Ising Quantum Chain and Biological Evolution

Holger Wagner[1] and **Ellen Baake**[2]

[1] Institut für Theoretische Physik, Universität Tübingen,
Auf der Morgenstelle 14, D-72076 Tübingen, Germany

[2] Zoologisches Institut, Universität München,
Luisenstr. 14, D-80333 München, Germany

Abstract

A sequence space model which describes the interplay of mutation and selection in molecular evolution is shown to be equivalent to an Ising quantum chain. Three explicit examples with representative fitness landscapes are discussed and exactly solved with methods from statistical mechanics.

Introduction

In the theory of (molecular) biological evolution, various sequence space models are well established, one of the best-known being Eigen's quasispecies model [1]. This model assumes mutations to originate as replication errors on the occasion of reproduction events. We investigate an alternative here which describes mutation and selection as going on in parallel; it is abbreviated as *para-muse* (parallel mutation-selection) model. Which is the more appropriate version from the biological point of view amounts to the question whether rates of molecular evolution are closer to constant per generation or constant in time – a long-standing, but still unresolved issue. Even in the former situation, however, the parallel version is an excellent approximation.

In this contribution, we will show that the para-muse model corresponds to the Hamiltonian of an Ising quantum chain. Methods of statistical mechanics, which exploit the translation or permutation invariance, respectively, will then be used to treat three sample fitness landscapes exactly.

Evolution and quantum chain

In the framework of population genetics, individuals are characterized by their genetic information. In sequence space models, a gene (or the whole genome) is identified with a point in (binary) sequence space, e.g. $\{-1,1\}^N$, where $N \in \mathbb{N}$ is the (fixed) length of the string considered; so there are 2^N different sequences (or types of individuals) termed A_i, $i = 1, \ldots, 2^N$. By identification of A_i with the ith element of the canonical basis of $\otimes^N \mathbb{C}^2$, the composition of an (infinite) population of haploid organisms (they

carry one copy of every gene only) under the influence of mutation and selection acting *independently* of each other is given by

$$\boldsymbol{p}(t) \;=\; \frac{\boldsymbol{x}(t)}{\sum_{j=1}^{n} x_j(t)}\,, \tag{1}$$

where $\boldsymbol{x}$ evolves according to

$$\frac{d}{dt}\,\boldsymbol{x}(t) \;=\; (\mathcal{M}+\mathcal{R})\,\boldsymbol{x}(t)\;. \tag{2}$$

Here, p_i denotes the relative frequency of A_i individuals ($1 \le i \le 2^N$), and $\mathcal{R} := \mathrm{diag}(r_1,\ldots,r_n)$ and $\mathcal{M} = (m_{ij})$ are the matrices of reproduction and mutation rates, respectively. In particular, m_{ij} is the rate at which A_j mutates to A_i. When every digit mutates independently at rate $\mu \ge 0$, we have

$$m_{ij} \;=\; \begin{cases} \mu\,, & d_{ij}=1 \\ -N\mu\,, & i=j \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{3}$$

where the Hamming distance $d_{ij} := d(A_i, A_j)$ is the number of positions where the two strings differ. The r_i are, as yet, unspecified; their choice determines the so-called fitness landscape.

The problem is completely solved as soon as the spectrum of $\mathcal{M}+\mathcal{R}$ is known. For $\mu > 0$, $\mathcal{M}+\mathcal{R}$ plus a suitable constant is primitive (i.e. some power of this sum has strictly positive entries only). Consequently, the largest eigenvalue and the corresponding Perron-Frobenius (PF) eigenvector of $\mathcal{M}+\mathcal{R}$ are unique.

Let us first observe that, owing to our two-letter alphabet, the possible occupation numbers are 0 and 1 at every site. We can thus rewrite $\mathcal{H} := \mathcal{M}+\mathcal{R}+N\mu$ in terms of Pauli's matrices:

$$\mathcal{H} \;=\; \mu \sum_{k=1}^{N} \sigma_k^x \;+\mathcal{R} \tag{4}$$

with $\mathcal{R} \in \mathrm{span}_{\mathbb{R}}\,\langle \sigma_k^z\,; 1 \le k \le N\rangle$. Here, $\sigma^x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma^z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and

$$\sigma_k^a \;:=\; \mathbb{1}_2 \otimes \ldots \otimes \mathbb{1}_2 \otimes \sigma^a \otimes \mathbb{1}_2 \otimes \ldots \otimes \mathbb{1}_2 \tag{5}$$

with $a \in \{x, z\}$ and σ^a in the k-th place. The σ_k^x flips the k-th digit ('mutation'), and σ_k^z measures the value of the variable (the spin) at site k.

Obviously, $\mathcal{H}$ is, at the same time, the Hamiltonian of an Ising quantum chain in a transverse magnetic field, with general interactions within the chain. Since energies appear with positive sign here, the PF eigenvalue belongs to the ground state. In the sequel we discuss three *exactly solvable* examples from this model class.

Mount Fujiyama landscape

Let us first consider the following simple Hamiltonian which is composed from non-interacting one-site Hamiltonians $\mathcal{H}_k$:

$$\mathcal{H} \;=\; \sum_{k=1}^{N} \mathcal{H}_k \;=\; \sum_{k=1}^{N} (\mu\, \sigma_k^x + \alpha\, \sigma_k^z)\;. \tag{6}$$

In the evolution model, this is the case of no interaction between sites, or purely additive fitness (cf. Eq. (2)):

$$r_i \;=\; \sum_{j=1}^{N} \alpha\, s_j^{(i)}\;, \tag{7}$$

where $s_j^{(i)}$ is the value of the variable at site j of sequence A_i. This fitness landscape is an N-dimensional "Fujiyama-peak"; at the same time, it yields an Ising quantum chain without spin-spin interaction.

To characterize the genetic structure of the population, the average surplus u of sites with value $+1$,

$$u(t) \;:=\; \frac{\sum_{i=1}^{n} u_i\, x_i(t)}{\sum_{j=1}^{n} x_j(t)}, \quad u_i \;:=\; \frac{1}{N} \sum_{j=1}^{N} s_j^{(i)}\;, \tag{8}$$

is appropriate. For $t \to \infty$, the composition of the population becomes stationary: $\boldsymbol{x}(t) \to \boldsymbol{v}$, where $\boldsymbol{v}$ is the PF eigenvector of $\mathcal{H}$. In this case, we find ($h := \mu/\alpha$)

$$u \;=\; \frac{1}{h + \sqrt{1+h^2}}. \tag{9}$$

Hence u vanishes smoothly when $\mu \to \infty$, and no phase transition occurs (see Fig. 1).

Onsager landscape

Let us now discuss the well-known Ising quantum chain in a transverse field, with constant coupling γ of nearest-neighbour spins and cyclic boundary conditions, i.e. $\sigma_{N+1}^z = \sigma_1^z$,

$$\mathcal{H} \;=\; \mu \sum_{k=1}^{N} \sigma_k^x + \gamma \sum_{k=1}^{N} \sigma_k^z \sigma_{k+1}^z\;. \tag{10}$$

The implied fitness landscape is interesting since there are flat ridges (sequences from which one cannot go to a fitter sequence by a single mutation event) and since there are, in general, many different sequences with the same fitness (a phenomenon known as 'neutrality').

Again, the surplus u (or order parameter) as defined in Eq. (8) is the proper quantity to characterize the biological population. It is similar but not directly related to the physical magnetization; the replacement of $\lim_{t\to\infty} \exp(t(\mathcal{M}+\mathcal{R}))$ by

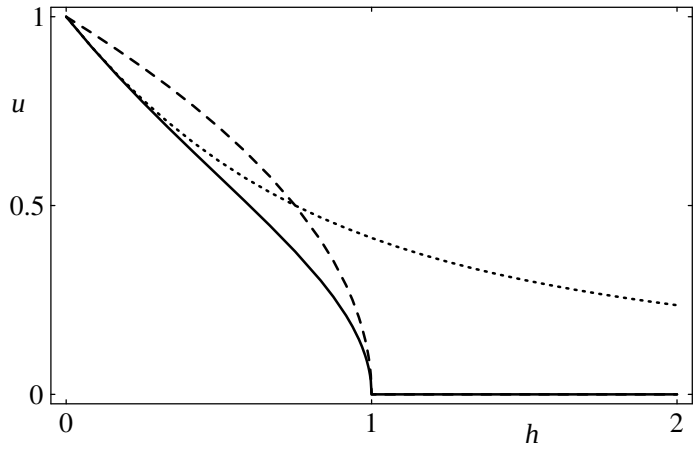

Figure 1: Average surplus of sites with value +1 (u) as defined in (8), for a stationary population in the macroscopic limit. Dotted: Fujiyama landscape; dashed: Onsager landscape; solid line: mean-field landscape.

$\lim_{M\to\infty} (\exp(\mathcal{R})\exp(\mathcal{M}))^M$ (this corresponds to the transition from the para-muse model to a discrete-time version of Eigen's quasispecies model [2]) maps the surplus to the surface magnetization of a classical 2D-Ising model. One finds that the surplus in the limit $\lim_{N\to\infty}\lim_{t\to\infty}(u(t))$ is given by

$$u = \begin{cases} (1-h)^{1/2} \ , & 0 \le h < 1 \\ 0 \ , & h \ge 1 \, . \end{cases} \tag{11}$$

A second-order phase transition occurs at $h = 1$ (see Fig. 1). Here, the population loses its genetic structure, a phenomenon known as 'error threshold'.

Mean-field landscape

Nearest neighbour interaction has provided us with a fitness landscape with ridges and neutrality, but it should not be taken too literally as a model of biological interaction. After all, DNA strings serve as templates for proteins, which then fold in three dimensions, thus giving rise to interactions that are long-range and very complicated in the sequence picture.

Let us therefore, as a representative model with long-range interaction, consider the mean-field Hamiltonian

$$\mathcal{H} = \mu \sum_{k=1}^{N} \sigma_k^x + \frac{\gamma}{2N} \sum_{k,\ell=1}^{N} \sigma_k^z \sigma_\ell^z \ . \tag{12}$$

In the context of evolution, the mean-field landscape is related to the quadratic-optimum model as used extensively in quantitative genetics. We use mean-field theory

to reduce the desired solution of a pure phase equilibrium state to that of a one-site Hamiltonian in a self-consistent field.

For the surplus of the stationary population, one obtains for this landscape, with $h := \mu/\gamma$,

$$u = \begin{cases} \left(\frac{1-h}{1+h}\right)^{1/2}, & 0 \leq h < 1 \\ 0 \quad , & h \geq 1\,. \end{cases} \tag{13}$$

Again, a second-order phase transition occurs at $h = 1$, and, as in Eq. (11), the exponent of the singular behaviour is 1/2 (see Fig. 1). At the same time, the mean fitness of the population (or ground state energy) reaches its minimum value; consequently, selection ceases to operate.

Concluding remarks

We have demonstrated how the Ising quantum chain can be employed to obtain exact solutions of muse models. The idealized fitness landscapes treated explicitly represent frustration-free situations without, with short-range, and with extremely long-range interaction. Genetic structure fades away gradually with increasing mutation in the (unrealistic) case of independent sites, but vanishes abruptly at finite μ when sites interact, thus imposing an upper limit on the mutation rate in an evolving population. Further details concerning the stationary situation, and results on the time-dependent behaviour, will be presented in forthcoming papers [3,4].

Acknowledgments

It is our pleasure to thank M. Baake and T. Gerisch for their cooperation.

References

[1] M. Eigen, J. McCaskill, and P. Schuster, "The molecular quasi-species", *Adv. Chem. Phys.* **75**, 149 (1989).

[2] I. Leuthäusser, "Statistical mechanics of Eigen's evolution model", *J. Stat. Phys.* **48**, 343 (1987).

[3] E. Baake, M. Baake, and H. Wagner, "Ising quantum chain is equivalent to a model of biological evolution", *Phys. Rev. Lett.*, in press (1996).

[4] H. Wagner, T. Gerisch, and E. Baake, in preparation.