoulis, 1966; Feller, 1972)

$$F_t = \rho F_{t-1} + \xi_t, \qquad -1 < \rho < 1 \tag{3.11}$$

where ξ_t is denotes Gaussian white noise with variance σ_ξ^2. The resulting time series is stationary and has the Markov property. Its autocorrelatation function is

$$r(s) = \varrho^s = \exp(-s/\ell), \qquad \ell \stackrel{\text{def}}{=\!=} -\frac{1}{\ln \varrho} \tag{3.12}$$

where λ is the *correlation length* as defined in (Weinberger, 1990; Fontana *et al.*, 1991). Conversely, any Gaussian stationary Markov process has an autocorrelation function of the form (3.12). The parameter ϱ measures the correlation of the time series. If $\varrho \approx 0$ then the time series is almost uncorrelated, i.e., $\{F_t\}$ is almost white noise. On the other hand, for $\varrho \approx 1$ the time series approximates Brownian motion.

Before we proceed let us briefly discuss the relation between ℓ and $\hat{\ell}$ defined in section 2. The RNA free energy landscapes and almost all of the model landscapes that have been investigated so far have correlation length that scale linearly with n, see e.g. (Schuster & Stadler, 1994) for a recent overview. In other words, we have $\varrho \stackrel{\text{def}}{=\!=} 1 - x$ where x scales as $1/n$ for large systems. If $r(s)$ is of the form (3.12), then we find

$$\begin{aligned} \ell &= \frac{1}{x} - \frac{1}{2} + \frac{1}{12}x + \mathcal{O}(x^2) \\ \hat{\ell} &= \frac{1}{x} - \frac{1}{2} \end{aligned} \tag{3.13}$$

Thus ℓ and $\hat{\ell}$ differ only by a contribution of order $(1-\varrho) \sim 1/n$ for the landscapes of interest.

The power spectrum of an AR(1) time series is

$$S(\omega) = \frac{\sigma^2(1-\varrho^2)}{2\pi(1+\varrho^2 - 2\varrho\cos\omega)}, \tag{3.14}$$

see, e.g., (Yaglom, 1986). In fact, the Wiener-Khinchin theorem, equ.(2.9), shows that equ.(3.14) hold for all elementary landscapes, irrespective of the distribution function of the fitness values.

Weinberger (Weinberger, 1990) called a landscape $f : \mathcal{C} \to \mathbb{R}$ an *AR(1) landscape* if the time series obtained by a random walk on the landscape is Gaussian and has an autocorrelation function of the form (3.14). The parameter ϱ describes the ruggedness of the landscape. The landscapes with exponential autocorrelation functions are *exactly* the elementary landscape discussed above. An AR(1) landscape in the sense of Weinberger is thus an elementary landscape with a Gaussian fitness distribution. A number of model landscapes have been shown to be elementary (Grover, 1992; Weinberger & Stadler, 1993; Stadler, 1995c). Most of them have in fact a Gaussian distribution of fitness values, at least asymptotically as a consequence of the central limit theorem. The best known examples are the p-spin models, the graph bipartitioning problem, graph matching, graph coloring, and symmetric traveling salesman problems. Kauffman's Nk models are *approximately* AR(1); their decomposition into elementary components is discussed in detail in (Stadler & Happel, 1995). The class of landscapes that are *approximately AR(1)* includes a variety of landscapes based on RNA secondary structures (Fontana *et al.*, 1991; Fontana *et al.*, 1993; Bonhoeffer *et al.*, 1993; Tacker *et al.*, 1994).

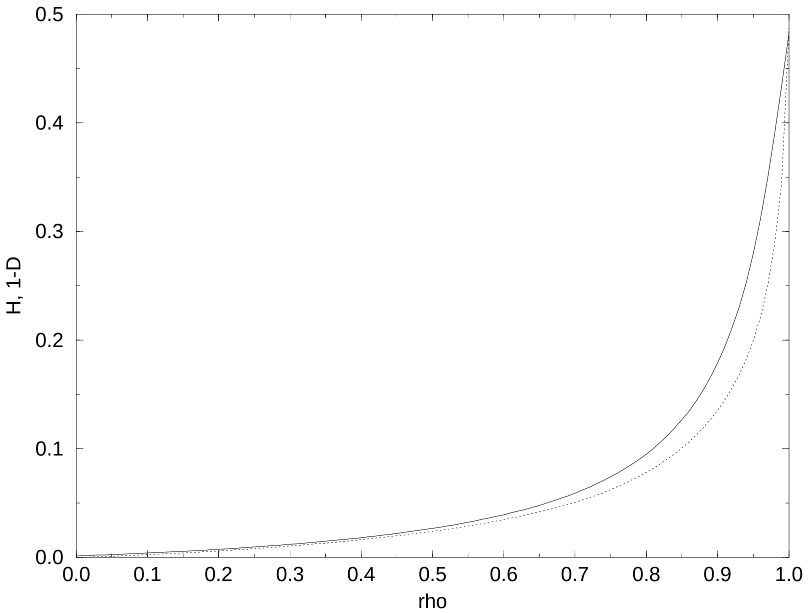

Figure 1: Approximated scaling exponents H (solid line) and $1-D$ (dotted line) as a function of the correlation ρ of an AR(1) process. The deviations are of the order of 10%

The following considerations, like equ.(3.14), depend only on the form of correlation function $r(s)$, not on the distribution function of the fitness values. They are therefore valid for any elementary landscape. The linear approximation

$$r(s) \approx 1 - s/\ell, \qquad s \ll n \tag{3.15}$$

of equ.(3.12) is a good approximation for highly correlated landscapes, i.e., for landscapes with correlation lengths $\ell = \mathcal{O}(n)$. By comparing equ.(3.15) with equ.(3.9) we observe that elementary landscapes with large correlation length are locally self-affine, with scaling parameter H=1/2, i.e., time series obtained from such landscapes behave locally like ordinary Brownian motion.

Surprisingly, however, we find that even AR(1) time series with small correlation length show approximate power laws for $J(s)$, equ.(3.2), and for the curve lengths $L(s)$. Numerical simulations show that we have $H \to 0$ for $\varrho \to 0$, while $\rho \to 1$ yields $H \to 1/2$. Data obtained from direct measurement of H, according to equ.(3.3), and estimates of the scaling exponent D of the curve length obtained from (3.5) are consistent with each other. Best fits of the characteristic exponents H and $1 - D$ as functions of ϱ are shown in Figure 1. It is interesting to note in this context that certain log-normal distributions can also mimic $1/f$ spectra in a limited frequency domain (Montroll & Shlesinger, 1983).

4. RNA Free Energy Landscapes

Folding biopolymer sequences into structures is a central problem in molecular biology research. Both robustness and accessibility of structures, as functions of mutational change in the underlying sequence, are crucial to natural as well as molecular evolution applied to biotechnology. RNA molecules are an excellent model system. In fact, they are the only class of biopolymers for which the folding problem has been solved at least at the level of secondary structures.

An RNA sequence is a string of length n composed of an alphabet of size α. In nature the alphabet consists of the $\alpha = 4$ bases **G**uanine, **C**ytosine, **A**denosine, and **U**racile. In this paper we shall also consider the restricted alphabet {**G**,**C**} with $\alpha = 2$. A natural distance between sequences is the Hamming distance measuring the number of positions in which two sequences differ (Hamming, 1950). The configuration space is hence a generalization of the Boolean hypercube known as the sequence space.

A secondary structure is tantamount to a list of Watson-Crick type and **GU** base pairs. Such a structure can be uniquely decomposed into structural elements that are (i) base pair stacks, (ii) loops differing in size (number of unpaired bases) and branching degree: hairpin loops (degree one), internal loops (degree two or more), and (iii) bases which are not part of a stack or a loop are termed external (freely rotating joints and unpaired ends). Each stack or loop element contributes additively to the overall free energy of the structure. These energy terms are empirically determined parameters that depend on the nucleotide sequence (Freier *et al.*, 1986). The folding process considered here maps an RNA sequence into a *secondary structure* minimizing free energy. This structure can be computed using a dynamic programming algorithm (Zuker & Stiegler, 1981; Zuker & Sankoff, 1984). The implementation used in this contribution is described in detail in (Hofacker *et al.*, 1994); it is available as a public domain package (Hofacker *et al.*, 1995).

In this contribution we focus not on the secondary structures themselves but rather on the free energies, ΔG, of structure formation. The bulk properties of these

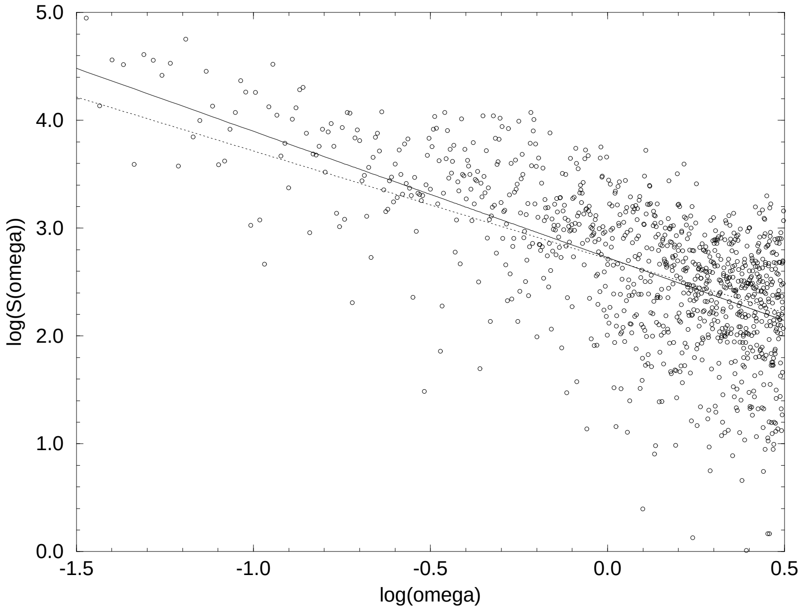

Figure 2: Raw dat of the power spectrum obtained for a **GC** landscape with chain length $n = 70$ at excursion level $\Delta G = 0$. Walk length is $N = 2048$. The solid line is the best fit to $S(\omega) = 1/\omega^a$, with $a = -1.17$. The dotted line is *1/f*-noise.

minimum free energy landscapes haven been studied extensively in the past (Bonhoeffer *et al.*, 1993; Fontana & Schuster, 1987; Fontana *et al.*, 1989; Fontana *et al.*, 1991; Fontana *et al.*, 1993; Fontana *et al.*, 1993; Schuster *et al.*, 1994; Schuster & Stadler, 1994). They are typical representants of rugged landscapes.

Figure 2 shows a sample power spectrum obtained along a random walk as described in section 2. The data are rather noisy. In order to smooth them we break the walk into pieces of 128 steps, calculate the power spectrum for each of them, and then we average the power spectra (128 steps is about twice the diameter of the sequence space in this case, thus significantly longer walks are not meaningful because the range of local self-affinity is necessarily restricted to a small multiple of the the geometrical relaxation time τ of the random walk $\{x_t\}$):

$$\langle d(s)\rangle = \langle d(\infty)\rangle\,(1 - e^{-s/\tau}) \tag{4.1}$$

For a free random walk on a sequence spaces we find

$$\tau = \frac{\alpha - 1}{\alpha} n + \mathcal{O}(1)\,. \tag{4.2}$$

Table 1. Power spectrum index a for time series obtained from RNA landscapes. The values of a as obtained directly from the power spectrum are compared to the values calculated from jump exponent H and the scaling exponent D, $a_H = 1+2H$, and $a_D = 3-2D$.

n	50			70			90		
ΔG^+	a	a_H	a_D	a	a_H	a_D	a	a_H	a_D
	AUGC								
0	1.32	1.36	1.24	1.46	1.41	1.36	1.56	1.42	1.35
5	1.19	1.25	1.16	1.45	1.35	1.29	1.58	1.42	1.35
10	1.06	1.18	1.12	1.27	1.29	1.22	1.50	1.35	1.28
15	1.07	1.17	1.15	1.18	1.15	1.14	1.29	1.24	1.15
20	*	1.15	1.11	1.11	1.15	1.11	1.13	1.16	1.13
25	*	1.15	1.13	*	1.15	1.11	1.13	1.15	1.13
	GC								
0	1.15	1.16	1.12	1.25	1.21	1.19	1.24	1.27	1.20
5	1.15	1.16	1.12	1.25	1.21	1.19	1.24	1.27	1.20
10	1.15	1.16	1.12	1.25	1.21	1.19	1.24	1.27	1.20
15	1.15	1.15	1.12	1.25	1.21	1.19	1.24	1.27	1.20
20	1.09	1.17	1.12	1.25	1.21	1.19	1.24	1.27	1.20
25	1.00	1.10	1.09	1.25	1.21	1.19	1.24	1.27	1.20

$^+$ ΔG in kcal/mol.
* indicates insufficient data.
Systematic errors are estimated to be of the order of 10%, compare Figure 1.

This expression will be derived in the Appendix.

The data are consistent with a ω^{-a} spectrum with a not much larger than 1. Numerical values are shown in Table 1. It turns out, however, that the data are also consistent with the power spectrum of an AR(1) process, see Figure 3.

The additivity of the energy contributions implies a certain degree of neutrality in the landscape, for details see (Fontana *et al.*, 1993). Several structures which consist of identical sets of substructures map onto the same selective values, although their phenotypic appearances are different. In fact, there are very large neutral networks on the level of secondary structures themselves (Schuster *et al.*, 1994; Reidys *et al.*, 1995; Grüner, 1994). This implies that even at fairly high excursion

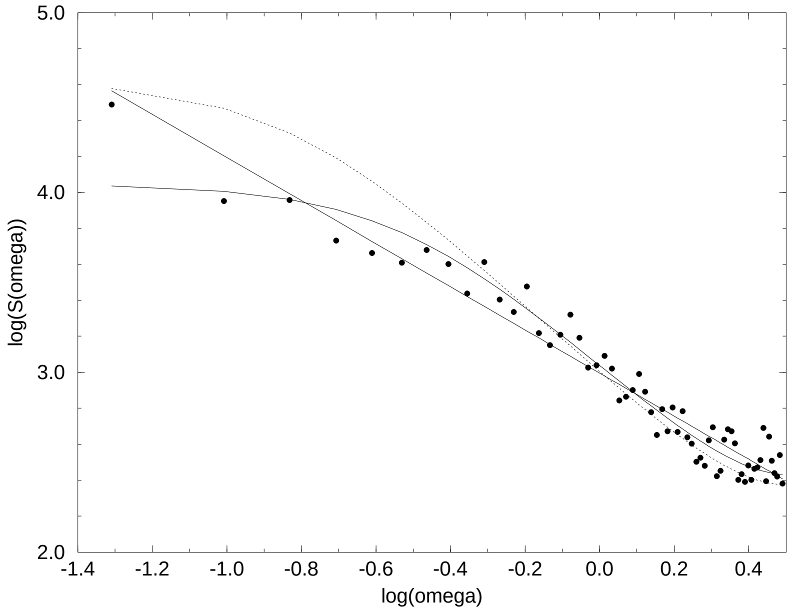

Figure 3: The dots are spectral data for walks on a **GC** landscape with n=70, excursion level 0, averaged over 10 walks. The solid curve is the best fit to the AR(1) power spectrum, equ.(3.15), with an estimate for the parameter $\varrho \approx 0.73$, corresponding to a correlation length $\ell \approx 3.2$. The correlation length estimated directly from the autocorrlation function is about 6.7 as shown in ref. (Fontana *et al.*, 1993). The corresponding power spectrum is shown as dotted line. The solid straight line is a least square fit with a power law ω^{-a} we find $a \approx -1.20$.

levels (a couple of standard deviation above the mean) the excursion sets are still large.

We find that the scaling properties do not depend strongly on the excursion level. There is, however, a systematic trend towards smaller values of α for higher excursion levels. Our data indicate that mountainous regions of the landscape are not drastically different from the average. Our data are biassed by the geometry of the cycles C_E, however, and hence a detailed quantitative analysis is not possible at present.

5. Conclusions

The structure of the mountainous parts of RNA free landscapes was studied by random walks confined to excursion sets at given energy levels.

Spectral data and local scaling analysis of the series generated by simple random walks show self-affinity consistent with the low-frequency behavior of an AR(1) time series. We find that in general an AR(1) processe appears to be approximately self-affine on length scales smaller than a few correlation lengths. The data obtained from RNA free energy landscapes indicate that a fractal-like structure is present at length scales up to the diameter of the sequence space. This is a consequence of the fact that the correlation length of the RNA free energy landscapes is comparable to the sequence length.

Our computer experiments exhibit no significant dependence of the statistical properties of the excursion set confined walk on the energy level. Hence, at least qualitatively, the statistical properties of the mountains do not differ from the low-lands. This is true at least as long as the excursion set does not break up in very small cycles.

The present study suggests that a detailed investigation of the percolation of excursion sets, of the geometry of excursion sets, and of the geometrical relaxation of random walks confined to cycles will be necessary before a complete understanding of the structure of the mountain ranges of fitness landscape is possible.

Acknowledgments

The work was funded by ÖAD proj.no. Z1.:20721/EH/91 EH-Project 894/91. S.B. gratefully thanks Prof. Murali Sheshadri and Prof. A. Nadarajan for their interest and support. P.F.S. thanks the Inst. Ciencias Nucleares and the Inst. de Fisica of the Universidad Nacional Autonoma de Mexico for their hospitality in September 1995, when this paper was finished.

References

Adler, D. (1981). *The Geometry of Random Fields.* New York: John Wiley & Sons.

Attkinson, W., Fey, L., & Newman, J. (1963). Spectrum analysis of extremely low frequency variation of quartz ocillators. *Proc.IEEE,* **51**, 379.

Azencott, R. (1992). *Simulated Annealing — parallelization techniques.* New York: John Wiley & Sons.

Bonhoeffer, S., McCaskill, J., Stadler, P., & Schuster, P. (1993). Temperature dependent RNA landscapes, a study based on partition functions. *European Biophysics Journal,* **22**.

Bonhoeffer, S. & Stadler, P. F. (1993). Errortreshold on complex fitness landscapes. *J.Theor.Biol.* **164**, 359–372.

Brouwer, A., Cohen, A., & Neumaier, A. (1989). *Distance-regular Graphs.* Berlin, New York: Springer Verlag.

Buldyrev, S., Goldberger, A., & Stanley, H. (1995). Long-range correlation properties of coding and noncoding dna sequences: Genbank analysis. *Phys.Rev.E,* **51**, 5084–5091.

Burlaga, L. & Klein, L. (1986). Fractal structure of the interplanetary magnetic field. *J.Geophys.Res.* **91**, 347–???

Eigen, M., McCaskill, J., & Schuster, P. (1989). The molecular Quasispecies. *Adv. Chem. Phys.* **75**, 149 – 263.

Feller, W. (1972). *An Introduction to Probability Theory and its Applications.* New York: Wiley.

Fontana, W., Griesmacher, T., Schnabl, W., Stadler, P., & Schuster, P. (1991). Statistics of landscapes based on free energies, replication and degredation rate constants of RNA secondary structures. *Monatshefte der Chemie,* **122**, 795–819.

Fontana, W., Konings, D. A. M., Stadler, P. F., & Schuster, P. (1993). Statistics of rna secondary structures. *Biochemistry,* **33**, 1389–1404.

Fontana, W., Schnabl, W., & Schuster, P. (1989). Physical aspects of evolutionary optimization and adaption. *Physical Review A,* **40** (6), 3301–3321.

Fontana, W. & Schuster, P. (1987). A computer model of evolutionary optimization. *Biophysical Chemistry,* **26**, 123–147.

Fontana, W., Stadler, P. F., Bornberg-Bauer, E. G., Griesmacher, T., Hofacker, I. L., Tacker, M., Tarazona, P., Weinberger, E. D., & Schuster, P. (1993). RNA folding and combinatory landscapes. *Phys. Rev. E,* **47** (3), 2083 – 2099.

Freidlin, M. & Wentzell, A. (1984). *Random Perturbations of Dynamical Systems.* New York: Springer-Verlag.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986). Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci., USA,* **83**, 9373–9377.

Grover, L. (1992). Local search and the local structure of NP-complete problems. *Oper.Res.Lett.* **12**, 235–243.

Grüner, W. (1994). *Evolutionary Optimization on RNA Folding Landscapes.* PhD thesis Inst. of Theoretical Chemistry, Uni. Vienna, Austria.

Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell Syst.Tech.J.* **29**, 147–160.

Happel, R. & Stadler, P. F. (1995). Canonical approximation of fitness landscapes. Santa Fe Institute Preprint 95-07-068.

Higuchi, T. (1988). Approach to an irregular time series on the basis of fractal theory. *Physica D,* **31**, 277–282.

Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, L. S., Tacker, M., & Schuster, P. (1995). `Vienna RNA Package`. `pub/RNA/ViennaRNA-1.03 @ ftp.itc.univie.ac.at`. (Public Domain Software).

Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, S., Tacker, M., & Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. *Monatsh. Chemie,* **125** (2), 167–188.

Hordijk, W. (1995). A measure of landscapes. Santa Fe Institute Preprint 95-05-049.

Huynen, M. A., Stadler, P. F., & Fontana, W. (1995). Evolution of RNA and the Neutral Theory. *Proc.Natl.Acad.Sci.* . in press, Santa Fe Institute Preprint 95-01-006.

Kauffman, S. (1993). *The Origin of Order.* New York, Oxford: Oxford University Press.

Li, W. & Kaneko, K. (1992). Long-range correlation and partial $1/f^\alpha$ spectrum in a noncoding dna sequence. *Europhys.Lett.* **17**, 655–660.

Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature.* New York: Freeman.

Mandelbrot, B. B. & vanNess, J. W. (1968). Fractional brownian motion, fractional noise, and applications. *SIAM Rev.* **10**, 422–437.

Mohar, B. (1991). The laplacian spectrum of graphs. In: *Graph Theory, Combinatorics, and Applications,* (Alavi, Y., Chartrand, G., Ollermann, O., & Schwenk, A., eds) pp. 871–897, New York: John Wiley & Sons.

Montroll, E. & Shlesinger, M. (1983). Maximum entropy formalism, fractals, scaling phenomena, and 1/f noise. A tale of tails. *J.Stat.Phys.* **32**, 209–230.

Musha, T. & Higuchi, H. (1976). The 1/f fluctuation of a traffic current on an expressway. *Jap.J.Appl.Phys.* **15**, 1271–1275.

Nolan, P. L., Gruber, D. E., Matteson, J. L., Peterson, L. E., Rothschild, R. E., Doty, J. P., Levine, A. M., Lewin, W. H. G., & Primini, F. A. (1981). Rapid variability of 10-140 kev X-rays from Cygnus X-1. *Astrophys.J.* **246**, 494–501.

Osborne, A. & Provenzale, A. (1989). Finite correlation dimension for stochastic systems with power law spectra. *Physica D,* **35**, 357–381.

Palmer, R. (1991). Optimization on rugged landscapes. In: *Molecular Evolution on Rugged Landscapes: Proteins, RNA, and the Immune System*, (Perelson, A. S. & Kauffman, S. A., eds) pp. 3–25. Addison Wesley Redwood City, CA.

Papoulis, A. (1966). *Probability, Random Variables, and Stochastic Processes.* New York: McGraw Hill.

Reidys, C., Schuster, P., & Stadler, P. F. (1995). Generic properties of combinatory maps: Neutral networks of RNA secondary structures. Santa Fe Institute Preprint 95-07-058.

Schuster, P., Fontana, W., Stadler, P. F., & Hofacker, I. L. (1994). From sequences to shapes and back: A case study in RNA secondary structures. *Proc.Roy.Soc.Lond.B,* **255**, 279–284.

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

Sorkin, G. B. (1988). Combinatorial optimization, simulated annealing, and fractals. Technical Report RC13674 (No.61253) IBM Research Report.

Spitzer, F. (1971). Markov random fields and gibbs ensembles. *Amer. Math. Monthly,* **78**, 142–154.

Stadler, P. F. (1992). Correlation in landscapes of combinatorial optimization problems. *Europhys. Lett.* **20**, 479–482.

Stadler, P. F. (1995a). Random walks and orthogonal functions associated with highly symmetric graphs. *Disc. Math.* . in press, Santa Fe Institute Preprint 93-06-31.

Stadler, P. F. (1995b). Towards a theory of landscapes. In: *Complex Systems and Binary Networks*, (López Peña, R., ed). Springer-Verlag New York. in press, Santa Fe Institute Preprint 95-03-30.

Stadler, P. F. (1995c). Landscapes and their correlation functions. Santa Fe Institute Preprint 95-07-067.

Stadler, P. F. & Happel, R. (1992). Correlation structure of the landscape of the graph-bipartitioning-problem. *J. Phys. A.: Math. Gen.* **25**, 3103–3110.

Stadler, P. F. & Happel, R. (1995). Random field models for fitness landscapes. Santa Fe Institute Preprint 95-07-069.

Stadler, P. F. & Schnabl, W. (1992). The landscape of the traveling salesman problem. *Phys. Letters A,* **161**, 337–344.

Tacker, M., Fontana, W., Stadler, P., & Schuster, P. (1994). Statistics of RNA melting kinetics. *Eur. J. Biophys.* **23**, 29–38.

Taft, B., Hickey, B., Wunsch, C., & Baker, D. (1974). Equatorial undercurrent and deeper flows in the central pacific. *Deep Sea Res.* **21**, 403–430.

Verveen, A. A. & Derkson, H. E. (1968). Fluctuation phenomena in nerve membranes. *Proc.IEEE,* **56**, 906–916.

Voss, R. F. (1992). Evolution of long-range fractal correlations and 1/f noise in DNA base sequences. *Phys.Rev.Lett.* **68**, 3805–3808.

Voss, R. F. & Clarke, J. (1975). 1/f noise in music and speech. *Nature,* **258**, 317–318.

Weinberger, E. D. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biol.Cybern.* **63**, 325–336.

Weinberger, E. D. (1991a). Local properties of Kauffman's N-k model: A tunably rugged energy landscape. *Phys. Rev. A,* **44(10)**, 6399–6413.

Weinberger, E. D. (1991b). Fourier and Taylor series on fitness landscapes. *Biological Cybernetics,* **65**, 321–330.

Weinberger, E. D. & Stadler, P. F. (1993). Why some fitness landscapes are fractal. *J. Theor. Biol.* **163**, 255–275.

Weissman, M. (1988). 1/f noise and other slow, non-exponential kinetics in condensed matter. *Rev.Mod.Phys.* **60**, 537–571.

Yaglom, A. (1986). *Correlation Theory of Stationary and Related Random Functions*, volume 1,2. New York: Springer-Verlag.

Zuker, M. & Sankoff, D. (1984). RNA secondary structures and their prediction. *Bull.Math.Biol.* **46** (4), 591–621.

Zuker, M. & Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamic and auxilliary information. *Nucl.Acid Res.* **9**, 133–148.

Appendix: Relaxation of Random Walks in Sequence Spaces

For a sequence space a detailed analysis of the geometric relaxation of a simple random walk is possible. The probabilities φ_{sd} as defined in sect.2 can be obtained recursively from

$$\begin{aligned} &\varphi_{00} = 1 \\ &\varphi_{sd} = 0 \text{ for } s < d \\ &\varphi_{sd} = w^{+}(d-1)\varphi_{s-1,d-1} + w^{0}(d)\varphi_{s-1,d} + w^{-}(d+1)\varphi_{s-1,d+1} \end{aligned} \tag{A.1}$$

where coefficients $w^{+}(d)$, $w^{0}(d)$, and $w^{-}(d)$ are the probabilities for making a step forwards, backwards or sidewards given one is in distance d from the origin of the walk. For sequence spaces we have (Fontana *et al.*, 1991)

$$w^{+}(d) = \frac{n-d}{n} \qquad w^{0}(d) = \frac{d}{n}\frac{\alpha-2}{\alpha-1} \qquad w^{-}(d) = \frac{d}{n}\frac{1}{\alpha-1} \tag{A.2}$$

Define the moments of the distribution φ_{sd} by

$$\Delta_m(s) = \langle d(s)^m \rangle = \sum_d \varphi_{sd} d^m \tag{A.3}$$

$\Delta_1(s)$ is then the average distance after s steps. Inserting the recursion (A.1) into the definition of $\Delta_m(s)$ yields after considerable algebra the following recursion for the m-th moment

$$\begin{aligned} \Delta_m(s) = 1 + &\sum_{\ell=0}^{m-1} \binom{m}{m-\ell}\left[1 - \frac{m-\ell}{\ell+1}\frac{\alpha}{\alpha-1}\frac{1}{n}\right] \cdot \Delta_{m-\ell}(s-1) \\ &+ \frac{2}{\alpha n}\sum_{\ell=0}^{m/2-1} \binom{m}{m-2\ell}\Delta_{m-2\ell-1}(s-1) \end{aligned} \tag{A.3}$$

This recursion is of the form $\vec{\Delta}(s) = 1 + A \cdot \vec{\Delta}(s-1)$, where A is lower triangular. Hence the eigenvalues λ_m of A are given by the diagonal elements of A:

$$\lambda_m = 1 - m\frac{1}{n}\frac{\alpha}{\alpha-1} \tag{A.4}$$

The m-th moment is therefore of the form

$$\Delta_m(s) = \Delta_m(\infty)\left[1 - a_k\lambda_k^s\right]. \tag{A.5}$$

The slowest mode corresponds to the eigenvalue λ_1. The corresponding relaxation time is

$$\tau_1 = -\frac{1}{\ln\lambda_1} = \frac{\alpha-1}{\alpha}n + \mathcal{O}(1) \tag{A.6}$$

for large n. Explicit expressions for the long time limits $\Delta_m(\infty)$ of the moments are obtained as non-zero fixed points of the recursions (A.3).

Table of Contents