J. Math. Biol. (1994) 32:193–218

Journal of Mathematical Biology

Mutation load and mutation-selection-balance in quantitative genetic traits

Reinhard Bürger, Josef Hofbauer

Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Wien, Austria

Received 18 November 1992; received in revised form 17 February 1993

Abstract. Haldane (1937) showed that the reduction of equilibrium mean fitness in an infinite population due to recurrent deleterious mutations depends only on the mutation rate but not on the harmfulness of mutants. His analysis, as well as more recent ones (cf. Crow 1970), ignored back mutation. The purpose of the present paper is to extend these results to arbitrary mutation patterns among alleles and to quantitative genetic traits. We derive first-order approximations for the equilibrium mean fitness (and the mutation load) and determine the order of the error term. For a metric trait under mutation-stabilizing-selection balance our result differs qualitatively from that of Crow and Kimura (1964), whose analysis is based on a Gaussian assumption. Our general approach also yields a mathematical proof that the variance under the usual mutation-stabilizing-selection model is, to first order, μ/s (the house-of-cards approximation) as μ/s tends to zero. This holds for arbitrary mutant distributions and does not require that the population mean coincide with the optimum. We show how the mutant distribution determines the order of the error term, and thus the accuracy of the house-of-cards approximation. Upper and lower bounds to the equilibrium variance are derived that deviate only to second order as μ/s tends to zero. The multilocus case is treated under the assumption of global linkage equilibrium.

Key words: Selection – Mutation – Mutation load – Quantitative genetic traits

1 Introduction

The genetic load is usually defined as the proportion by which the fitness of the average genotype in a population is reduced in comparison with the best genotype, i.e. $L = (w_{\max} - \bar{w})/w_{\max}$. One of the most important factors creating genetic load is mutation. The corresponding load is called the mutation load (see Muller 1950, and Crow 1970, for a comprehensive treatment of genetic loads). It was Haldane (1937) who showed that the reduction of equilibrium mean fitness in an infinite

population due to recurrent deleterious mutations depends only on the mutation rate but not on the harmfulness of the mutants. Haldane (1937, 1957) considered the special case of two alleles: "... *the loss of fitness to the species depends entirely on the mutation rate and not at all on the effect of the gene upon the fitness of the individual carrying it* ...". His work has been extended in several directions (e.g., Crow and Kimura 1964, King 1966, Kimura and Maruyama 1966, Crow 1970, Fraser and Mayo 1974, Kondrashov and Crow 1988). In all these investigations, back mutations to the optimal type were ignored. Also, quantitative traits have received little attention, most likely because of mathematical difficulties. The only exception seems to be the final section in Crow and Kimura (1964).

It is the purpose of the present paper to explore possible generalizations and limitations of Haldane's result. Throughout, "arbitrary" mutation patterns among alleles are admitted. For mathematical simplicity, our analysis is based on a continuous-time model. Since, in general, continuous and the corresponding discrete-time models have the same equilibrium behavior, these results should apply also to the latter case. It should be noticed that in continuous time the mutation load has to be defined as the absolute reduction in Malthusian fitness, i.e., $L = m_{\max} - \bar{m}$. Our populations are assumed to be infinitely large.

Section 2 treats the classical case of n alleles at one locus in a haploid and in a diploid population. It is shown that under very general assumptions Haldane's result that the equilibrium mean fitness depends only on the mutation rate is true as a first-order approximation for small μ. The order of the error term is also calculated, and it may depend on the selection intensity s.

In Sect. 3 the continuum-of-alleles model for a haploid population is investigated in detail and the first-order approximation of the mutation load is derived. In many important special cases, it is simply the mutation rate. For various particular fitness functions and mutant distributions, the order of the error term is calculated.

In Sect. 4 the well-known mutation-selection-balance model (see Kimura 1965, Lande 1975, Turelli 1984, Bürger 1986) is investigated. Results of Sect. 3 are used to prove that Turelli's (1984) house-of-cards approximation is accurate to first order (as $\mu/s \to 0$) for any mutant distribution. It is investigated how different mutant distributions affect the accuracy of this approximation. Upper and lower bounds to the equilibrium variance that differ only to second order as μ/s tends to zero are derived. Under the assumption of global linkage equilibrium, the genetic variance in the multilocus model is, to first order, $2\sum_i \mu_i/s$, even if the population mean deviates from the optimum.

The main results are summarized, and compared with the existing literature, in Sect. 5.

2 The classical case: *n* alleles

2.1

For motivation we start with the *n-allelic haploid mutation-selection equation.*

$$\dot{p}_i = p_i(m_i - \bar{m}) + \sum_{j=1}^{n} (\mu_{ij} p_j - \mu_{ji} p_i), \quad i = 1, \ldots, n\,. \tag{2.1}$$

Here p_i denotes the frequency of the ith allele A_i, m_i its fitness, and $\bar{m} = \sum_i m_i p_i$ the average fitness. For $i \neq j$, μ_{ij} is the mutation rate from A_j to A_i, and we assume $\mu_{ii} = 0$ for all i. Let n be the fittest allele, i.e. $m_n > m_i$ for all $i = 1, \ldots, n-1$. Then for $\mu_{ij} = 0$, the point $\hat{p}^o = (0, 0, \ldots, 1)$ is the global attractor for (2.1). Hence, the perturbed system (2.1) (with μ_{ij} small) has an attracting fixed point $\hat{p}(\mu)$ nearby. Since $\hat{p}^o - \hat{p}(\mu) = O(\mu)$, in first order we have

$$\hat{p}_i(\mu)(m_i - \bar{m}) + \mu_{in} = 0$$

or

$$\hat{p}_i(\mu) \approx \frac{\mu_{in}}{m_n - m_i} \tag{2.2}$$

for $i < n$. Therefore

$$\bar{m} = \sum_{i=1}^{n} m_i \hat{p}_i(\mu) = m_n \left(1 - \sum_{i=1}^{n-1} \hat{p}_i(\mu)\right) + \sum_{i=1}^{n-1} m_i \hat{p}_i(\mu) \approx m_n - \sum_{i=1}^{n-1} \mu_{in} .$$

Hence, in the *first approximation*, the mutation load is given by

$$L = \sum_{i=1}^{n-1} \mu_{in} \tag{2.3}$$

which is the total mutation rate from the fittest allele to all the others.

If there is no back mutation to the fittest allele, i.e. $\mu_{nj} = 0$ for $j < n$, then the formula (2.3) is even exact: take $i = n$ in (2.1) and cancel through p_n. This is a well-known result (cf. Crow 1970, p. 148).

If there are two or more alleles with maximal fitness, the mutation load may strongly depend on the pattern of mutation between optimal alleles. Consider, for example, three alleles A_1, A_2, A_3 with fitness values $-s$, 0, 0 and mutation rates $\mu_{ij} = \mu u_{ij}$ such that $u_{31} = u_{21} = 0$, i.e. there is no back mutation from A_1 to A_2 and A_3. Moreover, we assume that $u_{12} + u_{32} = 1$, $u_{13} + u_{23} = 1$, and $u_{21} + u_{31} = 1$, and recall that $u_{ii} = 0$. Then it is easy to see that equilibrium mean fitness is $\bar{m} = -\mu(1 - \sqrt{u_{32} u_{23}})$. Therefore, the load can attain any value between 0 ($u_{23} = u_{32} = 1$) and μ ($u_{32} = 0$ or $u_{23} = 0$), but it is still independent of the selection coefficient s (because there is no back mutation).

2.2

The *diploid mutation-selection equation* is given by

$$\dot{p}_i = p_i(m_i - \bar{m}) + \sum_{j=1}^{n} (\mu_{ij} p_j - \mu_{ji} p_i), \quad i = 1, \ldots, n , \tag{2.4}$$

with

$$m_i = \sum_{j=1}^{n} m_{ij} p_j \quad \text{and} \quad \bar{m} = \sum_{i=1}^{n} m_i p_i = \sum_{i,j=1}^{n} m_{ij} p_i p_j .$$

For the equilibrium mean fitness with back mutations ignored, we refer to Crow and Kimura (1970, Chap. 6.4) in continuous time, and Nagylaki (1992, Sect. 4.9) in discrete time. We generalize these results to arbitrary mutation patterns among alleles.

For simplicity, we consider more generally

$$\dot{p}_i^o = p_i^o f_i(p^o), \quad i = 1, \ldots, n \tag{2.5}$$

and its perturbation

$$\dot{p}_i = F_i(p, \varepsilon) = p_i f_i(p) + \varepsilon v_i(p), \quad i = 1, \ldots, n\,. \tag{2.6}$$

Suppose, $\hat{p}^o$ is a regular (i.e. the Jacobian $D_pF(\hat{p}^o, 0)$ is invertible) and externally stable, or saturated (see Hofbauer and Sigmund 1988, Chap. 19.4), equilibrium point of (2.5): After renumbering the indices, this means

$$\hat{p}_i^o = 0, \quad f_i(\hat{p}^o) < 0 \quad \text{for} \quad 1 \leqq i \leqq k \tag{2.7.a}$$

(the case $f_i(\hat{p}^o) = 0$ is excluded by the regularity assumption on $\hat{p}^o$; see also Remark 3 below) and

$$\hat{p}_i^o > 0, \quad f_i(\hat{p}^o) = 0 \quad \text{for} \quad k < i \leqq n\,. \tag{2.7.b}$$

Note that (2.7.a) states that the growth rates at $\hat{p}^o$, $\dot{p}_i^o/p_i^o = f_i(\hat{p}^o)$, are negative, which means that none of the absent alleles can invade the equilibrium state $\hat{p}^o$. If only one allele is present at $\hat{p}^o$, we have $k = n - 1$ and (2.7.a) implies stability of $\hat{p}^o$. Furthermore, every (stable or unstable) interior fixed point (polymorphism) $\hat{p}^o$ is saturated because $k = 0$ and (2.7.a) is an empty condition.

By the implicit function theorem, there is a smooth curve $\hat{p}(\varepsilon)$, defined for small ε, of equilibrium points of (2.6), such that $\hat{p}(0) = \hat{p}^o$. Moreover,

$$\hat{p}(\varepsilon) = \hat{p}^o + \varepsilon q + O(\varepsilon^2)\,, \tag{2.8}$$

where q is the solution of the linearized equation

$$D_pF(\hat{p}^o, 0)q + D_\varepsilon F(\hat{p}^o, 0) = 0\,.$$

For the first k components, this simplifies to

$$f_i(\hat{p}^o)q_i + v_i(\hat{p}^o) = 0 \quad (i \leqq k)\,,$$

so that (2.8) turns into

$$\hat{p}_i(\varepsilon) = -\varepsilon \frac{v_i(\hat{p}^o)}{f_i(\hat{p}^o)} + O(\varepsilon^2) \quad (i \leqq k)\,. \tag{2.9}$$

Equation (2.9) shows also (the well-known fact) that a regular saturated fixed point of (2.5) moves into the interior of the state space after introduction of mutation terms, whereas a non-saturated fixed point (with at least one of the inequalities (2.7.a) reversed) will move out for $\varepsilon > 0$, thus becoming unacceptable. Hence, the question of mutation load makes sense only at saturated equilibria.

Let us now return to (2.4), setting $f_i(p) = m_i(p) - \bar{m}(p)$ and $\mu_{ij} = \varepsilon u_{ij}$, so that

$$v_i(p) = \sum_{j=1}^{n} (u_{ij}p_j - u_{ji}p_i)\,.$$

Then

$$\bar{m}(p) = \bar{m}(\hat{p}^o) + 2\sum_{i=1}^{k} (p_i - \hat{p}_i^o)(m_i(\hat{p}^o) - \bar{m}(\hat{p}^o)) + \bar{m}(p - \hat{p}^o)$$

and $\hat{p}^o$ is the equilibrium frequency vector of the unperturbed equation. Together with (2.7.b), (2.8), and (2.9), this implies

$$\bar{m}(\hat{p}(\varepsilon)) = \bar{m}(\hat{p}^o) - 2\varepsilon \sum_{i=1}^{k} \sum_{j=k+1}^{n} u_{ij} \hat{p}_j^o + O(\varepsilon^2) . \tag{2.10}$$

Again, the *first-order approximation* of the mutation load

$$L = 2 \sum_{i=1}^{k} \sum_{j=k+1}^{n} \mu_{ij} \hat{p}_j^o \tag{2.11}$$

has the same biological meaning as the total rate of mutations leading to new (less fit) alleles, occurring at the equilibrium $\hat{p}^o$, with a factor 2 because of diploidy.

Remarks and examples. 1) If we assume in (2.4) that selection terms are of order s and mutation terms of smaller order μ, then this situation reduces to the above with $\varepsilon = \mu/s$. After multiplying (2.10) by s, we see that the load is given again by (2.11), and the error term is of the order

$$O(\mu^2/s) . \tag{2.12}$$

2) In the case of two alleles ($A_1 = a$ = mutant, $A_2 = A$ = wild type), the above applies to the state $\hat{p}_1^o = 0$, whenever the heterozygote Aa has smaller fitness than AA: the mutation load then always is given by 2μ (in first order approximation). This corresponds to the discussion in Crow (1970, pp. 136–138).

3) If the mutant gene a is recessive, i.e., fitness of Aa = fitness of AA, the above argument does not apply, since the equilibrium $\hat{p}^0 = (0, 1)$ is not regular ($f_i(\hat{p}^0) = 0$ and the fitness has zero slope at $\hat{p}^0$). Direct calculations show that

$$p_1(\varepsilon) = \varepsilon^{1/2} + O(\varepsilon) \quad \text{(with } \varepsilon = \mu/s) \quad \text{and} \quad L = \mu$$

instead of (2.9) and (2.11) (cf. Crow 1970, p. 137).

4) If $k = 0$, i.e. p is an interior equilibrium, then $L = 0$ in (2.11). This means that the mutation load in (2.10) is actually of smaller order $O(\varepsilon^2)$, or $O(\mu^2/s)$ (following Remark 1). As an illustration, consider again the case of two alleles A_1 and A_2, with fitnesses of A_1A_1, A_1A_2, A_2A_2 being $-s_1$, 0, $-s_2$, respectively, and mutation rates $\mu_{12} = \mu_1$, $\mu_{21} = \mu_2$. Then the second-order approximation of the mutation load can be calculated as:

$$L = (s_1 + s_2) \left(\frac{\mu_1}{s_2} - \frac{\mu_2}{s_1} \right)^2 = \frac{(\mu_1/p_1 - \mu_2/p_2)^2}{s_1 + s_2} .$$

Here $p = (p_1, p_2) = (s_2, s_1)/(s_1 + s_2)$ is the overdominant selective equilibrium. If mutation alone leads to the same equilibrium as selection, then $L = 0$ of course. Otherwise, L is of order μ^2/s.

5) The above argument applies to generalizations of (2.5) if they are Shahshahani gradients (see Sigmund 1984, Hofbauer and Sigmund 1988, p. 241) of a homogeneous potential V (of degree d), i.e.

$$f_i(p) = \frac{\partial V}{\partial p_i} - dV(p) .$$

If $V(p) = \sum_i m_i p_i$, we have $d = 1$ and recover the haploid case 2.1. If $V(p) = \bar{m}/2 = \frac{1}{2}\sum m_i p_i = \frac{1}{2}\sum_{i,j} m_{ij} p_i p_j$, we have $d = 2$ and the diploid case.

3 Continuum of alleles: The asexual case

3.1 The general model and result

The model assumes an effectively infinite population that reproduces asexually and has overlapping generations. To cover not only the continuum-of-alleles case, but also models with discrete alleles, we need some abstract terminology. The types in the population are identified with points in a so-called locally compact space X. Typical special cases for applications are $X = \mathbb{R}$ (the real numbers), X some interval in $\mathbb{R}$ (if types are constrained in some way), $X = \{1, 2, \ldots, n\}$ (thus including the model of Sect. 2.1), or $X = \{0, \pm 1, \pm 2, \ldots\}$. Type densities and mutant distributions are taken with respect to a fixed positive, σ-finite Borel measure λ on X. Naturally, λ is simply the Lebesgue measure in cases one and two above ($X = \mathbb{R}$, or X some interval), and λ is the counting measure if X is a discrete set (cases three and four). In the first and second case, each $x \in X$ may be considered as an allelic effect on a quantitative trait, whereas in the third and fourth case, each $x \in X$ may be interpreted as a (possible) allele at a given gene locus.

To model selection in continuous time, a Malthusian fitness value $m(x)$ is assigned to each type $x \in X$. For technical simplicity, we assume that the function m is continuous. Throughout this paper we consider the case where the fitness function is bounded above, since we are only interested in models where a stable mutation-selection balance occurs. In particular, we assume that

$$m(x) \leqq 0 \quad \text{for all } x\,, \tag{3.1.a}$$

and

$$\text{there exists at least one } x, \text{ e.g., } x = 0, \text{ such that } m(x) = 0\,, \tag{3.1.b}$$

that is, the set of optimal types

$$M_{\text{Opt}} = \{x \in X : m(x) = 0\} \text{ is nonempty}\,, \tag{3.1.c}$$

and the maximal fitness is $m_{\max} = 0$. The reader may observe that the fitness functions m and $m + c$ (c any constant) give rise to the same allele-frequency dynamics; see (3.3) below.

To introduce mutation, assume that $\mu u(x, y)dt$ is the fraction of individuals of type x originating through mutation from individuals of type y during the time interval dt. Here, μ denotes the mutation rate and $u : X \times X \to \mathbb{R}_+$ is a nonnegative, Borel-measurable function that describes the conditional probability that a y-allele mutates to an x-allele, i.e.,

$$\int_X u(x, y)\lambda(dx) = 1 \quad \text{for all } y \in X\,. \tag{3.2}$$

(This was not imposed in Sect. 2). If a type remains unchanged, this is not considered as a mutation. Therefore, we assume $u(x, x) = 0$ for all $x \in X$. This assumption may and will be omitted in the continuum-of-alleles case when the set of all (x, x) has measure zero in $X \times X$.

Let $p(x, t)$ denote the type density (with respect to λ) at time t. Then it is well known that the evolution of type densities is given by

$$\dot{p}(x, t) = \frac{\partial}{\partial t} p(x, t)$$

$$= [m(x) - \bar{m}(t)]p(x, t) + \mu \left[\int_X u(x, y)p(y, t)\lambda(dy) - p(x, t) \right], \tag{3.3}$$

where $\bar{m}(t) = \int_X m(x)p(x, t)\lambda(dx)$ is the mean fitness (e.g., Kimura 1965).

In the following we shall be interested in equilibrium solutions of this equation. We shall consider the fitness function m and the mutant distribution u as fixed, and treat the mutation rate μ as a parameter. Our aim is to derive asymptotic estimates for the equilibrium mean fitness $\bar{m}_\mu$ as μ tends to zero. For simplicity, we use the operator notation

$$Uf(x) = \int_X u(x, y)f(y)\lambda(dy)$$

(see Appendix A for details). For technical reasons, an additional compactness condition has to be imposed on this operator U. Appendix A gives a list of simple conditions that ensure this assumption, which is satisfied in all our applications.

It follows from Fubini's theorem and (3.2) that

$$\int_X Up(x)\lambda(dx) = \int_X \left(\int_X u(x, y)\lambda(dx) \right) p(y)\lambda(dy) = 1 . \tag{3.4}$$

At equilibrium, we obtain from (3.3)

$$m(x)p_\mu(x) + \mu U p_\mu(x) = a_\mu p_\mu(x) , \tag{3.5.a}$$

where p_μ denotes the equilibrium density (depending on the parameter μ) and

$$a_\mu = \bar{m}_\mu + \mu = \int_X [m(x)p_\mu(x) + \mu U p_\mu(x)]\lambda(dx) . \tag{3.5.b}$$

Equation (3.5.a) yields

$$p_\mu(x) = \mu \frac{U p_\mu(x)}{a_\mu - m(x)} . \tag{3.6}$$

In Bürger (1988) simple sufficient conditions for the existence of an equilibrium solution p_μ (that is a density) were derived (see Appendix B), and it was shown that the corresponding "eigenvalue" a_μ is strictly positive. Because $\bar{m}_\mu \leqq 0$, (3.5.b) yields

$$0 < a_\mu = \bar{m}_\mu + \mu \leqq \mu . \tag{3.7.a}$$

In particular, we have

$$a_\mu \to 0 \text{ as } \mu \to 0 . \tag{3.7.b}$$

For the equilibrium mean fitness, we obtain from (3.6)

$$\begin{aligned}\bar{m}_\mu &= \int_X m(x)p_\mu(x)\lambda(dx)\\ &= \mu\int_X Up_\mu(x)\frac{m(x)}{a_\mu - m(x)}\lambda(dx)\\ &= -\mu\int_X Up_\mu(x)\lambda(dx) + \mu\int_X \frac{a_\mu}{a_\mu - m(x)}Up_\mu(x)\lambda(dx)\\ &= -\mu + \mu\int_{M_{\mathrm{Opt}}} Up_\mu(x)\lambda(dx) + \mu\int_{X\setminus M_{\mathrm{Opt}}} \frac{a_\mu}{a_\mu - m(x)}Up_\mu(x)\lambda(dx)\end{aligned}$$

($X\setminus M_{\mathrm{Opt}}$ denotes the complementary set of M_{Opt} in X). It follows that

$$\frac{\bar{m}_\mu + \mu(1 - \int_{M_{\mathrm{Opt}}} Up_\mu(x)\lambda(dx))}{\mu} = \int_{X\setminus M_{\mathrm{Opt}}} \frac{a_\mu}{a_\mu - m(x)}Up_\mu(x)\lambda(dx)\,. \tag{3.8}$$

In Appendix C it is proved that the expression on the right-hand side tends to zero as $\mu \to 0$. Therefore, we obtain from (3.8)

$$\lim_{\mu\to 0} \frac{\bar{m}_\mu + \mu(1 - \int_{M_{\mathrm{Opt}}} Up_\mu(x)\lambda(dx))}{\mu} = 0\,. \tag{3.9}$$

This shows that, in general, the equilibrium mean fitness, and thus the mutation load, may depend on details of the equilibrium distribution. A simple example was given at the end of Sect. 2.1. In that case, the load depended on the mutation pattern among the optimal alleles.

However, for a variety of important special cases, the equilibrium mean fitness is in fact independent of the fitness function m and of the particular form of the (conditional) probability distribution u of mutant effects. In the examples below, we assume that $X = \mathbb{R}$ or X is some interval in $\mathbb{R}$, λ is the Lebesgue measure (unless otherwise stated), and we will write dx instead of $\lambda(dx)$.

3.2 A unique optimal type

Suppose that there exists a unique optimal type or, more generally, $\lambda(M_{\mathrm{Opt}}) = 0$. Then (3.9) reduces to

$$\lim_{\mu\to 0} \frac{\bar{m}_\mu + \mu}{\mu} = 0\,. \tag{3.10.a}$$

This means that Haldane's result holds asymptotically, i.e., the mean fitness equals $-\mu(= m_{\max} - \mu)$ to first order, and is independent of the selection intensity and of the details of the mutant distribution. Therefore, the *first-order approximation* of the mutation load is

$$L = \mu\,. \tag{3.10.b}$$

Typical mutant distributions will be considered below, and second-order terms will be derived.

3.3 Kingman's (1978) house-of-cards model

This assumes that

$$u(x, y) = u(x) \text{ (by abuse of notation) if } x \neq y, \text{ and } \int_X u(x)dx = 1$$

(cf. (3.2)). It follows that $Up_\mu(x) = u(x)$. Therefore, (3.9) becomes

$$\lim_{\mu \to 0} \frac{\bar{m}_\mu + \mu \int_{X \setminus M_{\text{Opt}}} u(x)dx}{\mu} = 0 , \tag{3.11.a}$$

and the *first-order approximation* of the mutation load is

$$L = \mu \int_{X \setminus M_{\text{Opt}}} u(x)dx , \tag{3.11.b}$$

which is just the total mutation rate toward suboptimal alleles. Here we did not assume that $\lambda(M_{\text{Opt}}) = 0$.

For certain fitness functions and mutant distributions, the rate at which $(\bar{m}_\mu + \mu)/\mu$ tends to zero, and hence the order of the error term, can be determined easily. Suppose that $X = \mathbb{R}$. Let

$$m(x) = -s|x|^q \quad \text{and} \quad u(x) = |x|^{-1/n} g(x) ,$$

where

(i) g is an arbitrary, bounded, nonnegative function such that $\int u = 1$ and $g(x) \geqq g_0 > 0$ for $x \in [-c, c]$ (with $g_0 > 0$, $c > 0$ fixed but arbitrary), and

(ii) $n \geqq 1$ (including ∞) and $q > \beta = 1 - (1/n)$.

Thus, if $n = \infty$, we have $u = g$ and u may be a Gaussian or an exponential distribution reflected about 0. A Γ-distribution (reflected about 0) is obtained for

$$g(x) = \frac{a^\beta}{2\Gamma(\beta)} \exp(-a|x|) .$$

We denote $\|g\|_\infty = \sup_x g(x)$, and have due to (i), $\|g\|_\infty < \infty$.

It follows from (3.3) and (3.5) that with $m(x) = -s|x|^q$ the equilibrium distribution and its mean fitness depend only on μ/s and not on μ and s separately. Therefore, we may define

$$\alpha_{\mu/s} = a_\mu/s = (\bar{m}_\mu + \mu)/s .$$

Since we have $\int p_\mu = 1$ and $Up_\mu = u$, (3.6) shows that $\alpha_{\mu/s}$ can be calculated from the condition

$$\frac{\mu}{s} \int_{-\infty}^{\infty} \frac{|x|^{-1/n} g(x)}{\alpha_{\mu/s} + |x|^q} dx = 1 . \tag{3.12.a}$$

Let

$$r(\alpha) = \int_{-\infty}^{\infty} \frac{|x|^{-1/n} g(x)}{\alpha + |x|^q} dx . \tag{3.12.b}$$

Then (3.12.a) becomes

$$\frac{\mu}{s} r(\alpha_{\mu/s}) = 1 . \tag{3.12.c}$$

We obtain, using formula 431.16.a of Gröbner and Hofreiter (1975),

$$r(\alpha) \leqq 2\|g\|_\infty \int_0^\infty \frac{|x|^{-1/n}}{\alpha + |x|^q}\, dx$$

$$= 2\|g\|_\infty \frac{\pi}{q \sin\frac{\pi\beta}{q}} \cdot \alpha^{\frac{\beta}{q}-1} . \tag{3.13.a}$$

Furthermore, for $\alpha \in (0, c^q)$ we obtain, using property (i) of u and g, and the transformation $y = x^\beta$,

$$r(\alpha) \geqq \int_{-\alpha^{1/q}}^{\alpha^{1/q}} \frac{u(x)}{\alpha + |x|^q} dx$$

$$\geqq 2g_0 \int_0^{\alpha^{1/q}} \frac{|x|^{-1/n}}{\alpha + x^q} dx$$

$$= \frac{2g_0}{\beta} \int_0^{\alpha^{\beta/q}} \frac{dy}{\alpha + y^{q/\beta}}$$

$$\geqq \frac{g_0}{\alpha\beta} \int_0^{\alpha^{\beta/q}} dy$$

$$= \frac{g_0}{\beta} \alpha^{\frac{\beta}{q}-1} . \tag{3.13.b}$$

Therefore, we obtain from (3.12.c) and (3.13) the asymptotic equality

$$\alpha_{\mu/s} \sim (\text{const.}) \left(\frac{\mu}{s}\right)^{q/(q-\beta)} \tag{3.14}$$

as $\mu/s \to 0$. It follows that

$$\frac{\bar{m}_\mu + \mu}{\mu} \sim (\text{const.}) \left(\frac{\mu}{s}\right)^{\frac{\beta}{q-\beta}} , \tag{3.15.a}$$

as $\mu/s \to 0$, and the order of the (absolute) error term of L in (3.10.b) is

$$(\text{const.})\, \mu \left(\frac{\mu}{s}\right)^{\frac{\beta}{q-\beta}} . \tag{3.15.b}$$

Note that $\beta/(q-\beta) > 0$ holds under our assumptions. The constant in (3.15.b) depends on the function g, and estimates may be obtained from (3.13).

It is interesting to notice that the error term of the load L is $O(\mu^2/s)$, as for n alleles, if and only if $q = 2\beta$. If $u = g$ is a Gaussian or a reflected exponential distribution (thus, $\beta = 1$) and $q > 1$, then

$$\frac{\bar{m}_\mu + \mu}{\mu} \sim (\text{const.}) \left(\frac{\mu}{s}\right)^{\frac{1}{q-1}} \tag{3.16}$$

follows from (3.15.a), and the order of the error term of L is $O(\mu^2/s)$ if and only if $m(x) = -sx^2$. In Appendix D the asymptotic equality (3.16) is extended to non-HC-type mutant distributions. In the limiting case $q \to \infty$, the first-order approximation

of the mutation load is not necessarily μ because then the order of the error term (3.16) is $O(\mu)$. In fact, then (3.11.b) must be applied because all types $x \in [-1, 1]$ have optimal fitness 0 and all other types are lethal.

If $\beta \geqq q$, the estimates leading to (3.15) do not apply because the integrals do not necessarily exist and the equilibrium distribution may have an atom of probability at $x = 0$. A necessary and sufficient condition for the existence of an equilibrium solution with an absolutely continuous part (with respect to λ) and a singular part (in the present case an atom of probability at $x = 0$) is

$$\mu \int_X \frac{u(x)}{-m(x)} dx < 1 . \tag{3.17}$$

Moreover, if

$$\mu \int_X \frac{u(x)}{-m(x)} dx \leqq 1 , \tag{3.18}$$

then

$$\bar{m}_\mu = -\mu \tag{3.19}$$

(see Theorem 4.1 in Bürger and Bomze 1992).

The following example treats a critical case, where the method leading to (3.15) does not apply. Let $X = [0, 1]$, λ be the Lebesgue measure, $u(x) \equiv 1$, and $m(x) = -sx^q$ with $q \leqq 1$. For $q = 1$, we obtain instead of (3.12.a)

$$1 = \int_0^1 p_\mu(x)\lambda(dx) = \frac{\mu}{s} \ln\left(1 + \frac{s}{a_\mu}\right) .$$

It follows that

$$\frac{\bar{m}_\mu + \mu}{\mu} = \frac{1}{\frac{\mu}{s} e^{s/\mu} - 1} . \tag{3.20.a}$$

Therefore,

$$\frac{\bar{m}_\mu + \mu}{\mu} = O\left(\left(\frac{\mu}{s}\right)^n\right) \quad \text{for all } n \geqq 1 . \tag{3.20.b}$$

For $q < 1$, we obtain (3.19) because (3.18) is satisfied for sufficiently small μ.

3.4 "Translational" mutation

(See Crow and Kimura 1964)

This assumes that the effect of new alleles simply adds to the effect of the parental allele, i.e.

$$u(x, y) = u(x - y) \text{ if } x \neq y, \text{ with } \int_X u(x)dx = \int_X U p_\mu(x)dx = 1 .$$

This has become the most popular mutation model in quantitative genetics, and has been used with various mutant distributions u. Again (3.9) and, for $\lambda(M_{\text{Opt}}) = 0$, Eqs. (3.10.a, b) hold. In particular, it is again possible to derive estimates analogous to (3.16), but the derivation uses methods from functional analysis. In this case the integral $r(\alpha)$ has to be replaced by the spectral radius of a certain operator. The proof is given in Appendix D.

3.5

In Bürger and Bomze (1992) it was shown under very general assumptions that (3.19) always holds if the stationary distribution consists of an absolutely continuous *and* a singular part. This is the case if the fitness function m has a cusp at its optimum, and/or if the mutation rate to types near the optimal type is extremely low (cf. (3.18) and Appendix B).

3.6 Discrete alleles

In this classical case, we have $X = \{0, 1, 2, \ldots, n\}$ or $X = \{0, \pm 1, \pm 2, \ldots\}$, and λ is simply the counting measure. For finitely many alleles the results of Sect. 2.1 are easily obtained from (3.9). However, the "ladder", or "stepwise mutation model", where $X = \{0, \pm 1, \pm 2, \ldots\}$, is also included, and results parallel to those of Sect. 2.1 can be derived. If all mutations are deleterious and arise in a Poisson fashion, then the result of Kimura and Maruyama (1966) that the load is $L = 1 - e^{-\mu}$ follows immediately from (3.3) by considering only the optimal type (cf. the last paragraph of Sect. 2.1).

4 The equilibrium variance of metric traits under mutation-selection balance

4.1 The haploid case

Many authors have provided approximations for the equilibrium variance of a quantitative genetic character subject to mutation-selection balance. Two qualitatively different types of approximations were obtained: the Gaussian approximation (Kimura 1965, Lande 1975, Fleming 1979, Nagylaki 1984) and the so-called house-of-cards (HC-) approximation (Turelli 1984, Turelli and Barton 1990). Kimura and Turelli obtained their approximations by considering the asexual model of Sect. 3 and subsequently extrapolating to the multilocus case (assuming global linkage equilibrium). Lande, Fleming, Nagylaki, and Turelli and Barton, included linkage in their theoretical analyses. All these authors assumed that mutation is "translational" and the mutant distribution is Gaussian, i.e., $u(x, y) = u(x - y)$ and $u(x) = (2\pi\gamma^2)^{-1/2} \exp(-x^2/2\gamma^2)$. Kimura used a continuous-time model with fitness given by $m(x) = -sx^2$, whereas the others considered the discrete-time analog $w(x) = \exp(-x^2/2V_s)$, with the correspondence $s = 1/2V_s$. Kimura's Gaussian approximation states that (in the haploid case) the equilibrium variance is

$$\sigma^2 \approx \sqrt{\frac{\mu}{s}\frac{\gamma^2}{2}}\,. \tag{4.1}$$

Turelli performed an HC-approximation, i.e., he assumed that the mutant distribution depends only on the target gene (Kingman's 1978, HC-mutation model), and is Gaussian. Intuitively, this should be a good approximation if the variance of mutant effects γ^2 is much larger than the equilibrium variance. He obtained the HC-approximation

$$\sigma^2 \approx \frac{\mu}{s}\,. \tag{4.2}$$

In fact, numerical simulations showed that the Gaussian approximation applies if γ^2 is small compared with μ/s, whereas (4.2) applies in the opposite case (cf. Turelli 1984, Bürger 1986). In Bürger (1986) and, more generally, in Bürger (1988) it was shown that μ/s is an upper bound to the equilibrium variance in this model. This holds for *any* mutant distribution.

Here, we will use the results of Sect. 3 to prove that the variance in this mutation-stabilizing-selection model approaches the HC-approximation for *any* mutant distribution $u(x, y)$ (that satisfies our general assumptions and a mild additional one) as mutation becomes weak relative to selection. We will also see how different mutant distributions affect the accuracy of this approximation.

It follows from Sect. 3.2 that for

$$m(x) = -s(x - x_0)^2 ,$$

with fixed s and arbitrary (but fixed) mutant distribution $u(x, y)$, (3.10.a) holds. Let $\bar{x}$ and σ^2 denote the mean and the variance at equilibrium, respectively. Then, the equilibrium mean fitness is $\bar{m} = -s(\sigma^2 + (\bar{x} - x_0)^2)$ (we omit the index μ of Sect. 3 here). Therefore, (3.10.a) implies

$$\lim_{\mu \to 0} \frac{\sigma^2 + (\bar{x} - x_0)^2 - \mu/s}{\mu/s} = 0 .$$

In fact, it is easy to see from the derivation in Sect. 3 that it is sufficient to fix u, and the slightly stronger result

$$\lim_{\mu/s \to 0} \frac{\sigma^2 + (\bar{x} - x_0)^2 - \mu/s}{\mu/s} = 0 \tag{4.3.a}$$

holds. This is true *for any mutant distribution* that satisfies our general assumptions.

If the mutant distribution is such that the mean is at the optimum, i.e., $\bar{x} = x_0$, then we obtain

$$\lim_{\mu/s \to 0} \frac{\sigma^2 - \mu/s}{\mu/s} = 0 . \tag{4.3.b}$$

The mean is at the optimum, for example, if $u(x_0 + x, x_0 + y) = u(x_0 - x, x_0 - y)$, i.e., if u is symmetric around the optimum.

We want to prove (4.3.b) without imposing symmetry assumptions on u. Instead, we require that u be bounded, i.e.,

$$\| u \|_\infty = \sup_{x,y} u(x, y) < \infty . \tag{4.4.a}$$

As a consequence, we have

$$\| Up \|_\infty \leqq \| u \|_\infty < \infty \quad \text{if} \quad \int |p(x)|\, dx = 1 . \tag{4.4.b}$$

Assumption (4.4.a) is obviously met for Gaussian and (reflected) exponential distributions, in the HC and in the translational mutation model.

To prove (4.3.b) if $\bar{x} \neq x_0$, we observe that, on account of (4.3.a), it is sufficient to show

$$\lim_{\mu/s \to 0} \left(\frac{\mu}{s}\right)^{-1} (\bar{x} - x_0)^2 = 0 \,. \tag{4.5}$$

Let $\hat{p}$ be the equilibrium density. Then, using (3.6), the definition of $\alpha_{\mu/s}$, the transformation $x \mapsto x_0 + x$, a decomposition of the integral, (4.4.b), and (3.4), we obtain

$$\begin{aligned}
|\bar{x} - x_0| &\leqq \int |x - x_0| \hat{p}(x) dx \\
&= \frac{\mu}{s} \int \frac{|x - x_0|}{\alpha_{\mu/s} + (x - x_0)^2} U\hat{p}(x) dx \\
&= \frac{\mu}{s} \int \frac{|x|}{\alpha_{\mu/s} + x^2} U\hat{p}(x + x_0) dx \\
&= \frac{\mu}{s} \left(\int_{-1}^{1} \frac{|x|}{\alpha_{\mu/s} + x^2} U\hat{p}(x + x_0) dx + \int_{|x|>1} \frac{|x|}{\alpha_{\mu/s} + x^2} U\hat{p}(x + x_0) dx \right) \\
&\leqq \frac{\mu}{s} \left(\| U\hat{p} \|_\infty \int_{-1}^{1} \frac{|x|}{\alpha_{\mu/s} + x^2} dx + \int U\hat{p}(x) dx \right) \\
&\leqq \frac{\mu}{s} \left(\| u \|_\infty \ln \left(1 + \alpha_{\mu/s}^{-1}\right) + 1 \right) \,.
\end{aligned} \tag{4.6.a}$$

We know from (D.10) that $\alpha_{\mu/s} \geqq (\text{const.})(\mu/s)^2$, and thus arrive at

$$\left(\frac{\mu}{s}\right)^{-1} (\bar{x} - x_0)^2 \leqq (\text{const.}) \left(\frac{\mu}{s}\right) \left(\ln \frac{\mu}{s}\right)^2 . \tag{4.6.b}$$

This proves (4.5) and, hence (4.3.b) in the general case. Therefore, combining (4.3.b), (D.11), and (4.6.b), we obtain

$$\sigma^2 = \frac{\mu}{s} + O\left(\left(\frac{\mu}{s}\right)^2 \left(\ln \frac{\mu}{s}\right)^2 \right) . \tag{4.7}$$

(Recall that $\sigma^2 < \mu/s$, so the constant in the O-term is negative.)

Remark In the HC-case, a similar result can be obtained without the assumption (4.4.a) for the distributions considered below (3.11) if $\beta > 1$; the (reflected) gamma distribution is an example. Instead of (4.6.a), it can be proved that

$$|\bar{x} - x_0| \leqq \left(\frac{\mu}{s}\right) \left(c_1 \alpha_{\mu/s}^{(\beta-1)/2} + c_2 \right) ,$$

and (4.8) below is obtained.

For the remainder of Sect. 4.1, we assume that $\bar{x} = x_0$. This allows us to improve (4.3) and (4.7). First consider the HC-mutation model with mutant distributions defined below (3.11). Then we obtain

$$\sigma^2 = \frac{\mu}{s} + O\left(\left(\frac{\mu}{s}\right)^{\frac{2}{2-\beta}} \right) \tag{4.8}$$

from (3.15). The exponent of the second-order term is two (as for discrete-allelic models; see Sect. 2) if and only if $\beta = 1$, e.g., for a Gaussian or an exponential

mutant distribution (reflected about zero). Since for highly leptokurtic Γ-distributions ($\beta \to 0$) the exponent of the second term tends to one, the accuracy of the HC-approximation decreases with increasing kurtosis (small β).

If the mutant distribution u is Gaussian with mean x_0 and variance γ^2, then (4.8) can be further improved. For the HC-mutation model, formula 314.8.b of Gröbner and Hofreiter (1975) shows that (3.12.a) becomes

$$\sqrt{\frac{\pi}{2\gamma^2}}\,\frac{\mu}{s}\,\frac{\exp\left(\frac{\alpha_{\mu/s}}{2\gamma^2}\right)}{\sqrt{\alpha_{\mu/s}}}\left[1-\operatorname{erf}\sqrt{\frac{\alpha_{\mu/s}}{2\gamma^2}}\,\right]=1\,,$$

where $\operatorname{erf}(x)=2\pi^{-1/2}\int_0^x \exp(-t^2)dt$ denotes the error function. For $\mu/s \to 0$ (hence $\alpha_{\mu/s} \to 0$) or $\gamma^2 \to \infty$, the left-hand side is, to first order,

$$\sqrt{\frac{\pi}{2\gamma^2}}\,\frac{\mu}{s}\,\frac{1}{\sqrt{\alpha_{\mu/s}}}\,.$$

Noticing

$$\sigma^2=-\frac{\bar{m}}{s}=\frac{\mu}{s}-\alpha_{\mu/s}\,,$$

we obtain

$$\sigma^2 \sim \frac{\mu}{s}-\frac{\pi}{2\gamma^2}\left(\frac{\mu}{s}\right)^2, \tag{4.9}$$

and the HC-approximation (4.2) applies as $\mu/(s\gamma^2) \to 0$.

If mutation is translational (Sect. 3.4) and u is Gaussian, as in the original model of Crow and Kimura (1964), and in Kimura (1965), then (D.14) in Appendix D yields the following bounds for the equilibrium variance for arbitrary μ/s:

$$\frac{\mu}{s}-\frac{\pi}{2\gamma^2}\left(\frac{\mu}{s}\right)^2 \leqq \sigma^2 \leqq \frac{\mu}{s}-\frac{\exp\left(-\frac{4b^2}{\gamma^2}\right)}{2\pi\gamma^2}\left(\frac{\mu}{s}\right)^2, \tag{4.10}$$

with

$$b^2=\min\left(\frac{\mu}{s},\,\frac{\pi}{2\gamma^2}\left(\frac{\mu}{s}\right)^2\right)\,,$$

as in (D.12). Note that the constants associated with the second-order term $(\mu/s)^2$ differ approximately by a factor π^2 for small μ/s. Essentially the same inequality (with $4b^2$ on the right-hand side replaced by b^2) may be obtained directly from (3.13) for the HC-mutation model by observing (D.12) and putting

$$c=b \quad \text{and} \quad g_0=\frac{1}{\sqrt{2\pi\gamma^2}}\exp\left(-\frac{b^2}{2\gamma^2}\right)\,.$$

The estimates (4.7) to (4.10) improve a result in Bürger (1986), which states that the equilibrium variance is always less than μ/s. If μ/s becomes large compared with γ^2, the lower bound becomes negative (and thus useless) and the upper bound is almost identical to μ/s. The upper and the lower bounds (4.10) are of most interest if μ/s is small (compared with γ^2). Figure 1 displays the Gaussian approximation (4.1), the HC-approximation (4.2), the upper and the lower bounds (4.10), and numerical

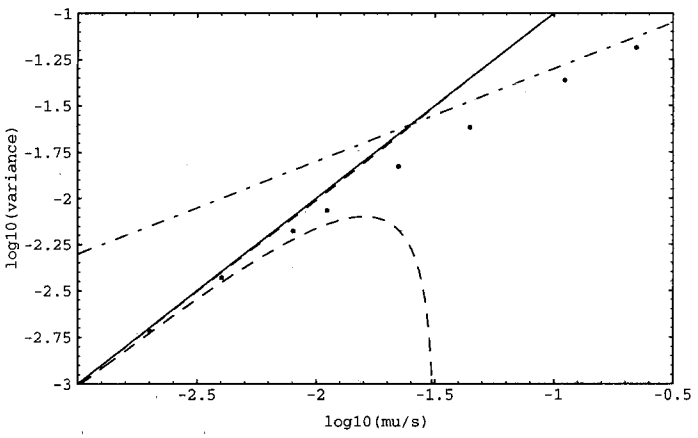

Fig. 1. The equilibrium variance and some of its approximations for mutation-stabilizing-selection balance. The plot is double logarithmic (to base 10), i.e., it shows the logarithm of the equilibrium variance, $\log_{10}(\sigma^2)$, as a function of $\log_{10}(\mu/s)$. The variance of the Gaussian mutant distribution is $\gamma^2 = 0.05$. The *solid line* represents the HC-approximation (4.2), the *dashed-dotted line* the Gaussian approximation, and the *dashed lines* the upper and lower bounds (4.10). The *dots* are data points obtained from numerical calculation (see main text). The Gaussian and the HC-approximation intersect at $\mu/s = \gamma^2/2$.

data for the equilibrium variance that either are taken from Bürger (1986) (mainly for large μ/s) or were obtained by numerical integration of (3.3), as described in detail in Bürger (1993). The Gaussian approximation performs better than the HC-approximation if and only if it is smaller, i.e., if $\mu/s > \gamma^2/2$. For the corresponding discrete-time model, Fleming (1979) obtained higher-order approximations than the Gaussian in the limit $\gamma^2(\mu/s)^{-1} \to 0$ (see also Nagylaki 1984, 1992). These are, in fact, lower than the Gaussian approximation (4.1), but it was not proved that the Gaussian always provides also an upper estimate, as does the HC-approximation.

4.2 The diploid multi-locus case

Here we extend the above results to quantitative traits by using the model and methods of Bürger (1991). We consider a quantitative character in a randomly mating diploid population with no sex differences that is affected by ℓ additive loci exhibiting neither dominance nor epistasis. Let $\mathcal{Z}$ denote the phenotypic value of this trait, $\mathcal{G}$ its geotypic value and $\mathcal{E}$ an independent, normally distributed environmental effect, so that

$$\mathcal{Z} = \mathcal{G} + \mathcal{E} \quad \text{and} \quad \mathcal{G} = \sum_{i=1}^{\ell} (x_i + y_i) ,$$

where $x_i(y_i)$ denotes the allelic effect of the maternally (paternally) inherited allele at the ith locus. The population is so large that random drift can be ignored. As in the previous sections, we assume a continuous-time model and, additionally, global linkage equilibrium (see Bürger 1991, for a discussion). However, we allow different mutation rates and mutant distributions at all loci, and make no assumptions about the position of mean values of allelic or genotypic distributions. Mutant distributions satisfy our general assumptions, as well as (4.4.a).

We start with the fitness of phenotypic values and assume quadratic stabilizing selection with optimum at $\mathcal{O}$, i.e.,

$$m_P(\mathcal{Z}) = a_0 - s(\mathcal{Z} - \mathcal{O})^2 .$$

Then the fitness of genotypic values is

$$m(\mathcal{G}) = b_0 - s(\mathcal{G} - \mathcal{O})^2 ,$$

where $b_0 = a_0 - sV_e$ and V_e denotes the environmental variance.

Let $\sigma_{\mathcal{G}}^2$ denote the genetic variance at equilibrium. We scale the per-locus mutation rates such that $\mu_i = \mu u_i$, where the u_i are strictly positive constants. Our aim is to prove

$$\lim_{\mu/s \to 0} \frac{\sigma_{\mathcal{G}}^2 - 2\sum_i \mu_i/s}{2\sum_i \mu_i/s} = 0 , \tag{4.11}$$

where the summation is over all ℓ loci.

The proof is by reduction to the haploid one-locus case. First, we notice that, in analogy to (3.6), the (haploid) equilibrium allelic density $\hat{p}_i$ at locus i is the (unique, positive) solution of

$$\hat{p}_i(x_i) = \mu_i \frac{\int u_i(x_i, \xi_i)\hat{p}_i(\xi_i)d\xi_i}{\bar{m} + \mu_i - \hat{m}_i^*(x_i)} \tag{4.12}$$

where $\hat{m}_i^*(x_i)$ denotes the marginal fitness of x_i at equilibrium, and $\bar{m}$ is the mean fitness of the population (cf. Eq. (2.1) in Bürger 1991). The expression for the marginal fitness may be obtained by specializing Eq. (3.5.a) in Bürger (1991), but we use the opportunity to present a simplification of some of the main formulas in that paper (see Appendix E). From (E.1) and (E.2) we obtain, using $\bar{\mathcal{G}}$ for the genotypic mean, and $\bar{x}_i$ for the allelic mean at locus i (at equilibrium),

$$\hat{m}_i^*(x_i) = b_0 - s(\sigma_{\mathcal{G}}^2 - \sigma_i^2) - s(x_i - [\mathcal{O} - (\bar{\mathcal{G}} - \bar{x}_i)])^2 \tag{4.13}$$

and

$$\bar{m} = b_0 - s(\sigma_{\mathcal{G}}^2 + (\bar{\mathcal{G}} - \mathcal{O})^2) . \tag{4.14}$$

Therefore, we can apply our results in Sect. 4.1 to $\hat{p}_i$ with

$$x_0 = \mathcal{O} - \left(\sum_{j \neq i} (\bar{x}_j + \bar{y}_j) + \bar{y}_i \right) = \mathcal{O} - (\bar{\mathcal{G}} - \bar{x}_i) ,$$

and $m_i(x_i) = -s(x_i - x_0)^2$. Thus, the fitness optimum "experienced" by the maternal alleles at locus i is the difference between the phenotypic fitness optimum and the mean contribution of all other loci together with the paternal alleles at locus i.) Then $\bar{x}_i - \bar{x}_0 = \bar{\mathcal{G}} - \mathcal{O}$, and (4.5) implies that

$$\frac{(\bar{\mathcal{G}} - \mathcal{O})^2}{\mu_i/s} \to 0 \quad \text{for each } i, \text{ as } \mu/s \to 0 .$$

Hence, (4.3.b) yields

$$\lim_{\mu/s \to 0} \frac{\sigma_i^2 - \mu_i/s}{\mu_i/s} = 0 \quad \text{for all } i .$$

This finishes the proof of (4.11) because

$$\left| \frac{\sigma_G^2 - 2\sum_i \mu_i/s}{2\sum_i \mu_i/s} \right| \leqq (\text{const.}) \sum_{i=1}^{\ell} \left| \frac{\sigma_i^2 - \mu_i/s}{\mu_i/s} \right| .$$

In the special case $\bar{G} = \mathcal{O}$, the above reasoning, together with (4.10), leads to

$$\frac{\sigma_G^2 - 2\sum_i \mu_i/s}{2\sum_i \mu_i/s} = \mathcal{O}\left(\left(\frac{\mu}{s}\right)^2\right) . \tag{4.15}$$

These results show that the multilocus HC-approximation is valid to first order, even if the mean deviates from the optimum. We notice that any equilibrium solution has to satisfy

$$\sigma_i^2 + (\bar{G} - \mathcal{O})^2 < \frac{\mu_i}{s} \tag{4.16}$$

for all i, because the denominator on the right-hand side of (4.12) must be strictly positive for all x_i (cf. (3.7.a)). In particular, this shows that

$$\sigma_G^2 < 2\sum_{i=1}^{\ell} \frac{\mu_i}{s} , \tag{4.17.a}$$

and

$$(\bar{G} - \mathcal{O})^2 < \min_i \frac{\mu_i}{s} \tag{4.17.b}$$

must hold for any equilibrium distribution, and for arbitrary parameters. The estimates (4.16) and (4.17) do not require assumption (4.4.a), i.e., boundedness of u.

Turelli and Barton (1990) derived approximations for the effects of linkage on the equilibrium distribution in the multilocus mutation-stabilizing-selection model. The influence of linkage turned out to be small. This is consistent with the exact analysis of a two-locus two-allele model (Bürger 1989). For investigations of the multilocus Gaussian approximation that include linkage, the reader is referred to Lande (1975), Fleming (1979), and Nagylaki (1984, 1992). Again, linkage causes only small deviations. In all of these investigations, it was assumed that the mean $\bar{G}$ is at the optimum $\mathcal{O}$, and that mutation is translational.

5 Summary and diccussion

Our paper generalizes earlier results on the mutation load (e.g., Haldane 1937, Crow and Kimura 1964, Crow 1970) in various directions. We prove that Turelli's (1984) HC-approximation applies to first order for any mutant distribution as $\mu/s \to 0$, and derive new estimates for the equilibrium variance under the well-known mutation-stabilizing-selection model. Below is a short summary of our main findings. We use a continuous-time model and derive approximations for the mutation load in the limit $\mu \to 0$ or $\mu/s \to 0$. Since continuous-time models and the corresponding discrete-time models usually have the same equilibrium behavior, our results should carry over to the latter case too.

(i) We treated the classical case with n alleles at a single locus (haploid and diploid) and allowed for arbitrary mutation rates among alleles, in particular for

back mutations. If there is one allele that has higher fitness than all the other alleles, the mutation load is – to the first-order approximation – the total mutation rate from the fittest allele to all the others (see (2.3) and (2.11)). This is in accordance with earlier results (see Crow 1970). In general, the load may depend on the selection coefficients, but only to second order. According to (2.12), the error term is $O(\mu^2/s)$. If there are two or more alleles with maximal fitness, the first-order approximation depends strongly on the mutation pattern among the optimal alleles.

(*ii*) For a continuum of possible alleles at a locus in a haploid population, we derived the general formula (3.9) for the first-order approximation of the equilibrium mean fitness. In the special but important case that exactly one allele has optimal fitness (e.g., $m(x) = -sx^2$) the mutation load is again – to first order – equal to the mutation rate. This holds independently of the fitness function and of the particular mutation pattern; see (3.10). For certain types of fitness functions ($m(x) = -s|x|^q$, $q \geqq 1$) and mutant distributions (Gaussian, exponential, reflected gamma) the order of the error term is calculated; see (3.15) and (3.16).

(*iii*) For the classical mutation-selection-balance model (e.g., Crow and Kimura 1964, Kimura 1965, Lande 1975, Turelli 1984) with Malthusian fitness $m(x) = -sx^2$ and Gaussian mutant distribution with mean $\bar{u}$ and variance γ^2, it follows that the load is μ with error term $O(\mu^2/s)$. This is in contrast to the result of Crow and Kimura (1964, Eq. (12)) that the load is $(\mu\bar{u}^2/(\bar{u}^2 + \gamma^2)) + \sqrt{\mu s\gamma^2(\bar{u}^2 + \gamma^2)/2}$. The reason for this difference is that their analysis is based on the Gaussian approximation, which does not apply for small $\mu/(s\gamma^2)$ (but does so for small $s\gamma^2/\mu$).

(*iv*) We proved that the equilibrium variance in that model satisfies the bounds (4.10), with b as in (D.12). For the HC-mutation model it was proved that the variance tends to the lower bound as $\mu/s \to 0$. More generally, the house-of-cards approximation for the equilibrium variance (Turelli 1984) applies to first order for small μ/s and arbitrary mutant distributions, even if the population mean deviates from the optimum. We derived also the order of the error term for mutant distributions such as reflected Γ-distributions; see (4.8). It follows that the accuracy of the HC-approximation decreases with increasing kurtosis of such mutant distributions.

(*v*) Under the assumption of global linkage equilibrium, we showed that the equilibrium genetic variance in the multilocus case is, to first order, $2\sum_i \mu_i/s$, as μ_i/s tends to zero for all i. This generalization to the multilocus model with quadratic stabilizing selection on the phenotype allows different mutation rates and mutant distributions at individual loci, and does not require that the population mean be at the optimum. More generally, we proved that for arbitrary mutant distributions and mutations rates, the genetic variance is bounded by $2\sum_i \mu_i/s$, as shown in (4.17.a), and that the deviation of the mean from the optimum satisfies $(\bar{\mathcal{G}} - \mathcal{O})^2 < \min_i(\mu_i/s)$, as shown in (4.17.b).

The inequalities (4.17.a, b) are in contrast to the equilibrium behavior of Barton's (1986) symmetric, biallelic, multilocus model, who found equilibria where the deviation of the mean from the optimum may reach the value of an allele's effect, and where the variance is higher than the HC-approximation. The reason for the discrepancy between the model used by Barton and the present one, and, in particular, for the fact that the deviation of the mean from the optimum is determined by the minimal per-locus mutation rate (4.17.b), is the following: without mutation, in the continuum-of-alleles model any given optimum (that lies in the range of variation of the trait) can be perfectly matched through the build-up of spikes in the distribution at each locus such that $2\sum_i \bar{x}_i = \mathcal{O}$, where the $\bar{x}_i$'s denote the positions

of the spikes. Thus, no polymorphism is maintained. Deviations from the optimum are introduced through asymmetries of the mutant distributions. These deviations decrease as the mutation rate decreases. In Barton's (1986) symmetric model, only half of the optimum values lead to monomorphic equilibria. This explanation is consistent with Barton's (1986) discussion, where he mentions that introduction of asymmetries drives populations much closer to the "optimal" equilibrium, and to the HC-approximation.

Another interesting point is that, in biallelic models (Wright 1935, Bulmer 1972, Barton 1986, Keightley and Hill 1988), stabilizing selection leads to underdominance, i.e., to disruptive selection, at each individual locus. This is not the case in the continuum-of-alleles model, where each locus experiences stabilizing selection, as may be seen from (4.12), (4.13), and (4.14). The phenomenon of underdominance is an artifact of the assumption of global linkage equilibrium together with the symmetries of that model. If one considers, for example, a full two-locus model, then each marginal locus experiences directional selection, and only one equilibrium that involves heterozygotes exists. It is unstable, and both loci are heterozygous at this equilibrium (see Bürger 1989).

Finally, we remark that we proved neither existence nor uniqueness of an equilibrium solution in the multilocus model. The results of Sect. 4.2 apply to any existing equilibrium. Only in the special case that the mutant distributions at individual loci are such that $2\sum_i \bar{x}_i = \bar{G} = \mathcal{O}$ (and each $\bar{x}_i$ coincides with the optimum and the mean for the corresponding haploid model at locus i) does existence of a multilocus equilibrium follow from the haploid existence result in a straightforward way. This includes the case $\bar{x}_i = \bar{G} = \mathcal{O} = 0$ and all mutant distributions symmetric around 0.

In fact, it turns out that uniqueness of equilibrium solutions may depend on the mutation model. It is easy to see that in the translational mutation model, an infinite number of equilibria exists. To show this, suppose that the equilibrium allelic densities $\hat{p}_1, \ldots, \hat{p}_\ell$ define an equilibrium solution, choose constants a_i such that $\sum_i a_i = 0$, and translate each $\hat{p}_i$ by a_i. This gives a new equilibrium solution with identical phenotypic distribution. For the Gaussian model, this property was discovered and discussed by Lande (1975, 1980). Kimura (1981) observed and discussed a similar phenomenon. It provides the possibility of extensive neutral evolution at individual loci. Other mutation models, such as the HC-model, do not have this property. For simplicity, consider the HC-model. Then the equation for the change of marginal allelic distributions (Bürger 1991, Eq. (2.1)) shows that per-locus means change according to $\dot{\bar{x}}_i = \mu_i a_i$, the phenotypic mean according to $\dot{\bar{G}} = 2\sum_i \mu_i a_i$, the per-locus variances according to $\dot{\sigma}_i^2 = \mu_i(a_i^2 - 2a_i\bar{u}_i)$, etc. Thus, the changes at individual loci are still very small, so that some neutral evolution in finite populations may still be possible. For the two-locus HC-model, it is easy to show directly (using (4.12)) that only a single equilibrium can exist. In general, mutant distributions may be different at different loci, and therefore allelic distributions may be "translatable" at some loci but not at others.

Appendix A

We denote by $L^1(\lambda)$ the Banach space of all Lebesgue-integrable functions such that

$$\|f\|_1 = \int |f(x)|\,\lambda(dx) < \infty\,.$$

Throughout the paper, we need the following technical assumption. The operator $U : L^1(\lambda) \to L^1(\lambda)$ defined by

$$Uf(x) = \int_X u(x, y) f(y) \lambda(dy) \tag{A.1}$$

is bounded on $L^1(\lambda)$, and compact as an operator of $L^1(\lambda_1) \to L^1(\lambda)$, where λ_1 denotes the measure $(1 - m)\lambda$. Thus, $L^1(\lambda_1) = \{f \in L^1(\lambda) : mf \in L^1(\lambda)\} \subseteq L^1(\lambda)$. It is known from Bürger (1988) that any equilibrium solution p_μ of (3.5) is in $L^1(\lambda_1)$. That the operator U is bounded follows from (3.2). The compactness condition is met in the following cases (see Bürger 1988):

(*i*) The house-of-cards model, i.e., $u(x, y) = u(x)$ (by abuse of notation) and $\int_X u = 1$, because then the operator U has rank one.

(*ii*) The state space X is a locally compact group and mutation is translational, i.e., $X = \mathbb{R}$ or $X = \{0, \pm 1, \pm 2, \ldots\}$, $u(x, y) = u(x - y)$ and $\int_X u = 1$. In this case, we need the additional assumption that $1/(1 - m(x))$ vanishes at infinity. The latter condition is satisfied if and only if $m(x) \to -\infty$ as $|x| \to \infty$.

(*iii*) X is compact, e.g., $X = [0, 1]$ or $X = \{1, 2, \ldots, n\}$, and $u(x, y)$ is continuous.

Appendix B

Existence and uniqueness of an equilibrium distribution that has a density with respect to λ requires the compactness condition of Appendix A (in fact, a weaker condition is sufficient; see Bürger 1988, and Appendix D) and the following:

Let $x_0 \in M_{\text{Opt}}$ and define

$$I_\alpha = \{x \in X : -\alpha \leqq m(x)\} \cap U(x_0)\,,$$

where $U(x_0)$ is a small neighbourhood of x_0. There exist $u_0 > 0$ and $\alpha > 0$ such that

$$\inf_{x \in I_\alpha} \int_{I_\alpha} u(x, y) \lambda(dy) \geqq \lambda(I_\alpha) u_0 \tag{B.1}$$

and

$$\frac{u_0 \lambda(I_\alpha)}{2\alpha} > 1\,. \tag{B.2}$$

Example. Let $X = \mathbb{R}$ and suppose that $u(x, y) \geqq u_0 > 0$ in the neighbourhood $U(x_0) \times U(x_0)$ of (x_0, x_0). Then (B.1) is satisfied. If $m(x) \geqq -c|x|$ in some neighbourhood of $x_0 = 0$ and $c < u_0$, it follows that $\lambda(I_\alpha) \geqq 2\alpha/c$, and (B.2) is satisfied.

Condition (B.2) may be interpreted as a cusp condition near the optimal type x_0. It is interesting to notice that for functions m with more than one zero, it is sufficient that conditions (B.1) and (B.2) be satisfied for one of these zeros.

If the conditions above are not satisfied, then an equilibrium distribution either does not exist or it may have an "atom of probability" at the optimal type. In the house-of-cards mutation model, a simple necessary and sufficient condition can be derived (see (3.17), (3.18), and Bürger and Bomze 1992).

Appendix C

Here we prove that the expression on the right-hand side of Eq. (3.8) tends to zero, i.e.,

$$\int_{X\setminus M_{\mathrm{Opt}}} \frac{a_\mu}{a_\mu - m(x)} Up_\mu(x)\lambda(dx) \to 0 \quad \text{as } \mu \to 0\,. \tag{C.1}$$

Since $m(x) \leqq 0$, $0 < a_\mu \leqq \mu$, and $Up_\mu \geqq 0$, we have

$$0 \leqq \frac{a_\mu}{a_\mu - m(x)} Up_\mu(x) \leqq \frac{\mu}{\mu - m(x)} Up_\mu(x)\,.$$

Therefore, it is sufficient to prove that

$$\int_{X\setminus M_{\mathrm{Opt}}} \frac{\mu_n}{\mu_n - m(x)} Up_{\mu_n}(x)\lambda(dx) \to 0 \quad \text{as } \mu_n \to 0$$

for any subsequence μ_n. Since $p_\mu \in L^1(\lambda_1)$ (Appendix A), $\|p_\mu\|_{\lambda_1} \leqq \|p_\mu\|_1 + \|mp_\mu\|_1 \leqq 1 + \mu$ (see (3.7.a)), and since $U : L^1(\lambda_1) \to L^1(\lambda)$ is compact, it follows that any subsequence of Up_μ is contained in a compact subset of $L^1(\lambda)$. Now we use the well-known fact that for an L^1-convergent sequence $(f_n) \subseteq L^1$, i.e., $\|f_n - f\|_1 \to 0$, and a uniformly bounded sequence (g_n), i.e., $g_n(x) \leqq$ const. λ-a.e. (λ-almost everywhere) for all n, such that $g_n(x) \to 0$ λ-a.e., the sequence $(f_n g_n)$ converges to 0 in L^1. It is easy to see that the same assertion holds for any sequence (f_n) that is contained in a compact subset of L^1. If we choose $f_n = Up_{\mu_n}$, $g_n(x) = \frac{\mu_n}{\mu_n - m(x)}$ if $x \in X \setminus M_{\mathrm{Opt}}$, and $g_n(x) = 0$ otherwise, we obtain (C.1).

Appendix D

Here we prove that (3.16), and as a consequence, the results in Sect. 4 hold for arbitrary mutant distributions $u(x, y)$, as long as the assumptions of Appendices A and B, as well as (4.4.a) hold. All these assumptions are satisfied for $X = \mathbb{R}$ if mutation is translational or of HC-type, and u is Gaussian or exponential reflected about zero (see Sects. 3.3, 3.4, Appendices A, B). For simplicity, we restrict our attention to the continuum-of-alleles case and assume that $X = \mathbb{R}$ (or some interval), λ is the Lebesgue measure, and $m(x) = -s|x - x_0|^q$, $q > 1$.

We define the following families of linear operators for $\alpha > 0$ (compare Bürger and Bomze 1992)

$$K_\alpha : L^1(\lambda) \to L^1(\lambda), \quad (K_\alpha f)(x) = \int_X \frac{u(x, y)}{\alpha + |y - x_0|^q} f(y)dy \tag{D.1}$$

and

$$\tilde{K}_\alpha : L^1(\lambda) \to L^1(\lambda), \quad (\tilde{K}_\alpha f)(x) = \int_X \frac{u(x, y)}{\alpha + |x - x_0|^q} f(y)dy\,. \tag{D.2}$$

Each of these operators is bounded, and the image of $\tilde{K}_\alpha$ is contained in $L^1(\lambda_1) = \{f \in L^1(\lambda) : \int |m(x)f(x)|dx < \infty\}$. We assume that there is some integer $n \geqq 1$ such that K_α^n is compact for all $\alpha > 0$ (which is weaker than the condition in Appendix A). We denote by $r(K_\alpha)$ and $r(\tilde{K}_\alpha)$ the spectral radii of these operators. In the HC-case of Sect. 3.3, we have $r(K_\alpha) = r(\alpha)$, with $r(\alpha)$ as in (3.12.b) (cf. Bürger 1988).

It follows from the assumptions (see Bürger 1988) that a unique equilibrium distribution exists. Its mean fitness is determined by the equation

$$\frac{\bar{m}_\mu + \mu}{s} = \alpha_{\mu/s}\,, \tag{D.3}$$

where $\alpha = \alpha_{\mu/s}$ is the unique solution of

$$\frac{\mu}{s}\, r(K_\alpha) = 1 \tag{D.4}$$

(cf. 3.6). Equation (3.7) implies that

$$\alpha_{\mu/s} \leqq \mu/s\,. \tag{D.5}$$

We need asymptotic upper and lower bounds for $r(K_\alpha)$ as $\alpha \to 0$. An upper bound is derived upon noticing that

$$r(K_\alpha) \leqq r(\tilde{K}_\alpha) \leqq \|\tilde{K}_\alpha\| \tag{D.6}$$

(see Lemma 3.9 of Bürger and Bomze 1992) and

$$\begin{aligned}
\|\tilde{K}_\alpha\| = \sup_{\|f\|_1 = 1} \|\tilde{K}_\alpha f\|_1 &= \sup_{\|f\|_1 = 1} \int\int \frac{u(x, y)}{\alpha + |x - x_0|^q}\, |f(y)|\, dy\, dx \\
&= \sup_{\|f\|_1 = 1} \int\int \frac{u(x + x_0, y)}{\alpha + |x|^q}\, |f(y)|\, dy\, dx \\
&\leqq \|u\|_\infty \int_{-\infty}^{\infty} \frac{1}{\alpha + |x|^q}\, dx \\
&= \frac{2\|u\|_\infty \pi}{q \sin(\pi/q)}\, \alpha^{\frac{1}{q} - 1}
\end{aligned} \tag{D.7}$$

(using formula 151.4 of Gröbner and Hofreiter 1975).

A lower bound is derived from the fact that

$$r(K_\alpha) \geqq \frac{u_0 \lambda(I_\alpha)}{2\alpha}, \quad \text{for all } \alpha, 0 < \alpha \leqq \mu/s\,, \tag{D.8}$$

holds if we choose I_α and u_0 as in Appendix B, with $m(x) = -|x - x_0|^q$ and $U(x_0) = \left(x_0 - (\mu/s)^{1/q},\, x_0 + (\mu/s)^{1/q}\right)$ (see the proof of Proposition 3.4 in Bürger 1988). From (D.5) we see that (D.8) applies for $\alpha = \alpha_{\mu/s}$. Since $\lambda(I_\alpha) = 2\alpha^{1/q}$, we obtain from (D.6), (D.7), and (D.8)

$$u_0 \alpha^{\frac{1}{q} - 1} \leqq r(K_\alpha) \leqq \frac{2\|u\|_\infty \pi}{q \sin(\pi/q)}\, \alpha^{\frac{1}{q} - 1}\,, \tag{D.9}$$

this corresponding to (3.13), and, by (D.4),

$$\left(u_0 \frac{\mu}{s}\right)^{\frac{q}{q-1}} \leqq \alpha_{\mu/s} \leqq \left(\frac{2\|u\|_\infty \pi}{q \sin(\pi/q)} \cdot \frac{\mu}{s}\right)^{\frac{q}{q-1}}\,. \tag{D.10}$$

This inequality holds for any $\mu/s > 0$ on account of (D.5). Equation (D.3) implies

$$\frac{\bar{m}_\mu + \mu}{\mu} \sim \text{(const.)} \left(\frac{\mu}{s}\right)^{\frac{1}{q-1}}, \tag{D.11}$$

which is just (3.16), and the bounds for the constant are those in (D.10).

For quadratic selection, i.e. $q = 2$, translational mutation, and a Gaussian mutant distribution with mean x_0 and variance γ^2, (D.10) yields

$$\alpha_{\mu/s} \leqq \frac{\pi}{2\gamma^2} \left(\frac{\mu}{s}\right)^2,$$

which, together with (D.5) gives

$$\alpha_{\mu/s} \leqq b^2 = \min\left(\frac{\mu}{s}, \frac{\pi}{2\gamma^2}\left(\frac{\mu}{s}\right)^2\right). \tag{D.12}$$

It follows that (D.8) still holds for $\alpha_{\mu/s}$ if we put $U(x_0) = (x_0 - b, x_0 + b)$. Therefore, the example below (B.2) shows that we may choose

$$u_0 = \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left(-\frac{4b^2}{2\gamma^2}\right).$$

Then (D.10) reduces to

$$\frac{\exp\left(-\frac{4b^2}{\gamma^2}\right)}{2\pi\gamma^2}\left(\frac{\mu}{s}\right)^2 \leqq \alpha_{\mu/s} \leqq \frac{\pi}{2\gamma^2}\left(\frac{\mu}{s}\right)^2. \tag{D.13}$$

Therefore, we obtain for the equilibrium variance $\sigma^2 = -\bar{m}/s = (\mu/s) - \alpha_{\mu/s}$ the estimate

$$\frac{\mu}{s} - \frac{\pi}{2\gamma^2}\left(\frac{\mu}{s}\right)^2 \leqq \sigma^2 \leqq \frac{\mu}{s} - \frac{\exp\left(-\frac{4b^2}{\gamma^2}\right)}{2\pi\gamma^2}\left(\frac{\mu}{s}\right)^2. \tag{D.14}$$

This holds for all $\mu/s > 0$.

Appendix E

Here we simplify some of the main formulas in Bürger (1991) for the marginal and mean fitness, (3.5.a) and (3.5.b), and for the dynamics of cumulants, (4.6). We use the same notation as in that article, and all references refer to it.

Denote by

$$B_j(t) = \int \cdots \int (X_1 + Y)^j \prod_{i=2}^{L} p(x_i, t) p(y_i, t) dx_i dy_i p(y, t) dy$$

the j-th moment around zero of the distribution of $\mathcal{G} - x = X_1 + Y$, remembering the notation $(x, y) = (x_1, y_1)$ for the locus under consideration. In terms of cumulants,

these are given by

$$\begin{aligned} B_0 &= 1 \\ B_1 &= C_1 - c_1 \\ B_2 &= C_2 - c_2 + (C_1 - c_1)^2 \\ B_3 &= C_3 - c_3 + 3(C_2 - c_2)(C_1 - c_1) + (C_1 - c_1)^3 \\ &\vdots \end{aligned} \tag{E.1}$$

(cf. Appendix B). Then, one obtains for the marginal fitness

$$\omega^*(x, t) = \sum_{k=0}^{K} a_k \sum_{l=0}^{k} \binom{k}{l} x^l B_{k-l}(t) \tag{E.2.a}$$

instead of (3.5.a). The mean fitness may be represented as

$$\bar{\omega}(t) = \sum_{k=0}^{K} a_k M_k^0 \,. \tag{E.2.b}$$

It follows that (4.6) may be replaced by

$$\dot{c}_n = \sum_{k=1}^{K} a_k \sum_{l=1}^{k} \binom{k}{l} B_{k-l} F_{nl} + \mu u_n, \quad \text{for all } n \geqq 1 \,, \tag{E.3}$$

which is easier to compute in concrete cases.

Another useful expression for $\bar{\omega}(t)$, directly obtained from (E.2.a), is

$$\bar{\omega}(t) = \sum_{k=0}^{K} a_k \sum_{l=0}^{k} \binom{k}{l} m_l^0 B_{k-l}(t) \,. \tag{E.2.c}$$

Acknowledgement. We are most grateful to Tom Nagylaki for pointing out an unproved assertion in Bürger (1986), for providing motivation, and for comments that led to considerable improvements. We thank Russ Lande for helpful discussions, S. Gavrilets, M. Turelli, and a reviewer for comments, and D. Butcher for assistance in preparing the figure. This work was finished while RB was visiting the Department of Biology at the University of Oregon, Eugene, OR, USA. Financial support by the "Max Kade Foundation" to RB, and by the Austrian "Fonds zur Förderung der wissenschaftlichen Forschung", Project P8043, is gratefully acknowledged.

References

Barton, N.: The maintenance of polygenic variation through a balance between mutation and stabilizing selection. Genet. Res. Camb. **47**, 209–216 (1986)

Bulmer, M. G.: The genetic variability of polygenic characters under optimising selection, mutation and drift. Genet. Res. **19**, 17–25 (1972)

Bürger, R.: On the maintenance of genetic variation: Global analysis of Kimura's continuum-of-alleles model. J. Math. Biol. **24**, 341–351 (1986)

Bürger, R.: Perturbations of positive semigroups and applications to population genetics. Math. Z. **197**, 259–272 (1988)

Bürger, R.: Linkage and the maintenance of heritable variation by mutation-selection balance. Genetics **121**, 175–184 (1989)

Bürger, R.: Moments, cumulants, and polygenic dynamics. J. Math. Biol. **30**, 199–213 (1991)

Bürger, R.: Predictions for the dynamics of a polygenic character under directional selection. J. Theor. Biol. **162**, 487–513 (1993)
Bürger, R., Bomze, I. M.: Mutation-selection balance and a perturbation problem on the space of measures. SIAM J. Math. Anal (submitted)
Crow, J. F.: Genetic loads and the coast of natural selection. In: Kojima, K. (ed.), Mathematical Topics in Population Genetics, pp. 128–177. Berlin Heidelberg New York: Springer 1970
Crow, J. F., Kimura, M.: The theory of genetic loads. In: Proc XI Int. Congr. Genetics, vol. 2, pp. 495–505. Oxford: Pergamon Press 1964
Crow, J. F., Kimura, M.: An Introduction to Population Genetics Theory. New York: Harper and Row 1970
Fleming, W. H.: Equilibrium distributions of continuous polygenic traits. SIAM J. Appl. Math. **36**, 148–168 (1979)
Fraser, G. R., Mayo, O.: Genetic load in man. Humangenetik. **23**, 83–110 (1974)
Gröbner, W., Hofreiter, N.: Integraltafel. Zweiter Teil. Berlin Heidelberg New York: Springer 1975
Haldane, J. B. S.: The effect of variation on fitness. Am. Nat. **71**, 337–349 (1937)
Haldane, J. B. S.: The cost of natural selection. J. Genet. **55**, 511–524 (1957)
Hofbauer, J.: The selection mutation equation. J. Math. Biol. **23**, 41–53 (1985)
Hofbauer, J., Sigmund, K.: The Theorie of Evolution and Dynamical Systems. Cambridge: Cambridge University Press 1988
Keightley, P. D., Hill, W. G.: Quantitative genetic variability maintained by mutation-stabilizing selection balance in finite populations. Genet. Res. Camb. **52**, 33–43 (1988)
Kimura, M.: A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proc. Natl. Acad. Sci., USA **54**, 731–736 (1965)
Kimura, M.: Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons. Proc. Natl. Acad. Sci., USA **78**, 5773–5777 (1981)
Kimura, M., Maruyama, T.: The mutational load with epistatic gene interaction in fitness. Genetics **54**, 1337–1351 (1966)
King, J. L.: The gene interaction component of the genetic load. Genetics **53**, 403–413 (1966)
Kingman, J. F. C.: A simple model for the balance between selection and mutation. J. Appl. Probab. **15**, 1–12 (1978)
Kondrashov, A. S., Crow, J. F.: King's formula for the mutation load with epistasis. Genetics **120**, 853–856 (1988)
Lande, R.: The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. Camb. **26**, 221–235 (1975)
Lande, R.: Genetic variation and phenotypic evolution during allopatric speciation. Am. Nat. **116**, 463–479 (1980)
Muller, H. J.: Our load of mutations. Am. J. Hum. Genet. **2**, 111–176 (1950)
Nagylaki, T.: Selection on a quantitative character. In: Chakravarti A. (ed.) Human Population Genetics: The Pittsburgh Symposium, pp. 275–306. New York: Van Nostrand 1984
Nagylaki, T.: Introduction to Theoretical Population Genetics. Berlin Heidelberg New York: Springer 1992
Sigmund, K.: The maximum principle for replicator equations. In: Ebeling, W., Peschel, M. (eds.) Lotka-Volterra Approach to Dynamical Systems, pp. 63–72. Berlin: Akademie Verlag 1984
Turelli, M.: Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Popul. Biol. **25**, 138–193 (1984)
Turelli, M., Barton, N.: Dynamics of polygenic characters under selection. Theor. Popul. Biol. **38**, 1–57 (1990)
Wright, S.: The analysis of variance and the correlation between relatives with respect to deviations from an optimum. J. Genet. **30**, 243-256 (1935)