Exact Solutions for Evolutionary Strategies on Harmonic Landscapes

Jörn Dunkel dunkel@physik.hu-berlin.de
Lutz Schimansky-Geier alsg@physik.hu-berlin.de
Werner Ebeling ebeling@physik.hu-berlin.de
Institute of Physics, Humboldt-Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany

Abstract
In this paper two different evolutionary strategies are tested by means of harmonic landscapes. Both strategies are based on ensembles of searchers, spreading over the search space according to laws inspired by nature. The main difference between the two prototypes is given by the underlying selection mechanism, governing the increase or decrease of the local population of searchers in certain regions of the search space. More precisely, we compare a thermodynamic strategy, which is based on a physically motivated local selection criterion, with a biologically motivated strategy, which features a global selection scheme (i.e., global coupling of the searchers). Confining ourselves to a special class of initial conditions, we show that, in the simple case of harmonic test potentials, both strategies possess particular analytical solutions of the same type. By means of these special solutions, the velocities of the two strategies can be compared exactly. In the last part of the paper, we extend the scope of our discussion to a mixed strategy, combining local and global selection.

Keywords
Evolutionary strategies, thermodynamic strategy, Fisher-Eigen model, global coupling, generalized Boltzmann distribution

1 Introduction

In various fields of sciences one has to deal with optimization problems of the type

$$U(x) \stackrel{!}{=} \min \qquad x \in \Omega \subseteq \mathbb{R}^n, \tag{1}$$

where U is some given scalar function, which we refer to as landscape or potential. In general, for a given function U defined on a high-dimensional search space Ω, it is impossible to calculate the exact position x^* of a global optimum and numerical methods must be applied. Modern numerical algorithms developed for optimization problems (Kirkpatrick et al., 1983; Rechenberg, 1973; Holland, 1975; Fogel, 1995; Koza, 1992) are often based on physical or biological principles (gradient descent, selection schemes) and also include stochastic elements (diffusion, temperature). Such algorithms are usually referred to as evolutionary algorithms.

While there are many proposals regarding diffusion or temperature control, e.g., cooling schedules in 'Simulated Annealing' strategies (Mahnig and Mühlenbein, 2001; Andresen and Gordon, 1994; Andresen and Gordon, 1993), the present paper exclusively focuses on the efficiency of different selection schemes at constant temperature.

More exactly, we compare the behavior of two prototypes of evolutionary strategies, referred to as the Boltzmann strategy and the Darwin strategy (Boseniuk et al., 1987; Asselmeyer and Ebeling, 1996; Asselmeyer et al., 1997). Both strategies are based on the ideas that (i) an ensemble of searchers moves on the landscape U, and (ii) this ensemble can be described by means of a time-dependent probability density $f(x,t)$. The main difference between the two models lies in the underlying principles (selection schemes) according to which the ensembles evolve. The Boltzmann strategy is physically motivated and only includes local selection, i.e., it only compares local properties of U and f, whereas the biologically oriented Darwin strategy (Fisher, 1930; Eigen, 1971; Ebeling et al., 1990; Schweitzer et al., 1998) is based on global selection, i.e., searchers survive or die according to their status in the overall ensemble (global coupling). At the foundation of the Boltzmann strategy is the concept of over-damped Brownian motion (Einstein and von Smoluchowski, 1999), which was well-investigated in various fields of physics over the past decades (Hänggi et al., 1990; Chandrasekhar, 1943). We are going to consider this strategy and related results as a guide for the investigation of the less common Darwin strategy.

Even though both strategies have been successfully applied to a number of complex optimization problems (Schweitzer et al., 1998; Asselmeyer and Ebeling, 1996; Asselmeyer et al., 1997; Feistel and Ebeling, 1989; Rosenkranz, 1996), we still do not know quantitatively which of them is more efficient with respect to certain properties of U. The main cause for this dilemma is that analytical investigations of the corresponding partial differential equations (PDE) for the evolution of the density f are difficult to handle if U is complicated. However, as we shall work out below, in the case of harmonic potentials, one can find, for both strategies, particular analytical solutions, jointly based on the ansatz of a generalized Boltzmann distribution (Mahnig and Mühlenbein, 2001). Even though harmonic landscapes must of course be considered as relatively simple examples, the results obtained from them can be used in order to elucidate some more general advantages and disadvantages of either strategies. In particular, we emphasize the fact that in the vicinity of their minima each sufficiently smooth landscape U can always be approximated by a harmonic function (by virtue of a Taylor expansion). Thus, the analysis of a strategy's behavior for a simple harmonic test function may also reveal some insight with regard to more complicated applications.

In order to simplify navigation through this paper, we will end the introduction with some remarks on its structure. Section 2.1 is dedicated to the Boltzmann strategy. Subsequently, we will discuss the Darwin strategy in Section 2.2, following an analogous sequence of steps as before for the Boltzmann strategy. In principle, we always begin by briefly reviewing some well-known results. Then, in the second part, a special class of time-dependent solutions is identified, which allows for an explicit comparison of the strategies' efficiency. On the basis of the results derived in Sections 2.1 and 2.2, we proceed by analyzing a mixed strategy in Section 3, combining selection schemes of Boltzmann and Darwin type. Finally, Section 4 contains a summary of the main results and some conclusions.

2 Boltzmann and Darwin strategy

In this paper we compare evolutionary strategies on a landscape $U : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto U(x)$. We assume that an ensemble of searchers on U can be described by a normalized probability density

$$f(x,t) \geq 0 , \qquad \forall\, x \in \Omega,\ t \in T = [0,\infty) ; \tag{2a}$$

$$\int_\Omega \mathrm{d}x\ f(x,t) = 1 , \qquad \forall\, t \in T. \tag{2b}$$

This means, that the total number of searchers is conserved, and that $f(x,t)\mathrm{d}x$ characterizes the fraction of searchers in the interval $[x, x+\mathrm{d}x]$ of the search space Ω at time t. Each of the strategies we are interested in can be described by a PDE of the following type

$$\partial_t f(x,t) = (\hat{S} + \hat{M})\, f(x,t) , \tag{3}$$

where the initial distribution

$$f(x,0) = f_0(x) \tag{4}$$

is given. In (3) the operator $\hat{S}$ is a so-called selection operator and $\hat{M}$ is the mutation operator. In principle, there exist several possible choices for $\hat{M}$; for example, it can be chosen explicitly time-dependent, as in 'Simulated Annealing' strategies. However, for the time being, we concentrate on the case where mutation is realized by a simple diffusion process on the search space Ω corresponding to

$$\hat{M} = D\,\nabla^2. \tag{5}$$

This automatically implies that mutations do not depend on the landscape U, but instead on some external parameter $D > 0$, referred to as the diffusion constant or the temperature of a strategy. In Sections 2.1.1 and 2.2.1, where we discuss simple algorithms which realize the respective strategies, we shall also clarify how the mutation operator (5) is connected with mutations of individual searchers in an ensemble.

In contrast to $\hat{M}$, the selection operator $\hat{S}$ has to depend on U, since it governs the increase or decrease of the local population density in $[x, x+\mathrm{d}x]$ during the time interval $[t, t+\mathrm{d}t]$.

2.1 Boltzmann strategy

In the Boltzmann model, which is the first model we are studying, the selection operator is given by

$$\hat{S}_B := (\nabla^2 U) + (\nabla U\ )\nabla. \tag{6}$$

For convenience, we apply the compact notation $f_x := \nabla f$, $f_{xx} := \nabla^2 f$, $U_{xx} := \nabla^2 U$, and $f_t := \partial_t f$ from now on. Inserting (6) into (3) yields the well-known Smoluchowski equation (Einstein and von Smoluchowski, 1999; **?**; Hänggi et al., 1990)

$$f_t = (U_x f)_x + D f_{xx}, \tag{7}$$

governing the motion of over-damped Brownian particles in an external potential U. Since the selection operator (6) only depends on local quantities, we may speak of local selection here. With the definitions

$$\tau := Dt , \qquad u(x) := \frac{U(x)}{D} , \qquad p(x,\tau) := f(x,t) , \tag{8}$$

we can rewrite (7) in the simplified form

$$p_\tau = (u_x p)_x + p_{xx}. \tag{9}$$

To distinguish u from U, we also refer to the function u as a reduced potential. As one can easily check by insertion, the stationary solution of (9) is given by the stationary Boltzmann distribution

$$p^{st}(x) = \frac{e^{-u(x)}}{\mathcal{Z}^{st}}\ , \tag{10}$$

where the stationary normalization constant reads

$$\mathcal{Z}^{st} = \int_\Omega \mathrm{d}x\ e^{-u(x)}\ . \tag{11}$$

Hence, a strategy described by (9) is also called a thermodynamical Boltzmann strategy. Since p^{st} possesses maxima at the minima of U, it can be considered as an evolutionary strategy (with local selection).

2.1.1 Numerical algorithms

There exist different numerical algorithms realizing this type of Boltzmann search. They all have in common that they are based on a finite ensemble of N searchers, and that in the limit $N \to \infty$, the density of searchers p is described by the Smoluchowski equation (9). For the sake of simplicity, here we only consider the $1d$-case $\Omega = \mathbb{R}$, and discuss a straightforward approach, which is based on discretized Langevin equations

$$x_i(\tau + \Delta\tau) = x_i(\tau) - \nabla u\big(x_i(\tau)\big)\Delta\tau + \sqrt{2\Delta\tau}\ \xi_i \tag{12}$$

governing the dynamics of each individual searcher ($i = 1, \ldots, N$). In (12) the quantity $x_i(\tau)$ is the coordinate of the ith searcher at time τ, and the ξ_i's are random variables taken from a Gaussian standard normal distribution. In the physical picture, (12) describes the stochastic motion of an over-damped Brownian particle in the potential u. It is a well-known textbook result of statistical physics (Chandrasekhar, 1943; Hänggi et al., 1990) that an ensemble of infinitely many, independent particles, which move according to (12), can also be described by the Smoluchowski equation (9), if additionally $\Delta\tau \to 0$ is assumed. Note, that (12) combines gradient descent via the ∇U-term and diffusion in the search space $\Omega = \mathbb{R}$ by virtue of the stochastic forces ξ_i. The diffusion process can also be interpreted as a mutation process.

For further possible implementations of the Boltzmann search, we refer the reader to (Asselmeyer et al., 1997; Rosé, 1996; Schweitzer et al., 1998).

2.1.2 General time-dependent solution

In this section, we briefly summarize some well-known analytical results concerning the Boltzmann strategy (Asselmeyer et al., 1997; Asselmeyer and Ebeling, 1996). In order to construct the general time-dependent solutions of (9), the ansatz

$$p(x,\tau) = \varrho(x,\tau)\ e^{-u(x)/2}\ , \tag{13}$$

is used leading to

$$\varrho_\tau = \varrho_{xx} - \left(\frac{1}{4}\,u_x^2 - \frac{1}{2}\,u_{xx}\right)\varrho\ . \tag{14}$$

Mathematically, this equation (14) can be treated like a Schrödinger equation. Thus, assuming a discrete, non-degenerate spectrum of eigenvalues[1], the complete solution of (9) can be expressed by the series expansion

$$p(x,\tau) = e^{-u(x)/2} \sum_{n=0}^{\infty} c_n \, \phi_n(x) \, e^{-E_n \tau} \tag{15}$$

where the $\phi_n(x)$ are the $L^2(\Omega)$-normalized time-independent eigenfunctions, i.e.,

$$(\phi_n, \phi_m) := \int_\Omega \mathrm{d}x \; \phi_n^*(x) \; \phi_m(x) = \delta_{nm} \; , \tag{16}$$

of the Hamilton operator

$$\hat{H}_B = -\nabla^2 + v(x) \; , \qquad\qquad v(x) = \frac{1}{4} \, u_x^2 - \frac{1}{2} \, u_{xx} \tag{17}$$

with corresponding eigenvalues E_n. The coefficients c_n in (15) are determined by the initial condition

$$c_n = (\phi_n, \varrho_0) = \int_\Omega \mathrm{d}x \; \phi_n^*(x) \; e^{u(x)/2} \, p(x,0), \tag{18}$$

where $\varrho_0(x) = \varrho(x,0)$ is known via $f_0(x)$. For $t \to \infty$ the series expansion (15) must converge to the stationary solution (10). Thus one finds, that $E_0 = 0$, $c_0 = 1/\sqrt{\mathcal{Z}^{st}}$ and

$$\phi_0(x) = \frac{e^{-u(x)/2}}{\left[\int_\Omega \mathrm{d}z \, e^{-u(z)}\right]^{1/2}} = \frac{e^{-u(x)/2}}{\sqrt{\mathcal{Z}^{st}}}. \tag{19}$$

As we can immediately see from (15), the eigenvalue difference $\Delta E := E_1 - E_0 = E_1$ dominates the dynamics of the density $p(x,\tau)$ for $\tau \gg 0$. In other words, for landscapes U with a large eigenvalue difference $\Delta E \gg 0$ or large eigenvalue $E_1 \gg 0$, respectively, the ensemble of the Boltzmann strategy approaches its stationary distribution faster compared to the case where $E_1 \approx 0$ (a detailed discussion of this problem with regard to the problem of overcoming barriers in a landscape U can be found in (Dunkel et al., 2003)).

We note, that although the solution of the Smoluchowski equation is formally known by (15), it is only in parts useful with regard to general applications. One reason for this is, that, apart from some very simple examples, for a given potential U the eigenvalues and eigenfunctions of the Hamilton operator (17) with the effective potential v are unknown. Hence, following the approach outlined in the introduction we shall have a closer look at harmonic landscapes. For this particular class of test potentials particular analytical representations of certain symmetric solutions can be found, which are not only more elegant than the series expansion (15), but also turn out to be very fruitful for a quantitative comparison with the biologically oriented Darwin strategy discussed in Section 2.2.

[1] We note that in principle the more complicated case of a degenerate eigenvalue spectrum can be treated in a very similar manner.

2.1.3 Particular solutions for harmonic landscapes

We begin with the simplest case of the one-dimensional harmonic oscillator

$$U(x) = \frac{1}{2}\, A\, x^2 \;, \qquad\qquad x \in \mathbb{R},\; A > 0, \tag{20}$$

since n-dimensional convex problems can be reduced to this case, as we shall outline at the end of this subsection. Defining the reduced frequency $\omega := \sqrt{A/D}$, the corresponding reduced potential reads

$$u(x) = \frac{\omega^2}{2}\, x^2. \tag{21}$$

Using the well-known results from the quantum harmonic oscillator, it is straightforward to write down the complete analytic solution in terms of Hermitian polynomials and corresponding eigenvalues. Indeed, for the Boltzmann and the Darwin strategy this has already been done earlier (Asselmeyer et al., 1997; Asselmeyer and Ebeling, 1996). Hence, as a kind of complement, we concentrate on a particular class of solutions of (9), which is based on the ansatz

$$p(x,\tau) \;=\; \frac{1}{\mathcal{Z}(\tau)}\; \exp\bigl[-\beta(\tau)\; u(x)\bigr]\;, \qquad\qquad \beta(\tau) > 0 \quad \forall\, \tau \geq 0\;; \tag{22a}$$

$$\mathcal{Z}(\tau) \;=\; \int_\Omega \mathrm{d}x\; \exp\bigl[-\beta(\tau)\; u(x)\bigr]\;, \tag{22b}$$

representing a generalized Boltzmann distribution (Mahnig and Mühlenbein, 2001). Later, in Sections 2.2.3 and 3.1 we will show that this ansatz also represents a solution for the Darwin strategy and the mixed strategy.

A general property of the distribution (22) is, that the ensemble average value $\overline{u}(\tau)$ is simply given by

$$\overline{u}(\tau) = -\frac{\mathrm{d}}{\mathrm{d}\beta(\tau)}\, \ln \mathcal{Z}(\tau)\,, \tag{23}$$

and the corresponding variance by

$$var[u](\tau) := \overline{u^2}(\tau) - \overline{u}(\tau)^2 = \frac{\mathrm{d}^2}{\mathrm{d}\beta^2}\, \ln \mathcal{Z}. \tag{24}$$

Additionally, we shall study the velocity

$$vel[u](\tau) := \frac{\mathrm{d}}{\mathrm{d}\tau}\, \overline{u}(\tau), \tag{25}$$

representing a measure for the change of the expectation value $\overline{u}(\tau)$, i.e., for the speed of the strategy. Inserting (21) into (22b) yields

$$\mathcal{Z}(t) = \sqrt{\frac{2\pi}{\omega^2 \beta(\tau)}}, \tag{26}$$

and one obtains

$$\overline{u}(\tau) = \frac{1}{2\;\beta(\tau)}\;, \qquad\qquad var[u](\tau) = \frac{1}{2\;\beta(\tau)^2}. \tag{27}$$

Furthermore, the velocity reads

$$vel[u](\tau) = -\frac{1}{2\,\beta(\tau)^2}\,\frac{\mathrm{d}}{\mathrm{d}\tau}\,\beta(\tau) = -var[u](\tau)\,\frac{\mathrm{d}}{\mathrm{d}\tau}\,\beta(\tau). \tag{28}$$

We note that thus far the expressions for $\overline{u}(\tau),\ var[u](\tau)$ and $vel[u](\tau)$ are general results that are valid for any distribution of type (22). In order to relate them to the Boltzmann strategy, we have to determine $\beta(\tau)$ by inserting (22) into Smoluchowski's equation (9). This leads to the following ordinary differential equation (ODE) for $\beta(\tau)$

$$0 = \frac{\mathrm{d}\beta}{\mathrm{d}\tau} + 2\omega^2(\beta - 1)\beta, \tag{29}$$

with some initial condition $\beta(0) > 0$. Although this ODE is nonlinear, it can be solved exactly and the solution is given by

$$\beta(\tau) = \frac{\beta(0)\ e^{2\omega^2\tau}}{1 + \beta(0)\ (e^{2\omega^2\tau} - 1)}. \tag{30}$$

In fact, due to the rather special ansatz (22), the respective solutions can only cover a relatively small subset of all possible initial conditions. However, the special case of a uniform initial distribution $p(x, 0) = const$ is asymptotically included in (22), if $\beta(0) \to 0$ is considered. The opposite case $\beta(0) \to \infty$ corresponds to an initial δ-distribution concentrated in the minimum $x = 0$. This special case has already been discussed by Smoluchowski in one of his fundamental papers on Brownian motion (Einstein and von Smoluchowski, 1999).

(a) (b)

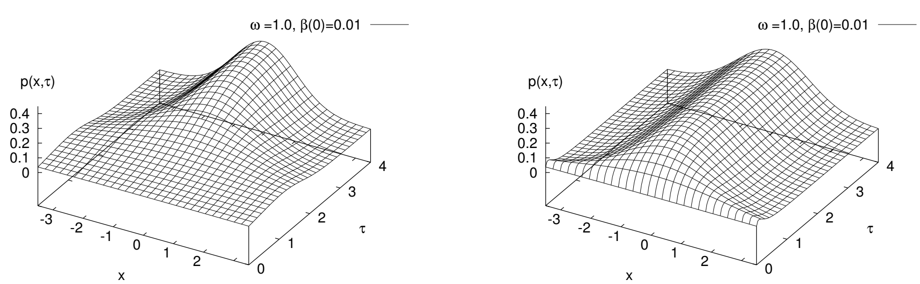

Figure 1: Spatio-temporal development of the probability density for (a) thermodynamical Boltzmann strategy and (b) Darwin strategy with identical parameters and initial conditions. As one can see, for the chosen parameters, the Darwin strategy (b) approaches the stationary state much 'faster' than the Boltzmann strategy, but the stationary distribution of the latter possesses better properties with regard to the concentration of searchers around the minimum of the landscape (at $x = 0$).

The stationary distribution of the Boltzmann strategy is readily obtained from (22) and (30)

$$p^{st}(x) = \sqrt{\frac{\omega^2}{2\pi}}\ \exp\left(-\frac{\omega^2}{2}\,x^2\right), \tag{31}$$

and leads to the stationary values

$$\overline{u}^{st} = \frac{1}{2}\,, \qquad\qquad var[u]^{st} = \frac{1}{2}. \tag{32}$$

We note, that both quantities are independent of ω for the Boltzmann strategy. The velocity $vel_B[u](\tau)$ of the Boltzmann strategy for the harmonic oscillator is given by

$$vel_B[u](\tau) = \frac{\beta(0) - 1}{\beta(0)}\,\omega^2\, e^{-2\omega^2 t}. \tag{33}$$

In Fig. 1 we plotted the spatio-temporal development of $p(x, \tau)$ for an almost uniform initial distribution.

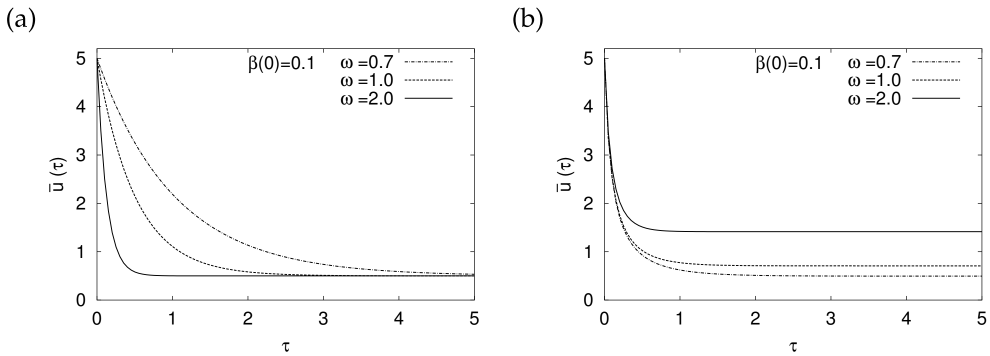

Figure 2: Ensemble average value $\overline{u}(\tau) = 1/(2\beta(\tau))$ of the harmonic oscillator for (a) thermodynamical Boltzmann strategy and (b) generalized Fisher-Eigen (Darwin) strategy. Only for the Boltzmann strategy is the stationary value $\overline{u}^{st}$ independent from ω.

We will complete this section with a short remark on n-dimensional problems. As n-dimensional harmonic landscape we consider

$$U(x_1, \ldots, x_n) = \frac{1}{2}\sum_{i=1}^{n} A_i\, x_i^2\,, \qquad\qquad x_i \in \mathbb{R},\ A_i > 0 \quad \forall\, i = 1, \ldots, n\,, \tag{34}$$

where the corresponding reduced potential is now

$$u(x_1, \ldots, x_n) = \frac{1}{2}\sum_{i=1}^{n} \omega_i^2\, x_i^2\,, \qquad\qquad \omega_i := \sqrt{\frac{A_i}{D}}\,. \tag{35}$$

For this case, a generalized Boltzmann distribution can be written in the product form

$$p(x_1, \ldots, x_n, \tau) = \prod_{i=1}^{n} p_i(x_i, \tau) \tag{36}$$

where

$$p_i(x, \tau) = \frac{1}{\mathcal{Z}_i(\tau)}\, \exp\bigl[-\beta_i(\tau)\, u_i(x_i)\bigr]\,, \qquad\qquad \beta_i(\tau) > 0 \quad \forall\, \tau \geq 0\,; \tag{37a}$$

$$u_i(x_i) = \frac{1}{2}\,\omega_i^2\, x_i^2\,, \tag{37b}$$

$$\mathcal{Z}_i(\tau) = \int_{\mathbb{R}} \mathrm{d}x_i\, \exp\bigl[-\beta_i(\tau)\, u_i(x_i)\bigr] = \sqrt{\frac{2\pi}{\omega_i^2\, \beta_i(\tau)}} \tag{37c}$$

for $i = 1, \ldots, n$. Furthermore, if each of the (normalized) p_i's satisfies a one-dimensional Smoluchowski equation, then p satisfies a n-dimensional Smoluchowski equation. Thus we can construct a solution of the n-dimensional problem from the above solutions of the one-dimensional problem.

2.2 Darwin strategy

Having discussed the physically motivated Boltzmann strategy with local selection in the previous sections, we will now study a biologically motivated model of an evolutionary strategy featuring non-local selection. After briefly introducing the corresponding PDE, we will outline the main features of a related numerical algorithm in 2.2.1. Section 2.2.2 reviews some known facts concerning general solutions, and in 2.2.3 we concentrate once again on harmonic test functions.

If instead of (6) we choose a selection operator defined by

$$\hat{S}_D := \overline{U}(t) - U(x), \tag{38}$$

where

$$\overline{U}(t) = \int_{\Omega} \mathrm{d}x \; U(x)\, f(x,t) \tag{39}$$

is the time-dependent ensemble average of U, we obtain as dynamical equation for the related ensemble density f a generalized Fisher-Eigen equation (Fisher, 1930; Eigen, 1971; Ebeling et al., 1990; Feistel and Ebeling, 1989)

$$f_t = [\overline{U}(t) - U]\; f + D f_{xx}. \tag{40}$$

The effect of the biologically motivated Fisher-Eigen operator (38) is obvious. It leads to an increase of the local population, if the value $U(x)$ is lower than the ensemble average $\overline{U}(t)$ and to a decrease, otherwise. In contrast to the purely local selection scheme represented by (6) we have a non-local selection criterion in the case of (38). In other words, the change of the local population in $[x, x + \mathrm{d}x]$ between t and $t + \mathrm{d}t$ can also be strongly influenced by those parts of the overall population which are located at far distances from x. In this sense, models with non-local selection are always based on the assumption that the corresponding system includes, as a fundamental feature, long-range information transfer mechanisms, i.e., communication between its constituents, as for example met in biological systems. In agreement with the terminology of previous articles (Boseniuk et al., 1987; Asselmeyer and Ebeling, 1996; Schweitzer et al., 1998), we speak of a Darwin strategy when considering the generalized Fisher-Eigen equation (40).

2.2.1 Numerical algorithm

As with the Boltzmann strategy, there exist several numerical algorithms realizing a Fisher-Eigen ensemble described by (40). A typical implementation is described in (Rosé, 1996; Rosenkranz, 1996). It is based on a finite ensemble with N searchers, evolving in the search space Ω. In order to explain the main features, it is again sufficient to confine ourselves to the simple $1d$-case $\Omega = \mathbb{R}$. As for the Boltzmann strategy, each searcher is described by a coordinate $x_i(t)$; and the mutation (diffusion) of a searcher is again realized by Langevin dynamics ($i = 1, \ldots, N$)

$$x_i(t + \Delta t) = x_i(t) + \sqrt{2D\Delta t}\; \xi_i, \tag{41}$$

where ξ_i is a random variable taken from a Gaussian standard normal distribution. The main difference compared with the Boltzmann search is given by the selection procedure applied in the Fisher-Eigen ensemble. It works as follows: For each particle i one picks randomly (with uniform probability) a second particle, say j, from the remaining $N-1$ particles. Then one applies the following selection rules

$$(i,j) \xrightarrow{|w(x_i)|} (i,i), \qquad \text{if} \quad w(x_i(t)) < 0; \tag{42a}$$

$$(i,j) \xrightarrow{|w(x_i)|} (j,j), \qquad \text{if} \quad w(x_i(t)) > 0, \tag{42b}$$

where the ('reaction') rate $|w|$ is determined by

$$w(x_i(t)) = U(x_i(t)) - \overline{U}(t), \qquad \overline{U}(t) = \frac{1}{N}\sum_{i=1}^{N} U(x_i(t)). \tag{43}$$

In the case of (42a) this means, for example, that particle j is likely to be replaced by a copy of particle i, if $U(x_i(t))$ is much smaller than the ensemble average $\overline{U}(t)$. Note that $|w|$ is not a selection probability, but the rate for the 'two-particle-reactions'. Thus, the actual numerical realization of this selection procedure includes some further technical subtleties, which are extensively discussed in (Asselmeyer et al., 1997). There, it was also shown by means of computer experiments that already relatively small ensembles of searchers (e.g., $N = 10$ or $N = 100$), which evolve according to the above algorithm, are well described by the Fisher-Eigen equation (40).

Moreover, applications of this Darwinian strategy to simple quantum-mechanical problems can also be found in (Rosenkranz, 1996).

2.2.2 General time-dependent solution

Before we study the solutions of the Fisher-Eigen equation (40), it is useful to apply the substitutions from (8), leading to the slightly simplified equation

$$p_\tau = [\overline{u}(\tau) - u]\, p + p_{xx}. \tag{44}$$

In contrast to the Smoluchowski equation (9), the generalized Fisher-Eigen equation (44) is a nonlinear PDE. Nevertheless, it can also be transformed into a PDE of Schrödinger type (14) by using the ansatz (Feistel and Ebeling, 1989)

$$p(x,\tau) = \varrho(x,\tau)\, \exp\left[\int_0^\tau \overline{u}(s)\, ds\right]. \tag{45}$$

By insertion one obtains

$$\varrho_\tau = \varrho_{xx} - u\varrho. \tag{46}$$

Compared with (14) the only difference is that instead of the effective potential v in (14), now the original function u appears in (46). Thus, assuming a discrete, non-degenerate spectrum of eigenvalues again, the formal solution of (44) is given by

$$p(x,\tau) = \exp\left[\int_0^\tau \overline{u}(s)\, ds\right] \sum_{n=0}^{\infty} c_n\, \phi_n(x)\, e^{-E_n \tau}, \tag{47}$$

where ϕ_n is a $L^2(\Omega)$-normalized eigenfunction of the Hamilton operator

$$\hat{H}_D = -\nabla^2 + u(x)\ , \tag{48}$$

and E_n the corresponding eigenvalue. The coefficients

$$c_n = \int_\Omega \mathrm{d}x\ \phi_n^*(x)\ \ p(x,0) \tag{49}$$

are again determined by the initial condition. In order to identify the pre-factor in (47), one integrates (47) over x and makes use of the fact that $p(x,\tau)$ is normalized. Then the final result is

$$p(x,\tau) = \frac{\sum_{n=0}^{\infty} c_n\ \phi_n(x)\ e^{-E_n\tau}}{\sum_{m=0}^{\infty} c_m\, l_m\, e^{-E_m\tau}}\ , \tag{50}$$

where

$$l_m = \int_\Omega \mathrm{d}x\ \phi_m(x). \tag{51}$$

Hence, the solution of the generalized Fisher-Eigen problem (44) can be expressed exclusively in terms of characteristic quantities of the eigenvalue problem (46).

Let us now have a closer look yet at the stationary situation. Assuming a time-independent solution $p^{st}(x)$ of (44), we find that the stationary value $\overline{u}^{st}$ of $\overline{u}(\tau)$ is given by the lowest eigenvalue

$$\overline{u}^{st} = E_0, \tag{52}$$

and the stationary solution $p_{st}(x)$ is proportional to $\phi_0(x)$, i.e.,

$$p^{st}(x) = \frac{\phi_0(x)}{\mathcal{Z}^{st}}, \tag{53}$$

where $\mathcal{Z}^{st} = l_0$.

Despite the fact that the time-dependent solution of the generalized Fisher-Eigen equation (44) is formally known, for complicated functions u only approximate analytical results can be derived (due to the same reasons as for the Boltzmann strategy). Hence, we concentrate in the following again on special solutions obtainable for harmonic landscapes. These results are compared with those of the Boltzmann strategy. Furthermore, in the remainder of the paper we confine the discussion to one-dimensional problems, since the Darwin strategy solutions of n-dimensional problems can also be constructed as products, in a similar manner as outlined in Section 2.1.3 for the Boltzmann strategy.

2.2.3 Particular solutions for 1d-harmonic landscapes

It is a very fortunate fact[2] that the generalized Boltzmann distribution (22)

$$\begin{aligned} p(x,\tau) &= \frac{1}{\mathcal{Z}(\tau)}\ \exp\bigl[-\beta(\tau)\ u(x)\bigr]\ , \qquad \beta(\tau) > 0 \quad \forall\,\tau \geq 0\ ; \\ \mathcal{Z}(\tau) &= \int_\Omega \mathrm{d}x\ \exp\bigl[-\beta(\tau)\ u(x)\bigr]\ , \end{aligned}$$

can also be used to solve the Fisher-Eigen equation (44). Inserting this ansatz into the Fisher-Eigen equation (44) leads to the following ODE for $\beta(\tau)$

$$0 = \frac{\mathrm{d}\beta}{\mathrm{d}\tau} - 1 + 2\,\omega^2\,\beta^2. \tag{55}$$

[2] We are grateful to H. Mühlenbein for bringing this to our attention.

Note that this equation slightly differs from (29), which was obtained when we used the same ansatz in order to solve the Smoluchowski equation (9). A solution of the nonlinear ODE (55) is given by

$$\beta(\tau) = \left(2\omega^2\right)^{-1/2} \tanh\left\{\sqrt{2\omega^2}\,\tau + \tanh^{-1}\left[\sqrt{2\omega^2}\,\beta(0)\right]\right\}, \tag{56}$$

where $\tanh^{-1}$ denotes the inverse function of $\tanh$. Again, $\beta(0) \to 0$ corresponds to an initial uniform distribution.

The velocity $vel_D[u](\tau)$ of the Darwin strategy in the harmonic oscillator is given by

$$vel_D[u](\tau) = \frac{-\omega^2}{\sinh\left\{\sqrt{2\omega^2}\,\tau + \tanh^{-1}\left[\sqrt{2\omega^2}\,\beta(0)\right]\right\}^2}. \tag{57}$$

In Fig. 2 we plotted $\overline{u}(\tau)$ for both strategies, and in Fig. 3 one can see the ratio

$$R(\tau) := \frac{vel_B[u](\tau)}{vel_D[u](\tau)} \tag{58}$$

between the velocities of the two strategies. The stationary distribution of the Darwin

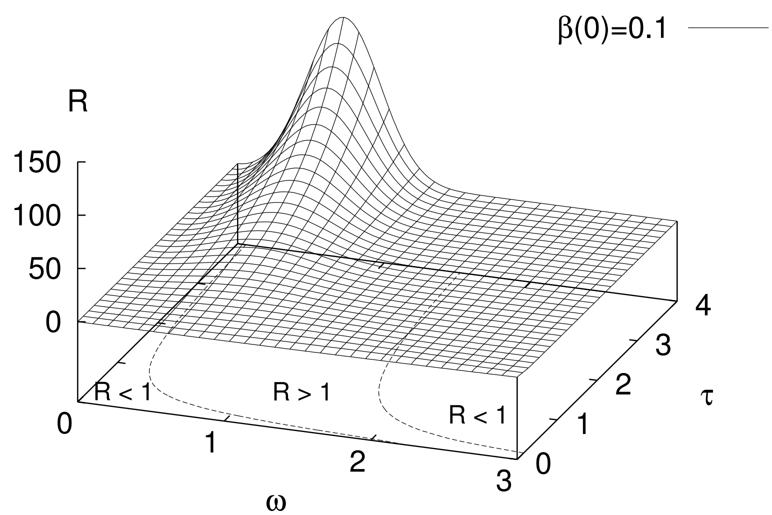

Figure 3: In regions with $R < 1$ the thermodynamical Boltzmann strategy is 'slower' than the Fisher-Eigen strategy, and for $R > 1$ the opposite is true. We note, that an evaluation of the efficiency of the strategies must always be based on both, the velocity $vel[u](\tau)$ and the respective value $\overline{u}(\tau)$.

strategy is obtained from (22) and (56)

$$p^{st}(x) = \sqrt{\frac{\omega}{2\pi\sqrt{2}}}\;\exp\left(-\frac{\omega}{\sqrt{8}}\,x^2\right), \tag{59}$$

and leads us to

$$\overline{u}^{st} = \frac{\omega}{\sqrt{2}} , \qquad\qquad var[u]^{st} = \frac{\omega^2}{2}. \tag{60}$$

In contrast to the thermodynamical Boltzmann strategy, $\overline{u}^{st}$ and also $var[u]^{st}$ are ω-dependent for the Darwin strategy. Thus, for small values $\omega < \omega_c = 1/\sqrt{2}$, corresponding to weak curvature of u, the non-local Darwin strategy possesses a lower stationary expectation value $\overline{u}^{st}$ than the Boltzmann strategy with local selection. On the other hand, for a given parameter A of the original potential $U = Ax^2/2$ and a sufficiently small temperature parameter D, i.e., more exactly for

$$D < 2\,A, \tag{61}$$

the stationary distribution (31) of the Boltzmann strategy is always more concentrated around the minimum than the stationary distribution (59) of the Darwin strategy.

3 Mixed strategy

Although the stationary distributions of the thermodynamical Boltzmann strategy and biological Darwin strategy are qualitatively similar, i.e., they are both concentrated around the minima of the landscape U, the quantitative behavior of the two strategies might be significantly different depending on the special structure of a given U. With regard to applications of the strategies to real optimization problems, one certainly wants to know:

1. Which strategy converges faster towards its stationary distribution?
2. Which stationary distribution is more concentrated around the minima of U?

In the general case, one should expect that neither one nor the other strategy is preferable. Thus, the idea of combining the two strategies to one so-called mixed strategy is obvious and some work on this subject was started some years ago (Asselmeyer et al., 1997; Schweitzer et al., 1998).

Here, we intend to deal with a mixed strategy realized by a one-parametric coupling. The corresponding selection-operator reads

$$\hat{S} := \alpha\,\hat{S}_B + (1-\alpha)\,\hat{S}_D \qquad\qquad \alpha \in [0,1]\,, \tag{62}$$

and leads to the evolutionary equation

$$f_t = \alpha\,(U_x f)_x + (1-\alpha)\,[\overline{U}(t) - U]\;f + Df_{xx}, \tag{63}$$

which, by virtue of (8), is equivalent to

$$p_\tau = \alpha\,(u_x p)_x + (1-\alpha)\,[\overline{u}(\tau) - u]\;p + p_{xx}. \tag{64}$$

Using the ansatz

$$p(x,\tau) = \varrho(x,\tau)\;\exp\left[(1-\alpha)\int_0^\tau \overline{u}(s)\,ds - \alpha\,\frac{u(x)}{2}\right] \tag{65}$$

we get

$$\varrho_\tau = \varrho_{xx} - \left[(1-\alpha)\,u + \frac{\alpha^2}{4}\,u_x^2 - \frac{\alpha}{2}\,u_{xx}\right]\varrho. \tag{66}$$

We merely note that, on the analogy of (14) and (46), the last equation can also be treated like a Schrödinger equation; i.e., in principle it is again possible to write down the exact solution for the combined strategy in terms of a series expansion based on the corresponding eigenfunctions and eigenvalues.

3.1 Harmonic oscillator

Considering the one-dimensional harmonic oscillator as before, we may again look for solutions based on the generalized Boltzmann distribution (22)

$$\begin{aligned} p(x,\tau) &= \frac{1}{\mathcal{Z}(\tau)} \exp\bigl[-\beta(\tau)\, u(x)\bigr]\,, \qquad \beta(\tau) > 0 \quad \forall\, \tau \geq 0\,; \\ \mathcal{Z}(\tau) &= \int_\Omega \mathrm{d}x \, \exp\bigl[-\beta(\tau)\, u(x)\bigr]\,. \end{aligned}$$

Inserting this ansatz now into the mixed equation (64) leads again to a nonlinear ODE for $\beta(\tau)$

$$0 = \frac{\mathrm{d}\beta}{\mathrm{d}\tau} - 1 + \alpha - 2\omega^2 \beta(\alpha - \beta) \tag{68}$$

generalizing ODE (29) and (55). With the abbreviation

$$\gamma := \sqrt{2\alpha - \omega^2\alpha^2 - 2} \tag{69}$$

a solution of (68) reads

$$\beta(\tau) = \frac{1}{2}\left\{\alpha - \frac{\gamma}{\omega}\, \tan\left[\omega\gamma\,\tau + \tan^{-1}\left[\frac{\omega\alpha - 2\omega\beta(0)}{\gamma}\right]\right]\right\}, \tag{70}$$

where $\beta(0) > 0$ is the initial condition and $\tan^{-1}$ the inverse function of $\tan$. In the limit cases $\alpha = 1$ and $\alpha = 0$ the solution (70) reduces to (30) and (56), respectively.

In Fig. 4 we plotted the average value $\overline{u}(\tau) = [2\,\beta(\tau)]^{-1}$ of the mixed strategy for different values of ω and α. In diagram 4 (a) one can see that for high values $\omega > \omega_c = 1/\sqrt{2}$ the parameter value $\alpha = 1$, corresponding to a pure Boltzmann strategy, gives the best (that is the lowest) stationary value $\overline{u}(\tau)$. In contrast to this, for subcritical values $\omega < \omega_c$ as shown in Fig. 4 (b), the parameter value $\alpha = 0$, corresponding to a pure Fisher-Eigen (Darwinian) search, yields the best result. Since ω was earlier defined as $\omega = \sqrt{A/D}$, where A is the parameter of the potential U and D the temperature of the strategy, one can apparently change ω by varying the temperature D (note, that in practice the parameter A must be considered as fixed for a given function U, whereas D can be varied during the course of the simulation).

4 Conclusions

Even though we confined our discussion to simple harmonic test functions, we can use the results obtained above, to predict some general properties of the two strategies. One reason for this is, for example, that in the vicinity of a minimum each sufficiently smooth landscape U can be approximated by a harmonic potential. Thus, we may conclude that compared with the Darwin strategy, the stationary distribution of the Boltzmann strategy possesses better properties, if the temperature parameter D is chosen sufficiently small.

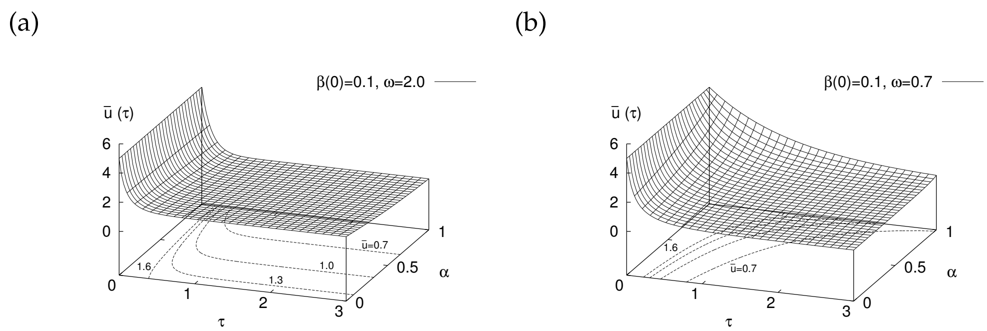

Figure 4: Ensemble average value $\overline{u}(\tau) = 1/(2\beta(\tau))$ of the harmonic oscillator for the mixed strategy with different parameter values: (a) $\omega > \omega_c$, and (b) $\omega < \omega_c$. The contour levels not labeled in (b) correspond to those in (a). The initial condition $\beta(0) = 0.1$ is the same as used in Fig. 2. One can see that for small ω the Darwin strategy (corresponding to $\alpha \to 0$) is preferable.

We also want to note that from an investigation of the symmetric double-well potential (Dunkel et al., 2003), one can learn that for low temperatures D corresponding to high barriers of the reduced potentials U/D, searchers following a Darwin strategy with global selection are more efficient, if a barrier must be surmounted. One should expect that this property is kept for high-dimensional problems too.

With regard to these results, we would therefore recommend the following combination of the two strategies:

1. Use the Darwin strategy (with global selection) in the beginning of the search process.

2. Switch to the Boltzmann strategy (with small $D \to 0$) at the final stages.

In terms of a combined strategy as introduced in Section 3, this proposal (except from the variation of D) could also be expressed using a time-dependent coupling constant

$$\alpha(t) = \Theta(t - t') \qquad\qquad t' \gg 0, \tag{71}$$

where Θ denotes Heaviside's step-function

$$\Theta(y) = \begin{cases} 0, & \qquad y < 0, \\ 1, & \qquad y \geq 0. \end{cases} \tag{72}$$

Of course, instead of using the step-function, one could also think of a continuous variation of $\alpha(t)$, starting at $\alpha(0) = 0$ and converging to $\alpha = 1$ for $t \gg 0$. Such a variation of α could then also be directly connected with a continuous decrease of the temperature D as widely applied in the field of 'Simulated Annealing' .

Finally, due to the special shape of the generalized Boltzmann ansatz, the relaxation of related solutions to the stationary state could also be interpreted as a self-consistent change of temperature.

Acknowledgments

The authors are very grateful to H. Mühlenbein, T. Mahnig and S. Hilbert for valuable discussions. They also want to thank two unknown referees for their suggestions. This work was financially supported by the DFG via the Sfb 555 (L. S.-G. and J. D.) and by the Studienstiftung des deutschen Volkes (J. D.).

References

Andresen, B. and Gordon, J. M. (1993). Analytic constant thermodynamic speed-cooling strategies in simulated annealing. *Open Systems & Information Dynamics in Physical and Life Sciences*, 2(1):1–12.

Andresen, B. and Gordon, J. M. (1994). Constant thermodynamic speed for minimizing entropy production thermodynamic processes and simulated annealing. *Physical Review E*, 50:4346–4351.

Asselmeyer, T. and Ebeling, W. (1996). Unified description of evolutionary strategies over continuous parameter spaces. *BioSystems*, 41:167–178.

Asselmeyer, T., Ebeling, W., and Rosé, H. (1997). Evolutionary strategies of optimization. *Physical Review E*, 56:1171–1180.

Boseniuk, T., Ebeling, W., and Engel, A. (1987). Boltzmann and Darwin strategies in complex optimization. *Physics Letters*, 125:307–310.

Chandrasekhar, S. (1943). Stochastic problems in physics and astronomy. *Review of Modern Physics*, 15(1):1–89.

Dunkel, J., Ebeling, W., Schimansky-Geier, L., and Hänggi, P. (2003). The Kramers problem in evolutionary strategies. *Physical Revue E*. Accepted for publication.

Ebeling, W., Engel, A., and Feistel, R. (1990). *Physik der Evolutionsprozesse*. Akademie-Verlag Berlin.

Eigen, M. (1971). The Selforganization of Matter and the Evolution of Biological Macromolecules. *Die Naturwissenschaften*, 58(10):465–523.

Einstein, A. and von Smoluchowski, M. (1999). *Untersuchungen über die Theorie der Brownschen Bewegung/Abhandlungen über die Brownsche Bewegung und verwandte Erscheinungen*, volume 199. Harri Deutsch, Frankfurt, 3 edition.

Feistel, R. and Ebeling, W. (1989). *Evolution of Complex Systems*. Kluwer Academic Publisher.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford University Press.

Fogel, D. B. (1995). *Evolutionary computation - towards a new philosophy of machine intelligence*. IEEE Press, Piscataway NJ.

Hänggi, P., Talkner, P., and Borkovec, M. (1990). Reaction rate theory: fifty years after Kramers. *Review of Modern Physics*, 62(2):251–341.

Holland, J. H. (1975). *Adaption in Natural and Artificial Science*. University of Michigan Press, Ann Arbor.

Kirkpatrick, S., Jr., C. D. G., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220:671–680.

Koza, J. R. (1992). *Genetic Programming: On the programming of computers by means of natural selection*. MIT Press, Cambridge.

Mahnig, T. and Mühlenbein, H. (2001). A new adaptive Boltzmann selection schedule SDS. In *Proceedings of the 2001 Congress on Evolutionary Computation*, pages 183–190, Piscataway. IEEE Press.

Rechenberg, I. (1973). *Evolutionsstrategien - Optimierung technischer Systeme nach Prinzipien der biologischen Information*, volume Stuttgart-Bad Cannstatt. Friedrich Frommann Verlag (Günter Holzboog K.G.).

Rosé, H. (1996). Von Boltzmann und Darwin zum Self Annealing - Strategien evolutionärer Optimierung. In Freund, J. A., editor, *Dynamik, Evolution, Strukturen - Nichtlineare Dynamik und Statistik komplexer Strukturen*, volume Wissenschaftliche Schriftenreihe Physik 50, pages 29–38. Verlag Dr. Köster, Berlin, 1 edition.

Rosenkranz, D. (1996). Anwendung evolutionärer Algorithmen auf quantenmechanische Probleme. In Freund, J. A., editor, *Dynamik, Evolution, Strukturen - Nichtlineare Dynamik und Statistik komplexer Strukturen*, volume Wissenschaftliche Schriftenreihe Physik 50, pages 79–88. Verlag Dr. Köster, Berlin, 1 edition.

Schweitzer, F., Ebeling, W., Rosé, H., and Weiss, O. (1998). Optimization of road networks using evolutionary strategies. *Evolutionary Computation*, 5(4):419–438.