Error Propagation in Reproduction of Diploid Organisms

A Case Study on Single Peaked Landscapes *

THOMAS WIEHE[1], ELLEN BAAKE[2], and PETER SCHUSTER[1,3]

[1]Institut für Molekulare Biotechnologie, Beutenbergstr. 11, D-07745 Jena, Germany.
[2]Max-Planck-Institut für Entwicklungsbiologie, Spemannstr. 35/IV, D-72076 Tübingen, Germany.
[3]Institut für Theoretische Chemie der Universität Wien, Währingerstraße 17, 1090 Wien, Austria, and The Santa Fe Institute, Santa Fe, NM 87501, USA

(Running head: Error Threshold and Diploid Reproduction)

Abstract

We study two versions of the diploid selection mutation equation as adapted to sequence space. Focussing on diploid generalizations of the well-established single peaked landscape, we find quantitative effects of dominance on error thresholds, as well as unexpected qualitative features like multiple equilibria, in infinite populations. Analogues of these phenomena are recovered in stochastic versions for finite populations as well.

1 Introduction

Mutation is an unavoidable side reaction of reproduction or the inevitable consequence of some external impact. Accordingly, the combined effect of mutation and selection on the distribution of allelic frequencies received detailed attention already in the early theoretical investigations by Fisher, Haldane and Wright (e.g. Haldane, 1927, Fisher, 1929, Wright, 1929; review in Ewens, 1979). The development of the well known formulae of mutation selection balance in these papers was from the very beginning related to the discussion of the evolution of dominance.

However, the theoretical models received fundamental revision with the discovery of the molecular structure of DNA (Watson & Crick, 1953) and the deciphering of the genetic code in the sixties. In particular, the latter event had an immediate impact on models in theoretical population genetics, which now started to reflect molecular properties of the genetic material. In this vein, the infinite allele (Kimura & Crow, 1964) and the infinite site model (Kimura, 1969) were introduced.

The process of producing offspring could now also be interpreted as a replication process of macromolecules with a certain sequence of nucleotides. This view has been formalized in the language of biochemical reaction kinetics by Eigen (1971). His model may be viewed as a haploid

*Dedicated to Prof. A. Gierer on occasion of his 65th birthday selection mutation equation which takes account of the primary structure of the underlying polymers. Eigen (1971) considered correct replication and mutation as parallel biochemical reactions within the same general mechanism. In this framework, polynucleotide strings such as DNA or RNA are identified with points in sequence space which in the case of binary strings is the discrete hypercube of dimension ν, where the integer ν is the (fixed) length of the string considered. A natural metric on the hypercube is the Hamming metric $d(A_i, A_j)$, which counts the number of different positions in two aligned strings A_i and A_j. Mutation frequencies were introduced in a way that was later characterized as the uniform error rate model (Eigen & Schuster, 1977). They are fully determined by an accuracy parameter (q), the chain length (ν) of the polynucleotide, and the distance in sequence space between the correct (i.e., wild type) copy and the mutant which is tantamount to the Hamming distance between the two sequences.

The dynamics of the frequency distribution of alleles in a population under the interplay of correct and erroneous replication as well as selection may thus be viewed as a process of *error propagation*, an outstanding feature of which is the loss of genetic information when a critical error rate is surpassed.

For diploid organisms, selection mutation equations have been studied in the more general context of population genetics (i.e., without reference to sequence space) as "Fisher's equation with mutation" in two different versions (Crow & Kimura, 1970; Hadeler, 1981).

In this paper, we focus on these equations, with the mutation mechanism adapted to sequence space. We will show that, depending on the properties of the selection matrices, properties of the haploid sequence space models carry over, or unexpected behaviour may emerge.

The paper is organized as follows. In Section 2, we comment on the (deterministic) model equations and discuss their possible equilibrium states in Section 3. In Sections 4 and 6, we analyze effects of diploidy on error thresholds in infinite and finite populations, respectively. Section 5 treats convergence times to equilibria.

2 Model equations

One–locus multiple–allele models describing the interplay of selection and mutation were formulated in two different versions, depending on the assumptions on the mutation mechanism. If mutations originate as replication errors on the occasion of reproduction events, the appropriate ODE system for the composition of an (infinite) diploid population with overlapping generations in Hardy-Weinberg equilibrium reads

$$\dot{x}_i = \sum_j \left(m_{ij}\, x_j (Wx)_j\right) - x_i(x, Wx)\ , \qquad i = 1, \ldots, n\,. \tag{1}$$

Here, n is the number of alleles ($n = \kappa^\nu$, where e.g. $\kappa = 2$ in binary sequence space, $\kappa = 4$ in case of natural DNA or RNA sequences), and the x_i denote the relative frequencies of alleles A_i (with $\sum_i x_i = 1$). W is the symmetric $n \times n$ matrix of fitnesses (to be interpreted as reproduction rates) of (ordered) $\mathrm{A}_i\mathrm{A}_j$ genotypes. $(Wx)_i := \sum_{j=1}^n w_{ij} x_j$ is the marginal fitness of allele A_i and $(x, Wx) := \sum_{j,k=1}^n w_{jk} x_j x_k$ denotes the average fitness of the population. It is assumed that all genotypes have identical death rates; so, due to normalization, these do not contribute to the equations. M is the mutation matrix with elements $m_{ij} := m_{i \leftarrow j}$ denoting the *probability* that an A_j allele is replicated as A_i, with $\sum_i m_{ij} = 1$.

Let us recall that the assumption of Hardy-Weinberg equilibrium is an approximation in the case of overlapping generations, even if random mating is assumed (Hadeler, 1974). Equation (1), called the 'coupled selection mutation equation' in what follows, was first studied by Hadeler (1981).

With the transformation of time: $d\tau := (x(t), Wx(t))dt$, $\tau = \int_0^t (x(\theta), Wx(\theta))d\theta$ may be interpreted as the time scale of generations, with 'generation' to be understood as the time span it takes the population to grow to e times its size (e is Euler's constant), where deaths are not counted. This yields

$$\frac{dx_i}{d\tau} = \frac{\sum_j m_{ij}\, x_j (Wx)_j}{(x, Wx)} - x_i\,. \tag{2}$$

A formally identical equation has been arrived at in passing from the difference equation for non-overlapping generations to an ODE (Ewens, 1979). (In this case, however, fitnesses must be interpreted as viability, or number of offspring contributed, to the next – discrete – generation.)

If one relaxes the assumption that mutation occurs at the time of replication but stipulates that it may occur at any time in the life cycle of a gamete, mutation and selection must be considered as independent events. One is then led to the following decoupled version of the selection mutation equation

$$\dot{x}_i = x_i \left((\tilde{W}x)_i - (x, \tilde{W}x) \right) + (\tilde{M}x)_i \qquad i = 1, \ldots, n\,, \tag{3}$$

which was formulated by Crow & Kimura (1970), and thoroughly investigated by Akin (1979). Here, $\tilde{m}_{ij}$ denotes the mutation *rate* from A_j to A_i for $i \neq j$, and $\tilde{m}_{ii} = -\sum_j \tilde{m}_{ji}$. The entries in the symmetric fitness matrix $\tilde{W}$ are now to be interpreted as Malthusian fitnesses; in particular, the $\tilde{w}_{ij}$ need no longer be positive. However, the equation is invariant under the transformation $\tilde{w}_{ij} \to \tilde{w}_{ij} + c$ for an arbitrary constant c, if performed for all index pairs i, j simultaneously. Thus, it is no restriction to consider $\tilde{W}$ to be a positive matrix as well.

Hofbauer (1985) discusses the relationship between the versions. He shows in particular that eqn (3) is the limiting version of eqn (1) in the limit of small selective differentials and small mutation rates. As far as error thresholds are concerned, mutation rates need, however, not be sufficiently small as to neglect the difference between the two models.

¿From the biological point of view, the relative contributions of the underlying mutation mechanisms are controversial. A mechanism described by eqn (1) implies constancy of mutation rates per generation, whereas eqn (3) is suited to describe systems with mutation rates which are constant per year. 'Real life' evolution does, of course, have contributions from both replication errors and independent processes like radiation damage that become manifest at replication time but occur with constant intensity in time. Experimental evidence concerning the proportions of these contributions is as controversial (see, e.g., Kimura, 1987, Ohta, 1993) as is the existence of a molecular clock (Gillespie, 1991). This issue still being unresolved, we will analyze both versions here in a parallel fashion.

As mentioned already, we identify 'alleles' with nucleotide sequences of length ν. The mutation matrices M and $\tilde{M}$ are specified to be of sequence space type. For the coupled model this means

$$m_{ij} = \left(\frac{p}{\kappa - 1}\right)^{d_{ij}} (1-p)^{\nu - d_{ij}}\,, \tag{4}$$

where $d_{ij} := d(A_i, A_j)$ and p is the mutation probability per nucleotide. We assume that mutations occur independently at any site and that, given a mutation at a site occurs, any other nucleotide is equally likely to be incorporated instead of the original one.

The mutation matrix for the decoupled equation, on the other hand, must exclude double mutations, i.e.:

$$\tilde{m}_{ij} = \begin{cases} \frac{\mu}{\kappa-1} & , \quad d_{ij} = 1 \\ -\nu\mu & , \quad i = j \\ 0 & , \quad \text{otherwise} \end{cases} \tag{5}$$

Here, μ is the per nucleotide mutation *rate*.

In the sequel, we will concentrate on binary sequence space and assume $\kappa = 2$ for the rest of our considerations. This assumption does not severely restrict the validity of our investigations since any alphabet may be translated into a binary one.

Let us briefly comment on how to compare p and μ. In the case $w_{ij} \equiv r$ (no selection), eqns (1) and (3) reduce to

$$\dot{x} = r(M - I)x \tag{6}$$

and

$$\dot{x} = \tilde{M}x\,, \tag{7}$$

respectively, where I denotes the identity matrix. $\tilde{M}$ and $M - I$ are symmetric, they commute and may thus be diagonalized simultaneously. From the explicit knowledge of the spectral properties of M as analyzed by Rumschitzky (1987), it follows that $r(M - I)$ and $\tilde{M}$ have corresponding eigenvalues $r\left((1-2p)^{\nu-\rho} - 1\right)$, and $2\mu(\rho-\nu)$, respectively, both with multiplicity $\binom{\nu}{\rho}$, $\rho = 0, \ldots, \nu$. So, obviously, eqn (7) is the first-order perturbation expansion in p of eqn (6) when p is identified with μ/r.

In order to be able to compare haploid and diploid models, it is essential to know for which fitness schemes the diploid equations reduce to their haploid counterparts. As noted in (Hofbauer & Sigmund, 1988), the dynamics of eqn (3) reduces to that of a haploid population if the fitness of heterozygotes is (*arithmetically*) intermediate between the corresponding homozygotes, i.e., $\tilde{w}_{ij} = \frac{1}{2}(\tilde{w}_{ii} + \tilde{w}_{jj})$. For, in this case, the difference between marginal fitness of allele A_i and mean fitness is

$$(\tilde{W}x)_i - (x, \tilde{W}x) = \frac{1}{2}(\tilde{w}_{ii} - \sum_j \tilde{w}_{jj}\, x_j), \tag{8}$$

so

$$\dot{x}_i = x_i(\tilde{r}_i - (\tilde{r}, x)) + (\tilde{M}x)_i\,. \tag{9}$$

Here, $\tilde{r}_i := \frac{1}{2}\tilde{w}_{ii}$ may be interpreted as the fitnesses of single alleles which make up a haploid population and $(\tilde{r}, x)$ as the corresponding mean fitness $\sum_i \tilde{r}_i x_i$.

For the coupled equation (in the form (2)), on the other hand, the diploid dynamics reduces to the haploid case, if heterozygote fitness equals the *geometric* mean of the corresponding homozygote fitness values, i.e. $w_{ij} = \sqrt{w_{ii}w_{jj}} =: r_i \cdot r_j$. In this case, we have

$$\frac{dx_i}{d\tau} = \frac{\sum_j m_{ij} r_j x_j}{(r, x)} - x_i\ . \tag{10}$$

3 Equilibria

It has been shown (Thompson & McBride, 1974; Jones *et al.*, 1976) that, via the transformation $z_i := x_i \exp(\int_0^t (r, x(\tau))d\tau)$, the second degree ODE's describing the coupled haploid replication dynamics can be transformed into a set of linear differential equations. Thus, the problem of solving the haploid system reduces to the spectral problem of the matrix $M \cdot R$ (where $R := \mathrm{diag}\{r_1, \ldots, r_n\}$). Furthermore, for $p > 0$ it follows from the positivity of $M \cdot R$ and the Perron-Frobenius theorem (see for example Karlin & Taylor, 1975, Heuser, 1992) that the ODE system has a unique, globally stable equilibrium in the interior of the $n-1$ dimensional simplex S_n of relative frequencies.

The same results hold for the decoupled equation, as the above considerations go through for $\tilde{M} + \tilde{R}$, $\tilde{R} := \mathrm{diag}\{\tilde{r}_1, \ldots, \tilde{r}_n\}$ instead of $M \cdot R$ ($\tilde{M} + \tilde{R}$ itself is not a positive matrix, but $(\tilde{M} + \nu\mu I + \tilde{R})^\nu$ is, if $\mu > 0$).

A similar transformation to reduce the degree of the system is not known for the diploid case. Even worse, the number of equilibria and their stability properties depend crucially upon the entries of the fitness matrix W and $\tilde{W}$, respectively (detailed analyses may be found in e.g. Karlin, 1980, Kingman, 1988). For the decoupled equation, fitness matrices may be constructed such that oscillating behaviour of the trajectories may emerge (Akin, 1979). The resulting fitness landscapes have more than one maximum on S_n in all cases considered. In Section 4, we will show that, even for fitness landscapes with a unique maximum, multiple equilibria may occur, which rules out the possibility of global convergence even for such situations.

A perturbation analysis of a general replication–mutation system has been studied by Stadler & Schuster (1992). Their rest point migration theorem gives quantitative and qualitative features of equilibria in the interior of the simplex once those of the unperturbed system (i.e. $p = 0$) are known, in particular transversality properties. However, the perturbation approach is satisfactory only for small p, which is not fulfilled in the cases we are interested in. In particular, numerical calculations for a single peaked fitness landscape show that second order perturbation expansions of eqn (1) are highly inaccurate for mutation probabilities close to the error threshold. Moreover, we will see that, in diploid situations, perturbation theory does not reveal all relevant equilibria.

4 Effects of Dominance on Error Thresholds

In haploid sequence space models, error thresholds were studied extensively by means of single-peaked fitness landscapes, as characterized by one 'fit' master allele (denoted by A_0) and competitors whose fitness is lower but equal for all of them. We use the term "master allele" synonymously with "allele with highest fitness".

It is well known that in this case the (unique) stationary distribution of relative frequencies exhibits a sharp transition into a distribution which is approximately uniform as the nucleotide mutation rate surpasses a critical value p_{max}. Such a transition has been identified as an "error threshold" by Swetina & Schuster (1982), Nowak & Schuster (1989), and elsewhere.

As a well-suited diploid analogue to the haploid single-peaked landscape, we propose the following model. All genotypes are classified according to whether they are

a. homozygous of (A_0)–(A_0) type with fitness f_0,

b. heterozygous of (A_0)–(non-A_0) type with fitness f_1 or

c. homozygous of (non-A_0)–(non-A_0) type with fitness f_2.

We express the f_i in terms of the selective advantage s and a dominance parameter h:

$$
\begin{aligned}
f_0 &= w_{11} = (1+s)^2\,, \\
f_1 &= w_{1i} = w_{i1} = (1+s)^{2h}\,,\ i \neq 1 \ \text{ and} \\
f_2 &= w_{ij} = 1\,,\ i,j \neq 1
\end{aligned}
$$

and for the decoupled equation

$$
\begin{aligned}
f_0 &= \tilde{w}_{11} = 1+2s\,, \\
f_1 &= \tilde{w}_{1i} = \tilde{w}_{i1} = 1+2hs\,,\ i \neq 1 \ \text{ and} \\
f_2 &= \tilde{w}_{ij} = 1\,,\ i,j \neq 1\,.
\end{aligned}
$$

For $0 \leq h \leq 1$, this defines a single-peaked landscape in the above sense. In particular, for $h = \frac{1}{2}$ the fitness regimes reduce to the multiplicative and additive cases, respectively. It is worth mentioning that both fitness schemes coincide to first order in s.

Like its haploid counterpart, the diploid single-peaked landscape allows reduction of systems (1) and (3) to ODE systems of $\nu + 1$ equations for $y_0, \ldots, y_\nu$, where y_i denotes the relative frequency of alleles carrying i mutations with respect to the master allele. In general, this reduction is possible if $d_{i1} = d_{k1}$ together with $d_{j1} = d_{l1}$ imply $w_{ij} = w_{kl}$ for all $i,j,k,l \in \{1,\ldots,n\}$. This *homogeneity condition* is clearly satisfied for our single peaked landscape. The mutation matrices for the Hamming class formulation, M^H and $\tilde{M}^H$, read

$$
m_{ij}^H = (1-p)^\nu \sum_{k=j-\nu+i}^{\min(i,j)} \binom{j}{k} \binom{\nu-j}{i-k} \left(\frac{p}{1-p}\right)^{i+j-2k} \qquad (i,j = 0,\ldots,\nu) \tag{11}
$$

(cf. Swetina & Schuster, 1982), and

$$
\tilde{M}^H = \mu \begin{pmatrix} -\nu & 1 & & & 0 \\ \nu & -\nu & 2 & & \\ & \ddots & \ddots & \ddots & \\ & & 2 & -\nu & \nu \\ 0 & & & 1 & -\nu \end{pmatrix}. \tag{12}
$$

Let us resort to the single-peaked example to discuss the equilibrium properties mentioned previously. Bürger (1983) has shown that, in cases where the selection vector field alone yields a globally stable stationary state, the decoupled selection mutation equation may have more than one stable equilibrium if the mutation matrix is non-symmetric. We will show now that counterexamples to global stability with single-peaked landscapes may also be found when mutation matrices are of sequence space type (and therefore symmetric in the formulation of eqns (1) or (3)). Let us start with the slightly simpler decoupled equation.

With the choice $h = 0$, the equilibrium condition (in Hamming class notation) reduces to

$$
(\tilde{M}^H y)_i = 2sy_0^2 \cdot \begin{cases} (y_0 - 1)\,, & i = 0 \\ y_i\,, & \text{otherwise.} \end{cases} \tag{13}
$$

Now consider the inhomogeneous system of linear equations arising from eqn (13) by formally replacing $2sy_0^2$ by α:

$$
\tilde{M}^H y = \alpha(y - e_1) \tag{14}
$$

with parameter α, where $e_1 := (1, 0, \ldots, 0)^T$. As long as $\alpha > 0$, it has a unique solution $\bar{y}(\alpha)$ (since the eigenvalues of $\tilde{M}^H$ are ≤ 0). From Cramer's rule, its components have the form

$$y_i(\alpha) = \frac{P_i(\alpha)}{Q(\alpha)}\,,\ i = 0, \ldots, \nu\,, \tag{15}$$

where $P_i(\alpha), Q(\alpha)$ are polynomials in α of degree at most $\nu + 1$. Specializing to $\nu = 4$ and substituting $\alpha = 2sy_0^2$, we obtain equilibrium values of y_0 as roots of the following polynomial $P(y_0)$, defined as

$$\begin{aligned} P(y_0) &:= y_0 Q_0(2sy_0^2) - P_0(2sy_0^2) \\ &= -3 + 48y_0 - 32\sigma y_0^2 + 100\sigma y_0^3 - 40\sigma^2 y_0^4 \\ &\quad +70\sigma^2 y_0^5 - 16\sigma^3 y_0^6 + 20\sigma^3 y_0^7 \\ &\quad -2\sigma^4 y_0^8 + 2\sigma^4 y_0^9\,, \end{aligned}$$

where $\sigma := s/\mu$. We have thus reformulated the problem of finding solutions of eqn (14) to finding roots of a 9th degree polynomial. The remaining entries of $(y_0, \ldots, y_\nu)^T$ are then uniquely determined by eqn (15). We have $P(0) < 0$ and $P(1) > 0$, independently of σ. Choosing $\sigma = 8$, $P(y_0)$ is easily seen to have at least three roots in $[0, 1]$. We have thus constructed a single-peaked situation with (at least) three equilibria.

For the coupled equation, the situation is hardly more complicated. The equilibrium condition now reads (with $\alpha := (2 + s)sy_0^2$)

$$(M^H y)_i = (1 + \alpha)y_i - \alpha\, m_{i0}^H\,,\ i = 0, \ldots, \nu\,, \tag{16}$$

and we obtain, again, situations with multiple equilibria, e.g., for $\nu = 4$, $p = 1/20$, and $s = 9/20$.

We remark that even if the reduced system (describing Hamming class dynamics) is in equilibrium, this need not be the case for the full systems (1), (3). Since our concern here is the Hamming class dynamics, this fact does, however, not pose a problem for the treatment of the following questions.

In order to study diploid error threshold behaviour, we numerically follow branches of equilibria as the mutation rate is increased. We start with $p = 0$ ($\mu = 0$, respectively) and a homogeneous population consisting of master strings only. The curves in Figure 1 represent stationary frequencies of Hamming classes. A transition into the (approximate) uniform distribution for $p > p_{max}$ is evident. In analogy to the haploid case we call such a transition an "error threshold". However, in the diploid situation the location of the transition depends on the dominance parameter h.

For $h = \frac{1}{2}$, haploid and diploid thresholds coincide. On the other hand, recessivity ($h < \frac{1}{2}$) or dominance ($h > \frac{1}{2}$) of the master cause shifts of the error threshold with respect to the haploid calibration in opposite directions; see Figure 2, and Table 1 for numerical values (for the latter, we used equilibrium master frequency $< 1\%$ as the threshold criterion). In the case of complete dominance, the critical mutation rate is roughly doubled with respect to the haploid case, and the transition becomes smoother; recessivity, on the other hand, shifts the threshold to lower values and sharpens it. So a master which is dominant in heterozygotes can tolerate much higher mutation rates before it is lost from the population than when it is intermediate or even recessive in heterozygotes. We will come back to this point later.

Approximate values for the location of the threshold can be derived in an way similar to that employed by Eigen (1971). We consider a caricature of the original model by neglecting back mutation and concentrate on the differential equation for the master frequency x_1. At stationarity, we are left with

$$0 = x_1(Wx)_1 m_{11} - x_1(x, Wx)\,. \tag{17}$$

Its nontrivial solution $\bar{x}_1 \neq 0$ is characterized by

$$m_{11} = \frac{(x, Wx)}{(Wx)_1}\,, \tag{18}$$

where $m_{11} = (1-p)^\nu$, $(Wx)_1 = x_1(1+s)^2 + (1-x_1)(1+s)^{2h}$, and $(x, Wx) = 1 + 4hsx_1 + 2sx_1^2(1-2h)$. We find that a branch of $\bar{x}_1$ intersects the p-axis when

$$p = p_{max} := 1 - \left(\frac{1}{(1+s)^{2h}}\right)^{\frac{1}{\nu}}\,. \tag{19}$$

The bifurcation structure will be analyzed in a forthcoming paper.

The corresponding quantity for the haploid case is

$$p_{max} = 1 - \left(\frac{1}{1+s}\right)^{\frac{1}{\nu}}\,. \tag{20}$$

Similarly, for the decoupled system (3) one obtains

$$\mu_{max} = \frac{2sh}{\nu}\,, \tag{21}$$

which makes sense only for $h \geq 0$.

Crude as these approximations certainly are, they demonstrate that, in the coupled case, the ratio of homo- and heterozygote non-master fitness is a relevant quantity, whereas it is their difference in the decoupled case. With fitness parameters to be interpreted as reproduction rates in both cases, the above relations are consistent with the interpretation of μ as a mutation *rate* (and μ_{max} a difference of rates) and p as a mutation *probability* (and p_{max} a ratio of rates, hence dimension-less). Also, it is consistent with the invariance properties of the equilibria (equilibria of eqn (1) are invariant with respect to $w_{ij} \to w_{ij} \cdot c\ \forall i, j$, whereas those of (3) are invariant under $w_{ij} \to w_{ij} + c$).

We give a list of threshold values for a choice of parameters in Table 1. Dominance not only affects the location of the error threshold, but also tunes the variability of a population. If population quantities, like *average Hamming distance from master* or *variance of Hamming distance from master* are compared for different values of h and plotted as functions of p (similar results hold for the decoupled equation), one observes an inflection at $p = \widehat{p}$ and three characteristic parameter regions (see Figures 2 and 3): A region of very exact replication $p \in I_0 = [\,0, \widehat{p}\,]$, an intermediate one with $p \in I_1 = [\,\widehat{p}, p_{max}\,]$ and the region beyond the error threshold I_2 with $p > p_{max}$. For small p, variability (measured in terms of expectation and variance of Hamming distance) at equilibrium is higher for dominant master alleles and lower for recessive ones. For intermediate p ($p \in I_2$) the situation is reversed. Here, dominance leads to a relatively lower variability and a higher concentration of the population around the master allele. Whereas for $p > \widehat{p}$ dominance 'shields' the master from a quick loss from the population, for values $p < \widehat{p}$ recessivity is needed to keep the master allele at highest possible concentrations. We calculated the inflection point $\widehat{p}$, where the effects of dominance and recessivity are reversed to be approximately

$$\widehat{p} = 1 - \left(\frac{\sqrt{1+2s}-1}{s}\right)^{\frac{1}{\nu}}\,. \tag{22}$$

Expanding in a Taylor series and retaining terms of order up to $O(s^2)$ one derives

$$\widehat{p} \approx \frac{s}{2\nu}(1+\frac{s}{4}) \ . \tag{23}$$

Comparing *mean fitnesses* for different values of h yields a different picture, however. In a way similar to Chao (1988), we define a quantity $\bar{w}_{rel}$, the relative difference between *equilibrium* mean fitness as a function of h and the intermediate calibration $h = 1/2$ (corresponding to the haploid reference case)

$$\bar{w}_{rel} = \bar{w}_{rel}(h) := \frac{(\bar{x}, W(h)\bar{x}) - (\bar{x}, W(1/2)\bar{x})}{(\bar{x}, W(1/2)\bar{x})} \ . \tag{24}$$

The results as plotted in Figure 4 suggest that the advantage of dominance ($h > 1/2$) increases almost linearly with increasing mutation rates and reaches a maximal advantage at the threshold value $p_{max}|_{h=1/2}$, after which the advantage decreases. The disadvantage of recessivity ($h < 1/2$, where we choose the upper branch when multiple equilibria are present) on the other hand, increases in a nonlinear way. After its threshold value $p_{max}(h)$, it decreases and vanishes entirely at $p_{max}|_{h=1/2}$. The fact that, for $h < 1/2$ and small mutation rates, mean fitness is almost identical to that of $h = 1/2$ may explain why recessive alleles may be found at all in a population and why they are not eliminated right away. On the other hand, the advantage of dominance for small mutation rates (as one has to hypothesize for natural diploid populations) might give another argument for the preponderance of dominant alleles, as throughout assumed in the discussions of evolution of dominance (Fisher, 1928).

5 Time dependent behaviour

For the haploid dynamics the eigenvalue spectrum of the matrix $M \cdot R$, as a function of p, shows a series of 'avoided crossings' as pointed out by Nowak & Schuster (1989). In particular, the avoided crossing of the first and second largest eigenvalues coincides with the error threshold, indicating a singularity of convergence times of the dynamics in this region, since the ratio of λ_1 and λ_2 approaches unity [1]. If one identifies the error threshold with a first-order phase transition (as has been done by Leuthaeusser (1987)), this corresponds to the singularity of the correlation length that goes together with the phase transition.

In the diploid replication regime, a similar explosion of convergence times is observed. Due to the inherent non–linearity of systems (1), (2) and (3) an eigenvalue characterization of convergence times is, however, not readily available in the diploid case.

We present results of numerical integrations of system (2) (for consistency with the haploid case) and analytical approximations. For the latter we again neglect back mutation. We concentrate on the coupled equation in the following; very similar results are obtained when the decoupled equation is treated along the same lines. For the master in eqn (2), one has the following rational first order autonomous ODE

$$\frac{dx_1}{d\tau} = \frac{x_1 (Wx)_1 m_{11}}{(x, Wx)} - x_1 \ . \tag{25}$$

We concentrate on the initial value $x_1(0) = 1$. To derive convergence times, it suffices to integrate the reciprocal of the right hand side of (25) with respect to x_1 (separation of variables). For the

[1] These eigenvalues must *avoid* each other because of the uniqueness of the largest one.

time $\bar{\tau}$ it takes for the master to reach its equilibrium value $\bar{x}_1 + \epsilon$ starting from $x_1(0) = 1$, this yields

$$\bar{\tau} = \int_1^{\bar{x}_1+\epsilon} \frac{(\xi, W\xi)}{\xi(Wx)_1 m_{11} - \xi(\xi, W\xi)} d\xi \,. \tag{26}$$

Although the integral in eqn (26) can be evaluated for arbitrary h, for simplicity we present here only the result for $h = \frac{1}{2}$ explicitly. In this case we derive

$$\bar{\tau} = \frac{m_{11}(1+s)\log(\frac{1+s(\bar{x}_1+\epsilon)-m_{11}(1+s)}{1+s-m_{11}(1+s)}) - \log(\bar{x}_1+\epsilon)}{1-m_{11}(1+s)} \,. \tag{27}$$

The right hand side of (27) has a singularity at the threshold value $p = p_{max}$. This fact again emphasizes the important role of the error threshold. For the mutation force being that strong, evolution comes practically to a standstill, no matter how the composition of the population in detail is. This is true also for other choices of h.

In Figure 5, we show convergence times for different initial conditions as obtained by numerical integration of the entire system (2). To this end, we first integrated the ODE system over a long period of time ($t = 10^5$) to determine equilibria $\bar{x}$. In a second run, we measured the time the system needed to approach $\bar{x}$ s.t. $|x_i - \bar{x}_i| \leq \epsilon$ for all i, with $\epsilon = 10^{-5}$. Important appears the fact that convergence times have a minimum for some value p_{min} with the property $0 < p_{min} < p_{max}$, if 'worst case' initial conditions are assumed, i.e. $x_n(0) = 1$. A new advantageous allele, out-competing all other alleles and at a maximum distance from the current – homogeneous – population composition is established most quickly for definitely positive mutation rates around p_{min}. In a rapidly changing environment such a fast convergence to new equilibria is an essential requirement for quick adaptation.

Intuitively, one could expect that a dominant master allele is faster at establishing itself in the population than is an intermediate or even recessive one. In contrast to this supposition we find that convergence times – for the caricature as well as for the entire system – do in fact depend only very weakly on any dominance properties, except near the poles.

6 Finite Populations

So far, the effect of random genetic drift has been totally neglected. Dealing with reproduction of diploid organisms in a general setting, the effects of finite population size (N in the haploid case, $2N$ in the diploid case) on the composition of the gene pool can in general not be ignored, especially when considering sequence space. Here, even at moderate chain lengths, the number of possible individuals by far exceeds any realistic population size. We give accounts of the influence on the error threshold due to finite population size. Adapted to the two systems (1) and (3) are two different approaches to incorporate random drift. As pointed out previously, the coupled equation with mutation occurring at replication is closely related to the discrete generation model, which motivates introduction of stochasticity by classical multinomial Wright–Fisher sampling and the according treatment as a time homogeneous diffusion process. For the decoupled equation, however, stochastic fluctuations are more suitably described by a time homogeneous birth and death process. As in previous chapters, we perform numerical simulations of the entire system, along with an analytic treatment of a simpler caricature. In both cases, we are interested in stationary distributions. For these to exist, none of the states may be absorbing. Instead of neglecting back-flow altogether, we now resort to the caricature introduced by Nowak & Schuster (1989): We collect all Hamming classes except the master into a single ODE, called 'error tail', and assume alleles to be equidistributed within it. We thus obtain reduced mutation matrices

$$M^R = \begin{pmatrix} (1-p)^\nu & \frac{1-(1-p)^\nu}{2^\nu-1} \\ 1-(1-p)^\nu & 1-\frac{1-(1-p)^\nu}{2^\nu-1} \end{pmatrix}, \tag{28}$$

and

$$\tilde{M}^R = \begin{pmatrix} -\nu\mu & \frac{\nu\mu}{2^\nu-1} \\ \nu\mu & -\frac{\nu\mu}{2^\nu-1} \end{pmatrix}. \tag{29}$$

We emphasize that the former reduction from originally n to $\nu+1$ equations leaves the dynamics of the Hamming classes unaffected as long as the homogeneity condition (chapter 4) is met. On the other hand, this second reduction (as suggested by the single peaked landscape) is an approximation.

6.1 Coupled case: Diffusion

We consider a one-dimensional diffusion process, representing the relative frequency X_1 of the master allele in the population. The frequency of the error tail is obtained by $X_2 = 1 - X_1$. X_1 and X_2 are random variables on the unit interval $[0,1]$. Infinitesimal drift and diffusion coefficients $a(x)$ and $b(x)$ are those for a two-allele model in population genetics and their derivation is standard (see e.g. Ewens, 1979). For our landscape $w_{11} = (1+s)^2, w_{12} = w_{21} = (1+s)^{2h}$, $w_{22} = 1$ (as well as for its approximation $w_{11} = 1+2s$, $w_{12} = w_{21} = 1+2hs$, $w_{22} = 1$) they read

$$a(x) = 2sx(1-x)(x+h(1-2x)) + (1-x)m_{12}^R - xm_{21}^R \tag{30}$$

as drift and

$$b(x) = \frac{1}{2N}x(1-x) \tag{31}$$

as diffusion coefficient. Specializing to $h = \frac{1}{2}$ and substituting N for $2N$ yields drift and diffusion for the haploid case with fitness scheme $r_1 = 1+s$ and $r_2 = 1$. Constraints on the order of magnitude of the parameters s, m_{12}^R and m_{21}^R as discussed by Ethier & Kurtz (1986), Ch. 10.1, are met, if $w_{ij} = 1 + O(N^{-1})$ and $m_{ij} = O(N^{-1})$. These are technical requirements in order to establish the diffusion process with coefficients $a(x)$ and $b(x)$ as the limit in distribution of the underlying discrete Markov chain. For practical purposes, however, it turns out that these constraints can be relaxed. The diffusion approximation will be satisfactory if $(1+2s)^{-1} \approx 1-2s$. The dominance parameter h is not affected by this type of constraint. Error thresholds are determined by analysis of the stationary distribution of the random variable X_1. The stationary density function $f_p(x)$ is obtained as equilibrium solution of the Kolmogorov forward equation (the index p emphasizes its dependence on nucleotide mutation probability p via m_{12}^R and m_{21}^R):

$$f_p(x) = cx^{4Nm_{12}^R-1}(1-x)^{4Nm_{21}^R-1}e^{4Nsx(x+2h(1-x))}, \tag{32}$$

where c is the normalization constant (see e.g. Ewens (1979); his parameter s has to be replaced by $2s$ due to the different selection regime used here).

f_p may be shown to perform a transition into a function monotonously decreasing in p as a certain value p_0 is surpassed, i.e., a local maximum vanishes. For a similar problem (Nowak & Schuster, 1989), this was taken as a criterion to diagnose an error threshold in finite populations. There is, however, no a priori reason why this heuristic criterion should correlate with an error threshold. In fact, in our case, we find this criterion to severely misestimate thresholds for wide ranges of parameters. We therefore propose to study directly the expectation of the random variable X_1,

$$E(X_1) = \int_0^1 x f_p(x) dx \, . \tag{33}$$

In analogy with the deterministic case, we locate the threshold where – at equilibrium – the master is 'lost' from the population. In the case of finite populations 'loss of the master' is naturally identified with $E(X_1) < \frac{1}{2N}$. For consistency with the deterministic calculations, however, we prefer to consider the master to be lost as soon as its expected relative frequency is smaller than 1%.

For certain values of h, $E(X_1)$ may be calculated explicitly:

$$h = 0: \qquad E(X_1) = \gamma \, \frac{{}_2F_2((\frac{1+\beta_2}{2}, 1 + \frac{\beta_2}{2}), (\frac{1+\beta_1+\beta_2}{2}, \frac{2+\beta_1+\beta_2}{2}), \alpha)}{{}_2F_2((\frac{1+\beta_2}{2}, \frac{\beta_2}{2}), (\frac{1+\beta_1+\beta_2}{2}, \frac{\beta_1+\beta_2}{2}), \alpha)} \tag{34}$$

$$h = \frac{1}{2}: \qquad E(X_1) = \gamma \, \frac{{}_1F_1(1+\beta_2, 1+\beta_1+\beta_2, \alpha)}{{}_1F_1(\beta_2, \beta_1+\beta_2, \alpha)} \tag{35}$$

$$h = 1: \qquad E(X_1) = \gamma \, \frac{{}_2F_2((\frac{1+\beta_1}{2}, \frac{\beta_1}{2}), (\frac{1+\beta_1+\beta_2}{2}, \frac{2+\beta_1+\beta_2}{2}), -\alpha)}{{}_2F_2((\frac{1+\beta_1}{2}, \frac{\beta_1}{2}), (\frac{1+\beta_1+\beta_2}{2}, \frac{\beta_1+\beta_2}{2}), -\alpha)} \tag{36}$$

where $\gamma = \frac{\beta_2}{\beta_1+\beta_2}$, $\alpha = 4Ns$, $\beta_1 = 4N(1-(1-p)^\nu)$ and $\beta_2 = 4N\frac{(1-(1-p)^\nu)}{2^\nu-1}$. Furthermore, ${}_1F_1$ denotes the confluent and ${}_2F_2$ the generalized hypergeometric function, ${}_2F_2$ is defined as

$${}_2F_2((a_1,a_2),(b_1,b_2),c) := \sum_{k=0}^{\infty} \frac{c^k}{k!} \frac{(a_1)_k (a_2)_k}{(b_1)_k (b_2)_k} \tag{37}$$

(compare e.g. Erdelyi, 1953).

For the full system, we carried out numerical simulations according to the multinomial sampling scheme. Starting from an initially homogeneous population of master individuals, we allowed for equilibration for 10^4 generations and then averaged the frequencies over another $2 \cdot 10^3$ generations. We would like to mention that the averages obtained in this way need not be estimates of frequency expectation in the strict sense. Although, if $p > 0$, the state space of the underlying Markov chain is irreducible without absorbing states (i.e. every state is recurrent), single trajectories need not be self-averaging on the time scale of a simulation, nor on evolutionary time scales. Consequently, dependence on initial conditions is observed under certain circumstances. For $h = 0$ and starting with a 'worst case initial condition' (see above), the master is usually not recovered. The existence of such meta-stable states is the stochastic analogue of the deterministic multiple stability which we observed for the differential equation.

In Figure 6 simulation results are compared with the approximate results from eqns (34) to (36). To determine thresholds, we numerically evaluate $p = p_{max}$ s.t. $E(X_1) \leq 0.01$. See columns 'coupled' in Table 1. We would like to remark that the sharp decrease observed in $E(X_1)$ at the threshold goes together with a peak in the variance, $V(X_1)$, as was also observed in (Nowak & Schuster, 1989). That is, fluctuations are largest around the error threshold, which, again, parallels the situation with phase transitions.

6.2 Decoupled case: Birth and Death

For the decoupled case, stochastic effects are naturally described by a time-homogeneous birth-death master equation (see Gardiner, 1990). As in the previous paragraph, let us first consider the two-allele caricature. Population size is kept fixed (at N), and K is a random variable representing the number of master alleles in the population and taking integer values k, $0 \leq k \leq 2N$. Then, we have the processes $k \to k+1$ and $k \to k-1$ with transition probabilities per unit time t^+ and t^-, respectively:

$$t^+(k) = \frac{\mu\nu}{2^\nu - 1}(2N-k) + \frac{2N-k}{2N}k\left(\frac{k-1}{2N}(1+2s) + \frac{2N-k}{2N}(1+2hs)\right) \tag{38}$$

and

$$t^-(k) = \mu\nu k + \frac{k}{2N}(2N-k)\left(\frac{k}{2N}(1+2hs) + \frac{2N-k-1}{2N}\right) . \tag{39}$$

The master equation for the conditional probability $P(k,t|k_0,t_0)$ corresponding to the initial condition $K(t_0) = k_0$ reads:

$$\begin{aligned} \frac{\partial}{\partial t}P(k,t|k_0,t_0) &= t^+(k-1)P(k-1,t|k_0,t_0) + t^-(k+1)P(k+1,t|k_0,t_0) \\ &\quad -(t^+(k)+t^-(k))P(k,t|k_0,t_0)\ , k = 0,\ldots,2N\,, \end{aligned}$$

where $t^+(2N) = t^-(0) = 0$ ensures reflecting boundary conditions.

For a haploid population (with population size $2N$ and selective advantage s for comparability with the diploid case), the corresponding transition rates read

$$t^+(k) = \frac{\mu\nu}{2^\nu - 1}(2N-k) + \frac{2N-k}{2N}k(1+s)\,, \tag{40}$$

$$t^-(k) = \mu\nu k + \frac{k}{2N}(2N-k)\,. \tag{41}$$

The diploid transition rates do *not* exactly reduce to the haploid ones in the case $h = \frac{1}{2}$.

Following Gardiner (1990), the stationary distribution P_μ is obtained via

$$P_\mu(k) = \frac{Q_\mu(k)}{Z}\,, k = 0,\ldots,2N\,, \tag{42}$$

where $Q_\mu(k) := \prod_{j=1}^{k}\frac{t^+(j-1)}{t^-(j)}$, and $Z := \sum_{j=0}^{2N} Q_\mu(j)$. For the haploid case, we obtain explicitly:

$$Q_\mu(k) = \binom{2N}{k}(1+s)^k\frac{\Gamma(u+k)\Gamma(v+2N+k)}{\Gamma(u)\Gamma(v+2N)}\,, \tag{43}$$

$$Z = \frac{\Gamma(u+v+2N)\Gamma(v)}{\Gamma(u+v)\Gamma(v+2N)}\,{}_2F_1(-2N,u,u+v;-s)\,, \tag{44}$$

and for the expectation

$$E(\frac{K}{2N}) = \frac{{}_2F_1(-(2N-1),u+1,u+v+1;-s)}{{}_2F_1(-2N,u,u+v;-s)}\,. \tag{45}$$

Here, $u = 2N\frac{\mu\nu}{(2^\nu-1)(1+s)}$, $v = 2N\mu\nu$, and ${}_2F_1$ denotes the hypergeometric function (see Abramowitz & Stegun, 1970).

For the diploid case, no closed formula is available for the stationary distribution. Expectations from numerical evaluation of eqn (42) are collected in Table 1.

For the entire system, stochastic simulations were performed with the method proposed by Gillespie (1976). Estimates for master frequency expectation were arrived at by averaging over $2000 \cdot 2N$ birth-death events after allowing for equilibration for another $2000 \cdot 2N$ such events. For both approaches, error thresholds are evaluated from the expectation and averages as described in the previous paragraph. The results are collected in Table 1.

Comparing these with the predictions obtained from the error tail caricature, we find the latter to underestimate the threshold position, especially when selective differentials and populations are small. This is due to the underestimation of back mutations, which certainly occurs in the caricature as compared with the full system. The birth-death process, due to its accessible boundaries, is much more sensitive to deviations near the boundaries than is the diffusion approximation.

7 Discussion

Error thresholds, viewed as a relationship between maximum genome length and maximum mutation rate, have been found to be important for the evolution of haploid populations of quickly-reproducing species with relatively small genomes, like in vitro evolution of RNA, evolution of bacteria and viruses. The latter were found to live close to that threshold (Eigen, 1993).

A similar phenomenon might be relevant for higher organisms as well, where mutation rates are lower but genomes are larger both by a factor of at least 10^3. To match the situation in higher organisms, however, additional features must be considered which concern the mutation as well as the selection mechanism.

As to the former, the assumption of mutation occurring exclusively as replication errors is probably a good approximation for organisms which replicate very quickly. This is, however, not the case for higher organisms where generation times may be years instead of minutes or hours. Instead, independent mutations—due to radiation effects etc.—are expected to play an important part, but the relative proportions of both contributions are unresolved.

Our analyses of the two extreme cases in the form of the coupled and decoupled selection mutation equation reveals that the qualitative feature of the error threshold is not affected by the details of the mutation mechanism. The similarity for the parameter values considered should, however, not obscure the fact that the relevant quantities for the threshold position are the ratio (coupled equation) versus difference (decoupled equation) in reproduction rates. It is not astonishing that the behaviour is intermediate when both mechanisms are combined via the sum of both right hand sides with a convex combination of the respective mutation parameters, αp and $(1-\alpha)\mu$, which is probably closer to the real situation.

As to the selection mechanism, diploidy effects are an immediate concern in higher organisms. We focussed on dominance here, which is easily accessible within the framework adopted. We find that diploid replication introduces new qualitative features, like multiple equilibria, even in the case of the simple single-peaked landscape.

Additionally, remarkable shifts in equilibria are observed: A dominant advantageous mutation is able to persist in the population at mutation rates much higher than in the intermediate or even recessive case. This might add an aspect to the old discussion of evolution of dominance. Relying on the observation that the wild type is often dominant, theories were put forward in the twenties which explain dominance as the result of the action of some modifier locus (e.g. Fisher, 1928). An opposite view (see Wright, 1934) understands dominance as a genuine character of a gene, brought about by the properties of the gene products. Modern versions of the modifier theory are discussed in a series of papers by Bürger and coworkers (see Wagner & Bürger, 1985 and the references therein). Our analysis suggests that dominant genes might be the prevailing wild type ones just because they can tolerate higher mutation rates than corresponding recessive ones. This is, however, highly speculative and depends on whether or not evolution of diploid organisms takes (or took) place close to an error threshold. Similarly, our account of "advantage of dominance" (see Figure 4) reveals that population mean fitness (as a function of mutation rate) at equilibrium increases with the degree of dominance. Positive values of the quantity $\bar{w}_{rel}$ indicate an advantage of diploid over haploid reproduction. The latter is equivalent to the diploid intermediate case. This result, of course, depends on the chosen fitness landscape, but it provides an aspect of evolution of diploidy which does not rely on recombinational mechanisms. With an analysis of mutation and selection mechanisms only, as in the present paper, we refrain from embarking on the discussion of evolution of sex. A detailed analysis of the effects of intragenic and/or intergenic recombination would be necessary to comment on this problem.

For fitness landscapes different from single peaked (multiplicative and epistatic landscapes) some investigations of this point have already been carried out (Kondrachov, 1988; Higgs, 1994), in which an advantage of sexual reproduction, in particular for epistatic landscapes, is reported. Interestingly, our dominance effects for a single-peaked landscape without recombination turn out to be very similar to those observed in the latter paper, where the fitness landscape is muliplicative and reproduction includes recombination.

Our concern in this paper is to concentrate on diploidy effects with a fitness distribution according to the single peaked landscape. All our results arrived at so far depend on this assumption. We are well aware of the fact that sophisticated multi-peaked landscapes are needed to describe real-life evolution. However, apart from the fact that this landscape allows at least some analytical caricatures, it is relevant on its own right. In many cases, evolution may probably be conceived as a process with only small selective differentials between the genotypes present for most of the time, with an advantageous mutation introduced occasionally as the result of a stochastic process (deleterious mutations will be eliminated very quickly in most cases). Then we are in a situation close to the single-peaked one, and the new allele will only manage to establish itself if mutation rates are not too high. Similarly, all cases of functional dichotomy (e.g. recognition of binding sites by enzymes) may well be modelled by such a two–class landscape.

We remark that two–class landscapes differ fundamentally from fitness landscapes discussed in connection with the advance of Muller's ratchet (Haigh, 1978). In both models, deleterious alleles may accumulate in the population. However, in the former case the *whole* population is – beyond the error threshold – a randomly reproducing ensemble, with no allele superior to any other allele. Accordingly, the mutational load (at equilibrium) is an increasing function of the mutation rate, saturating at the error threshold (similar to Average Hamming Distance, cf. Fig. 2). In contrast to this, the mutational load does not saturate in model landscapes of Muller's ratchet.

Our findings carry over to a stochastic description of the mutation selection process in finite populations. The shift of error thresholds to lower values of the mutation parameters is the more pronounced the smaller the population is, as was observed for a haploid selection regime before (Nowak & Schuster, 1989). However, we adopt a different way of diagnosing error thresholds by concentrating on expectations rather than local maxima of the stationary distribution. Within this framework, dominance effects are recovered which strongly resemble the deterministic picture: Error thresholds are shifted, and bistability re-emerges as metastability.

It is quite astonishing that the effects of finite population, for a given population size, saturate with sequence length. Anyway, stochastic effects must be kept in mind when estimates of mutation rates and genome lengths are related, which is usually done on the basis of the deterministic equations.

Let us close with a general remark on error thresholds. They are by no means a consequence of the assumption of single-peaked landscapes but have been shown to arise in other instances of fitness distributions as well. Tarazona (1992) described error thresholds for spin-glass landscapes. In addition, they were obtained in numerical studies on RNA folding landscapes (Fontana & Schuster, 1987). Furthermore, a transition from an ordered quasi-species into a randomly replicating ensemble as the nucleotide mutation probability is increased has been found for the case of non symmetric selection matrices. This transition may be associated with a bifurcation of super- or sub-critical type (Stadler *et al.*, 1994). On the other hand, there are fitness landscapes which do not exhibit threshold behaviour, like the multiplicative landscape as discussed by Haigh (1978). Wagner & Krall (1993) give a sufficient condition for such cases. However, these results are only valid for models in which chain length is infinite.

The general problem of classifying which landscapes produce which kind of error thresholds is still unresolved. It will be the subject of a forthcoming paper.

Acknowledgments

It is our pleasure to thank A. Gierer, W. Grüner, C. Reidys and S. Rotter for many stimulating discussions, M. Baake for his persistent support, and an unknown referee for helpful comments. We acknowledge gratefully F. Haubensak's patient assistance in solving various computer problems.

References

Abramowitz, M. & Stegun, I. (1970). *Handbook of Mathematical Functions.* New York: Dover.

Akin, E. (1979). *The Geometry of Population Genetics*, volume 31 of *Lect. Notes Biomath.* New York: Springer.

Bürger, R. (1983). On the evolution of dominance modifiers I. A nonlinear analysis. *J. Theor. Biol.* **101**, 585–598.

Chao, L. (1988). Evolution of sex in RNA viruses. *J. Theor. Biol.* **133**, 99–112.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory.* New York: Harper and Row.

Eigen, M. & Schuster, P. (1977). The hypercycle A: A principle of natural self–organization: Emergence of the hypercycle. *Naturwissenschaften* **64**, 541–565.

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

Eigen, M. (1993). Viral quasispecies. *Scientific American* pages 32–39.

Erdelyi, A. (1953). *Higher Transcendental Functions*, volume 1. New York: McGraw-Hill.

Ethier, S. N. & Kurtz, T. G. (1986). *Markov Processes.* New York: John Wiley & Sons.

Ewens, W. J. (1979). *Mathematical Population Genetics.* Berlin: Springer.

Fisher, R. A. (1928). The possible modification of the response of the wild type to recurrent mutations. *Amer. Nat.* **62**, 115–126.

Fisher, R. A. (1929). The evolution of dominance; reply to Professor Sewall Wright. *Amer. Nat.* **63**, 553–556.

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

Gardiner, C. W. (1990). *Handbook of Stochastic Methods*, volume 13 of *Springer Series in Synergetics.* Berlin: Springer.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp. Phys.* **22**, 403–434.

Gillespie, J. H. (1991). *The Causes of Molecular Evolution.* New York: Oxford University Press.

Hadeler, K. P. (1974). On the equilibrium states in certain selection models. *J. Math. Biol.* **1**, 51–56.

Hadeler, K. P. (1981). Stable polymorphisms in a selection model with mutation. *SIAM J. Appl. Math.* **41**, 1–7.

Haigh, J. (1978). The accumulation of deleterious genes in a population - Muller's ratchet. *Theor. Pop. Biol.* **14**, 251–267.

Haldane, J. B. S. (1927). A mathematical theory of natural and artificial selection. Part V: Selection and mutation. *Proc. Camb. Phil. Soc.* **23**, 838–844.

Heuser, H. (1992). *Funktionalanalysis.* Stuttgart: Teubner.

Higgs, P. G. (1994). Error thresholds and stationary mutant distributions in multi–locus diploid genetics models. *Genet. Res., Camb.* **63**, 63–78.

Hofbauer, J. & Sigmund, K. (1988). *The Theory of Evolution and Dynamical Systems.* Cambridge: Cambridge University Press.

Hofbauer, J. (1985). The selection mutation equation. *J. Math. Biol.* **23**, 41–53.

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

KARLIN, S. & TAYLOR, H. M. (1975). *A First Course in Stochastic Processes.* San Diego: Academic Press.

KARLIN, S. (1980). The number of stable equilibria for the classical one-locus multi-allele selection model. *J. Math. Biol.* **9**, 189–192.

KIMURA, M. & CROW, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738.

KIMURA, M. (1969). The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* **61**, 893–903.

KIMURA, M. (1987). Molecular evolutionary clock and the neutral theory. *J. Mol. Evol.* **26**, 24–33.

KINGMAN, J. F. C. (1988). Typical polymorphisms maintained by selection at a single locus. *J. Appl. Prob.* **25**, 113–125.

KONDRACHOV, A. S. (1988). Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440.

LEUTHAEUSSER, I. (1987). Statistical mechanics of Eigen's evolution model. *J. Stat. Phys.* **48**, 343–360.

NOWAK, M. & SCHUSTER, P. (1989). Error thresholds of replication in finite populations. Mutation frequencies and the onset of Muller's ratchet. *J. Theor. Biol.* **137**, 375–395.

OHTA, T. (1993). An examination of the generation–time effect on molecular evolution. *Proc. Natl. Acad. Sci. USA* **90**, 10676–10680.

RUMSCHITZKY, D. S. (1987). Spectral properties of Eigen evolution matrices. *J. Math. Biol.* **24**, 667–680.

STADLER, P. & SCHUSTER, P. (1992). Mutation in autokatalytic reaction networks. An analysis based on perturbation theory. *J. Math. Biol.* **30**, 597–632.

STADLER, P. F., SCHNABL, W., FORST, C. V. & SCHUSTER, P. Dynamics of small autocatalytic reaction networks. II: Replication, mutation and catalysis. Bull. Math. Biol. (submitted), (1994).

SWETINA, J. & SCHUSTER, P. (1982). Self–replication with errors. A model for polynucleotide replication. *Biophys. Chem.* **16**, 329–345.

TARAZONA, P. (1992). Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin–glass models. *Phys. Rev. A* **45(8)**, 6038–6050.

THOMPSON, C. J. & MCBRIDE, J. L. (1974). On Eigen's theory of the self-organization of matter and the evolution of biological macromolecules. *Math. Biosci.* **21**, 127–142.

WAGNER, G. P. & BÜRGER, R. (1985). On the evolution of dominance modifiers ii: A nonequilibrium approach to the evolution of genetic systems. *J. Theor. Biol.* **113**, 475–500.

WAGNER, G. P. & KRALL, P. (1993). What is the difference between models of error thresholds and Muller's ratchet. *J. Math. Biol.* **32**, 33–44.

WATSON, J. D. & CRICK, F. H. C. (1953). Molecular structure of nucleic acid. A structure of deoxyribose nucleic acid. *Nature* **171**, 737.

WRIGHT, S. (1929). The evolution of dominance. *Am. Nat.* **63**, 556–561.

WRIGHT, S. (1934). Physiological and evolutionary theories of dominance. *Amer. Nat.* **68**, 24–53.

Figure captions

Figure 1:

Stationary distribution of Hamming classes $i = 0$ to $i = 15$ as function of nucleotide mutation probability. Diploid case. Population size = ∞. Parameters: $s = 1/2$; $h = 1/2$; $\nu = 30$. The curves are equilibrium frequencies y_i of Hamming classes. Already for moderate chain lengths the transition into an (almost) uniform distribution (error threshold) is very sharp.

Figure 2:

Average Hamming distance from master $E = E(p)$ at stationarity. Parameters: $s = 1/2$; $\nu = 30$; $h = 0,\ 1/2,\ 1,\ 3/2$. Thresholds are located at p values where average Hamming distance (as function of p) becomes constant. E saturates at $\nu/2$.

Figure 3:

Variance of Hamming distance from master ($V = V(p)$) at stationarity. Parameters: $s = 1/2$; $\nu = 30$; $h = 0,\ 1/2,\ 1,\ 3/2$.

Figure 4:

Relative difference of mean fitness $\bar{w}_{rel} = \frac{(\bar{x},W(h)\bar{x})-(\bar{x},W(1/2)\bar{x})}{(\bar{x},W(1/2)\bar{x})}$ for $h = 0, 1$ and $3/2$ as function of p. Parameters: $\nu = 30,\ s = 1/2$. $\bar{w}_{rel} > 0$ indicates an advantage of dominance over the intermediate reference case ($h = 1/2$). In case of multiple equilibria (here for $h = 0$), the solid line belongs to the equilibrium attained when the initial distribution consists of wild type alleles only ($y_0(0) = 1$), the dotted line to that with initial condition $y_\nu(0) = 1$.

Figure 5:

Time $\bar{\tau}$ to reach equilibrium. Solid line: Analytical expression (27). Initial condition $x_1(0) = 1$. Dotted: Numerical integration. Initial condition $x_1(0) = 1$. Dashed: Numerical integration. Initial condition $x_n(0) = 1$ ('worst case condition'). Numerical and analytical curves agree very well for $p < p_{max} = 0.013$. For p beyond the threshold back-flow mutations become important, which have been neglected in the analytical treatment. This explains the discrepancy of numerical and analytical solutions in this region. Parameters: $\nu = 30,\ s = 1/2,\ h = 1/2,\ \epsilon = 10^{-5}$.

Figure 6:

Expectation of master frequency in the diffusion process. Parameters: $\nu = 30$; $s = 1/20$; $h = 0,\ 1/2,\ 1$; $N = 1000$. Gray: Simulation results. Black: $E(X_1)$ according to expressions (34) to (36).

Table 1: Comparison of threshold values for different parameters

Parameters				1		2		3	
ν	N	s	h	coupled	decoupled	coupled	decoupled	coupled	decoupled
10	100	0.05	0.5	$2.7\ 10^{-3}$	$1.4\ 10^{-3}$	$4.3\ 10^{-3}$	$3.3\ 10^{-3}$	56	76
10	∞	0.05	0.5	$6.1\ 10^{-3}$	$6.0\ 10^{-3}$	$4.9\ 10^{-3}$	$5.0\ 10^{-3}$	/	/
30	100	0.05	0	$3.9\ 10^{-4}$	$3.3\ 10^{-4}$	$4.1\ 10^{-4}$	$< 10^{-5}$	52	61
30	100	0.05	0.5	$4.6\ 10^{-4}$	$4.0\ 10^{-4}$	$6.0\ 10^{-4}$	$< 10^{-5}$	73	78
30	100	0.05	1	$5.2\ 10^{-4}$	$5.5\ 10^{-4}$	$1.1\ 10^{-3}$	$< 10^{-5}$	84	84
30	1000	0.05	0	$7.4\ 10^{-4}$	$8.1\ 10^{-4}$	$6.5\ 10^{-4}$	$5.3\ 10^{-4}$	10	4
30	1000	0.05	0.5	$1.3\ 10^{-3}$	$1.4\ 10^{-3}$	$1.3\ 10^{-3}$	$9.0\ 10^{-4}$	24	22
30	1000	0.05	1	$2.5\ 10^{-3}$	$2.9\ 10^{-3}$	$2.6\ 10^{-3}$	$1.7\ 10^{-3}$	22	15
30	10000	0.05	0	$8.2\ 10^{-4}$	$8.3\ 10^{-4}$	$6.9\ 10^{-4}$	$6.6\ 10^{-4}$	0	1
30	10000	0.05	0.5	$1.6\ 10^{-3}$	$1.7\ 10^{-3}$	$1.6\ 10^{-3}$	$1.4\ 10^{-3}$	6	6
30	10000	0.05	1	$3.1\ 10^{-3}$	$3.3\ 10^{-3}$	$3.2\ 10^{-3}$	$2.7\ 10^{-3}$	3	3
30	∞	0.05	0	$8.2\ 10^{-4}$	$8.4\ 10^{-4}$	+	+	/	/
30	∞	0.05	0.5	$1.7\ 10^{-3}$	$1.8\ 10^{-3}$	$1.6\ 10^{-3}$	$1.7\ 10^{-3}$	/	/
30	∞	0.05	1	$3.2\ 10^{-3}$	$3.4\ 10^{-3}$	$3.2\ 10^{-3}$	$3.3\ 10^{-3}$	/	/
100	100	0.05	0.5	$1.2\ 10^{-4}$	$7.0\ 10^{-5}$	$1.8\ 10^{-4}$	$< 10^{-5}$	76	86
100	1000	0.05	0.5	$3.6\ 10^{-4}$	$4.8\ 10^{-4}$	$3.8\ 10^{-4}$	$7.0\ 10^{-5}$	27	4
100	10000	0.05	0.5	$4.7\ 10^{-4}$	$4.9\ 10^{-4}$	$4.7\ 10^{-4}$	$3.4\ 10^{-4}$	4	2
100	∞	0.05	0.5	$4.9\ 10^{-4}$	$5.0\ 10^{-4}$	$4.9\ 10^{-4}$	$5.0\ 10^{-4}$	/	/
30	100	0.5	0.5	$8.8\ 10^{-3}$	$1.1\ 10^{-2}$	$*$	$7.1\ 10^{-3}$	37	38
30	1000	0.5	0.5	$1.2\ 10^{-2}$	$1.6\ 10^{-2}$	$*$	$1.3\ 10^{-2}$	14	11
30	10000	0.5	0.5	$1.3\ 10^{-2}$	$1.7\ 10^{-2}$	$*$	$1.6\ 10^{-2}$	7	6
30	∞	0.5	0.5	$1.4\ 10^{-2}$	$1.8\ 10^{-2}$	$1.3\ 10^{-2}$	$1.7\ 10^{-2}$	/	/

Column 1: Simulation results (N finite) and results by numerical integration (N infinite), respectively. Column 2: Analytical results according to formulas (34) to (36) and (42) (N finite) and (19), (21) (N infinite), respectively. Column 3: Relative shift in percent of threshold values of col.1 for finite N with respect to deterministic value (N infinite).

$*$) The diffusion approximation is not valid in this case.

+) Eq (21) is not relevant for these initial conditions.

Figure 1:

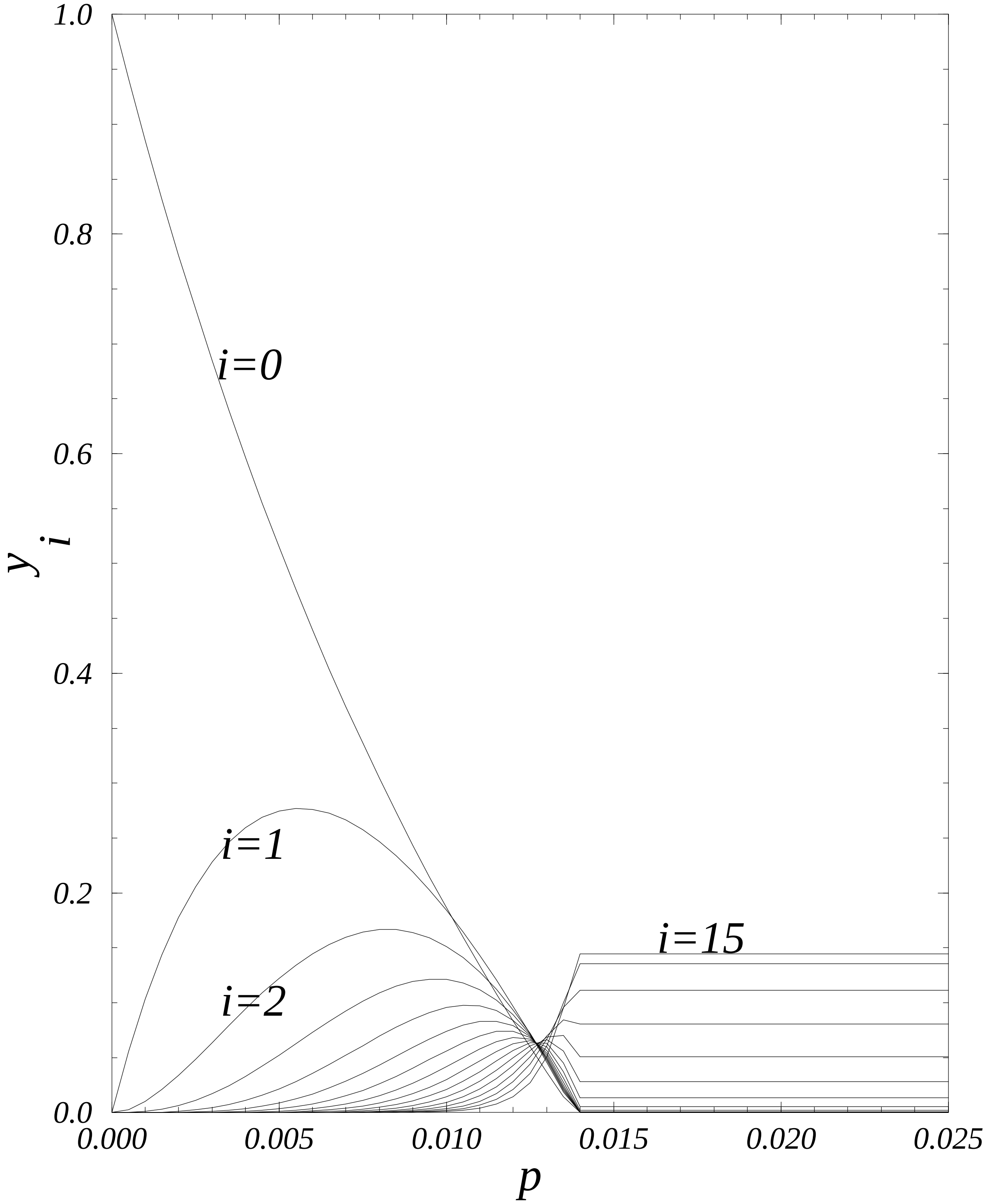


Figure 2:

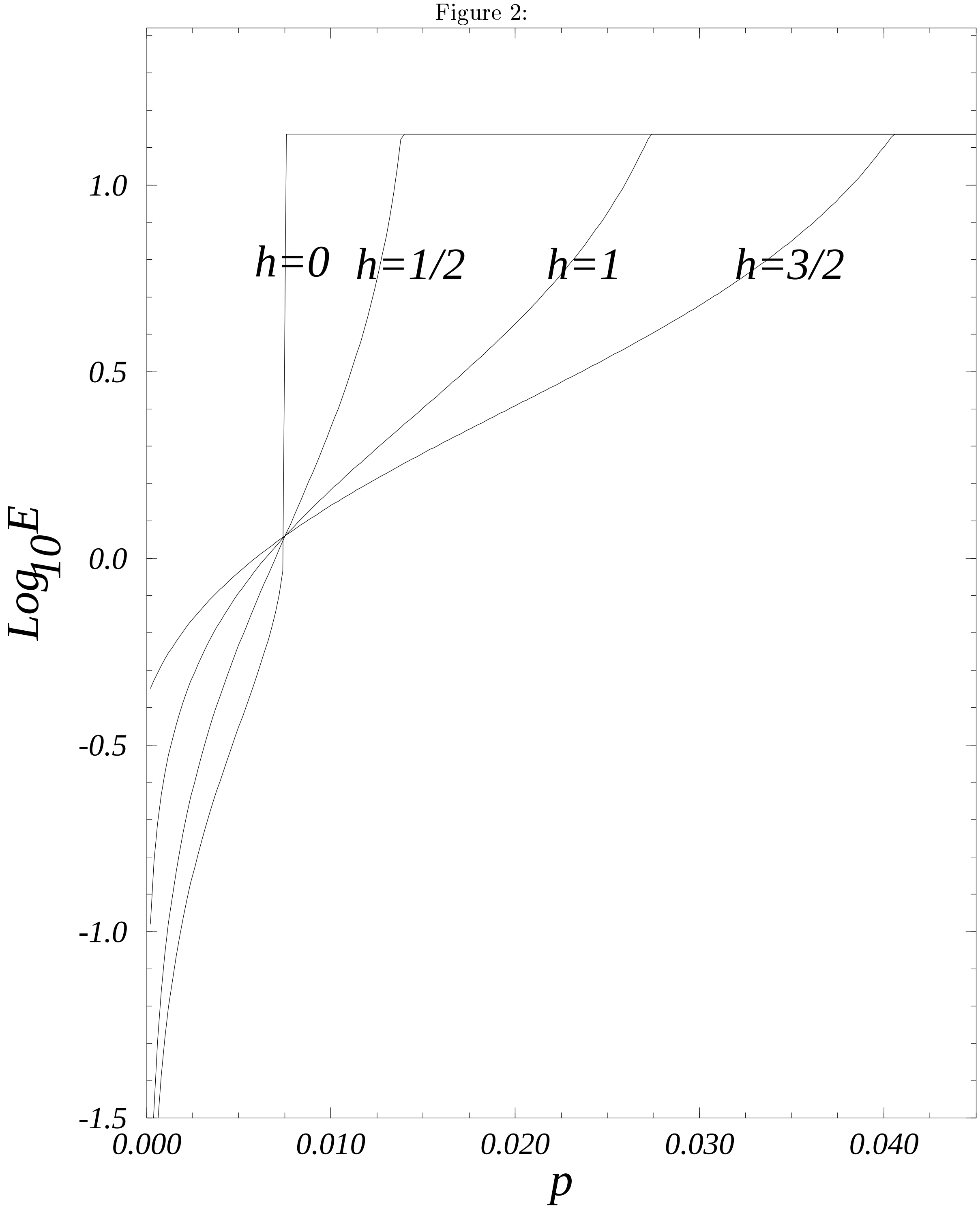


Figure 3:

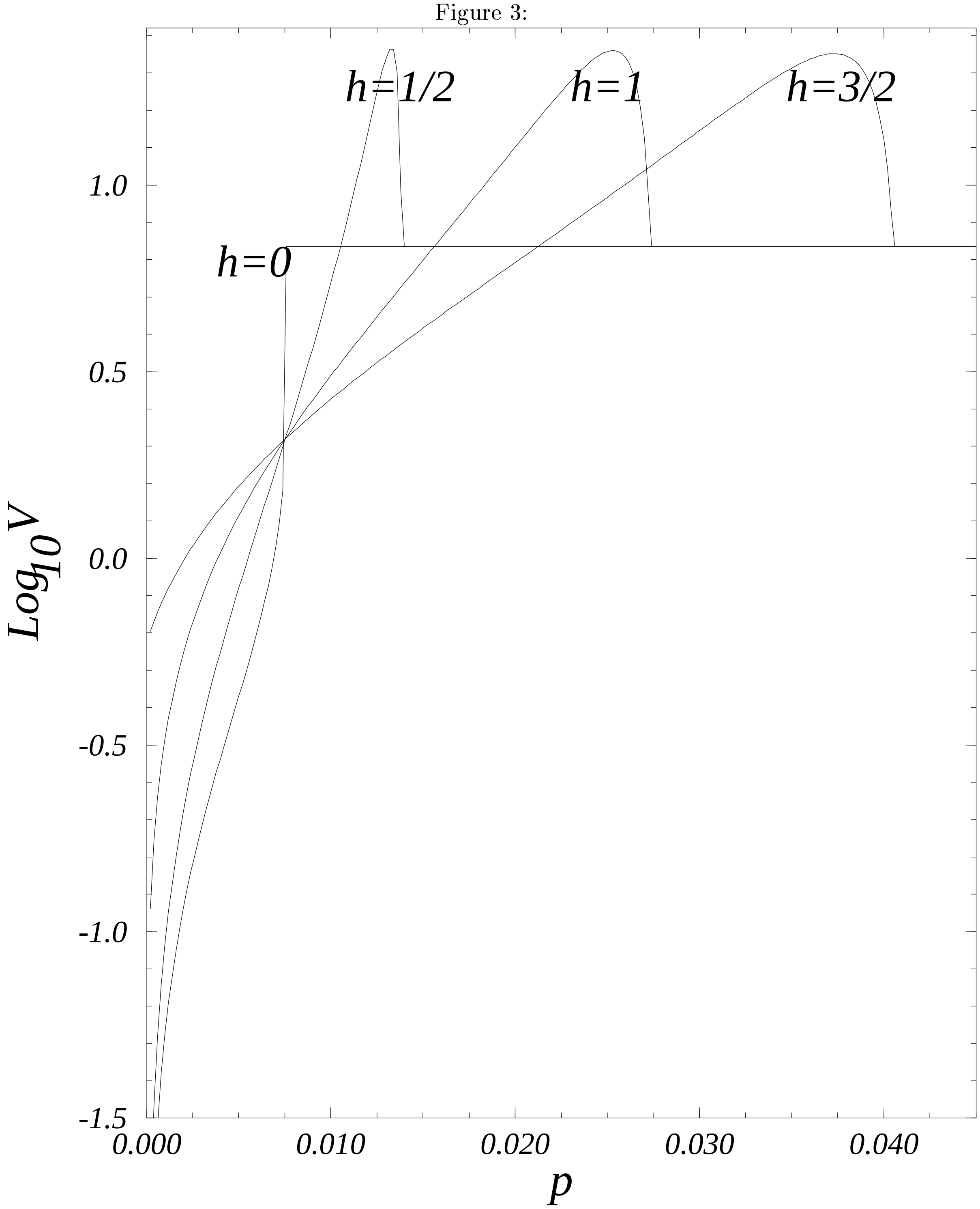
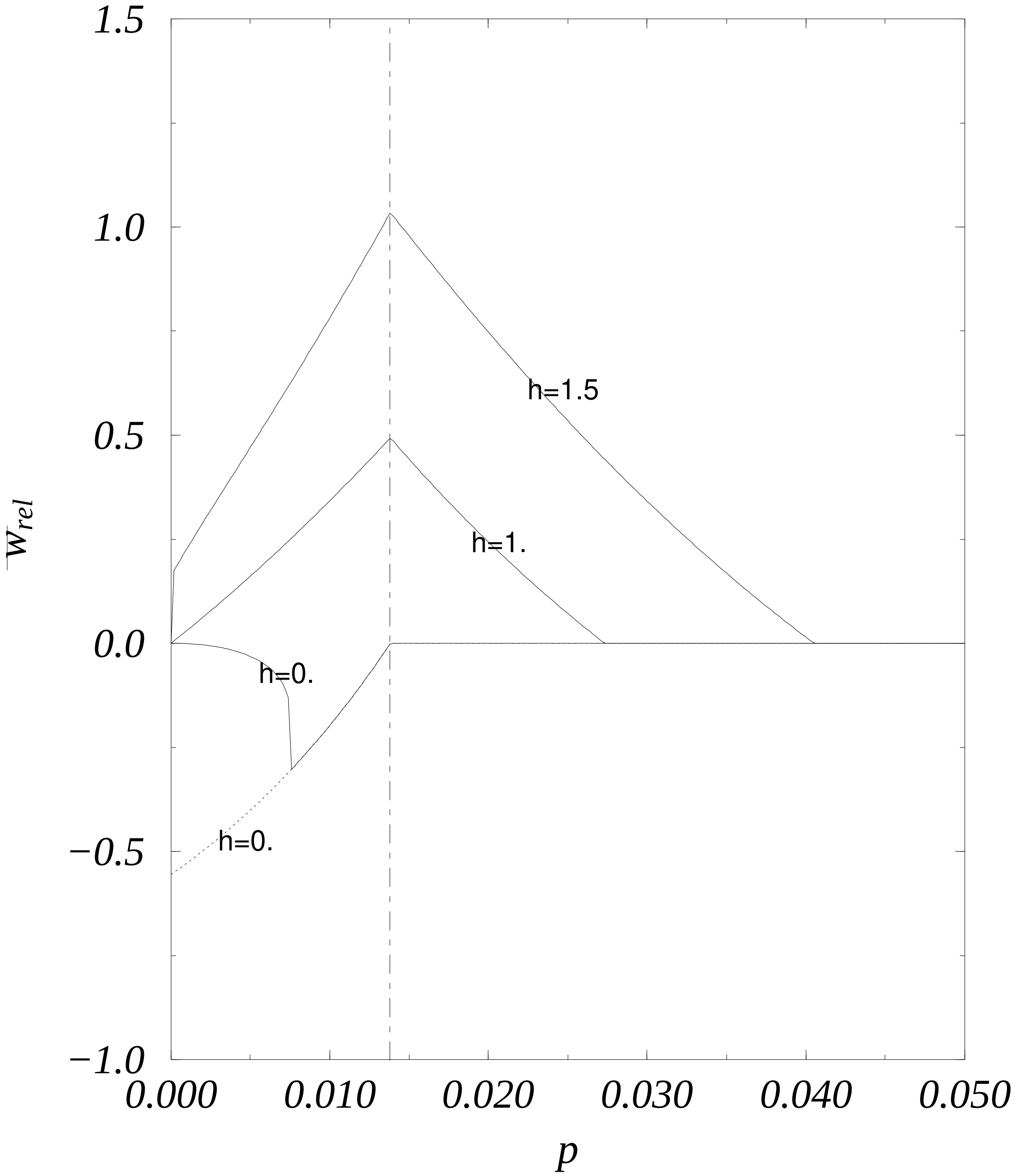
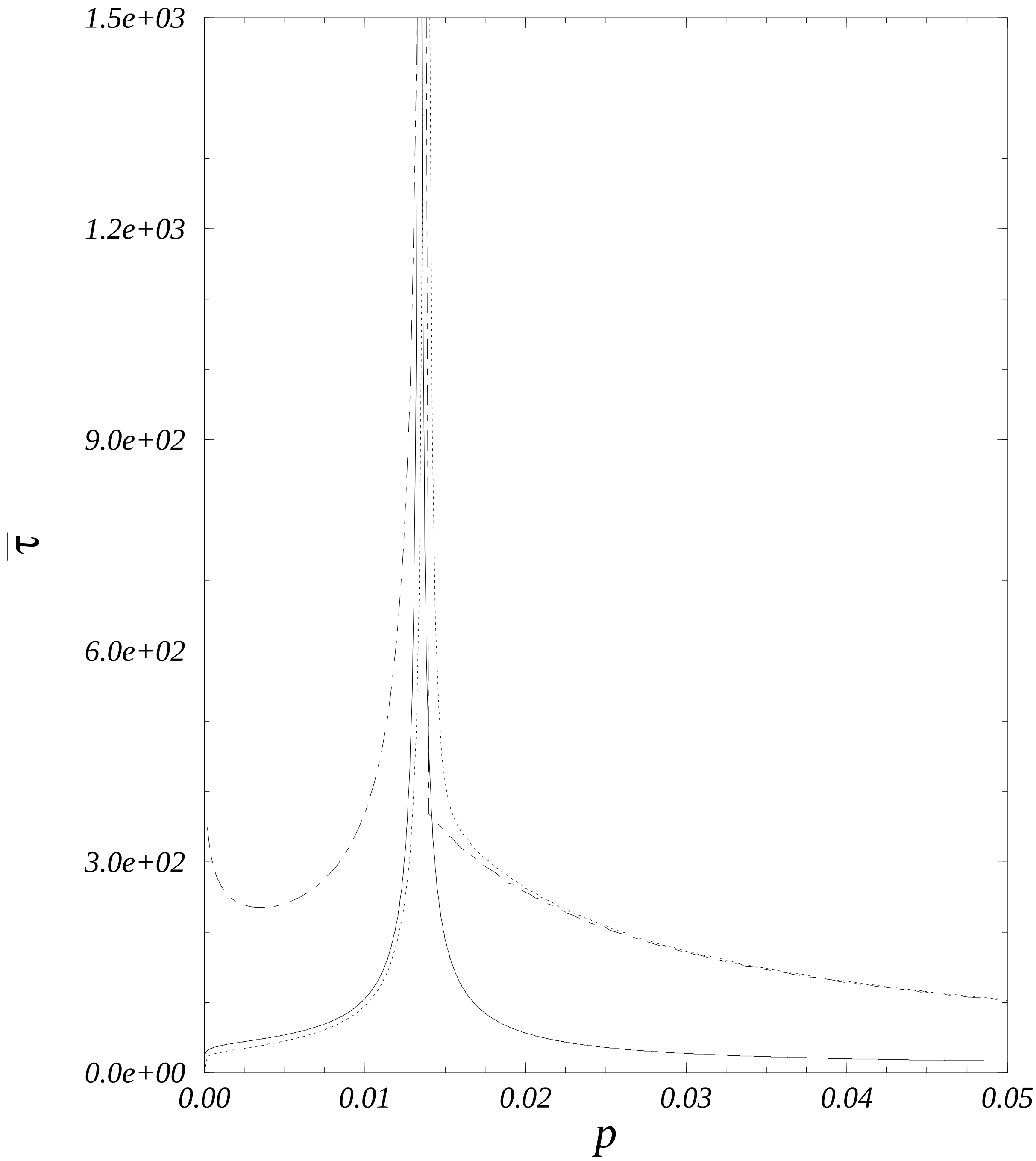


Figure 4:



1.5
1.0
0.5
0.0
−0.5
−1.0
0.000
0.010
0.020
0.030
0.040
0.050
p
w_{rel}
h=1.5
h=1.
h=0.
h=0.

Figure 5:



1.5e+03
1.2e+03
9.0e+02
6.0e+02
3.0e+02
0.0e+00
0.00
0.01
0.02
0.03
0.04
0.05
p
τ

Figure 6: