The mathematics of Darwin's theory of evolution: 1859 and 150 years later

Peter Schuster

Abstract. A mathematical formulation of Darwin's theory of evolutionary optimization through variation and selection is derived in terms of conventional ODEs that can be interpreted as chemical kinetics of evolution. Variation in form of mutation and recombination operates on genotypes being DNA or RNA sequences, whereas phenotypes, which are represented by organisms or molecular structures, are the target of selection. The impact of recombination on optimization is briefly sketched. Differential equations modeling selection in populations with correct replication and mutation are derived from the molecular mechanisms of polynucleotide replication. The analysis of these ODEs reveals restrictions of the optimization principle caused by mutation. Error propagation over generations sets a limit to mutation rates in evolution, which manifests itself in the form of a phase transition-like phenomenon characterized as error threshold. Conditions on fitness landscapes for the occurrence of error thresholds derived from numerical investigations are presented: Smooth fitness landscapes show no error thresholds but gradual transitions, sufficiently steep landscapes and rugged landscapes sustain error thresholds. Sharp transitions are also found with realistic landscapes combining ruggedness and neutrality. Lethal mutants may lead to extinction of populations and set another upper limit to mutation rates in form of an extinction threshold through lethal mutagenesis.

Mathematics Subject Classification (2010). Primary 92D25; Secondary 80A30.

Keywords. Error threshold, genotype-phenotype mapping, lethal mutagenesis, population dynamics, neutrality, quasispecies, sequence space.

This contribution reports work that was performed at Vienna University and at the Santa Fe Institute with financial support from the *Austrian Fonds zur Förderung der wissenschaftlichen Forschung (FWF).*

1. Preamble

This paper has been presented at a meeting that celebrated the 150 years anniversary of Charles Darwin's famous book on the *Origin of Species* [1]. Darwin was presumably the greatest naturalist that has ever lived, a highly talented observer, and a genius in intuition. He was certainly not a fan of mathematics and his centennial treatise of evolution does not contain a single formula. Here we shall make a *gedankenexperiment*: How might Charles Darwin have formulated his theory if he were a mathematician? The knowledge on mathematics related to evolution was sparse or non existing at Darwin's time. The mechanism of inheritance and the origin of variation were completely unknown to Charles Darwin and his contemporary biologists. Darwin's speculations on blending of the parental properties in the offspring were completely off the point. Gregor Mendel did the first careful experiments on plants [2] and drew the right conclusions concerning inheritance: The genomes of both parents are split into pieces and recombined in the offspring thereby conserving the order of genes but choosing more or less randomly from father or mother. Mendel had an education in mathematics and physics and with this background he was in the position to discover a statistical law – a regularity that becomes evident only when sufficiently many experiments are superimposed and correctly evaluated. Recombination is one common source of phenotypic variation in sexually reproducing species. Mutation, the second source of variation that is occurring in both, asexual and sexual species, and was identified as a change in the nucleotide sequence of the genetic message. Although Gregor Mendel gave the correct interpretation of his experiments and discovered the idealized principle of recombination, a full understanding of sexual reproduction without the insights from cellular and molecular biology is not possible. Mutation cannot be understood at all without molecular knowledge.

2. Selection and optimization

Charles Darwin's principle of natural selection is a powerful abstraction from observations, which provides insight into the origin of changing species. Species or populations don't multiply but individuals do, either directly in asexual species, like viruses, bacteria or protists, or in sexual species through pairings of individuals with opposite sex. Variability of individuals in populations is an empirical fact that can be seen easily in everyday life. Within populations the variants are subjected to natural selection and those having more progeny prevail in future generations. The power of Darwin's abstraction lies in the fact that neither the shape and the structure of individuals nor the mechanism of inheritance are relevant for selection unless they have an impact on the number of offspring. Otherwise Darwin's approach had been

doomed to fail since his imagination of inheritance was incorrect. Indeed Darwin's principle holds simultaneously for highly developed organisms, for primitive unicellular species like bacteria, for viruses and even for reproducing molecules in cell-free assays.

Molecular biology provided a powerful possibility to study evolution in its simplest form outside biology: Replicating ribonucleic acid molecules (RNA) in cell-free assays [3] play natural selection in its purest form: In the test tube, evolution, selection, and optimization are liberated from all unnecessary complex features, from obscuring details, and from unimportant accessories. Hence, *in vitro* evolution can be studied by the methods of chemical kinetics. The parameters determining the "fitness of molecules" are replication rate parameters, binding constants, and other measurable quantities, which can be determined independently of *in vitro* evolution experiments, and constitute an alternative access to the determination of the outcome of selection. Thereby "survival of the fittest" is unambiguously freed from the reproach of being the mere tautology of "survival of the survivor". In addition, *in vitro* selection turned out to be extremely useful for the synthesis of molecules that are tailored for predefined purposes. A new area of applications called evolutionary biotechnology branched off evolution in the test tube. Examples for evolutionary design of molecules are [4, 5] for nucleic acids, [6, 7] for proteins, and [8] for small organic molecules.

The section starts by mentioning a few examples of biological applications of mathematics before Darwin (subsection 2.1), we derive and analyze an ODE describing simple selection with asexual species (subsection 2.2), and consider the effects of variable population size (subsection 2.3). The next subsection 2.4 analyzes optimization in the Darwinian sense, subsection 2.5 presents a brief account of Fisher's selection equation and his fundamental theorem of natural selection and eventually we consider generic properties of typical growth functions (subsection 2.6).

2.1. Counting and modeling before Darwin

The first mathematical model that seems to be relevant for evolution was conceived by the medieval mathematician Leonardo Pisano also known as Fibonacci. His famous book *Liber abaci* has been finished and published in the year 1202 and was translated into modern English eight years ago [9]. Among several other important contributions to mathematics in Europe Fibonacci discusses a model of rabbit multiplication in *Liber abaci*. Couples of rabbits reproduce and produce young couples of rabbits according to the following rules:

(i) Every adult couple has a progeny of one young couple per month,
(ii) a young couple grows to adulthood within the first month and accordingly begins producing offspring in the second months,
(iii) rabbits live forever, and

(iv) the number of rabbit couples is updated every month.

The model starts with one young couple (**1**), nothing happens during maturation of couple **1** in the first month and we have still one couple in the second month. In the third month, eventually, a young couple (**2**) is born and the number of couples increases to two. In the fourth month couple **1** produces a new couple (**3**) whereas couple **2** is growing to adulthood, and we have three couples now. Further rabbit counting yields the Fibonacci sequence:[1]

month	0	1	2	3	4	5	6	7	8	9	...
# couples	0	1	1	2	3	5	8	13	21	34	...

It is straightforward to derive a recursion for the rabbit count. The number of couples in month $(n+1)$, f_{n+1}, is the sum of two terms: The number of couples in month n, because rabbits don't die, plus the number of young couples that is identical to the number of couples in month $(n-1)$:

$$f_{n+1} = f_{n-1} + f_n \quad \text{with} \quad f_0 = 0 \;\text{ and }\; f_1 = 1\,. \tag{2.1}$$

With increasing n the ratio of two subsequent Fibonacci numbers converges to the golden ratio, $f_{k+1}/f_k = (1+\sqrt{5})/2$ (For a comprehensive discussion of the Fibonacci sequence and its properties see [12, pp.290-301] or, e.g., [13]).

In order to proof this convergence we make use of a matrix representation of the Fibonacci model:

$$\mathrm{F}^n \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} = \begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix} \;\text{ with }\; \mathrm{F} = \begin{pmatrix} f_0 & f_1 \\ f_1 & f_2 \end{pmatrix} \text{ and } \mathrm{F}^n = \begin{pmatrix} f_{n-1} & f_n \\ f_n & f_{n+1} \end{pmatrix}.$$

The matrix representation transforms the recursion into an expression that allows for direct computation of the elements of the Fibonacci sequence.

$$f_n = \begin{pmatrix} 1 & 0 \end{pmatrix} \mathrm{F}^n \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} f_{n-1} & f_n \\ f_n & f_{n+1} \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \end{pmatrix}. \tag{2.2}$$

Theorem 2.1 (Fibonacci convergence). *With increasing n the Fibonacci sequence converges to a geometric progression with the golden ratio as factor,* $q = (1+\sqrt{5})/2$.

Proof. The matrix F is diagonalized by the transformation $T^{-1} \cdot \mathrm{F} \cdot \mathrm{T} = \mathrm{D}$ with $\mathrm{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. The two eigenvalues of F are: $\lambda_{1,2} = (1 \pm \sqrt{5})/2$. Since F is a symmetric matrix the $\mathbb{L}^2$-normalized eigenvectors of F, $(\mathbf{e}_1, \mathbf{e}_2) = \mathrm{T}$, form an orthonormal set,

$$T = \begin{pmatrix} \frac{1}{\sqrt{1+\lambda_1^2}} & \frac{1}{\sqrt{1+\lambda_2^2}} \\ \frac{\lambda_1}{\sqrt{1+\lambda_1^2}} & \frac{\lambda_2}{\sqrt{1+\lambda_2^2}} \end{pmatrix} \quad \text{and} \quad T \cdot T' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

[1] According to Parmanand Singh [10] the Fibonacci numbers were invented earlier in India and used for the solution of various problems (See also Donald Knuth [11]).

with T' being the transposed matrix, and $\mathrm{T}^{-1} = \mathrm{T}'$. Computation of the n-th power of matrix F yields

$$\mathrm{F}^n = T \cdot \mathrm{D}^n \cdot \mathrm{T}' = \mathrm{T} \cdot \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \cdot \mathrm{T}' = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^{n-1} - \lambda_2^{n-1} & \lambda_1^n - \lambda_2^n \\ \lambda_1^n - \lambda_2^n & \lambda_1^{n+1} - \lambda_2^{n+1} \end{pmatrix} ,$$

from which the expression for f_n is obtained by comparison with (2.2)

$$f_n = \frac{1}{\sqrt{5}}(\lambda_1^n - \lambda_2^n) \; . \tag{2.3}$$

Because $\lambda_1 > \lambda_2$ the ratio converges to zero: $\lim_{n\to\infty} \lambda_2^n/\lambda_1^n = 0$, and the Fibonacci sequence is approximated well by $f_n \approx \frac{1}{\sqrt{5}}\, q^n$ with $q = (1+\sqrt{5})/2$. □

Since λ_2 is negative the Fibonacci sequence alternates around the geometric progression. Expression (2.3) is commonly attributed to the French mathematician Jacques Binet [14] and named after him. As outlined in ref. [12, p.299] the formula has been derived already hundred years before by the great Swiss mathematician Leonhard Euler [15] but was forgotten and rediscovered.

Thomas Robert Malthus was the first who articulated the ecological and economic problem of population growth following a geometric progression [16]: Animal or human populations like every system capable of reproduction grow like a geometric progression provided unlimited resources are available. The resources, however, are either constant or grow – as Malthus assumes – according to an arithmetic progression if human endeavor is involved. The production of nutrition, says Malthus, is proportional to the land that is exploitable for agriculture and the gain in the area of fields will be a constant in time – the increase will be the same every year. An inevitable result of Malthus' vision of the world is the pessimistic view that populations will grow until the majority of individuals will die premature of malnutrition and hunger. Malthus could not foresee the *green revolutions* but he was also unaware that population growth can be faster than exponential – sufficient nutrition for the entire human population is still a problem. Charles Darwin and his younger contemporary Alfred Russel Wallace were strongly influenced by Robert Malthus and took form population theory that in the wild, where birth control does not exist and individuals fight for food, the major fraction of of progeny will die before they reach the age of reproduction and only the strongest will have a chance to multiply.

Leonhard Euler introduced the notions of the exponential function in the middle of the eighteenth century [17] and set the stage for modeling populations by means of differential equations. Simple reproduction results

in exponential growth of a population with $N(t)$ individuals:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN \quad \text{and} \quad N(t) = N_0 \exp(rt) \quad \text{with} \quad N_0 = N(0) \ , \tag{2.4}$$

where the parameter r is commonly called Malthus or growth parameter.

Presumably not known to Darwin, the mathematician Jean François Verhulst complemented the concept of exponential growth by the introduction of finite resources [18, 19, 20]. The Verhulst equation is of the form[2]

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN\left(1 - \frac{N}{K}\right) \ , \tag{2.5}$$

where $N(t)$ again denotes the number of individuals of a species $\mathcal{X}$, and K is the *carrying capacity* of the ecological niche or the ecosystem. Equ. (2.5) can be integrated by means of partial fractions and reads

$$N(t) = N_0 \, \frac{K}{N_0 + \big(K - N_0\big)\exp(-rt)} \ . \tag{2.6}$$

Apart from the initial condition N_0, the number of individuals $\mathcal{X}$ at time $t = 0$, the logistic equation has two parameters: (i) the Malthusian parameter or the growth rate r and (ii) the carrying capacity K of the ecological niche or the ecosystem. A population of size N_0 grows exponentially at short times: $N(t) \approx N_0 \exp(rt)$ for $K \gg N_0$ and t sufficiently small. For long times the population size approaches the carrying capacity asymptotically: $\lim_{t\to\infty} N(t) = K$.

The two parameters r and K are taken as criteria to distinguish different evolutionary strategies: Species that are r-selected exploit ecological niches with low density, produce a large number of offspring each of which has a low probability to survive, whereas K-selected species are strongly competing in crowded niches and invest heavily in few offspring that have a high probability of survival to adulthood. The two cases, r- and K-selection, are the extreme situations of a continuum of mixed selection strategies. In the real world the r-selection strategy is an appropriate adaptation to fast changing environments, whereas K-selection pays in slowly varying or constant environments.

2.2. The selection equation

The logistic equation can be interpreted in a different way that is useful for the forthcoming analysis: In the second term — $-(N/K)\,rN$ — the expression rN/K is identified with a constraint for limiting growth: $rN/K \equiv \phi(t)$,

$$\frac{\mathrm{d}N}{\mathrm{d}t} = N\Big(r - \phi(t)\Big) \ , \tag{2.5'}$$

[2]The Verhulst equation is also called logistic equation and its discrete analogue is the logistic map, a standard model to demonstrate the occurrence of deterministic chaos in a simple system. The name *logistic equation* was coined by Verhulst himself in 1845.

The introduction of $\phi(t)$ gives room for other interpretations of constraints than carrying capacities of ecosystems. For example, $\phi(t)$ may be a dilution flux in laboratory experiments on evolution in flow reactors [21, pp.21-27].

Equ. (2.5') can be used now for the derivation of a selection equation in the spirit of Darwin's theory. The single species $\mathcal{X}$ is replaced by several variants forming a population, $\Pi = \{\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n\}$; in the language of chemical kinetics competition and selection are readily cast into a reaction mechanism consisting of n independent, simple replication reactions:

$$(\mathsf{M}) + \mathcal{X}_j \xrightarrow{f_j} 2\,\mathcal{X}_j \ , \quad j = 1, 2, \ldots, n \ . \tag{2.7}$$

The symbol M denotes the material from which $\mathcal{X}_j$ is synthesized (It is put in parentheses, because we assume that it is present in access and its concentration is constant therefore). The numbers of individuals of the variants are denoted by $N_j(t)$, or in vector notation $\mathbf{N}(t) = \big(N_1(t), N_2(t), \ldots, N_n(t)\big)$ with $\sum_{i=1}^{n} N_i(t) = C(t)$. The same common carrying capacity is defined for all n variants:

$$\lim_{t\to\infty} \sum_{i=1}^{n} N_i(t) = \lim_{t\to\infty} C(t) = K \ .$$

The Malthus parameters are given here by the fitness values $f_1, f_2, \ldots, f_n$, respectively. For individual species the differential equations take on the form

$$\begin{aligned} &\frac{\mathrm{d}N_j}{\mathrm{d}t} = N_j \left(f_j - \frac{C}{K}\phi(t) \right) ; \ j = 1, 2, \ldots, n \quad \text{with} \\ &\phi(t) = \frac{1}{C} \sum_{i=1}^{n} f_i N_i(t) \end{aligned} \tag{2.8}$$

being the mean fitness of the population. Summation over all species yields a differential equation for the total population size

$$\frac{\mathrm{d}C}{\mathrm{d}t} = C \left(1 - \frac{C}{K} \right) \phi(t) \ . \tag{2.9}$$

Stability analysis is straightforward: From $\mathrm{d}C/\mathrm{d}t = 0$ follow two stationary states of equ. (2.9): (i) $\bar{C} = 0$ and (ii) $\bar{C} = K$.[3] For conventional stability analysis we calculate the (1×1) Jacobian and obtain for the eigenvalue

$$\lambda = \frac{\partial\big(\mathrm{d}C/\mathrm{d}t\big)}{\partial C} = \phi(t) - \frac{C}{K} \left(2\phi(t) - K \frac{\partial \phi}{\partial C} \right) - \frac{C^2}{K} \frac{\partial \phi}{\partial C} \ .$$

[3]There is also a third stationary state defined by $\phi = 0$. For strictly positive fitness values, $f_i > 0 \forall\, i = 1, 2, \ldots, n$, this condition can only be fulfilled by $N_i = 0 \forall\, i = 1, 2, \ldots, n$, which is identical to state (i). If some f_i values are zero – corresponding to lethal variants – the respective variables vanish in the infinite time limit because of $\mathrm{d}N_i/\mathrm{d}t = -\phi(t)\, N_i$ with $\phi(t) > 0$.

Insertion of the stationary values yields $\lambda_{(i)} = \phi > 0$ and $\lambda_{(ii)} = -\phi < 0$, state (i) is unstable and state (ii) is asymptotically stable. The total population size converges to the value of the carrying capacity, $\lim_{t\to\infty} C(t) = \bar{C} = K$.

Equ. (2.9) can be solved exactly yielding thereby an expression that contains integration of the constraint $\phi(t)$:

$$C(t) = C(0)\frac{K}{C(0) + \big(K - C(0)\big)\exp(-\Phi)} \quad \text{with} \quad \Phi = \int_0^t \phi(\tau)\mathrm{d}\tau \ ,$$

where $C(0)$ is the population size at time $t = 0$. The function $\Phi(t)$ depends on the distribution of fitness values within the population and its time course. For $f_1 = f_2 = \ldots = f_n = r$ the integral yields $\Phi = rt$ and we retain equation (2.6). In the long time limit Φ grows to infinity and $C(t)$ converges to the carrying capacity K.

At constant population size $C = \bar{C} = K$ equ. (2.8) becomes simpler

$$\frac{\mathrm{d}N_j}{\mathrm{d}t} = N_j\Big(f_j - \phi(t)\Big)\,; \quad j = 1, 2, \ldots, n\,. \tag{2.8'}$$

and can be solved exactly by means of the integrating factor transformation [22, p. 322ff.]:

$$Z_j(t) = N_j(t)\exp\left(\int_0^t \phi(\tau)\,\mathrm{d}\tau\right) \quad \text{or} \quad \mathbf{Z}(t) = \mathbf{N}(t)\exp\left(\int_0^t \phi(\tau)\,\mathrm{d}\tau\right),$$

solving the linear differential equation $\mathrm{d}\mathbf{Z}/\mathrm{d}t = \mathrm{F}\cdot\mathbf{Z}$, where F is a diagonal matrix containing the fitness values f_j $(j = 1, 2, \ldots, n)$ as elements, and using the trivial equality $\mathbf{Z}_0 = \mathbf{N}_0$ we obtain for the components of $\mathbf{N}$:

$$N_j(t) = N_j(0)\exp(f_j t)\,\frac{C}{\sum_{i=1}^{n} N_i(0)\exp(f_i t)}; \ j = 1, 2, \ldots, n\ . \tag{2.10}$$

Equ. (2.10) encapsulates Darwinian selection and will be discussed in detail in section 2.4.

2.3. Variable population size

Now we shall show that the solution of equ. (2.8) describes internal equilibration for constant **and** variable population sizes as long as the population does neither explode nor die out [23]. The validity of theorem 2.2 that will be proven below is not restricted to constant fitness values f_j and hence we can replace them by general growth functions $G_j(N_1, \ldots, N_n) = G_j(\mathbf{N})$ or fitness functions $F_j(\mathbf{N})$ with $G_j(\mathbf{N}) = F_j(\mathbf{N})N_j$ in the special case of replicator equations [24]: $\mathrm{d}N_j/\mathrm{d}t = N_j\big(F_j(\mathbf{N}) - \Psi(t)\big)$ where $\Psi(t)$ comprises both, variable total concentration and constraint.

Time dependence of the conditions in the ecosystem can be introduced in two ways: (i) variable carrying capacity, $K(t) = \bar{C}(t)$, and (ii) a constraint or flux[4] $\varphi(t)$, where flux refers to some specific physical device, for example

[4]There is a slight difference in the definitions of the fluxes ϕ and φ: $\phi(t) = \varphi(t)/C(t)$.

to a flow reactor. Constraints and fluxes may correspond to unspecific or specific migration.[5] Considering time dependent carrying capacity and variable constraints simultaneously, we obtain

$$\frac{\mathrm{d}N_j}{\mathrm{d}t} = G_j(\mathbf{N}) - \frac{N_j}{K(t)}\varphi(t);\ j = 1, 2, \ldots, n\ . \tag{2.11}$$

Summation over all variants $\mathcal{X}_j$ and restricting the analysis to an equilibrated total concentration $C \approx \bar{C} = K$ yields a relation between the time dependencies of flux and total concentration:

$$\begin{aligned} \varphi(t) &= \sum_{i=1}^{n} G_j(\mathbf{N}) \ - \ \frac{\mathrm{d}C}{\mathrm{d}t} \quad \text{or} \\ C(t) &= C(0) + \int_0^t \left(\sum_{i=1}^{n} G_j(\mathbf{N}) - \varphi(\tau) \right) \mathrm{d}\tau\ . \end{aligned} \tag{2.12}$$

Theorem 2.2 (Equilibration in populations of variable size). *Evolution in populations of changing size approaches the same internal equilibrium as evolution in populations of constant size provided the growth functions are homogeneous functions of degree γ in the variables N_j. Up to a transformation of the time axis, stationary and variable populations have identical trajectories provided the population size stays finite and does not vanish.*

Proof. Normalized variables, $x_i = N_i/C$ with $\sum_{i=1}^{n} x_i = 1$, are introduced in order to separate of population growth, $C(t)$, and population internal changes in the distribution of variants $\mathcal{X}_i$. From equs. (2.11) and (2.12) with $C = \bar{C} = K$ and $N_j = Cx_j$ follows:

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = \frac{1}{C}\left(G_j\big(C\mathbf{x}\big) - x_j \sum_{i=1}^{n} G_i\big(C\mathbf{x}\big) \right);\ j = 1, 2, \ldots, n\ . \tag{2.13}$$

The growth functions are assumed to be homogeneous of degree γ in the variables[6] N_j: $G_j(\mathbf{N}) = G_j(C\mathbf{x}) = C^{\gamma}\, G_j(\mathbf{x})$. and we find

$$\frac{1}{C^{\gamma-1}}\frac{\mathrm{d}x_j}{\mathrm{d}t} = G_j(\mathbf{x}) \ - \ x_j \sum_{i=1}^{n} G_i(\mathbf{x});\ j = 1, 2, \ldots, n\ ,$$

which is identical to the selection equation in normalized variables for $C = 1$. For $\gamma = 1$ the concentration term vanishes and populations of constant and

[5] *Unspecific migration* means that the numbers N_j of individuals for each variant $\mathcal{X}_j$ decrease (or increase) proportional to the numbers of individuals currently present in the population, $\mathrm{d}N_j = kN_j\mathrm{d}t$. *Specific migration* is anything else. In a flow reactor, for example, we have a dilution flux corresponding to unspecific emigration and an influx of one or a few molecular species corresponding to specific immigration into the reactor.

[6] The degree γ is determined by the mechanism of reproduction. For sexual reproduction according to Ronald Fisher's selection equation (2.19) [25] we have $\gamma = 2$. Asexual reproduction discussed here fulfils $\gamma = 1$.

variable size have identical trajectories and equilibrium points. In case $\gamma \neq 1$ the two systems are the same up to a transformation of the time axis:

$$\mathrm{d}\tilde{t} = C^{\gamma-1}\,\mathrm{d}t \quad \text{and} \quad \tilde{t} = \tilde{t}_0 + \int_{\tilde{t}_0}^{\tilde{t}} C^{\gamma-1}(t)\,\mathrm{d}t \ ,$$

where $\tilde{t}_0$ is the time corresponding to $t = 0$ (commonly $\tilde{t}_0 = 0$). From equ. (2.13) we expect instabilities at $C = 0$ and $C = \infty$. □

The instability at vanishing population size, $\lim C \to 0$, is of practical importance for modeling drug action on viral replication. In the case of lethal mutagenesis [26, 27] medication aims at eradication of the virus population, $C \to 0$, in order to terminate the infection of the host. At the instant of virus extinction equ. (2.8) is no longer applicable. More details will be discussed in section 3.7.

2.4. Optimization

Since systems with growing and stationary population size are identical for homogeneous growth function of degree $\gamma = 1$ by theorem 2.2, we shall use normalized or internal coordinates except in subsection 3.7. The ODE is of the form

$$\begin{aligned} &\frac{\mathrm{d}x_j}{\mathrm{d}t} = f_j x_j - x_j\phi(t) = x_j\Big(f_j - \phi(t)\Big);\ j = 1, 2, \ldots, n \quad \text{with} \\ &\phi(t) = \sum_{i=1}^{n} f_i x_i \ , \end{aligned} \tag{2.14}$$

the solution is derived in the same way as in case of equ.(2.8):

$$x_j(t) = \frac{x_j(0)\exp(f_j t)}{\sum_{i=1}^{n} x_i(0)\exp(f_i t)};\ j = 1, 2, \ldots, n \ . \tag{2.15}$$

The use of normalized variables, $\sum_{i=1}^{n} x_i = 1$, implies that the unit simplex, $\mathbb{S}_n^{(1)} = \{0 \leq x_i \leq 1 \forall\ i = 1, \ldots, n \wedge \sum_{i=1}^{n} x_i = 1\}$, is the physically accessible domain. All boundaries of the simplex — corners, edges, faces, etc. — are invariant sets, since $x_j = 0 \Rightarrow \mathrm{d}x_j/\mathrm{d}t = 0$ by equ. (2.14).

For a discussion of selection and optimization we shall assume here that all fitness values f_j are different and that without loosing generality we rank them:

$$f_1 > f_2 > \ldots > f_{n-1} > f_n \ . \tag{2.16}$$

The variables $x_j(t)$ fulfil two time limits:

$$\lim_{t\to 0} x_j(t) = x_j(0)\ \forall\ j = 1, 2, \ldots, n \ \text{ by definition, and}$$

$$\lim_{t\to\infty} x_j(t) = \begin{cases} 1 \text{ iff } j = 1 \\ 0\ \forall\ j = 2, \ldots, n \ . \end{cases}$$

In the long time limit the population becomes homogeneous and contains only the fittest genotype $\mathcal{X}_1$. The process of selection is illustrated best by differential fitness, $f_j - \phi(t)$, the second factor in the ODE (2.14): The constraint $\phi(t) = \sum_{i=1}^{n} f_i x_i = \bar{f}$ represents the mean fitness of the population. The population variables x_l of all variants with a fitness below average, $f_l < \phi(t)$, decrease whereas the variables x_h with $f_h > \phi(t)$ increase. As a consequence the average fitness $\phi(t)$ in increasing too and more sequences fall below the threshold for survival. The process continues until the fittest variant is selected.

Optimization of mean fitness can be proved also without referring to differential fitness:

Theorem 2.3 (Optimization of mean fitness). *The mean fitness* $\phi(t) = \bar{f} = \sum_{i=1}^{n} f_i x_i$ *with* $\sum_{i=1}^{n} x_i = 1$ *in a population as described by equ.* (2.14) *is non-decreasing.*

Proof. The time dependence of the mean fitness or flux ϕ is given by

$$\begin{aligned} \frac{\mathrm{d}\phi}{\mathrm{d}t} &= \sum_{i=1}^{n} f_i \dot{x}_i = \sum_{i=1}^{n} f_i \left(f_i x_i - x_i \sum_{j=1}^{n} f_j x_j \right) = \\ &= \sum_{i=1}^{n} f_i^2 x_i - \sum_{i=1}^{n} f_i x_i \sum_{j=1}^{n} f_j x_j \ = \\ &= \overline{f^2} - \left(\bar{f}\right)^2 = \mathrm{var}\{f\} \geq 0 \ . \end{aligned} \tag{2.17}$$

Since a variance is always nonnegative, equation (2.17) implies that $\phi(t)$ is a non-decreasing function of time. □

The condition $\mathrm{var}\{f\} = 0$ is met only by homogeneous populations. The one containing only the fittest variant $\mathcal{X}_1$ has the largest possible mean fitness: $\bar{f} = \phi_{\max} = f_1 = \max\{f_j; j = 1, 2, \ldots, n\}$. ϕ cannot increase any further and hence, it was been optimized by the selection process. The state of maximal fitness of population $\Pi = \{\mathcal{X}_1, \ldots, \mathcal{X}_n\}$, $\mathbf{x}|_{\max\{\phi(\Pi)\}} = \{x_1 = 1, x_i = 0 \,\forall\, i = 2, \ldots, n\} = \mathbf{P}_1$, is the unique stable stationary state, and all trajectories starting from initial conditions with nonzero amounts of $\mathcal{X}_1$, $x_1 > 0$, have $\mathbf{P}_1$ as ω-limit. An illustration of the selection process with three variants is shown in figure 2.1: The trajectories are plotted on the unit simplex $\mathbb{S}_3^{(1)}$.

Gradient systems [28, p.199] facilitate the analysis of the dynamics, they obey the equation

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = -\mathrm{grad}\{V(\mathbf{x})\} = -\nabla V(\mathbf{x}) \tag{2.18}$$

and fulfil criteria that are relevant for optimization:

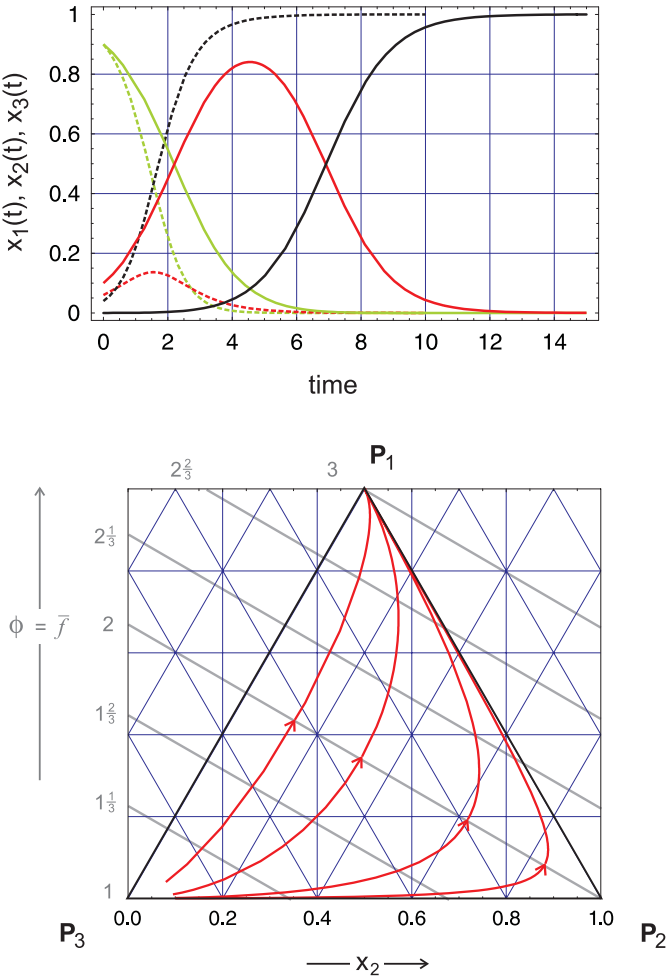

FIGURE 2.1. **Selection on the unit simplex.** In the upper part of the figure we show solution curves $\mathbf{x}(t)$ of equ. (2.15) with $n = 3$. The parameter values are: $f_1 = 3\,[\mathrm{t}^{-1}]$, $f_2 = 2\,[\mathrm{t}^{-1}]$, and $f_3 = 1\,[\mathrm{t}^{-1}]$, where $[\mathrm{t}^{-1}]$ is an arbitrary reciprocal time unit. The two sets of curves differ with respect to the initial conditions:
(i) $\mathbf{x}(0) = (0.02, 0.08, 0.90)$, dotted curves, and
(ii) $\mathbf{x}(0) = (0.0001, 0.0999, 0.9000)$, full curves.
Color code: $x_1(t)$ black, $x_2(t)$ red, and $x_3(t)$ green. The lower part of the figure shows parametric plots $\mathbf{x}(t)$ on the unit simplex $\mathbb{S}_3^{(1)}$. Constant level sets of $\phi(\mathbf{x}) = \bar{f}$ are shown in grey.

(i) The eigenvalues of of the linearization of (2.18) evaluated at the equilibrium point are real.
(ii) If $\bar{\mathbf{x}}_0$ is an isolated minimum of V then $\bar{\mathbf{x}}_0$ is an asymptotically stable solution of (2.18).
(iii) In $\mathbf{x}(t)$ is a solution of (2.18) that is not an equilibrium point, then $V\big(\mathbf{x}(t)\big)$ is a strictly decreasing function and the trajectories are perpendicular to the constant level sets of V.
(iv) Neither periodic nor chaotic solutions of (2.18) do exist.

As easily seen from figure 2.1 the trajectories of (2.14) are not perpendicular to the constant level sets of $\phi(\mathbf{x})$ and hence, equ. (2.14) is not a gradient system in the strict sense. With the definition of a generalized inner product corresponding to a Riemannian metric [29], however, the selection equation can be visualized as a generalized gradient and oscillations or deterministic chaos can be excluded [30].

2.5. Fisher's selection equation

Sexual reproduction introduces obligatory recombination of genotypes into the selection equation. A sexually reproducing organism carries two copies of every gene. Contrasting asexual reproduction where the whole genome is replicated and transferred to progeny in one piece, sexual reproduction is accompanied by partitioning the two parental genomes into pieces and by organized recombination of genes into the genomes of the progeny.[7] On the population level genes are chosen from a set of n variants called the gene pool: $\mathcal{A} = \{A_1, \ldots, A_n\}$. Every specific (single) locus on the genome is occupied by two alleles. Ronald Fisher, the great scholar of population genetics, presented the first mathematical unification of Darwin's theory of natural selection and Mendel's laws of inheritance [25]. Since population genetics will be treated extensively in other contributions, only a very brief account on Fisher's the selection equation and his fundamental theorem of natural selection will be given here. The selection equation describes the evolution of the

[7] *Organized* means here that each offspring carries two *alleles* for every gene. An *allele* A_j is a variant of a specific gene. One allele of the offspring is one of the two maternal alleles, the second one comes from the father who also carries two alleles of the gene. The position of a gene on the genome or chromosome is called the *locus*. The genome consists of several chromosomes. There are two classes of chromosomes: autosomes and sex chromosomes. In humans the males carry two different sex chromosomes, X and Y, with different genes and this constitutes an exception of the two alleles per gene rule.

allele distribution at a single locus:

$$\begin{aligned} \frac{\mathrm{d}x_j}{\mathrm{d}t} &= x_j \left(\sum_{i=1}^{n} a_{ji}\, x_i - \phi(t) \right) \; ; \; j = 1, \ldots, n \quad \text{with} \\ \phi(t) &= \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji}\, x_i\, x_j \quad \text{and} \quad \sum_{i=1}^{n} x_i \, = \, 1 \; . \end{aligned} \tag{2.19}$$

The (normalized) variables x_j represent the allele frequencies, the parameters a_{ji} represent the fitness values for the allele combination $[A_j \cdot A_i]$, and the constraint $\phi(t)$ conserves the normalization condition. Fisher's selection equation (2.19) is a replicator equation [24] with $F_j(\mathbf{x}) = \sum_{i=1}^{n} a_{ji} x_i$ being a linear function and $G_j(\mathbf{x})$ is homogeneous with $\gamma = 2$. Since the fitness of an allele combination is assumed to be independent of the descendence of the allele on an autosome – it does not matter whether a particular allele stems from the paternal or the maternal chromosome, the allele combinations $[A_j \cdot A_i]$ and $[A_i \cdot A_j]$ have identical fitness, and the matrix $\mathrm{A} = \{a_{ji}; i, j = 1, 2, \ldots, n\}$ is symmetric.

The introduction of mean rate parameters $\bar{a}_i = \sum_{j=1}^{n} a_{ij} x_j$ facilitates the forthcoming calculation. The time dependence of ϕ is now given by

$$\begin{aligned} \frac{\mathrm{d}\phi}{\mathrm{d}t} &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \left(\frac{\mathrm{d}x_i}{\mathrm{d}t} \cdot x_j + x_i \cdot \frac{\mathrm{d}x_j}{\mathrm{d}t} \right) = 2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} \cdot x_i \cdot \frac{\mathrm{d}x_j}{\mathrm{d}t} = \\ &= 2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} \cdot x_i \left(\sum_{k=1}^{n} a_{jk} x_j x_k - x_j \sum_{k=1}^{n} \sum_{\ell=1}^{n} a_{k\ell} x_k x_\ell \right) = \\ &= 2 \sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ji} x_i \sum_{k=1}^{n} a_{jk} x_k - 2 \sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ji} x_i \sum_{k=1}^{n} x_k \sum_{\ell=1}^{n} a_{k\ell} x_\ell = \\ &= 2 \left(\langle \bar{a}^2 \rangle - \langle \bar{a} \rangle^2 \right) = 2 \, \mathrm{var}\{\bar{a}\} \geq 0 \; . \end{aligned} \tag{2.20}$$

Again we see that the constraint $\phi(t)$ is a non-deceasing function of time, and it approaches an optimal value on the simplex. This result is often called Fisher's fundamental theorem of evolution (See, for example, [31]).

The physically relevant part of $\mathbb{R}^n$ with non-negative values for all variables is the unit simplex $\mathbb{S}_n^{(1)}$ and – as in case of the selection equation (2.14) for asexual reproduction – all boundary sets of the simplex are invariant for Fisher's selection equation (2.19). Both equations, (2.14) and (2.19), have also in common that the selection constraint $\phi(t)$ is a non-decreasing function of time t, and accordingly $\phi(t)$ is optimized. There is, however, also an important difference between the two selection equations: The selected state is unique in the asexual case, whereas Fisher's selection equation may converge to different states for different initial conditions. A straightforward example is sufficiently large fitness of the homozygotes $[A_i \cdot A_i]$ and $[A_j \cdot A_j]$ compared

to the heterozygote $[A_j \cdot A_i]$, $(a_{ii}, a_{jj}) > a_{ij}$. Then, either of the two homozygotes – corresponding to the corners $\mathbf{P}_i$ and $\mathbf{P}_j$ of the simplex – may be selected and the outcome of the selection process depends on the initial state $\mathbf{x}(0) = \mathbf{x}_0$. In case the heterozygote has higher fitness than the homozygotes,[8] $a_{ij} > (a_{ii}, a_{jj})$, Fisher's selection equation (2.19) predicts that the selected stationary allele distribution becomes 0.5:0.5. Heterozygote selection in real populations is more complicated, since a homogeneous population of heterozygotes is not compatible with random mating. In this case the resulting diploid combinations, $[A_i \cdot A_i]$, $[A_j \cdot A_i]$ and $[A_j \cdot A_j]$, are formed with a ratio of 0.25:0.5:0.25. Homozygotes will not vanish completely unless they are lethal. A real world example of overdominance with one practically lethal homozygote is sickle-cell anemia (For an overview of the highly complex phenotypes of this disease see [32, 33]).

In the simple version presented here, Fisher's fundamental theorem of natural selection is identified with optimization of $\phi(t)$ in the selection equation (2.19). Unfortunately – but fortunately for population geneticists and theoretical biologists, because a whole plethora of interesting mathematical problems emerged and still emerges from the mathematics of recombination – Fisher's selection equation is a single locus model and holds on the genome level only for the unrealistic assumption of independent genes. Two and more locus models with gene interaction turned out to be much more complicated and no generally valid optimization principle was reported so far: Natural selection in the sense of Charles Darwin is an extremely powerful optimization heuristic but no theorem (see also subsection 3.3). Nevertheless, Fisher's fundamental theorem is much deeper than the toy version that has been presented here. The interested reader is referred to a few, more or less arbitrarily chosen references from the enormous literature on this issue [34, 35, 36].

2.6. Growth functions and selection

It is worth considering different classes of growth functions $z(t)$ and the behavior of long time solutions of the corresponding ODEs. An intimately related problem concerns population dynamics: What is the long time or equilibrium distribution of genotypes in a normalized population, $\lim_{t\to\infty} \mathbf{x}(t)$ provided the initial distribution has been $\mathbf{x}_0$? Is there a universal long time behavior, for example selection, coexistence or cooperation, that is characteristic for certain classes of growth functions?

The differential equation describing unlimited growth,

$$\frac{\mathrm{d}z}{\mathrm{d}t} = f \cdot z^n \tag{2.21}$$

[8]The two situations are called *underdominance* or homozygote advantage and *overdominance* or heterozygote advantage, respectively.

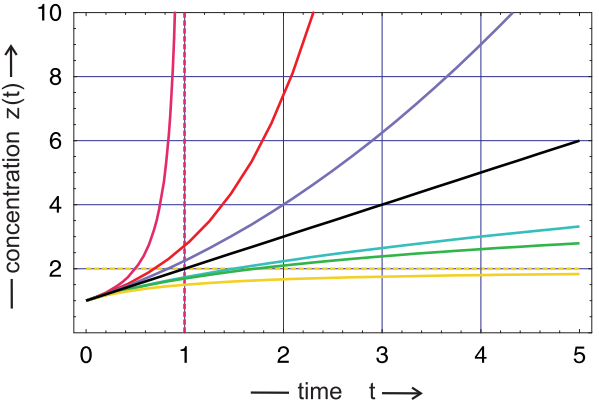

FIGURE 2.2. **Typical functions describing unlimited growth**. All functions are normalized in order to fulfil the conditions $z_0 = 1$ and $\mathrm{d}z/\mathrm{d}t|_{t=0} = 1$. The individual curves show hyperbolic growth ($z(t) = 1/(1-t)$; magenta; the dotted line indicates the position of the instability), exponential growth ($z(t) = \exp(t)$; red), parabolic growth ($z(t) = (1 + t/2)^2$; blue), linear growth ($z(t) = 1 + t$; black), sublinear growth ($z(t) = \sqrt{1+2t}$; turquoise), logarithmic growth ($z(t) = 1 + \log(1+t)$; green), and sublogarithmic growth ($z(t) = 1 + t/(1+t)$; yellow; the dotted line indicates the maximum value $z_{\max}$: $\lim_{t\to\infty} z(t) = z_{\max}$).

yields two types of general solutions for the initial value $z(0) = z_0$

$$z(t) = \left(z_0^{1-n} + (1-n)ft\right)^{1/(1-n)} \qquad \text{for} \quad n \neq 1 \quad \text{and} \tag{2.21a}$$

$$z(t) = z_0 \cdot e^{ft} \qquad \text{for} \quad n = 1\,. \tag{2.21b}$$

In order to make the functions comparable we normalize them in order to fulfil $z_0 = 1$ and $\mathrm{d}z/\mathrm{d}t|_{t=0} = 1$. According to equs. (2.21) this yields $z_0 = 1$ and $f = 1$. The different classes of growth functions indicated by different colors in figure 2.2 are characterized by the following behavior:

(i) Hyperbolic growth requires $n > 1$; for $n = 2$ it yields the solution curve of the $z(t) = 1/(1-t)$. Characteristic is the existence of an instability in the sense that $z(t)$ approaches infinity at some critical time, $\lim_{t\to t_{\mathrm{cr}}} z(t) = \infty$ with $t_{\mathrm{cr}} = 1$. The selection behavior of hyperbolic growth is illustrated by the Schlögl model:[9] $\mathrm{d}z_j/\mathrm{d}t = f_j z_j^2$;

[9]The Schlögl model is tantamount to Fisher's selection equation with diagonal terms only: $f_j = a_{jj}$; $j = 1, 2, \ldots, n$ [37].

$j = 1, 2, \ldots, n$. Depending on the initial conditions each of the replicators $\mathcal{X}_j$ can be selected. $\mathcal{X}_m$ the species with the highest replication parameter, $f_m = \max\{f_i; i = 1, 2, \ldots, n\}$ has the largest basin of attraction and the highest probability to be selected. After selection has occurred a new species $\mathcal{X}_k$ is extremely unlikely to replace the current species $\mathcal{X}_m$ even if its replication parameter is substantially higher, $f_k \gg f_m$. This phenomenon is called *once-for-ever selection*.

(ii) Exponential growth is observed for $n = 1$ and described by the solution $z(t) = e^t$. It represents the most common growth function in biology. The species X_m having the highest replication parameter, $f_m = \max\{f_i; i = 1, 2, \ldots, N\}$, is always selected, $\lim_{t\to\infty} z_m = 1$. Injection of a new species $\mathcal{X}_k$ with a still higher replication parameter, $f_k > f_m$, leads to selection of the fitter variant $\mathcal{X}_k$.

(iii) Parabolic growth occurs for $0 < n < 1$ and for $n = 1/2$ has the solution curve $z(t) = (1 - t/2)^2$. It is observed, for example, in enzyme free replication of oligonucleotides that form a stable duplex, i.e. a complex of one plus and one minus strand [38]. Depending on parameters and concentrations coexistence or selection may occur [39].

(iv) Linear growth follows from $n = 0$ and takes on the form $z(t) = 1 + t$. Linear growth is observed, for example, in replicase catalyzed replication of RNA at enzyme saturation [40].

(v) Sublinear growth occurs for $n < 0$. In particular, for $n = -1$ gives rise to the solution $y(t) = (1 + 2t)^{1/2} = \sqrt{1 + 2t}$.

In addition we mention also two additional forms of weak growth that do not follow from equation (2.21):

(vi) Logarithmic growth that can be expressed by the function $z(t) = z_0 + \ln(1 + ft)$ or $z(1) = 1 + \ln(1 + t)$ after normalization, and

(vii) sublogarithmic growth modeled by the function $z(t) = z_0 + ft/(1 + ft)$ or $z(t) = 1 + t/(1 + t)$ in normalized form.

Hyperbolic growth, parabolic growth, and sublinear growth constitute families of solution curves that are defined by a certain parameter range (see figure 2.2), for example a range of exponents, $n_{\text{low}} < n < n_{\text{high}}$, whereas exponential growth, linear growth and logarithmic growth are critical curves separating zones of characteristic growth behavior: Logarithmic growth separates growth functions approaching infinity in the limit $t \to \infty$, $\lim_{t\to\infty} z(t) = \infty$ from those that remain finite, $\lim_{t\to\infty} z(t) = z_\infty < \infty$, linear growth separates concave from convex growth functions, and exponential growth eventually separates growth functions that reach infinity at finite times from those that don't.

3. Mutation, selection, and optimization

Molecular biology was initiated when Watson and Crick published their centennial paper on the structure of deoxyribonucleic acid (DNA) [41]. Further development provided information on the chemistry of life at a breathtaking pace [42]. A closer look on the structure of DNA revealed the discrete nature of base pairing – two nucleotides make a base pair that fits into the double helix or they don't. With this restriction the natural nucleobases allow for only four pairings: AT, TA, GC, and CG. This fact is already sufficient for an rough understanding of the molecular basis of genetics: Genetic information is of digital nature and multiplication of information is tantamount to copying. Mutation, the process that leads to innovation in evolution, was identified with imperfect reproduction or erroneous copying. Molecular insights into recombination occurring during meiotic cell division[10] provided straightforward explanations for the deviations from Mendel's idealized ratios of offspring with different appearance.

3.1. Complementary replication of nucleic acids

The device for cellular DNA replication is highly involved as the replication complex consists of more than twenty protein enzymes performing a concerted reaction that makes two double stranded molecules from one double stranded molecule. Simpler DNA replication and common RNA replication use the principle of complementary strand template completion: A single strand is completed to a double helix nucleotide after nucleotide (figure 3.1) whereby the complementary strand is synthesized leading to the mechanism for complementary synthesis (in the unlimited growth case):

$$\begin{aligned} (M) + \mathcal{X}^- &\xrightarrow{f^-} \mathcal{X}^+ + \mathcal{X}^- \quad \text{and} \\ (M) + \mathcal{X}^+ &\xrightarrow{f^+} \mathcal{X}^- + \mathcal{X}^+ \\ \text{lead to} \qquad \frac{\mathrm{d}z^+}{\mathrm{d}t} &= f^- z^- \quad \text{and} \quad \frac{\mathrm{d}z^-}{\mathrm{d}t} = f^+ z^+ \,. \end{aligned} \tag{3.1}$$

Transformation of variables,

$$\zeta = \frac{z^+}{\sqrt{f^+}} + \frac{z^-}{\sqrt{f^-}} \quad \text{and} \quad \eta = \frac{z^+}{\sqrt{f^+}} - \frac{z^-}{\sqrt{f^-}} \,,$$

[10] Two different classes of cell divisions are distinguished: (i) *Mitosis* leading from one *diploid* cell to two diploid cells, where diploid defines the fact that the cells carries two sets of chromosomes, and (ii) *meiosis* leading from one diploid cell to four *haploid* cells constituting the germ line. Haploid expresses the fact that the cell carries only one set of chromosomes. Diploidy of the cells is restored when one paternal and one maternal haploid germ line cell fuse to yield a diploid zygote.

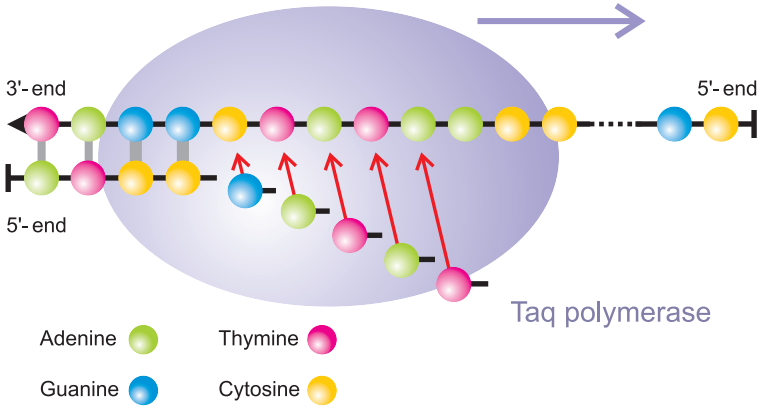

FIGURE 3.1. **Sketch of template induced DNA replication**. Template induced polymerization of nucleic acids (DNA or RNA) follows the same logical principle: The polymerase binds to the 3'-end of the template molecule and synthesizes the complementary strand in the direction from the 5'-end to the 3'-end by adding nucleotide after nucleotide. Considering the incorporation of individual nucleotides as independent reaction steps, the accuracy of correct replication of template $\mathcal{X}_j$ is simply: $Q_{jj} = q_1 \cdot q_2 \cdot \ldots \cdot q_\ell$, where ℓ denotes the chain length of the polynucleotide and q_j is the accuracy of correct incorporation at position "j". An important example of a DNA polymerizing single enzyme is the thermostable DNA polymerase isolated from the bacterium *Thermus aquaticus*. It replicates single stranded DNA and is used commonly for DNA amplification in the polymerase chain reaction (PCR) technique [43].

yields the solutions

$$\eta(t) = \eta(0)\exp(-ft) \quad \text{and} \quad \zeta(t) = \zeta(0)\exp(+ft) \tag{3.1'}$$

with $f = \sqrt{f^+ f^-}$. The combined variable $\eta(t)$ describes the internal equilibration of plus- and minus-strand and $\zeta(t)$ is the growth function of the plus-minus ensemble. It is worth noticing that the fitness value of the ensemble is the geometric mean of the individual fitness values of the two strands, f^+ and f^-.

Complementary replication of DNA is the basis of the polymerase chain reaction (PCR) [43], which is a standard technique in molecular genetics. RNA replication in cells infected by several classes of RNA viruses and

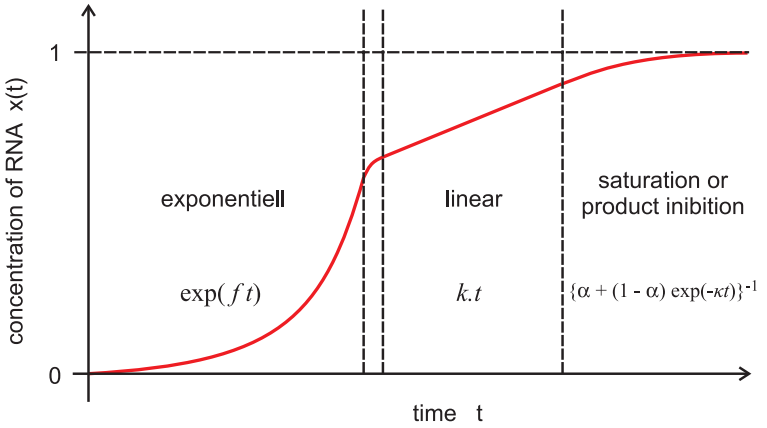

FIGURE 3.2. **Kinetics of RNA replication in a closed system**. The time course of specific RNA replication by Qβ-replicase shows three distinct growth phases: (i) an exponential phase, (ii) a linear phase, and (iii) a phase characterized by saturation through product inhibition [40, 44, 45]. The experiment is initiated by transfer of a very small sample of RNA suitable for replication into a medium containing Qβ-replicase (R) and the activated monomers, ATP, UTP, GTP, and CTP in excess (Consumed materials are not replenished in this experiment). In the phase of exponential growth there is shortage of RNA templates, every free RNA molecule is instantaneously bound to an enzyme molecule and replicated, and the corresponding over-all kinetics follows $\mathrm{d}x/\mathrm{d}t = f \cdot x$ resulting in $x(t) = x_0 \cdot \exp(ft)$. In the linear phase the concentration of template is exceeding that of enzyme, every enzyme molecule in engaged in replication, and over-all kinetics is described by $\mathrm{d}x/\mathrm{d}t = k' \cdot e_0^{(\mathrm{R})} = k$, wherein $e_0^{(\mathrm{R})}$ is the total enzyme concentration, and this yields after integration $x(t) = x_0 + kt$. Further increase in RNA concentration slows down the dissociation of product (and template) RNA from the enzyme-RNA complex and leads to a phenomenon known as product inhibition of the reaction. It can be approximated by a Verhulst-type saturation term: $x(t) = 1 \big/ \big(\alpha + (1 - \alpha) \exp(-\kappa t)\big)$. At the end, all enzyme molecules are blocked by RNA in complexes and no more RNA synthesis is possible, $x(t) \to 1$.

replication in cell free media follow also a complementary replication mechanism and usually involves only one or very few enzymes. The kinetic reaction mechanism of RNA replication *in vitro* has been studied in great detail [40, 44, 45]: Under suitable conditions, excess replicase and nucleotide triphosphates (ATP, UTP, GTP, and CTP), the concentration of the RNA plus-minus ensemble grows exponentially (figure 3.2). The population maintains exponential growth if consumed materials are replenished either by a suitable flow device or by serial transfer, which consists of repeated transfer of small quantities of the current reaction mixture into fresh reaction medium [3]. The condition of exponential growth is trivially fulfilled for cellular life, because replication and cell divisions are strong coupled mechanistically through cellular metabolism, and selection relevant multiplication occurs at the level of cells. With virus replication in the host cell the situation is more involved: Inside the cell replication is exhausting the reservoir of building blocks for the virus specific biomolecules and conditions of exponential growth exist only in the early stages of cellular infection. Formation of virus particles, so-called virions, however, introduces exponential growth and selection of fitter variants.

The repertoire of naturally occurring changes in genomes is very rich and ranges from point mutations, insertions and deletions to duplications of genes and whole genomes or other large scale genome rearrangements. For the sake of simplicity only single point mutations will be considered here. This, however, is sufficient to be able to reach every sequence from every other sequence of the same length through a finite number of mutation steps. Subsection 3.2 introduces the kinetic differential equation for replication and mutation, and derives the solution. Different from selection based on error free replication, optimization of mean fitness $\bar{f} = \phi$ is no longer a global property but restricted to some region on the unit simplex $\mathbb{S}_n^{(1)}$ (subsection 3.3). Correct reproduction and mutation at the molecular level are seen as parallel chemical reactions (figure 3.3). In order to guarantee inheritance, correct copying must occur more frequently than mutation. In subsection 3.4 we shall cast this intuitive statement into a quantitative expression and in subsection 3.5 the analytical results are supplemented by data from numerical computations. Eventually, we consider neutrality in replication (subsection 3.6) and lethal mutations (subsection 3.7).

3.2. The mutation selection equation

Replication leading to correct copies and mutations is properly described by the overall mechanism

$$(\mathsf{M}) + \mathcal{X}_j \xrightarrow{Q_{ij} f_j} \mathcal{X}_i + \mathcal{X}_j \; ; \quad i, j = 1, 2, \ldots, n \; , \tag{3.2}$$

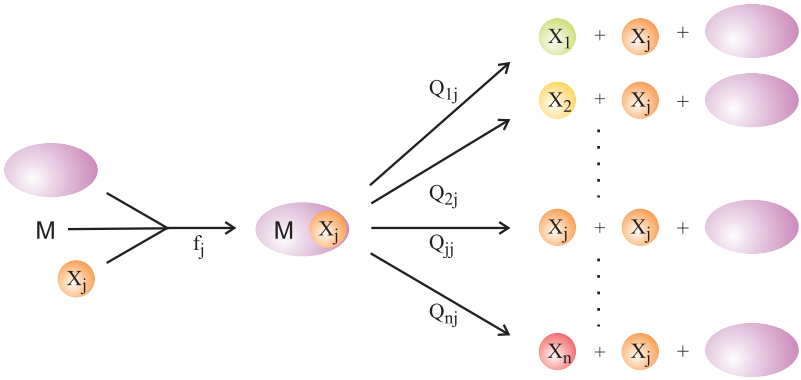

FIGURE 3.3. **A schematic view of replication and mutation**. The replication device (violet) binds the template nucleic acid molecule (DNA or RNA; $\mathcal{X}_j$, orange) and initiates copying of the genetic information with a rate parameter f_j. The reaction has n different channels and yields a correct copy with frequency Q_{jj} or leads through mutation to one of the $(n-1)$ variants $\mathcal{X}_k$ (spectrum of colors) with frequency Q_{kj} whereby $Q_{jj} \gg Q_{kj}\,\forall\, k \neq j$ is required for stable inheritance. Stoichiometry of replication requires $\sum_{i=1}^{n} Q_{ij} = 1$, since the product has to be either correct or incorrect. Replication requires activated monomers M – in nature in form of the deoxynucleoside-triphosphates (dATP, dTTP, dGTP, and dCTP in DNA) or nucleoside-triphosphates (ATP, UTP, GTP, and CTP in RNA) – and suitable reaction conditions.

that is cast by chemical kinetics into the differential equation:[11]

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = \sum_{i=1}^{n} Q_{ji} f_i x_i - \phi(t) x_j \,, \quad j = 1, 2, \ldots, n \qquad \text{with}$$

$$\phi(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} Q_{ji} f_i x_i = \sum_{i=1}^{n} f_i x_i \,, \tag{3.2'}$$

$$\text{or} \quad \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = (\mathrm{Q} \cdot \mathrm{F} - \phi(t))\,\mathbf{x} \quad \text{in vector notation} \,,$$

where $\mathbf{x}$ is an n-dimensional column vector; Q and F are $n \times n$ matrices. The matrix Q contains the mutation probabilities — Q_{ji} referring to the production of $\mathcal{X}_j$ as an error copy of template $\mathcal{X}_i$ — and F is a diagonal matrix

[11] The building blocks M are put in parentheses because they are assumed to be present in excess and therefore do not appear as variables in the ODEs.

whose elements are the replication rate parameters or fitness values f_j. The product $\mathrm{Q} \cdot \mathrm{F} = \mathrm{W}$ is defined as the *value matrix* W, since it encapsulated the selective values of variants.

The form of the mutation selection equation portrays the schematic mechanism shown in figures 3.1 and 3.3: Initiation of replication, propagation along the template $\mathcal{X}_j$, and termination are modeled by an overall rate parameter f_j that accounts also for the constant concentrations of building blocks M. After initiation the enzyme progresses stepwise along the polynucleotide chain, each nucleotide incorporation opens κ reaction channels where κ is the number of different nucleotide bases ($\kappa = 2$ holds for binary sequences and $\kappa = 4$ for natural DNA or RNA molecules) and hence, correct replication and mutation are parallel chemical reactions. One reaction channel incorporates the correct nucleotide whereas $\kappa - 1$ channels produce mutants, the matrix elements Q_{ji} with $j = 1, \ldots, n$ determine the probability to obtain $\mathcal{X}_j$ as a (correct or incorrect) copy of $\mathcal{X}_i$ and by conservation of probabilities we have $\sum_{j=1}^{n} Q_{ji} = 1$.

Equation (3.2) can be transformed into a linear ODE by means of integrating factor transformation and is thereby reduced to the following eigenvalue problem [46, 47]:

$$
\begin{aligned}
\boldsymbol{z}(t) &= \boldsymbol{x}(t) \cdot \exp\left(\int_0^t \phi(\tau)\mathrm{d}\tau\right) \ , \\
\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} &= \mathrm{Q} \cdot \mathrm{F}\,\boldsymbol{z} = \mathrm{W}\,\boldsymbol{z} \ , \quad \text{and} \\
\mathrm{W} &= \mathrm{B} \cdot \Lambda \cdot \mathrm{B}^{-1} \quad \text{or} \quad \Lambda = \mathrm{B}^{-1} \cdot \mathrm{W} \cdot \mathrm{B} \ ,
\end{aligned}
$$

with Λ being a diagonal matrix, whose elements are the ordered eigenvalues of W, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. The calculation of the solutions x_j yields by straightforward insertion:

$$
x_j(t) = \frac{\sum_{k=1}^{n} b_{jk} \sum_{i=1}^{n} h_{ki} x_i(0) \exp(\lambda_k t)}{\sum_{l=1}^{n} \sum_{k=1}^{n} b_{lk} \sum_{i=1}^{n} h_{ki} x_i(0) \exp(\lambda_k\, t)} \ , \ j = 1, 2, \ldots, n \ . \qquad (3.3)
$$

The new quantities in this equation are the elements of the two transformation matrices:

$$
\begin{aligned}
\mathrm{B} &= \{b_{jk}; j, k = 1, \ldots, n\} \ \text{ and} \\
\mathrm{B}^{-1} &= \{h_{kj}; k, j = 1, 2, \ldots, n\}
\end{aligned}
$$

The columns of B and the rows of B^{-1} represent the right hand and left hand eigenvectors of the matrix W. For example we have

$$\boldsymbol{\zeta}_1 = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix} .$$

Assuming a unique largest eigenvalue, $\lambda_1 > \lambda_2 \geq \lambda_3 \ldots \geq \lambda_n$ (see theorem 3.1), the stationary solution contains only the contributions of the largest eigenvector, $\boldsymbol{\zeta}_1$:

$$\lim_{t\to\infty} x_j(t) = \bar{x}_j = \frac{b_{j1}\sum_{i=1}^{n} h_{1i}x_i(0)}{\sum_{k=1}^{n} b_{k1}\sum_{i=1}^{n} h_{1i}x_i(0)} , \; j = 1, \ldots, n . \tag{3.4}$$

In other words, $\boldsymbol{\zeta}_1$ describes the stationary distribution of mutants and represents the genetic reservoir of an asexually reproducing species similarly to the gene pool of a sexual species. For this reason $\boldsymbol{\zeta}_1$ has be called *quasispecies*, its properties will be discussed now in subsection 3.3.

3.3. Mutation and optimization

Before considering the optimization problem we shall derive the biologically relevant properties of quasispecies.

Theorem 3.1 (Unique and strictly positive quasispecies). *If all single point mutations yield non-lethal variants, the largest eigenvector of matrix* W*, the quasispecies* $\boldsymbol{\zeta}_1$*, will be unique and all elements of* $\boldsymbol{\zeta}_1$ *will be strictly positive.*

Proof. The matrix F is a diagonal matrix with strictly positive elements. The matrix Q contains ℓ strictly positive diagonal elements representing the replication accuracies Q_{jj} $(j = 1, \ldots, \ell)$ and strictly positive elements for all single point mutations that can be reached from some specific (initial) sequence $\mathcal{X}_0$ by one replication event. These are all $(\kappa - 1)\ell$ sequences $\mathcal{X}_k$ with Hamming distance[12] $d_{\mathrm{H}}(\mathcal{X}_k, \mathcal{X}_0) = d^{\mathrm{H}}_{k0} = 1$ and accordingly, Q has $\ell(2\kappa - 1)$ strictly positive and $\ell(\ell - 2\kappa + 1)$ zero entries. Since multiplication with a diagonal matrix leaves zero entries unchanged, $\mathrm{W} = \mathrm{Q}\cdot\mathrm{F}$ has the same zero elements as Q. Cumulative consecutive replication events as described by the mutation matrices Q^2, Q^3, ..., provide nonzero mutation frequencies for all sequences $\mathcal{X}_k$ of Hamming distances $d^{\mathrm{H}}_{k0} = 2, 3, \ldots$, and eventually after ℓ replications the matrix Q^ℓ is strictly positive: Q is a primitive matrix, and so is W.

Accordingly, Perron-Frobenius theorem [49, pp.3-11] for primitive matrices applies for matrix W and there exists an largest eigenvalue λ_1 such that

[12] The Hamming distance counts the number of positions in which two aligned sequences differ[48].

(i) λ_1 is real and strictly positive,
(ii) associated with λ_1 are one strictly positive left and one strictly positive right eigenvector, $\tilde{\boldsymbol{\zeta}}_1$ and $\boldsymbol{\zeta}_1$,
(iii) the eigenvectors associated with λ_1 are unique to constant multiples,
(iv) $\lambda_1 > |\lambda_k|$ holds for all $\lambda_k \neq \lambda_1$, and
(v) λ_1 is a simple root of the the characteristic equation of W.
Items (ii), (iii), (iv), and (v) meet the conjectures of the theorem. □

In other words, the quasispecies theorem states that the mutation selection problem as modeled by equation (3.2) has a unique long time or equilibrium solution, the quasispecies, and no variant vanishes in the long time limit. Both conditions root the model safely in chemical kinetics and thermodynamics.

The eigenvectors belonging to the eigenvalues λ_k with $k \neq 1$ have no direct physical meaning but they are important for an understanding of optimization in the presence of mutation. For the purpose of illustration we consider $\mathbb{L}^1$ normalized variables on the unit simplex $\mathbb{S}_n^{(1)}$ and assume that the eigenvalues of W are real.[13] A point on $\mathbb{S}_n^{(1)}$ is defined by the vector $\mathbf{x}$, likewise we define a simplex $\boldsymbol{\Sigma}_n^{(1)}$ in the eigenspace of W and its points by the vectors $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$. The unit vectors $\mathbf{P}_j = (x_j = 1, x_i = 0 \,\forall\, i \neq j)$ in concentration space are replaced by unit vectors $\boldsymbol{\Xi}_k = (\xi_k = 1, \xi_i = 0 \,\forall\, i \neq k)$.

Definition 3.2. The **eigensimplex** $\boldsymbol{\Sigma}_n^{(1)}$ is a unit simplex of unit vectors $\boldsymbol{\Xi}_k$, $k = 1, \ldots, n$ in $\mathbb{R}^n$ spanned by the eigenvectors $\boldsymbol{\zeta}_k$ of the value matrix W: $\boldsymbol{\Sigma}_n^{(1)} = \{0 \leq \xi_i \leq 1 \,\forall\, i = 1, \ldots, n \wedge \sum_{i=1}^n \xi_i = 1\}$.

Theorem 3.3 (Restricted optimization of average fitness). *The average fitness* $\phi(t) = \bar{f} = \sum_{i=1}^n f_i x_i$ *with* $x_i \geq 0$ *and* $\sum_{i=1}^n x_i = 1$ *in a population described by the mutation selection equation* (3.2) *is nondecreasing on the intersection* $\mathbb{S}_n^{(1)} \cup \boldsymbol{\Sigma}_n^{(1)}$*, denoted as* optimization cone.

Proof. Points on the simplex $\mathbb{S}_n^{(1)}$ fulfil the conditions $x_i \geq 0$ and $\sum_{i=1}^n x_i = 1$. Equ. 3.2 transformed by diagonalization of $\mathrm{W} = \mathrm{B} \cdot \Lambda \cdot \mathrm{B}^{-1}$ becomes

$$\frac{\mathrm{d}\boldsymbol{\zeta}_k}{\mathrm{d}t} = \boldsymbol{\zeta}_k \Big(\lambda_k - \phi(t) \Big) \,, \;\; k = 1, \ldots, n \;,$$

which is identical to equ. (2.14) and hence, theorem 2.3 holds on the simplex $\boldsymbol{\Sigma}_n^{(1)}$. Both conditions are fulfilled on the intersection of the two simplices. □

The optimization principle does not hold outside the optimization cone, $\mathbb{S}_n^{(1)} \cup \boldsymbol{\Sigma}_n^{(1)}$. It can be shown that $\phi(t)$ is nonincreasing in the cone defined by $\boldsymbol{\Theta} = \{\xi_1 \geq 1, \xi_i \leq 0 \,\forall\, j = 2, \ldots, n\}$ [50]. This fact can be easily made

[13] The matrix W may have complex conjugate pairs of eigenvalues without violating Perron-Frobenius theorem. Special cases of W with complex eigenvalues can be constructed but are not compatible with realistic assumptions on mutation frequencies Q_{ji}.

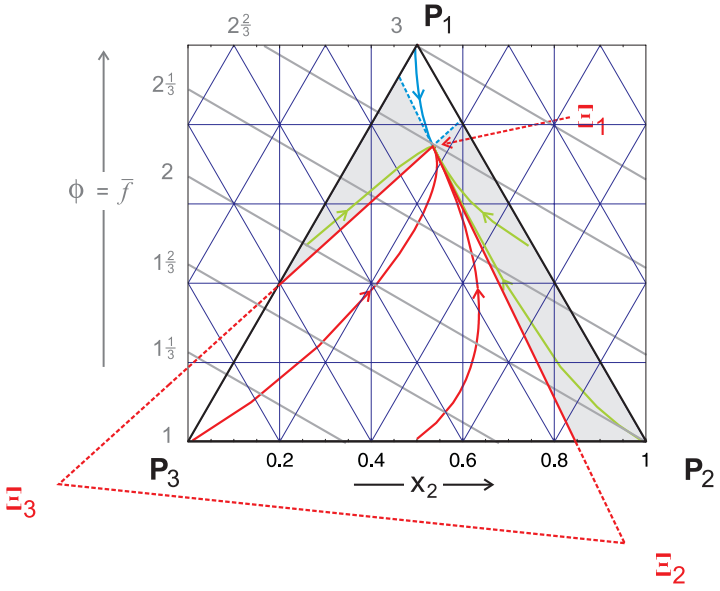

FIGURE 3.4. **The quasispecies on the unit simplex**. Shown is the case of three variables (x_1, x_2, x_3) on $\mathbb{S}_3^{(1)}$. The point representing pure quasispecies $\boldsymbol{\Xi}_1$, is shown together with the points for the other two eigenvectors, $\boldsymbol{\Xi}_2$ and $\boldsymbol{\Xi}_3$. The simplex is partitioned into an *optimization cone* ($\mathbb{S}_n^{(1)} \cup \boldsymbol{\Sigma}_n^{(1)}$; lower white area) where the mean replication rate $\bar{f}(t)$ is nondecreasing, and three other zones where $\bar{f}(t)$ may also decrease. Here, $\mathcal{X}_1$ is chosen to be the master sequence, the sequence with the highest fitness value. Solution curves are presented as parametric plots $\mathbf{x}(t)$. The mean replication rate $\bar{f}(t)$ is monotonously increasing along red trajectories, monotonously decreasing along the blue trajectory (upper white area), and not necessarily monotonous along green trajectories (grey areas). The parameter values are: $f_1 = 2.1\,[\mathrm{t}^{-1}]$, $f_2 = 2.0\,[\mathrm{t}^{-1}]$, and $f_3 = 1.9\,[\mathrm{t}^{-1}]$, the Q-matrix was assumed to be bistochastic with the elements $Q_{ii} = 0.98$ and $Q_{ij} = 0.01$ for $i, j = \{1, 2, 3\}$. The eigenvalues and eigenvectors of W are:

k	λ_k	b_{1k}	b_{2k}	b_{3k}
1	2.065	0.742	0.165	0.093
2	1.958	-0.248	1.078	0.170
3	1.857	-0.103	-0.224	1.327

plausible by assuming as initial condition a homogenous population of the *master sequence*, the sequence corresponding to the fittest phenotype. All variants have lower fitness and accordingly, mutation leads to a reduction in the mean fitness $\bar{f}$, which is decreasing as the quasispecies with $\bar{f} = \lambda_1 < f_1$ is approached in the long time limit. In the remaining sections of the unit simplex no predictions on $\phi(t)$ can be made, special cases with non-monotonous behavior can be constructed [50]. The somewhat sophisticated conditions of optimization in the replication mutation system are illustrated by means of a numerical example in figure 3.4.

3.4. Mutation rates and error thresholds

In general, the mutation rates are not tunable but they can be varied within certain limits by applying suitable experimental assays. In order to illustrate the mutation rate dependence of quasispecies and to subject it to mathematical analysis, a simplifying model called *uniform error rate* model is adopted [51]. The error rate per nucleotide and replication, p, is assumed to be independent of the position and the nature of the nucleotide exchange: A→U, A→G or A→C occur with the same frequency p and the total error rate at a given position is $3p$. Then the elements of the mutation matrix Q depend only on three quantities: The chain length of the sequence to be replicated, ℓ, the error frequency p and the Hamming distance between the template, $\mathcal{X}_i$, and the newly synthesized sequence, $\mathcal{X}_j$, denoted by d_{ij}^{H},

$$\begin{aligned} Q_{ji} &= \bigl(1-(\kappa-1)p\bigr)^{\ell-d_{ij}^{\mathrm{H}}}\cdot p^{d_{ij}^{\mathrm{H}}} = \bigl(1-(\kappa-1)p\bigr)^{\ell}\,\varepsilon^{d_{ij}^{\mathrm{H}}} \quad \text{with} \\ \varepsilon &= \frac{p}{1-(\kappa-1)p}\ , \end{aligned} \tag{3.5}$$

with κ the size of the nucleotide alphabet ($\kappa = 4$ for natural polynucleotides corresponding to $\{\mathsf{A}, \mathsf{U}(\mathsf{T}), \mathsf{G}, \mathsf{C}\}$). The explanation of equ. (3.5) follows from figure 3.1: The two sequences differ in d_{ij}^{H} positions and hence $\ell - d_{ij}^{\mathrm{H}}$ nucleotides have to be copied correctly, each one contributing a factor $1-(\kappa-1)p$, and d_{ij}^{H} errors with frequency p have to be made at certain positions. Since the Hamming distance is a metric, we have $d_{ij}^{\mathrm{H}} = d_{ji}^{\mathrm{H}}$, and within the approximation of the uniform error rate model the mutation matrix Q is symmetric.

For $p = 0$ we encounter the selection case (2.15): The species of highest fitness, the master sequence $\mathcal{X}_1$, is selected and all other variants disappear in the long time limit. The other extreme is random replication where correct and incorrect incorporations of digits are equally probable and occur with frequency $\tilde{p} = \kappa^{-1}$. Then all elements of matrix Q are equal to $\kappa^{-\ell}$, and the uniform distribution $\boldsymbol{\Upsilon} = \{\bar{x}_j = n^{-1}\,\forall\, j = 1, 2, \ldots, n \text{ with } n = \kappa^{\ell}\}$ is the eigenvector corresponding to the largest eigenvalue $\lambda_1 = \kappa^{-\ell}\sum_{i=1}^{n} f_i$ (all other eigenvalues of W vanish). In the whole range $0 \le p \le \kappa^{-1}$ the stationary

distribution changes from the homogeneous population, $\boldsymbol{\Xi}_1 = \{\bar{x}_1 = 1, \bar{x}_j = 0 \,\forall\, j = 2, \ldots, n\}$ to the uniform distribution $\boldsymbol{\Upsilon}$.

Between the two extremes the function $\bar{x}_1(p)$ was approximated by Manfred Eigen through neglect of back-flow from mutants to the master sequence. He obtained for $\mathrm{d}x_1/\mathrm{d}t = 0$ [51]:

$$\bar{x}_1 = \frac{Q_{11} f_1 - \bar{f}_{-1}}{f_1 - \bar{f}_{-1}} = \frac{Q_{11} - \sigma_1^{-1}}{1 - \sigma_1^{-1}} \quad \text{with} \quad \bar{f}_{-1} = \frac{\sum_{i=1}^{n} f_i \bar{x}_i}{1 - \bar{x}_1} \; . \tag{3.6}$$

The quantity $\sigma_1 = f_1/\bar{f}_{-1}$ is denoted as the *superiority* of the master sequence $\mathcal{X}_1$. In this rough, zeroth order approximation the frequency of the master sequence becomes zero at a critical value of the mutation rate parameter, $p_{\max}$, for constant chain length ℓ or at a maximal chain length $\ell_{\max}$ for constant replication accuracy p,

$$p_{\max} \;\approx\; \frac{\ln \sigma_1}{(\kappa - 1)\,\ell} \quad \text{or} \quad \ell_{\max} \;\approx\; \frac{\ln \sigma_1}{(\kappa - 1)\,p} \; ,$$

respectively. The critical replication accuracy has been characterized as the *error threshold* of replication. Numerically computed error thresholds remind of a phase transition in which the quasispecies changes from a mutant distribution centered around a master sequence to the uniform distribution. In other words the solution that becomes exact at $p = \tilde{p}$ is closely approached at $p_{\max}$ already. For the purpose of illustration for a superiority of $\sigma_1 = 1.1$ and a chain length of $\ell = 100$ we obtain $p_{\max} = 0.00032$ compared to $\tilde{p} = 0.5$.

Both relations for the error threshold, the maximum replication accuracy and the maximum chain length, were found to have practical implications: (i) RNA viruses replicate at mutation rates close to the maximal value [52]. A novel concept for the development of antiviral drugs makes use of this fact and aims at driving the virus population to mutation rates above the error threshold [53].
(ii) There is a limit in chain length for faithful replication that depends on the replication machinery: The accuracy limit of enzyme-free replication is around one error in one hundred nucleotides, RNA viruses with a single enzyme and no proof reading can hardly exceed accuracies of one error in 10 000 nucleotides, and DNA replication with repair on the fly reaches one error in 10^8 nucleotides. For prokaryotic DNA replication with post-replication repair the accuracy increases to 10^{-9}–10^{-10}, which is roughly one mutation in 300 duplications of bacterial cells, and for eukaryotes similar fractions of mutations were reported despite much longer genomes [54].

3.5. Numerical computations on error thresholds

The approximation of the error threshold through neglect of mutational backflow (3.6) caused the results to be largely independent of the distribution of the fitness values of mutants, since only the mean fitness, $\bar{f}_{-m}$, enters the

expressions, in contrast to numerical computations, which suggested that the appearance of an error threshold and its shape are strongly dependent on details of the fitness landscape (see [55] and [21, pp.51-60]). The influence of the distribution of fitness values will be considered in two steps:
(i) different fitness values are applied for error classes being sets of sequences with the same Hamming distances from the master sequence $\mathcal{X}_1$,

$$\begin{aligned} \mathcal{Y}_k &= \{\mathcal{X}_j \,|\, d_{j1}^{\mathrm{H}} = k\} \ , \\ y_k &= \sum\nolimits_j x_{j\,|\,\mathcal{X}_j \in \mathcal{Y}_k} \quad \text{and} \end{aligned} \tag{3.7}$$

(ii) different fitness values are assigned to individual sequences. In the first case all sequences $\mathcal{X}_j$ with Hamming distance $d_{1,j}^{H} = k$ fall into the *error class* 'k'. Although the assumption that all sequences in a given error class have identical fitness is not justified on the basis of molecular data, it is frequently used in population genetics. Calculations applying this assumption turned out to be useful for an understanding of the threshold phenomenon.

Five model landscapes, which are characterized by their fitness matrices $\mathrm{F} = \{F_{ij} = f_i \cdot \delta_{ij}\}$, were applied here:

(i) the single-peak landscape: $f(\mathcal{Y}_0) = f_0$ and $f(\mathcal{Y}_j) = f_n \,\forall\, j \neq 1$,
(ii) the hyperbolic landscape: $f(\mathcal{Y}_j) = f_0 - (f_0 - f_n)(n+1)j/\big(n(j+1)\big) \,\forall\, j$,
(iii) the step-linear landscape: $f(\mathcal{Y}_j) = f_0 - (f_0 - f_n)j/k \,\forall\, j = 0, \ldots, k$ and $f(\mathcal{Y}_j) = f_n \,\forall\, j = k+1, \ldots, n$,
(iv) the multiplicative landscape: $f(\mathcal{Y}_j) = f_0(f_n/f_0)^{j/n} \,\forall\, j$, and
(v) the additive or linear landscape: $f(\mathcal{Y}_j) = f_0 - (f_0 - f_n)j/n \,\forall\, j$.

Examples for the dependence of the quasispecies distribution on the error rate, $\boldsymbol{\mathcal{Y}}(p)$, on different landscapes are shown in figure 3.5.

The analysis of error thresholds on different landscapes revealed three separable features: (i) a steep decay in the frequency of the master sequence – $x_1(p) \to 0$ in the zeroth order approximation (3.6), (ii) a phase transition-like sharp change in the mutant distribution, and (iii) the transition leads from the quasispecies to the uniform distribution. All three phenomena coincide on the single-peak landscape. Characteristic for most hyperbolic landscapes is a steep decay of the master sequence (i) and an abrupt transition in the distribution of sequences according to (ii) but – in contrast to the single-peak landscape – the transition does not lead to the uniform distribution but to another distribution that changes gradually into the uniform distribution, which becomes the exact solution at the point $p = \tilde{p}$. The step-linear landscape illustrates the separation of the decay range (i) and the phase transition to the uniform distribution (ii and iii). In particular, the phase transition point $p_{\max}$ shifts towards higher values of p when the position of the step (at error class k) moves towards higher error-classes, whereas the decay of the master sequence ($\mathcal{X}_1$) moves in opposite direction. The additive and the multiplicative landscape, the two landscapes that are often used in population genetics,

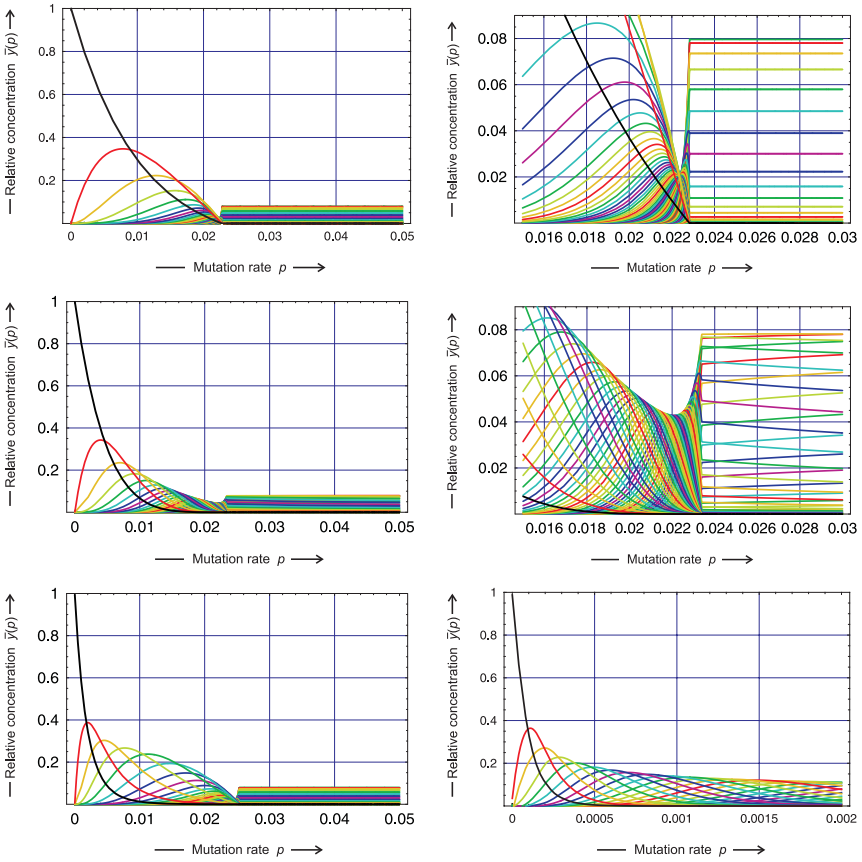

FIGURE 3.5. **Error thresholds on model landscapes**. Relative stationary concentrations of entire error classes $\bar{y}_k(p)$ are plotted as functions of the mutation rate p (Different error classes $\mathcal{Y}_k$ are color coded, $k = 0$: black, $k = 1$: red, $k = 2$: yellow, $k = 3$: chartreuse, k=4: green, etc). Top row: single-peak fitness landscape (enlarged on the right hand side), conditions (i), (ii), and (iii) coincide. Middle row: hyperbolic landscape (enlarged on the right hand side), the phase transition leads to a distribution that changes gradually into the uniform distribution, (i) has an offset to the left of (ii). Bottom row: step-linear landscape on the l.h.s. meets condition (ii) and (iii) but (i) has a large offset to the left, and additive landscape on the r.h.s. does not sustain an error threshold at all. Parameters used in the calculations: $\ell = 100$, $f_1 = f_0 = 10$, $f_n = 1$ (except the hyperbolic landscape where we used $f_n = 0.9091$ in order to have $\bar{f}_{-m} = 1$ as for the single peak landscape), and $k = 5$ for the step-linear landscape.

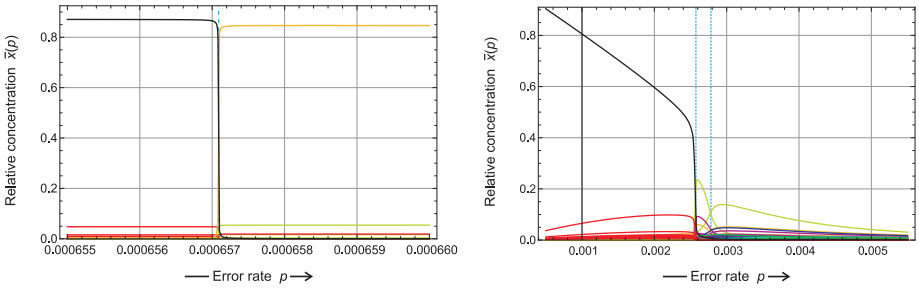

FIGURE 3.6. **Transitions between quasispecies**. Both figures show transitions on a landscape defined by equ. (3.8) for 1024 binary sequences of chain length $\ell = 10$. The transition on the l.h.s. involves the sequences **0** $\rightarrow$ **1003** (decimal equivalent of the binary sequence of chain length $\ell = 10$) with $d^{\mathrm{H}}_{\mathbf{0},\mathbf{1003}} = 8$ and is extremely sharp. The plot on the r.h.s shows to more transitions **0** $\rightarrow$ **923** and **923** $\rightarrow$ **247** ($d^{\mathrm{H}}_{\mathbf{0},\mathbf{923}} = 7$, $d^{\mathrm{H}}_{\mathbf{923},\mathbf{247}} = 6$). Parameters: $f_0 = 1.1$, $f_n = 1.0$, and the seed $s = 229$ for the random number generator *legacy* of *Mathematica*. The band width was chosen $d = 1.0$ (l.h.s.) and $d = 0.995$ (r.h.s.). The fitness values are: $f_{\mathbf{0}} = 1.1$, $f_{\mathbf{1003}} = 1.09999$ and 1.09949 for the two bandwidth, $f_{\mathbf{923}} = 1.09921$, and $f_{\mathbf{247}} = 1.09834$ for $d = 0.995$, respectively.

do not sustain threshold-behavior. On these two landscapes the quasispecies is transformed smoothly with increasing p into the uniform distribution.

Realistic fitness landscapes as derived from the properties of biomolecules are characterized by two features: (i) ruggedness and (ii) neutrality.[14] Error thresholds on realistic fitness landscapes can be modeled by the assumption of a randomly scattered distribution of fitness values within a given band of width d for all sequences except the master sequence:

$$f(\mathcal{X}_j) = \bar{f}_n + 2d(f_0 - f_n)\left(\eta_{\mathrm{rnd}}(j) - 0.5\right),\ j = 2, \ldots, \kappa^\ell\ . \qquad (3.8)$$

In this expression "$\eta_{\mathrm{rnd}}(j)$" is a random number drawn from a random number generator with uniform distribution of numbers in the range $0 \leq \eta_{\mathrm{rnd}}(j) \leq 1$ with j being the index of the consecutive calls of the random function, and d is the band width of fitness values. The two limiting cases

[14] *Ruggedness* implies that nearby sequences may lead to identical but also to very different structures. By the same token functions like fitness values may be the same or very different for close by lying genotypes. *Neutrality* means that a certain fraction of different genotypes measured by the degree of neutrality, λ, have properties that cannot be distinguished by selection.

are: the full band width $d = 1$, where individual fitness values may be as large as f_0, the value for the master sequence, and the single peak landscape with $d = 0$. The computational capacities of today allow for studies of error thresholds at the resolution of individual sequences up to chain lengths $n = 10$. Further increase in computational power raises expectation to be able to reach $n = 20$, which in case of binary sequences is tantamount to the diagonalization of $10^6 \times 10^6$ matrices.

Transitions between two quasispecies were found first by computational analysis [50]: The master sequence on quasispecies I with a larger fitness value, $f_1^{(I)} > f_1^{(II)}$, has a smaller mutational backflow than quasispecies II, $\sum_{i=2}^{n} Q_{1i}^{(I)} x_i^{(I)} < \sum_{k=2}^{n} Q_{1k}^{(II)} x_k^{(II)}$. Since the mutational backflow increases with increasing error rate p, there may exist a value $p = p_{\text{trans}} < p_{\max} \ll \tilde{p}$ at which the two quasispecies exchange stability. The sharpness of the transition depends primarily on the Hamming distance between the two master sequences, $d^{\text{H}}(\mathcal{X}_1^{(I)}, \mathcal{X}_1^{(II)}) = d_{I,II}^{\text{H}}$: The larger the distance the sharper is the transition – numerical computations with binary sequences of chain lengths $\ell = 10$ have shown that a Hamming distance of $d^{\text{H}} > 6$ is sufficient for hard transitions. Several transitions of this kind may occur before the error threshold. In figure 3.6 we show an example of a fitness landscape that sustains three transitions between four quasispecies. A sharp transition with Hamming distance $d_{I,II}^{\text{H}} = 8$ between the master sequences is presented on the l.h.s. of the figure. Two more consecutive transitions are shown on the l.h.s: A sharper one with $d_{II,III}^{\text{H}} = 7$ and a smoother one with $d_{III,IV}^{\text{H}} = 6$. A large mutational backflow requires a small difference in fitness between the master sequence and its mutants and hence, the transitions with increasing p between quasispecies lead from steeper to flatter region on the fitness landscape. Examples are common and constructed straightforwardly. Thirteen years later the phenomenon has been rediscovered with digital organisms and named *survival of the flattest* [56].

Two more questions are important in the context of fitness values at a resolution of individual sequences: (i) How does the dispersion of fitness values expressed by the band width d change the characteristics of the error threshold and (ii) what happens if two more sequences have the same maximal fitness value $f_1 = f_0$. The answer to question (i) follows readily from the computed results (figure 3.7): The position at which the frequency of the master sequence decays to zero, $x_1(p) \to 0$, migrates towards smaller f-values with increasing band width d. This observation agrees well with the expectations, since the fitness value closest to f_1 increases for broader bands (at constant $\bar{f}_{-1}$). In other words, the gap in fitness values between the fittest variant and the fittest but one variant becomes smaller. The scatter of fitness values at the same time broadens the band of curves for the sequences

belonging to one error class. Both phenomena are easily recognized in the three plots on the l.h.s of figure 3.7.

3.6. Neutral variants

Degeneracy of fitness values occurs when two or more genotypes have the same fitness and this is commonly denoted as *neutrality* in biology. An investigation of the role of neutrality requires an extension of equ. (3.8). A certain fraction of sequences, expressed by the degree of neutrality λ, is assumed to have the highest fitness value f_0 and the fitness values of the remaining fraction $1-\lambda$ are assigned as in the non-neutral case (3.8). This random choice of neutral sequences together with a random dispersion of the other fitness values yields an interesting result: Random selection in the sense of Motoo Kimura's neutral theory of evolution [57] occurs only for sufficiently distant fittest sequences. The case of vanishing mutation rate, $\lim p \to 0$, has been studied analytically [50] for two neutral genotypes, $\mathcal{X}_j$ and $\mathcal{X}_k$ and different Hamming distance d_{jk}^{H}. The exact results are

(i) $d_{jk}^{\mathrm{H}} = 1$: $\lim_{p\to 0} \frac{\bar{x}_j}{\bar{x}_k} = 1$ or $\lim_{p\to 0} \bar{x}_j = \lim_{p\to 0} = \bar{x}_k = 0.5$,
(ii) $d_{jk}^{\mathrm{H}} = 2$: $\lim_{p\to 0} \frac{\bar{x}_j}{\bar{x}_k} = \alpha$ or $\lim_{p\to 0} \bar{x}_j = \alpha/(1-\alpha)$, $\lim_{p\to 0} \bar{x}_k = 1/(1-\alpha)$, with some value α and
(iii) $d_{jk}^{\mathrm{H}} \geq 3$: $\lim_{p\to 0} \bar{x}_1 = 1$, $\lim_{p\to 0} \bar{x}_2 = 0$ or $\lim_{p\to 0} \bar{x}_1 = 0$, $\lim_{p\to 0} \bar{x}_2 = 1$.

In full agreement with the exact result [50] we find that two fittest sequences of Hamming distance $d^{\mathrm{H}} = 1$ – being two nearest neighbors in sequence space – are selected as a strongly coupled pair with equal frequency of both members. Numerical results demonstrate that this strong coupling occurs not only for small mutation rates but extends over the whole range of p-values from $p = 0$ to the error threshold $p = p_{\max}$ (figure 3.7). For clusters of more than two nearest neighbor sequences, the frequencies of the individual members of the cluster can be obtained from the largest eigenvector of the adjacency matrix. Pairs of fittest sequences with Hamming distance $d^{\mathrm{H}} = 2$ – being two next nearest neighbors with two sequences in between – are also selected together but the ratio of the two frequencies is different from one. Again strong coupling extends from zero mutation rates up to the error threshold $p = p_{\max}$. For fittest sequences with $d^{\mathrm{H}} \geq 3$ random selection chooses one sequence arbitrarily and eliminates all others as predicted by the Kimura's neutral theory of evolution. An example of a network of seven strongly coupled sequences is shown in figure 3.8. The stationary frequencies of the seven sequences are readily obtained from the largest eigenvector and have the simple form: $\bar{x}_{\mathrm{inner}} = 2\bar{x}_{\mathrm{outer}}$ where by the three inner sequences (black) and the four outer sequences have identical frequencies. Deviation from the exact result for $p \to 0$ can be seen only in the enlargement (plot at the bottom) and are small indeed. Strong coupling of fittest sequences manifests itself, for example, in virology in form of systematic deviations

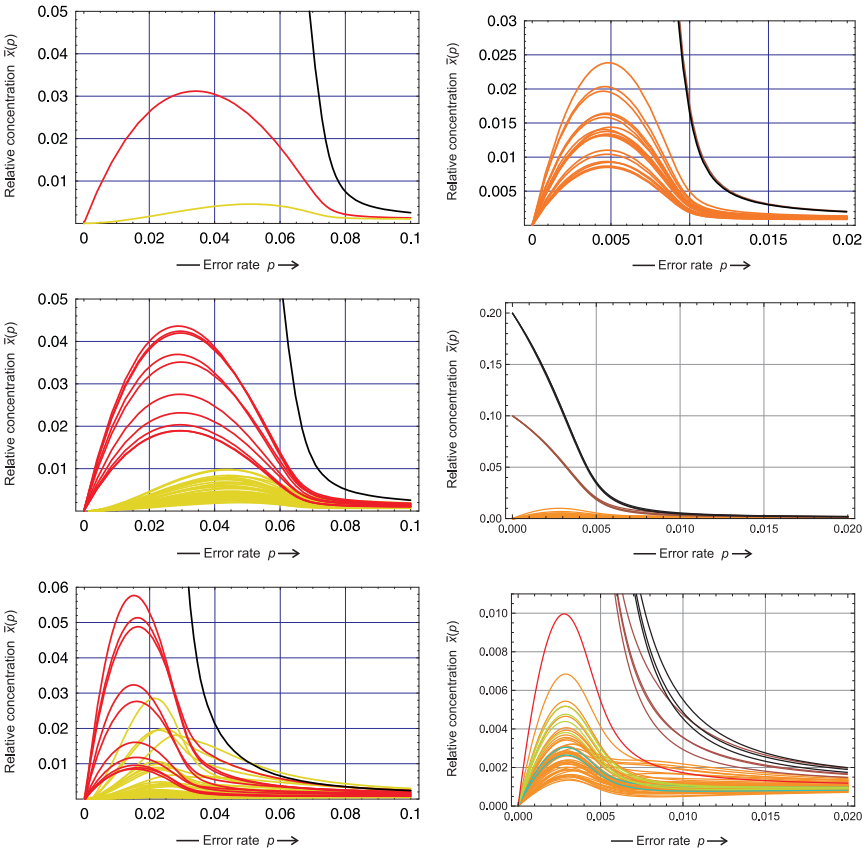

FIGURE 3.7. **Error thresholds on realistic landscapes**. The plots on the l.h.s. show the error threshold for a landscape defined by equ. (3.8) for 1024 binary sequences of chain length $\ell = 10$. With increasing d (top: $d = 0$, single peak landscape, middle: $d = 0.5$, and bottom: $d = 0.925$) the scatter of sequences within the same error class increases and the error threshold defined by sharp decline of the master sequence, $\bar{x}_1(p) \to 0$ migrates towards smaller error rates. The plots on the r.h.s. refers to landscapes with neutrality. The topmost plot shows the frequencies of the master sequences in a neutral network of two nearest neighbors, $\mathcal{X}_{\mathbf{0}}$ (black) and $\mathcal{X}_{\mathbf{6}4}$ (brown), with the 18 one-error mutants surrounding the two sequences (orange). The other plots refer to a network of seven nearest neighbors (figure 3.8; three inner sequences, $\mathcal{X}_{\mathbf{2}48}$, $\mathcal{X}_{\mathbf{7}60}$ and $\mathcal{X}_{\mathbf{7}28}$ (black), and four outer sequences, $\mathcal{X}_{\mathbf{1}84}$, $\mathcal{X}_{\mathbf{5}04}$, $\mathcal{X}_{\mathbf{6}00}$ and $\mathcal{X}_{\mathbf{7}29}$ (brown)). Strong coupling pertains until the population reaches the error threshold. The spectrum of one error mutants is shown in the plot at the bottom. Parameters, l.h.s: $f_0 = 2, f_n = 1, s = 491$, and r.h.s: $f_0 = 1.1, f_n = 1.0, d = 0.5$, topmost plot: $\lambda = 0.01, s = 367$, middle and bottom: $\lambda = 0.1, s = 229$, color code see figure 3.8.

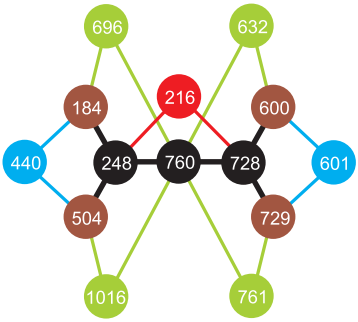

FIGURE 3.8. **A neutral network of replicating genotypes.** Shown is the network of seven neutral sequences being nearest neighbors (black and brown). Seven out of the 51 one error mutants of the neutral sequences occupy special positions that are coupled to two members of the network. As seen in the plot in figure 3.7 (r.h.s., bottom) only $\mathcal{X}_{\mathbf{216}}$ (red) has substantially higher frequency and the curves of the other six special sequences are embedded in the band of the one error class.

from consensus sequences of populations as they are indeed found in nature through systematic sequencing of populations.

3.7. Lethal mutants

Lethal mutants are mutants with zero fitness, $f = 0$, which are unable to replicate and which may lead to extinction of the population. The model that has been used to investigate mutation and selection at constant population size is not applicable as concentrations may go to zero and a more detailed physical setup is required for the description. Different setups were discussed and analyzed rigorously before [30] and the existence of an *extinction threshold* for autocatalytic processes like replication reactions was proved (see also [21, pp.18-27]). We choose here a continuously stirred flowreactor (CSTR) as an appropriate device: An influx of reactants, for example monomeric building blocks M at concentration m_0 and polymerase to compensate for enzyme degradation in case of RNA replication, is installed to replace consumed materials, the reactor is well stirred in order to guarantee instantaneous mixture of the reactor content, the influx is compensated by an unspecific outflux, and the flow rate r is assumed to be tunable. The mechanism for replication and mutation in the flowreactor with $\mathcal{X}_1$ being the only replicating variant is of

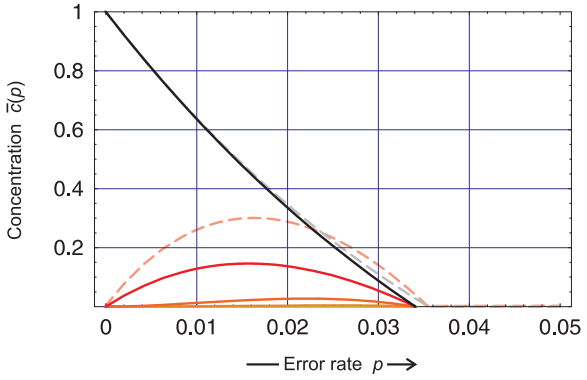

FIGURE 3.9. **Lethal mutants and replication errors.** The model for lethal mutants corresponding to a single peak landscape with $f_1 = 2$ and $f_2 = \ldots = f_n = 0$ is studied in the flowreactor. The concentrations of the master sequence (black) and the mutant classes (red, dark orange, light orange, etc.; full lines) are shown as functions of the error rate p. The parameters were chosen to be $\ell = 20$, $m_0 = 2$, and $\eta = 2$. The plots are compared to the curves for the master sequence and the one error class (grey, light red; broken curves) for a single peak landscape with $f_1 = 2$, $f_2 = \ldots = f_n = 1$, $\ell = 20$, $\sigma = 2$. The single peak landscape has been chosen such that the error threshold coincides with the extinction threshold at $\hat{p}_{\max} = p_{\max} = \ln 2/20$.

the form

$$\begin{aligned} \star &\xrightarrow{m_0\, r} \mathsf{M}\,, \\ \mathsf{M} + \mathcal{X}_1 &\xrightarrow{Q_{j1}k_1} \mathcal{X}_j + \mathcal{X}_1\,; \quad j = 1, 2, \ldots, n\,, \\ \mathsf{M} &\xrightarrow{r} \oslash\,, \quad \text{and} \\ \mathcal{X}_j &\xrightarrow{r} \oslash\,; \quad j = 1, 2, \ldots, n\,. \end{aligned} \tag{3.9}$$

where $\star$ and $\oslash$ refer to influx and outflux, respectively. Reaction kinetics leads to the differential equation

$$\begin{aligned} \frac{\mathrm{d}m}{\mathrm{d}t} &= -\left(\sum_{i=1}^{n} Q_{j1}k_1c_1\right) m + r(m_0 - m) \\ \frac{\mathrm{d}c_j}{\mathrm{d}t} &= Q_{j1}k_1mc_1 - rc_j\,,\ j = 1, 2, \ldots, n\,. \end{aligned} \tag{3.10}$$

In order to be able to handle lethal mutants properly we have to go back to absolute concentrations $c_j = x_j c(t)$ with $\sum_{j=1}^{n} c_j = c$, the variable m is the concentration of the building blocks M in the reactor, and the rate parameter k_1 is related to the fitness through $f_1 = k_1 m$.

Calculation of stationary states is straightforward and yields two solutions, (i) the state of extinction with $\bar{m} = m_0$ and $\bar{c}_j = 0 \,\forall\, j = 1, 2, \ldots, n$, and (ii) a state of quasispecies selection consisting of $\mathcal{X}_1$ and its mutant cloud at the concentrations $\bar{m} = r/(Q_{11}k_1)$, $\bar{c}_1 = Q_{11}m_0 - r/k_1$, and $\bar{c}_j = \bar{c}_1(Q_{j1}/Q_{11})$ for $j = 2, \ldots, n$. Stability analysis yields a straightforward result: A nonzero population is stable below a critical flow rate $r < r_{\mathrm{cr}} = Q_{11}k_1m_0$. The master sequence $\mathcal{X}_1$ and all its mutants vanish for the condition $r = r_{\mathrm{cr}}$ above which we have extinction. As an example for the distribution of genotypes we compute a maximum error rate for the uniform error rate model (3.5):

$$\begin{aligned} Q_{11} &= (1-p)^{\ell} \quad \text{and} \\ Q_{j1} &= p^{d_{j1}^{\mathrm{H}}} \cdot (1-p)^{\ell - d_{j1}^{\mathrm{H}}} \ , \end{aligned}$$

where d_{j1}^{H} is the Hamming distance between the two sequences $\mathcal{X}_j$ and $\mathcal{X}_1$. Instead of the superiority σ of the master sequence, which diverges since $\bar{k}_{-m} = 0$ because of $k_2 = \ldots = k_n = 0$, we use the carrying capacity of the flowreactor, η, which can be obtained straightforwardly as

$$\eta = \frac{k_1 m_0}{r} \ .$$

The value of p, at which the stationary concentration of the master sequence $\bar{c}_1(p)$ and all other mutants vanishes, represents the analogue to the error threshold (3.6), and for the sake of clearness it is referred to as *extinction threshold*. Using the approximation $\ln(1-p) \approx -p$ we obtain:

$$\hat{p}_{\max} \ \approx \ \frac{\ln \eta}{\ell} \quad \text{for small } p \ . \tag{3.11}$$

The major difference between the error threshold (3.6) and the extinction threshold (3.11) concerns the state of the population at values $p > p_{\max}$: Replication with non-zero fitness of mutants leads to the uniform distribution whereas the population goes extinct in the lethal mutant case. Accordingly, the transformation to relative concentrations fails and equ. (3.2) is not applicable. In figure 3.9 we show an example for the extinction threshold with $\ell = 20$ and $\eta = 2$. The extinction threshold is calculated from equ. (3.11) to occur at $\hat{p}_{\max} = 0.03466$ compared to an exact value of 0.03406. In the figure we see also a comparison of the curves for the master sequence and the one error class for the single peak landscape and the lethality model. The agreement of the two curves for the master sequences is not surprising, since the models were adjusted to coincide in the values $\bar{c}_1(0) = 1$ and $\hat{p}_{\max} = p_{\max} = \ln 2/20$.

TABLE 1. Optimization behavior in simple genetic systems.

Phenomenon	Optimization	Unique outcome
Selection	**yes**	**yes**
Mutation and selection	**no**	**yes**
Recombination and selection		
independent genes	**yes**	**no**
interacting genes	**no**	**no**

The curves for the one error class show some difference that is caused by the lack of mutational backflow from higher mutants in case of lethal variants.

It is important to note that a quasispecies[15] can exist also in cases where the Perron-Frobenius theorem is not fulfilled as in the current example of lethal mutants: Only genotype $\mathcal{X}_1$ has a positive fitness value, $f_1 > 0$ and $f_2 = \ldots = f_n = 0$, and hence only the entries $W_{k1} = Q_{k1} f_1$ of matrix W are nonzero and hence

$$\mathrm{W} = \begin{pmatrix} W_{11} & 0 & \ldots & 0 \\ W_{21} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ W_{n1} & 0 & \ldots & 0 \end{pmatrix} \quad \text{and} \quad \mathrm{W}^k = W_{11}^k \begin{pmatrix} 1 & 0 & \ldots & 0 \\ \frac{W_{21}}{W_{11}} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{W_{n1}}{W_{11}} & 0 & \ldots & 0 \end{pmatrix} .$$

Accordingly, W is not primitive in this example, but under suitable conditions $\bar{\mathbf{x}} = (Q_{11}, Q_{21}, \ldots, Q_{n1})$ is a stable stationary mutant distribution and for $Q_{11} > Q_{j1} \,\forall\, j = 2, \ldots, n$ (correct replication occurs more frequently than a particular mutation) genotype $\mathcal{X}_1$ is the master sequence. On the basis of a rather idiosyncratic mutation model consisting of a one-dimensional chain of mutants [58] the claim has been raised that no quasispecies can be stable in presence of lethal mutants, and hence no error thresholds can occur. A more recent paper [59, 60] used more realistic high-dimensional mutation models and presented numerically computed examples of perfect error thresholds in the presence of lethal mutants. A more detailed discussion of lethal mutagenesis is found in [26, 27, 60] but for a full understanding of the phenomenon investigations of realistic landscapes are indispensable.

[15] Although this distribution is not derived under constant population size, it will be called *quasispecies* here.

4. Perspectives

The Darwinian principle of variation and selection has been attributed with two features that were studied here by means of simple models: (i) optimization of mean fitness in populations and (ii) uniqueness of the result of a selection process. Both features are not necessarily fulfilled neither in model systems nor in nature. In table 1 a comparison is shown for the three models described in this chapter. Only the simple selection equation (2.14) sustains both features as mean fitness is optimized and in absence of neutrality the outcome of the selection process is unique. When mutation is included in equation (3.2') the unique outcome as represented by the quasispecies is guaranteed but the mean fitness may also decrease under certain circumstances. Fisher's selection equation (2.19) describes recombination on a single locus and mean fitness is optimized according to the fundamental theorem but no unique outcome is guaranteed – different initial conditions may lead to different allele combinations. Extrapolation of Fisher's model to multiple loci requires independence of genes and gene functions that is completely unrealistic. Interacting genes, in general, fulfil none of the two criteria. Therefore, the Darwinian principle is no theorem but a very powerful optimization heuristic that finds various applications in many disciplines from engineering to social sciences [61].

The mathematics of the ODE based model systems is fairly simple and straightforward. When applied to real situations the models gain enormous complexity essentially for two reasons: (i) any comprehensive model should cover a sufficiently large section of sequence space and this leads to gigantic dimensionality – the number of possible genotypes is 4^ℓ DNA or RNA sequences of chain length ℓ – that cannot be handled without drastic simplifications, and (ii) suitable models that allow for applications require realistic assumptions on the structure of fitness landscapes. Simple model landscapes assuming, for example, additivity of mutation effects or assigning equal fitness to all genotypes at the same distance from a reference sequence lead to wrong results. Ruggedness and neutrality shape fitness landscapes in reality and this can be taken into account properly only by models making use of random assignments, examples are the approach described here (subsection 3.5) and the frequently used NK-model [62, pp.40-60]. Direct calculation of fitness values for individual genotypes is possible at present only for RNA molecules within certain approximations [63, 64]. *In silico* evolution on these landscapes yields results that are in general agreement with those reported here. The great challenge for the future, however, is the construction of fitness landscapes that are based on solid experimental information. Apart from molecular *in vitro* systems, where such data are hard to get but within reach, systematic studies on virus infection in hosts are very promising for this goal.

References

[1] C. Darwin, *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London, 1859.

[2] G. Mendel, *Versuche über pflanzenhybriden (Experiments relating to plant hybridization)*. Verh. Naturforsch. ver Brunn **4** (1866), 3–17.

[3] S. Spiegelman, *An approach to the experimental analysis of precellular evolution*. Quart. Rev. Biophys. **4** (1971), 213–253.

[4] G. F. Joyce, *Directed evolution of nucleic acid enzymes*. Annu. Rev. Biochem. **73** (2004), 791–836.

[5] S. Klussmann (ed.), *The Aptamer Handbook. Functional Oligonucleotides and Their Applications*. Wiley-VCh Verlag, Weinheim, DE, 2006.

[6] S. Brakmann and K. Johnsson (eds.), *Directed Molecular Evolution of Proteins or How to Improve Enzymes for Biocatalysis*. Wiley-VCH, Weinheim, DE, 2002.

[7] C. Jäckel, P. Kast, and D. Hilvert, *Protein design by directed evolution*. Annu. Rev. Biophys. **37** (2008), 153–173.

[8] S. J. Wrenn and P. B. Harbury, *Chemical evolution as a tool for molecular discovery*. Annu. Rev. Biochem. **76** (2007), 331–349.

[9] L. E. Sigler, *Fibonacci's Liber Abaci: A Translation into Modern English of Leonardo Pisano's Book of Calculation*. Springer-Verlag, New York, 2002, a translation of Leonardo Pisano's book of 1202 from Latin into modern English.

[10] P. Singh, *The so-called Fibonacci numbers in ancient and medieval India*. Historia Mathematica **12** (1985), 229–244.

[11] D. E. Knuth, *The Art of Computer Programming. Vol.I: Fundamental Algorithms*. Third edition, Addison-Wesley Publishing Co., Reading, MA, 1997.

[12] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*. Second edition, Addison-Wesley Publishing Co., Reading, MA, 1994.

[13] R. A. Dunlap, *The Golden Ratio and Fibonacci Numbers*. World Sceintific, Singapore, 1997.

[14] J. Binet, *Mémoire sur l'intégration des équations linéaires aux différences finies, d'un ordre quelconque, à coefficients variables*. Comptes Rendue Hebdomadaires des Séances de l'Académie des Sciences (Paris) **17** (1843), 559–567.

[15] L. Euler, *Theoremata arithmetica nova methodo demonstrata*. Novi Comentarii Scientiarum Imperialis Petropolitanæ **8** (1760), 74–104, reprinted in Euler's *Commentationes Arithmeticæ Collectæ*, Vol.1, 274-286 and in Euler's *Opera Omnia*, Series 1, Vol.2, 531-555.

[16] T. R. Malthus, *An Essay of the Principle of Population as it Affects the Future Improvement of Society*. J. Johnson, London, 1798.

[17] L. Euler, *Introductio in Analysin Infinitorum, 1748. English Translation: John Blanton, Introduction to Analysis of the Infinite*, volume I and II. Springer-Verlag, Berlin, 1988.

[18] P. Verhulst, *Notice sur la loi que la population pursuit dans son accroisement*. Corresp. Math. Phys. **10** (1838), 113–121.

[19] P. Verhulst, *Recherches mathématiques sur la loi d'accroisement de la population.* Nouv. Mém. de l'Academie Royale des Sci. et Belles-Lettres de Bruxelles **18** (1845), 1–41.

[20] P. Verhulst, *Deuxième mémoire sur la loi d'accroisement de la population.* Mém. de l'Academie Royale des Sci., des Lettres et de Beaux-Arts de Belgique **20** (1847), 1–32.

[21] P. E. Phillipson and P. Schuster, *Modeling by Nonlinear Differential Equations. Dissipative and Conservative Processes*, volume 69 of *World Scientific Series on Nonlinear Science A*. World Scientific, Singapore, 2009.

[22] D. Zwillinger, *Handbook of Differential Equations.* Third edition, Academic Press, San Diego, CA, 1998.

[23] M. Eigen and P. Schuster, *The hypercycle. A principle of natural self-organization. Part B: The abstract hypercycle.* Naturwissenschaften **65** (1978), 7–41.

[24] P. Schuster and K. Sigmund, *Replicator dynamics.* J. Theor. Biol. **100** (1983), 533–538.

[25] R. A. Fisher, *The Genetical Theory of Natural Selection.* Clarendon Press, Oxford, 1930.

[26] J. J. Bull, L. Ancel Myers, and M. Lachmann, *Quasispecies made simple.* PLoS Comp. Biol. **1** (2005), 450–460.

[27] J. J. Bull, R. Sanjuán, and C. O. Wilke, *Theory for lethal mutagenesis for viruses.* J. Virology **81** (2007), 2930–2939.

[28] W. M. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, New York, 1974.

[29] S. Shahshahani, *A new mathematical framework for the study of linkage and selection.* Mem. Am. Math. Soc. **211** (1979), 34.

[30] P. Schuster and K. Sigmund, *Dynamics of evolutionary optimization.* Ber. Bunsenges. Phys. Chem. **89** (1985), 668–682.

[31] W. J. Ewens, *Mathematical Population Genetics*, volume 9 of *Biomathematics Texts.* Springer-Verlag, Berlin, 1979.

[32] J. B. Schnog, A. J. Duits, F. A. J. Muskiet, H. ten Cate, R. A. Rojer, and D. P. M. Brandjes, *Sickle cell disease: A general overview.* The Netherlands Journal of Medicine **62** (2004), 364–374.

[33] G. A. Barabino, M. O. Platt, and D. K. Kaul, *Sickle cell biomechanics.* Annu. Rev. Biomed. Eng. **12** (2010), 345–367.

[34] G. Price, *Fisher's "Fundamental Theorem" made clear.* Ann. Hum. Genet. **36** (1972), 129–140.

[35] A. W. Edwards, *The fundamental theroem of natural selection.* Biological Reviews **69** (1994), 443–474.

[36] S. Okasha, *Fisher's fundamental theorem of natural selection – A philosophical analysis.* Brit. J. Phil. Sci. **59** (2008), 319–351.

[37] F. Schlögl, *Chemical reaction models for non-equilibrium phase transitions.* Z. Physik **253** (1972), 147–161.

[38] G. von Kiedrowski, B. Wlotzka, J. Helbig, M. Matzen, and S. Jordan, *Parabolic growth of a self-replicating hexanucleotide bearing a 3'-5'-phosphoamidate linkage.* Angew. Chem. Internat. Ed. **30** (1991), 423–426.

[39] B. R. M. Stadler, P. F. Stadler, G. P. Wagner, and W. Fontana, *The topology of the possible: Formal spaces underlying patterns of evolutionary change.* J. Theor. Biol. **213** (2001), 241–274.

[40] C. K. Biebricher, M. Eigen, and J. William C. Gardiner, *Kinetics of RNA replication.* Biochemistry **22** (1983), 2544–2559.

[41] J. D. Watson and F. H. C. Crick, *A structure for deoxyribose nucleic acid.* Nature **171** (1953), 737–738.

[42] H. Judson, *The Eighth Day of Creation. The Makers of the Revolution in Biology.* Jonathan Cape, London, 1979.

[43] R. K. Saiki, D. H. G. amd Susanne Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich, *Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase.* Science **239** (1988), 487–491.

[44] C. K. Biebricher, M. Eigen, and J. William C. Gardiner, *Kinetics of RNA replication: Plus-minus asymmetry and double-strand formation.* Biochemistry **23** (1984), 3186–3194.

[45] C. K. Biebricher, M. Eigen, and J. William C. Gardiner, *Kinetics of RNA replication: Competition and selection among self-replicating RNA species.* Biochemistry **24** (1985), 6550–6560.

[46] C. J. Thompson and J. L. McBride, *On Eigen's theory of the self-organization of matter and the evolution of biological macromolecules.* Math. Biosci. **21** (1974), 127–142.

[47] B. L. Jones, R. H. Enns, and S. S. Rangnekar, *On the theory of selection of coupled macromolecular systems.* Bull. Math. Biol. **38** (1976), 15–28.

[48] R. W. Hamming, *Coding and Information Theory.* 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[49] E. Seneta, *Non-negative Matrices and Markov Chains.* Second edition, Springer-Verlag, New York, 1981.

[50] P. Schuster and J. Swetina, *Stationary mutant distribution and evolutionary optimization.* Bull. Math. Biol. **50** (1988), 635–660.

[51] M. Eigen, *Selforganization of matter and the evolution of biological macromolecules.* Naturwissenschaften **58** (1971), 465–523.

[52] J. W. Drake, *Rates of spontaneous mutation among RNA viruses.* Proc. Natl. Acad. Sci. USA **90** (1993), 4171–4175.

[53] E. Domingo, ed., *Virus entry into error catastrophe as a new antiviral strategy.* Virus Research **107** (2005), 115–228.

[54] J. W. Drake, B. Charlesworth, D. Charlesworth, and J. F. Crow, *Rates of spontaneous mutation.* Genetics **148** (1998), 1667–1686.

[55] T. Wiehe, *Model dependency of error thresholds: The role of fitness functions and contrasts between the finite and infinite sites models.* Genet. Res. Camb. **69** (1997), 127–136.

[56] C. O. Wilke, J. L. Wang, C. Ofria, R. E. Lenski, and C. Adami, *Evolution of digital organisms at high mutation rates leads to survival of the flattest*. Nature **412** (2001), 331–333.

[57] M. Kimura, *The Neutral Theory of Molecular Evolution*. Cambridge: University Press., UK, 1983.

[58] G. P. Wagner and P. Krall, *What is the difference between models of error thresholds and Muller's ratchet*. J. Math. Biol. **32** (1993), 33–44.

[59] N. Takeuchi and P. Hogeweg, *Error-thresholds exist in fitness landscapes with lethal mutants*. BMC Evolutionary Biology **7** (2007), 15.

[60] H. Tejero, A. Marín, and F. Moran, *Effect of lethality on the extinction and on the error threshold of quasispecies*. J. Theor. Biol. **262** (2010), 733–741.

[61] K. Lucas and P. Roosen (eds.), *Emergence, Analysis, and Evolution of Structures. Concepts and Strategies across Disciplines*. Understanding Complex Systems, Springer-Verlag, Berlin, 2010.

[62] S. A. Kauffman, *The Origins of Order. Self-Organization and Selection in Evolution*. Oxford University Press, New York, 1994.

[63] W. Fontana and P. Schuster, *Continuity in evolution. On the nature of transitions*. Science **280** (1998), 1451–1455.

[64] P. Schuster, *Prediction of RNA secondary structures: From theory to models and real molecules*. Reports on Progress in Physics **69** (2006), 1419–1477.

Peter Schuster
Institut für Theoretische Chemie, Universität Wien,
Währingerstraße 17, 1090 Wien, Austria and
Santa Fe Institute, 1399 Hyde Park Road,
Santa Fe, New Mexico 87501, USA
e-mail: pks@tbi.univie.ac.at