

An Alternative to the Breeder's and Lande's Equations

Bahram Houchmandzadeh¹

CNRS/University of Grenoble 1, LIPhy UMR 5588, Grenoble, F-38041, France

ABSTRACT The breeder's equation is a cornerstone of quantitative genetics, widely used in evolutionary modeling. Noting the mean phenotype in parental, selected parents, and the progeny by $E(Z_0)$, $E(Z_W)$, and $E(Z_1)$, this equation relates response to selection $R = E(Z_1) - E(Z_0)$ to the selection differential $S = E(Z_W) - E(Z_0)$ through a simple proportionality relation $R = h^2S$, where the heritability coefficient h^2 is a simple function of genotype and environment factors variance. The validity of this relation relies strongly on the *normal* (Gaussian) distribution of the parent genotype, which is an unobservable quantity and cannot be ascertained. In contrast, we show here that if the fitness (or selection) function is Gaussian with mean μ , an alternative, exact linear equation of the form $R' = j^2S'$ can be derived, regardless of the *parental genotype* distribution. Here $R' = E(Z_1) - \mu$ and $S' = E(Z_W) - \mu$ stand for the mean phenotypic lag with respect to the mean of the fitness function in the offspring and selected populations. The proportionality coefficient j^2 is a simple function of selection function and environment factors variance, but *does not* contain the genotype variance. To demonstrate this, we derive the exact functional relation between the mean phenotype in the selected and the offspring population and deduce all cases that lead to a linear relation between them. These results generalize naturally to the concept of G matrix and the multivariate Lande's equation $\Delta\bar{z} = GP^{-1}S$. The linearity coefficient of the alternative equation are not changed by Gaussian selection.

KEYWORDS

linear response
to selection
additive genetic
effects
quantitative
genetics

The breeder's equation for the evolution of quantitative traits for additive genetic effects, introduced by Lush (Lush 1943), is widely used both in artificial and natural selection theory and experiments (Lande 1976; Falconer and Mackay 1995; Lynch and Walsh 1998; Heywood 2005) and appears in all textbooks of quantitative genetic. This equation can be stated as follow: consider a continuous phenotypic trait Z subject to selection. Noting the mean phenotype in parental, selected parents and the progeny by $E(Z_0)$, $E(Z_W)$ and $E(Z_1)$, we can define the selection differential $S = E(Z_W) - E(Z_0)$ and the response $R = E(Z_1) - E(Z_0)$. The scalar breeder's equation reads $R = h^2S$ and ascertains that the response to selection and the selection differential are related through a proportionality relation that is the ratio of genotype to phenotype variances, h^2 . The equation naturally

extends to selection on multiple traits and its vectorial version reads $\Delta\bar{z} = GP^{-1}S$

Use of the breeder's equation and its underlying assumptions has been criticized by many authors (Kruuk 2004; Heywood 2005; Pigliucci 2006; Gienapp *et al.* 2008; Pemberton 2010). One fundamental assumption of the breeder's equation is the normal (Gaussian) distribution of the breeding value (genotype) and environment factors. Authors who demonstrate the linear relation (Kimura and Crow 1978; Lande 1979; Lande and Arnold 1983; Nagylaki 1992; Falconer and Mackay 1995; Lynch and Walsh 1998; Crow and Kimura 2009) assume normal distribution for the aforementioned quantities or the analogous hypothesis of linearity of the parent-offspring regression (see Appendix/Parent-offspring regression). When this assumption is relaxed, the breeder's equation is no longer valid, and one has to resort to a system of hierarchical moment (or alternatively, cumulant) equations to describe the changes; in general, this system is not closed, and the moments of a given order depend on moments of higher order (Turelli and Barton 1990).

The assumption of a Gaussian distribution of the genotype can be criticized on several grounds (Pigliucci and Schlichting 1997; Pigliucci 2006; Geyer and Shaw 2008). For example, the very act of selection causes the genotype distribution to deviate from a Gaussian (Turelli and Barton 1990; Turelli and Barton 1994) (see also equation 6

Copyright © 2014 Houchmandzadeh

doi: 10.1534/g3.113.008433

Manuscript received September 6, 2013; accepted for publication November 1, 2013; published Early Online November 8, 2013.

This is an open-access article distributed under the terms of the Creative Commons Attribution Unported License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

¹Corresponding author: CNRS/Univ. Grenoble 1, LIPhy UMR 5588, Grenoble, F-38041, France. E-mail: bahram.houchmandzadeh@ujf-grenoble.fr

below). Another important case is when the genotype is a cross between different breeds due to external gene flow or the breeder's scheme. In many cases, the phenotype can have a bell shape and thus is assumed to be Gaussian, when the genotype is indeed far from it (see, for example, Figure 2A). It is sometimes argued that even if the breeding value does not follow a normal distribution, a scale can be used to restore it to a normal distribution. Such a scale, however, will also distort the distribution of environment factors and the assumptions of breeder's equation are violated even in this case.

For *additive* genetic effects and in the absence of epistasis and dominance, I derive here a precise functional relation between the mean of the trait in the selected subpopulation and in their progeny for the general case. The mathematical formulation is close to the framework used by many authors such as Slatkin, Lande and Karlin (Slatkin 1970; Karlin 1979; Lande 1979). I then use a standard tool of functional analysis, the Fourier transform (FT), to deduce all the cases that lead to a linear relation between the response R and the selection differential S , regardless of the selection function. These cases imply a precise form of the distributions of genotype and environment factors, and I show that the proportionality factor between R and S is the heritability coefficient h^2 only if these distributions are *normal*.

The genotype, however, is not observable or controllable, and its normal distribution cannot be assumed *a priori*. I show that if instead of the genotype, the *fitness function* and environment factors are Gaussian, then a new proportionality relation can be obtained in the form of

$$R' = j^2 S' \quad (1)$$

regardless of the genotype distribution. Noting the mean of the Gaussian selection function by μ , $R' = E(Z_1) - \mu$ and $S' = E(Z_w) - \mu$ are the mean phenotypic lag with respect to the mean of the fitness function of the progeny and the selected population (Figure 1). As $E(Z_1)$, $E(Z_w)$, and μ are all measurable, R' and S' are both measurable in the same way as R and S are. The j^2 coefficient contains only the width of the fitness function and environment factors. The use of a Gaussian selection function, both in artificial and natural selection (as an approximation of stabilizing selection), is widespread (Lewontin 1964; Lande 1976; Kimura and Crow 1978; Zhang and Hill 2010) and the aforementioned relationship is potentially as useful as the standard breeder's equation.

The advantage is more critical when the breeder's or Lande's equations are used in long-term evolution, where the variance of the genotype (or the G matrix) also varies and h^2 cannot be assumed to remain constant (Gavrilets and Hastings 1995; Pigliucci and Schlichting 1997; Roff 2000); in contrast, the relation (1) remains valid if each round of selection uses a Gaussian fitness function.

The aforementioned results generalize naturally to multivariate trait selection where the alternative Lande's equation is

$$\mathbf{R}' = (\mathbf{\Omega} + E)\mathbf{\Omega}^{-1}\mathbf{S}' \quad (2)$$

where \mathbf{R}' and \mathbf{S}' are the vectorial phenotype lag, and $\mathbf{\Omega}$ and E are the covariance matrices of the fitness function and the environment respectively.

This article is organized as follows: in the *Results* section, I first derive the general functional relationship between R and S ; the second subsection is devoted to all the cases where these two quantities can be linearly related, including the special case of the breeder's equation. The alternative breeder's equation is derived in the third subsection, and all the results are generalized to selection on multiple traits in the

fourth subsection. The aforementioned results are put into perspective in the *Discussion* section. Technical details, such as the use of FTs, are treated in the Appendix.

RESULTS

General results

Consider a continuous phenotype Z , which is the result of additive genetic effect Y and the environment ξ (Fisher 1918; Lynch and Walsh 1998; Visscher *et al.* 2008)

$$Z = Y + \xi$$

The term *environment* encompasses here any source of *noise* that causes the observed phenotype z to deviate from the (unobserved) breeding value y (Wright 1920; Lynch and Walsh 1998; Raj and van Oudenaarden 2008). In the following, the population distribution of the breeding value (genotype) and its variance in the parental generation are denoted $p_0(y)$ and σ_A^2 . The environment effect is captured by the distribution law $f(z|y)$, the probability density of observing phenotype z with the given genotype y . We will suppose that f is a symmetric function of its argument of the form $f(z|y) = f(z - y)$ and denote its width by σ_E^2 .

A subpopulation among the parental generation is selected according to a fitness or selection function $W(z)$, the proportion of phenotypes in $[z, z + dz]$ to be selected for the production of the next generation. The selected individuals produce offspring which will constitute the next generation. As we will show herein, the response R (the mean of the phenotype trait in the offspring) and the selection differential S (the mean of the phenotype trait in the selected parents) are given by

$$R = E(Z_1) = \frac{1}{\bar{W}} \iint_{\mathbb{R}^2} y p_0(y) W(z) f(z - y) dy dz \quad (3)$$

$$S = E(Z_w) = \frac{1}{\bar{W}} \iint_{\mathbb{R}^2} z p_0(y) W(z) f(z - y) dy dz \quad (4)$$

where \bar{W} is the mean fitness of parental generation. Equations (3) and (4) are used, for example, by Lande (1979), although their derivation there depended on the normal distribution of the genotype. I derive these equations here for the more general case.

Before going into the details of calculations, note that the genotype distribution $p_0(y)$ and the selection function $W(z)$ play a symmetric role in the aforementioned expressions. In the following sections, we will explore specific functional forms of $p_0(y)$ and $W(z)$, which lead to a linear relationship between R and S . Because of the symmetric role of these two functions however, once a particular relation is obtained for a specific form of $p_0(y)$ regardless of $W(z)$, an analogous relationship can be obtained for a similar form of $W(z)$ regardless of $p_0(y)$. This is what leads us to an alternative form of the breeder's equation.

Let us now derive the equations (3,4). We note that the distribution of the phenotype Z in the parental generation is given by

$$q_0(z) = \int_{\mathbb{R}} p_0(y) f(z|y) dy \quad (5)$$

We will denote its variance by σ_p^2 .

The distribution of the phenotype z in the parental population selected according to the fitness function $W(z)$ is

$$q_w(z) = \frac{1}{\bar{W}} q_0(z) W(z)$$

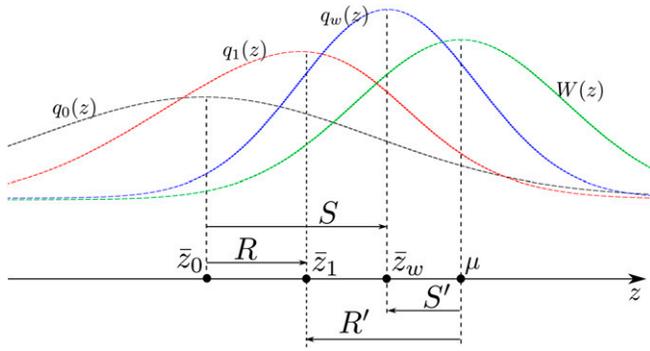


Figure 1 Schematic representation of the selection lag S' , the response lag R' , and their relation to the selection differential S and the response R . The mean phenotype of parental generation \bar{z}_0 , selected population \bar{z}_w , the progeny \bar{z}_1 , and the peak of selection function μ are represented on the phenotype axis z . Dashed curves represent a sketch of the distributions of parental phenotype $q_0(z)$, selected parents $q_w(z)$, the progeny $q_1(z)$, and the selection function $W(z)$.

where \bar{W} is the mean fitness of the parental generation

$$\begin{aligned}\bar{W} &= \int_{\mathbb{R}} q_0(z) W(z) dz \\ &= \iint_{\mathbb{R}^2} p_0(y) W(z) f(z|y) dy dz\end{aligned}$$

The genotype distribution of the selected population is (Turelli and Barton 1994)

$$p_w(y) = \frac{1}{\bar{W}^\dagger} \int_{\mathbb{R}} p_0(y) f(z|y) W(z) dz \quad (6)$$

$$= \frac{1}{\bar{W}^\dagger} p_0(y) W^\dagger(y) \quad (7)$$

where

$$W^\dagger(y) = \int_{\mathbb{R}} W(z) f(z|y) dz \quad (8)$$

is the *genotype* fitness function, *i.e.*, the convolution of the phenotype fitness function by the environment factors. \bar{W}^\dagger is the mean genotype fitness:

$$\begin{aligned}\bar{W}^\dagger &= \int_{\mathbb{R}} p_0(y) W^\dagger(y) dy \\ &= \iint_{\mathbb{R}^2} p_0(y) W(z) f(z|y) dy dz\end{aligned}$$

Note that $\bar{W} = \bar{W}^\dagger$ as both these quantities are defined by the same double integration over the domains of y and z .

For a large, randomly mating population, reproduction gives for the distribution of breeding values in the next generation (Slatkin 1970; Karlin 1979; Bulmer 1985; Turelli and Barton 1994)

$$p_1(y) = \iint_{\mathbb{R}^2} p_w(y_a) p_w(y_b) L(y - (y_a + y_b)/2) dy_a dy_b$$

The exact form of the probability density $L(y)$ that captures the inheritance process (recombination, segregation, ...) is not important here; Turelli and Barton (1994), for example, use a normal distribution for $L(y)$ in the framework of the infinitesimal model. For our purpose, it is enough to suppose that the mean of the distribution $L(y)$ is zero, *i.e.*, $\int_y y L(y) dy = 0$ which is valid in the absence of dominance and epistasis effects (Turelli and Barton 1990) (see also *Appendix/Segregation density function*).

The phenotype distribution of the progeny is

$$q_1(z) = \int_{\mathbb{R}} p_1(y) f(z|y) dy \quad (9)$$

We now make the further assumption that (1) the environment and genotype are independent random variables, so that $f(z|y) = f(z - y)$ and therefore the variances are additive: $\sigma_p^2 = \sigma_A^2 + \sigma_E^2$ and (2) environment effects are of zero mean ($\int_x x f(x) dx = 0$) and symmetric ($f(-x) = f(x)$). An environmental noise with such a distribution law does not change the mean of the random variable: $E(Z) = E(Y + \xi) = E(Y)$. Therefore, the mean phenotype of the offspring is

$$\begin{aligned}R &= E(Z_1) = E(Y_1) \\ &= \int_{\mathbb{R}} y p_1(y) dy \\ &= (1/2) \iint_{\mathbb{R}^2} (y_a + y_b) p_w(y_a) p_w(y_b) dy_a dy_b \\ &= \int_{\mathbb{R}} y p_w(y) dy\end{aligned} \quad (10)$$

$$= \frac{1}{\bar{W}} \iint_{\mathbb{R}^2} y p_0(y) W(z) f(z - y) dy dz \quad (11)$$

which is equation (3). Note that the first lines of the above equations merely state that the *expectations* of the breeding's value of parent and offspring are equal for purely additive traits.

On the other hand, the mean phenotype of the selected parents is

$$\begin{aligned}S &= E(Z_w) = \int_{\mathbb{R}} z q_w(z) dz \\ &= \frac{1}{\bar{W}} \int_{\mathbb{R}} z q_0(z) W(z) dz \\ &= \frac{1}{\bar{W}} \iint_{\mathbb{R}^2} z p_0(y) W(z) f(z - y) dy dz\end{aligned} \quad (12)$$

which is equation (4).

For an asexually reproducing organism, or for a sexually reproducing population which remains at Hardy-Weinberg equilibrium after selection-reproduction, we would have $p_1(y) = p_w(y)$; this would again lead to the same equation (10) and the same response (11). The conditions for the existence of multilocus Hardy-Weinberg equilibrium were analyzed by Karlin and Liberman (1979a,b), who concluded that for additive traits, the equilibrium is stable for a wide range of recombination distributions. The general relation between R and S can also be studied in the context of the Price equation. A detailed study of this relation has been performed by Heywood (2005).

Conditions for proportionality of R and S

The relations (3) and (4) show that the selection differential S and the response R to it are related through a functional equation involving three factors: genotype distribution, the selection function and the environmental noise. It is far from obvious that R and S could be proportional, a question we will investigate by using FTs.

FTs in functional analysis play a role analogous to logarithms in algebra. They are useful for clarifying the $R - S$ relation, where we can

transform the double integrations into simple ones. The FT of the function $u(x)$ is the function $\tilde{u}(k)$ defined as (see *Appendix/Fourier Transforms*)

$$\tilde{u}(k) = \int_{-\infty}^{\infty} e^{-ikx} u(x) dx.$$

For example, the FT of the function $u(x) = \exp(-\alpha|x|)$ is $\tilde{u}(k) = 2\alpha/(\alpha^2 + k^2)$. Part of the usefulness of FT is due to the fact that they transform convolution products into simple products: given two functions $u(x)$ and $v(x)$ and their convolution product $h(z)$:

$$h(z) = \int_{-\infty}^{\infty} u(x)v(z-x) dx$$

the relation between their FT is a simple product:

$$\tilde{h}(k) = \tilde{u}(k)\tilde{v}(k)$$

As the general relations (11) and (12) involve convolutions, FT proves to be very useful in their handling. Using the various properties of FT (see *Appendix/Fourier Transforms*), the relation between R and S in the Fourier space reads:

$$R = \frac{i}{2\pi\bar{W}} \int_{\mathbb{R}} \tilde{W}^*(k) \frac{d}{dk} [\tilde{p}_0(k)] \tilde{f}(k) dk \quad (13)$$

and

$$\begin{aligned} S &= \frac{i}{2\pi\bar{W}} \int_{\mathbb{R}} \tilde{W}^*(k) \frac{d}{dk} [\tilde{p}_0(k)\tilde{f}(k)] dk \\ &= R + \frac{i}{2\pi\bar{W}} \int_{\mathbb{R}} \tilde{W}^*(k)\tilde{p}_0(k) \frac{d}{dk} [\tilde{f}(k)] dk \end{aligned} \quad (14)$$

where the mean fitness \bar{W} is itself defined in Fourier space as

$$\bar{W} = \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{W}^*(k)\tilde{p}_0(k)\tilde{f}(k) dk.$$

Here a^* designate the complex conjugate of a , $i^2 = -1$ and we have set the origin of the breeding values at its mean in the parental population, i.e., $\int_{\mathbb{R}} y p_0(y) dy = 0$.

In general, the FT of a function is complex. However, as the function in direct space here are real, it can be shown that expressions (13) and (14) are indeed real; the fact that i appears in these expression insures this fact (see *Appendix/Fourier Transforms*). It is worthwhile to consider a particular case to clarify the above expressions. The detailed computations for a truncation selection in which breeding value and environmental factors are normally distributed are provided in *Appendix/Truncation selection*.

We see from equations (13) and (14) that S and R can be proportional if the second term of the r.h.s. of equation (14) is proportional to R ; this will be true, regardless of the selection function W , if

$$\tilde{p}_0(k) \frac{d\tilde{f}(k)}{dk} = a \frac{d\tilde{p}_0(k)}{dk} \tilde{f}(k) \quad (15)$$

where a is an arbitrary constant. Equation (15) is the necessary and sufficient condition that defines the functional shape of the genotype

distribution and the environment noise compatible with the proportionality of R and S regardless of the selection function. If condition (15) is fulfilled, then

$$R = (1+a)^{-1}S$$

On the other hand, equation (15) can be seen as a differential equation whose solution is given by

$$\tilde{f}(k) = b\tilde{p}_0(k)^a \quad (16)$$

where b is another arbitrary constant. Let us consider some particular case where the aforementioned relation is obeyed.

Normal distributions

If $\tilde{f}(k)$ and $\tilde{p}_0(k)$ are both Gaussians, i.e.,

$$\tilde{f}(k) = \exp(-\sigma_E^2 k^2/2)$$

$$\tilde{p}_0(k) = \exp(-\sigma_A^2 k^2/2)$$

then the relation (16) is satisfied by

$$a = \sigma_E^2/\sigma_A^2$$

and we retrieve the usual breeder's equation $R = h^2S$ where $h^2 = \sigma_A^2/(\sigma_A^2 + \sigma_E^2)$. Of course, if $\tilde{f}(k)$ and $\tilde{p}_0(k)$ are of the above form, their inverse FTs represent normal distributions of width σ_E and σ_A respectively (see *Appendix/Fourier Transforms*).

Stretched exponentials

We see, however, that even if the strict condition (16) is fulfilled, the proportionality constant need not be h^2 . Consider, for example, the class of stretched exponential functions $\phi(k) = \exp(-|k|^\alpha)$, which generalizes Gaussians (case $\alpha = 2$). Set $\tilde{f}(k) = \phi(\sigma_E k)$, $\tilde{p}_0(k) = \phi(\sigma_A k)$. The inverse FT of these functions gives the distribution of the genotype Y and environment effect E and it is straightforward to show that as for the Gaussian case, $\text{Var}(E)/\text{Var}(Y) = \sigma_E^2/\sigma_A^2$. Condition (16) however is satisfied this time with $a = \sigma_E^\alpha/\sigma_A^\alpha$ and therefore the realized heritability $h^\alpha = R/S$ is

$$h^\alpha = \frac{\sigma_A^\alpha}{\sigma_A^\alpha + \sigma_E^\alpha}$$

The aforementioned examples were to emphasize the fact that selection-independent proportionality is achieved only for particular pairs of genotype/environment distributions. In general, as shown in Figure 2, the realized heritability is not constant and depends critically on the selection function $W(z)$.

Alternative breeder's equation

Optimal phenotypic selection approximated by Gaussians has been considered by many authors both in artificial (as early as Lush 1943) and in natural selection (as early as Wright 1935; Haldane 1954) and it is widespread in the literature (Lewontin 1964; Lande 1976; Kimura and Crow 1978; Karlin and Liberman 1979a; Zhang and Hill 2010). If the selection function is Gaussian, a new linear relation can be extracted from the general relations (3) and (4), regardless of the (unobservable) breeding value distribution.

Note that a symmetric role is played by $W(z)$ and $p_0(y)$ in the general expressions (3) and (4). Hence permuting their role will lead

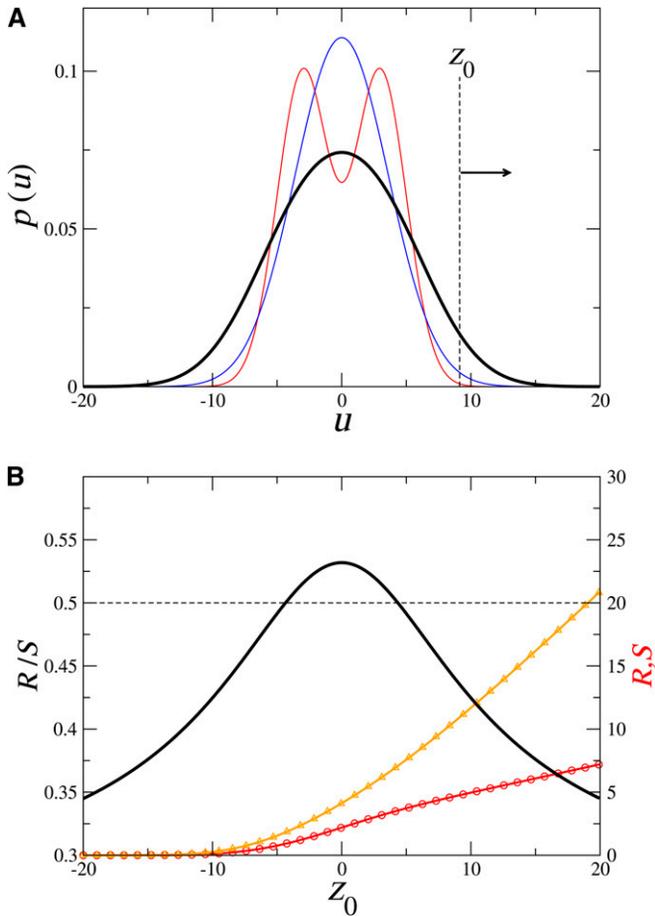


Figure 2 A simple example in which $R/S \neq h^2$. (A) The parental breeding value distribution (thin red line) is a double Gaussian $p_0(y) = (\mathcal{N}(m, s; y) + \mathcal{N}(-m, s; y))/2$; the environmental effects distribution (thin blue line) follows a normal distribution $f(x) = \mathcal{N}(0, \sigma_E; x)$. The phenotype distribution $q(z)$ (equation 5), (thick black line), has the appearance of a normal distribution. The result of a truncation selection, selecting only and all individuals with phenotype value $> z_0$ is shown in (B). (B) Right scale: The Response R (red line, circle) and the Selection differential S (orange line, triangle) as a function of the truncation selection z_0 . Left scale: the value of R/S (thick black line) as a function of z_0 and its comparison to h^2 (thin dashed line). All integrations (equations 11 and 12) can be performed exactly for this case: $R(z_0)$ and $S(z_0)$ are combination of Gaussian and $\text{erf}(z_0)$ functions, their exact expressions are given in Appendix/Computation of truncation selection. The parameters of the figures are $m = 3$, $s = 2$ and $\sigma_E = \sigma_A = \sqrt{m^2 + s^2} = \sqrt{13}$, therefore $h^2 = 1/2$.

us, following the same line of arguments, to deduce all linear cases regardless of genotype. Equations (3) and (4) are obtained by multiplying the function $F(y, z) = W(z)p_0(y)f(z - y)$ either by y or z and integrating over \mathbb{R}^2 . To obtain the breeder's equation of the previous section, we wrote the integration over the y variable as a convolution product and performed the FT on the z variable.

On the other hand, we could have proceeded by writing equations (3) and (4) first as a convolution product on z and then perform a FT on the variable y (see Appendix/Fourier Transform). In this case, we get

$$S = \frac{i}{2\pi W} \int_{\mathbb{R}} \tilde{p}^*(k) \frac{d}{dk} [\tilde{W}(k)] \tilde{f}(k) dk \quad (17)$$

and

$$R = \frac{i}{2\pi W} \int_{\mathbb{R}} \tilde{p}^*(k) \frac{d}{dk} [\tilde{W}(k) \tilde{f}(k)] dk \quad (18)$$

The arguments of the previous section can be repeated. Let us center the selection function by setting $W(z) = W_c(z - \mu)$ where

$$\mu = \int_{\mathbb{R}} z W(z) dz$$

Then

$$S' = (S - \mu) = \frac{i}{2\pi W} \int_{\mathbb{R}} \tilde{p}^*(k) e^{-ik\mu} \frac{d}{dk} [\tilde{W}_c(k)] \tilde{f}(k) dk \quad (19)$$

and

$$R' = (R - \mu) = \frac{i}{2\pi W} \int_{\mathbb{R}} \tilde{p}^*(k) e^{-ik\mu} \frac{d}{dk} [\tilde{W}_c(k) \tilde{f}(k)] dk \quad (20)$$

The quantities S' and R' are alternative selection differential and response and represent the lag with respect to the mean of the selection function (Figure 1). In the case in which the selection function and the environment factors are both normally distributed with width σ_W and σ_E , a repetition of the arguments of the previous sections leads to

$$R' = j^2 S' \quad (21)$$

where

$$j^2 = \frac{\sigma_W^2 + \sigma_E^2}{\sigma_W^2}$$

We stress that relation (21) is obtained *regardless* of the unknown genotype distribution $p_0(y)$.

The alternative breeder's equation (21) may seem unusual as it does not contain the genetic variance. Such a result may seem at first glance in contradiction with our basic understanding of the selection process. Fisher fundamental's theorem for example explicitly relates

the rate of increase in fitness to the genetic variance. There is, however, no contradiction: Both R' and S' are dependent on the genetic variance, as can be seen in the general equations (3) and (4); however, their ratio, i.e., the coefficient of the linear equation (21) relating them, is free of genetic variance. A similar situation occurs for the classical breeder's equation, where both R and S depend on the selection function $W(z)$ but their ratio contains only the heritability coefficient, independently of $W(z)$.

Equation (21) has been obtained through the tools of functional analysis and its demonstration may seem a little abstract. It is worthwhile to further illustrate this equation by considering few examples where the computations can be carried out explicitly. Let us designate the normally distributed selection function $W(z)$ and the environment factors as:

$$W(z) = \mathcal{N}(\mu, \sigma_W; z)$$

$$f(x) = \mathcal{N}(0, \sigma_E; x)$$

where $\mathcal{N}(a, b; u) = (1/\sqrt{2\pi b})\exp(- (u-a)^2/2b^2)$.

No genetic variance

The first example we consider is the extreme case in which there is no genetic variance ($\sigma_A = 0$) in the parental generation. The distribution of the breeding value then becomes a Dirac's delta function $p_0(y) = \delta(y)$. The basic rule of Dirac's delta, i.e., $\int_I \delta(y)\phi(y)dy = \phi(0)$ reduces the double integrations of equations (11) and (12) to simple integrations which involve only Gaussian functions. Note that for the general case, reduction of double integration to simple one was achieved by the use of FTs. The value of R and S are therefore readily obtained in this case:

$$\bar{W} = \int_{\mathbb{R}} W(z)f(z)dz = \frac{1}{\sqrt{2\pi}} \frac{\exp\left(-\frac{\mu^2}{2(\sigma_E^2 + \sigma_W^2)}\right)}{\sqrt{\sigma_E^2 + \sigma_W^2}} \quad (22)$$

$$R = 0 \quad (23)$$

$$S = \frac{1}{\bar{W}} \int_{\mathbb{R}} zW(z)f(z)dz = \mu \frac{\sigma_E^2}{\sigma_W^2 + \sigma_E^2} \quad (24)$$

As expected, in the absence of genetic variance, there is no response to selection. The response and selection lag R' and S' read:

$$R' = R - \mu = -\mu$$

$$S' = S - \mu = -\frac{\sigma_W^2}{\sigma_W^2 + \sigma_E^2} \mu$$

Therefore, $R' = j^2 S'$, and equation (21) is verified. This example shows explicitly that there is no contradiction between the alternative breeder's equation and Fisher's fundamental theorem.

Gaussian breeding values distribution

Let us now consider a less-extreme case in which there exists a normal genetic variability

$$p_0(y) = \mathcal{N}(0, \sigma_A; y)$$

The double integrations (11,12) giving S and R can again be carried out exactly, as all the integrands are Gaussian:

$$S = \alpha \mu \quad (25)$$

$$R = \alpha h^2 \mu \quad (26)$$

where $\alpha = (j^2 - 1)/(j^2 - h^2)$. Therefore,

$$\frac{R'}{S'} = \frac{R - \mu}{S - \mu} = \frac{\alpha h^2 - 1}{\alpha - 1} = j^2$$

and equation (21) is verified.

Note that in the aforementioned case, R and S are both proportional to the mean of the selection function μ . Applying a Gaussian selection function can therefore be used as a test of the normal distribution of the breeding values.

Non-Gaussian breeding values distribution

Let us now consider a case in which parental breeding values are not normally distributed but are concentrated around two particular values:

$$p_0(y) = \frac{1}{2}(\delta(y - \sigma_A) + \delta(y + \sigma_A)) \quad (27)$$

and therefore $E(Y_0) = 0$ and $\text{Var}(Y_0) = \sigma_A^2$. The computation of the expressions (11) and (12) can be again carried out exactly:

$$S = \frac{\mu\sigma_E^2 + \sigma_A\sigma_W^2 \tanh\left(\frac{\sigma_A\mu}{\sigma_W^2 + \sigma_E^2}\right)}{\sigma_W^2 + \sigma_E^2}$$

$$R = \sigma_A \tanh\left(\frac{\sigma_A\mu}{\sigma_W^2 + \sigma_E^2}\right)$$

We note that in this case, the ratio $R/S \neq h^2$ and the classical breeder's equation does not hold. The alternative breeder's equation however is again verified:

$$\frac{R'}{S'} = \frac{R - \mu}{S - \mu} = \frac{\sigma_W^2 + \sigma_E^2}{\sigma_W^2} = j^2$$

Figure 3 illustrates the accuracy of the alternative relation compared to the usual breeder's equation for this case.

The aforementioned few examples were to illustrate the alternative breeder's equation. Many other examples of the breeding value distributions, such as a double Gaussian or a rectangular function can be computed exactly and lead of course always to the alternative breeder's equation. The last example can indeed be generalized and used as an alternative demonstration of equation (21), as any function can be seen as a superposition of Dirac's deltas: $p_0(y) = \int_{\mathbb{R}} p_0(u)\delta(u - y)du$. The demonstration we provided using FT is, however, more straightforward.

Selection on multiple traits

The results of the aforementioned sections are naturally generalized to selection on multiple traits. Consider the vectors of parental breeding values $\mathbf{y}_0 = (y_1, y_2, \dots, y_N)$, environmental effects $\mathbf{e} = (e_1, \dots, e_N)$ and their phenotype $\mathbf{z}_0 = \mathbf{y}_0 + \mathbf{e}$, to which a selection function $W(\mathbf{z})$ is applied. Using the same notations as in the previous sections, we find without difficulty that

$$\bar{\mathbf{z}}_1 = \int_{\mathbb{R}^N \times \mathbb{R}^N} \mathbf{y} p_0(\mathbf{y}) W(\mathbf{z}) f(\mathbf{z} - \mathbf{y}) d\mathbf{y} d\mathbf{z}$$

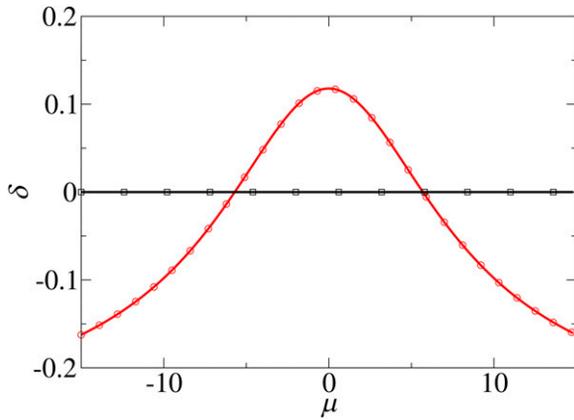


Figure 3 Gaussian selection function: deviation δ from theoretical prediction of the breeder's and alternative equations as a function of the mean of the selection function μ . The breeding value distribution corresponds to equation (27). Black line (squares): $(R - \mu)/(S - \mu) - j^2$; red line (circle): $R/S - h^2$. Parameters used here are $\sigma_A = 2$, $\sigma_W = 1$ and $\sigma_E = 3$ and therefore $h^2 = 4/13$ and $j^2 = 10$.

$$\bar{z}_w = \int_{\mathbb{R}^N \times \mathbb{R}^N} \mathbf{z} p_0(\mathbf{y}) W(\mathbf{z}) f(\mathbf{z} - \mathbf{y}) d\mathbf{y} d\mathbf{z}$$

As before, using FT, these relations transform into

$$\bar{z}_1 = \frac{i}{2\pi} \int_{\mathbb{R}^N} \tilde{W}^*(\mathbf{k}) (\nabla \tilde{p}_0(\mathbf{k})) \tilde{f}(\mathbf{k}) d^M \mathbf{k}$$

$$\bar{z}_w = \frac{i}{2\pi} \int_{\mathbb{R}^N} \tilde{W}^*(\mathbf{k}) \nabla (\tilde{p}_0(\mathbf{k}) \tilde{f}(\mathbf{k})) d^M \mathbf{k}$$

where ∇ is the gradient operator: $\nabla f = (\partial f / \partial x_1, \dots, \partial f / \partial x_N)$. We see again that \bar{z}_1 and \bar{z}_w are linearly related if

$$\tilde{p}_0(\mathbf{k}) (\nabla \tilde{f}(\mathbf{k})) = A (\nabla \tilde{p}_0(\mathbf{k})) \tilde{f}(\mathbf{k})$$

where A is a constant matrix. The linear relation is automatically satisfied if both p_0 and f follow a Gaussian distribution

$$p_0(\mathbf{y}) \propto \exp\left(-\frac{1}{2} \mathbf{y}^T G^{-1} \mathbf{y}\right)$$

$$f(\mathbf{x}) \propto \exp\left(-\frac{1}{2} \mathbf{x}^T E^{-1} \mathbf{x}\right)$$

where G and E are the covariance matrices for the genotype and environmental effects. Defining $P = G + E$ as the phenotype covariance matrix, it is straightforward to show that in this case $A = EG^{-1}$ and therefore (Lande 1979)

$$\bar{z}_1 = GP^{-1} \bar{z}_w$$

which is the usual breeder's equation for multiple traits. We stress that the limitation of this relation is the same as that of the scalar version: it relies on the normal distribution of the genotype. On the other hand, if the selection function $W(\mathbf{z})$ is Gaussian

$$W(\mathbf{z}) \propto \exp\left(-\frac{1}{2} (\mathbf{z} - \mu)^T \Omega^{-1} (\mathbf{z} - \mu)\right)$$

the arguments of the previous section 2 can be repeated and lead to the generalization of the alternative vectorial breeder's equation (21)

$$\bar{z}_1 - \mu = (\Omega + E) \Omega^{-1} (\bar{z}_w - \mu)$$

which, in analogy with equation (21) we write as

$$\mathbf{R}' = (\Omega + E) \Omega^{-1} \mathbf{S}'$$

DISCUSSION AND CONCLUSION

The breeder's equation is a cornerstone of quantitative genetics and appears as a fundamental equation in all the important textbooks of this field (Lynch and Walsh 1998; Falconer and Mackay 1995; Crow and Kimura 2009). It is widely used in artificial selection (Lush 1943; Hill and Kirkpatrick 2010); its usage in natural selection was popularized by Lande (1976), when he formalized the main idea of phenotypic evolution and it is now commonly used in many articles based on Lande's work (see, for example, Hansen *et al.* 2011; Manna *et al.* 2011; Svardal *et al.* 2011). The mathematical foundation of this equation rests upon the hypothesis that the breeding value is normally distributed. This hypothesis is plausible for a continuous trait in a population not subject to selection (see, however, *Appendix/Segregation density function*). The normal distribution of the breeding value is more fragile in populations subjected to selection on this trait (Turelli and Barton 1990), as the genotype of selected parents is given by (equation 7)

$$p_w(\mathbf{y}) = p_0(\mathbf{y}) \cdot W^\dagger(\mathbf{y}) / \bar{W}$$

where $W^\dagger(\mathbf{y})$ is the genotype fitness function defined by equation (8). Even if $p_0(\mathbf{y})$ were Gaussian, the very act of multiplying it by an arbitrary function makes $p_w(\mathbf{y})$, and hence $p_1(\mathbf{y})$ non-Gaussian. Therefore after the first round of selection, the normal distribution hypothesis of parental genotype cannot be sustained. Turelli and Barton (1994) have shown that for the infinitesimal model, the non-normality may not have large effects on the predictions of the breeder's equation, but they argued that when the number of loci is limited the discrepancy can grow much larger. Of course even $p_0(\mathbf{y})$ cannot be assumed to be Gaussian if different breeds are crossed to constitute the parental generation, which happens in artificial selection and in natural selection when gene flow from nearby patches is important.

The breeding value is not an observable quantity. The fitness or selection function $W(\mathbf{z})$ is more quantifiable and many authors have considered a Gaussian selection function. In artificial selection, it dates back at least to the work of Lush (Lush 1943, p140). In natural selection, it is used by most authors as a model for stabilizing selection. If Gaussian selection is used to evolve a population, then the alternative breeding equation (21) we derived is more precise and predictive and rests on more robust mathematical grounds while retaining the same simplicity of the standard breeder's equation. Note that the analysis of this article is not restricted to the infinitesimal model, but applies to all inheritance processes involving purely additive genetic effects. The alternative breeder's equation generalizes to selection on multiple traits in a way similar to the standard breeder's equation and can therefore be incorporated in the "adaptive landscape" formalism (Arnold *et al.* 2001) with the same ease.

In conclusion, we believe that in all cases where Gaussian selection functions are used to evolve a population, the alternative breeder's equation we develop above is a useful alternative approach to the standard method.

ACKNOWLEDGMENTS

I thank Jarrod Hadfield for discussions and comments on earlier drafts that helped to improve the manuscript. I am also grateful to M. Vallade, E. Geissler, O. Rivoire, and A. Dawid for careful reading of the manuscript and fruitful discussions.

LITERATURE CITED

- Arnold, S. J., M. E. Pfrender, and A. G. Jones, 2001 The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112–113: 9–32.
- Bulmer, M., 1985 *The Mathematical Theory of Quantitative Genetics*. Oxford University Press, Oxford, MA.
- Byron, F. W., and R. W. Fuller, 1992 *Mathematics of Classical and Quantum Physics*. Dover Publications Inc., New York, NY.
- Crow, J. F., and M. Kimura, 2009 *An Introduction to Population Genetics Theory*. The Blackburn Press, Caldwell, NJ.
- Falconer, D. S., and T. F. C. Mackay, 1995 *Introduction to Quantitative Genetics*, Ed. 4. Longman, Harlow, UK.
- Fisher, R., 1918 The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* 52: 399–433.
- Gavrilets, S., and A. Hastings, 1995 Dynamics of polygenic variability under stabilizing selection, recombination, and drift. *Genet. Res.* 65: 63–74.
- Geyer, C., and R. G. Shaw, 2008 *Commentary on Lande-Arnold Analysis*. Technical Report 670. School of Statistics/University of Minnesota, Minneapolis, MN.
- Gienapp, P., C. Teplitsky, J. Alho, J. Mills, and J. Merila, 2008 Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* 17: 167–178.
- Haldane, J., 1954 The measurement of natural selection. *Caryologia* 6(Suppl.): 480–487.
- Hansen, T. F., C. Pelabon, and D. Houle, 2011 Heritability is not evolvability. *Evol. Biol.* 38: 258–277.
- Heywood, J. S., 2005 An exact form of the breeder's equation for the evolution of a quantitative trait under natural selection. *Evolution* 59: 2287–2298.
- Hill, W. G., and M. Kirkpatrick, 2010 What animal breeding has taught us about evolution. *Annu. Rev. Ecol. Evol. Syst.* 41: 1–19.
- Karlin, S., 1979 Models of multifactorial inheritance. 1. Multivariate formulations and basic convergence results. *Theor. Popul. Biol.* 15: 308–355.
- Karlin, S., and U. Liberman, 1979a Central equilibria in multilocus systems. 1. Generalized non-epistatic selection regimes. *Genetics* 91: 777–798.
- Karlin, S., and U. Liberman, 1979b Representation of non-epistatic selection models and analysis of multilocus Hardy-Weinberg equilibrium configurations. *J. Math. Biol.* 7: 353–374.
- Kimura, M., and J. F. Crow, 1978 Effect of overall phenotypic selection on genetic change at individual loci. *Proc. Natl. Acad. Sci. USA* 75: 6168–6171.
- Kruuk, L. E. B., 2004 Estimating genetic parameters in natural populations using the 'animal model'. *Phil. Trans. R. Soc. Lond. Ser. B* 359: 873–890.
- Lande, R., 1976 Natural-selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314–334.
- Lande, R., 1979 Quantitative genetic-analysis of multivariate evolution, applied to brain-body size allometry. *Evolution* 33: 402–416.
- Lande, R., and S. J. Arnold, 1983 The measurement of selection on correlated characters. *Evolution* 37: 1210–1226.
- Lewontin, R. C., 1964 The interaction of selection and linkage. II. Optimum models. *Genetics* 50: 757–782.
- Lush, J. L., 1943 *Animal Breeding Plans*. The Iowa State College Press, Ames, IA.
- Lynch, M., and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
- Manna, F., G. Martin, and T. Lenormand, 2011 Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics* 189: 923–937.
- Nagylaki, T., 1992 Rate of evolution of a quantitative character. *Proc. Natl. Acad. Sci. USA* 89: 8121–8124.
- Pemberton, J. M., 2010 Evolution of quantitative traits in the wild: mind the ecology. *Phil. Trans. R. Society B* 365: 2431–2438.
- Pigliucci, M., 2006 Genetic variance-covariance matrices: a critique of the evolutionary quantitative genetics research program. *Biol. Philos.* 21: 1–23.
- Pigliucci, M., and C. D. Schlichting, 1997 On the limits of quantitative genetics for the study of phenotypic evolution RID A-1389–2010. *Acta Biotheor.* 45: 143–160.
- Raj, A., and A. van Oudenaarden, 2008 Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135: 216–226.
- Roff, D., 2000 The evolution of the G matrix: selection or drift? *Heredity* 84: 135–142.
- Slatkin, M., 1970 Selection and polygenic characters. *Proc. Natl. Acad. Sci. USA* 66: 87–93.
- Svardal, H., C. Rueffler, and J. Hermisson, 2011 Comparing environmental and genetic variance as adaptive response to fluctuating selection. *Evolution* 65: 2492–2513.
- Turelli, M., and N. H. Barton, 1990 Dynamics of polygenic characters under selection. *Theor. Popul. Biol.* 38: 1–57.
- Turelli, M., and N. H. Barton, 1994 Genetic and statistical-analyses of strong selection on polygenic traits—what, me normal. *Genetics* 138: 913–941.
- Visscher, P., W. Hill, and N. Wray, 2008 Heritability in the genomics era—concepts and misconceptions RID B-3198–2012. *Nat. Rev. Genet.* 9: 255–266.
- Wright, S., 1920 The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. *Proc. Natl. Acad. Sci. USA* 6: 320–332.
- Wright, S., 1935 The analysis of variance and the correlations between relatives with respect to deviations from an optimum. *J. Genet.* 30: 243–256.
- Zhang, X.-S., and W. G. Hill, 2010 Change and maintenance of variation in quantitative traits in the context of the Price equation. *Theor. Popul. Biol.* 77: 14–22.

Communicating editor: D.-J. De Koning

APPENDIX

FTs and convolutions

Logarithm was invented to simplify algebraic operations: a multiplication in the direct space transforms into an addition in the logarithm space, where it can be performed easily and the result brought back to direct space via the inverse transformation. FT plays a similar role in functional analysis, where derivation/integration of functions in direct space are transformed into multiplication/division by the variable in the Fourier space.

The FT of a function $f(x)$ is defined here as (Byron and Fuller 1992)

$$\tilde{f}(k) = \text{TF}[f(x)] = \int_{-\infty}^{+\infty} f(x)e^{-ikx} dx$$

where $i^2 = -1$. For example, the FT of the function $f(x) = \exp(-\alpha|x|)$ is

$$\begin{aligned} \tilde{f}(k) &= \int_{-\infty}^0 e^{(\alpha-ik)x} dx + \int_0^{\infty} e^{-(\alpha+ik)x} dx \\ &= \frac{1}{\alpha-ik} + \frac{1}{\alpha+ik} = \frac{2\alpha}{\alpha^2+k^2}. \end{aligned}$$

The main properties of FT we use here are

1. Parseval's theorem

$$\int_{-\infty}^{+\infty} f^*(x)g(x)dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{f}^*(k)\tilde{g}(k)dk$$

where a^* stands for the conjugate complex of a . Note that if $f(x)$ is a real function, the $\tilde{f}^*(k) = \tilde{f}(-k)$, which ensures that the right hand side of the above expression is always real if both functions f and g are real.

2. Derivation property

$$i \frac{d}{dk} \tilde{f}(k) = \int_{-\infty}^{+\infty} xf(x)e^{-ikx} dx = \text{TF}[xf(x)]$$

3. Convolution property

$$\text{FT}[(f * g)(x)] = \text{FT}[f(x)].\text{FT}[g(x)] = \tilde{f}(k).\tilde{g}(k)$$

where

$$(f * g)(x) = \int_{-\infty}^{\infty} f(u)g(x-u)du$$

4. Translation property

$$\text{FT}[f(x-\mu)] = e^{-ik\mu}\text{FT}[f(x)]$$

On the basis of the aforementioned properties, and the fact that all the aforementioned functions are real *i.e.*, for example $W^*(z) = W(z)$, we see that relation (3) can be written as

$$\begin{aligned} \iint_{\mathbb{R}^2} yp_0(y)W^*(z)f(z-y)dydz &= \int_{\mathbb{R}} W^*(z)(yp_0 * f)(z)dz \\ &= \frac{i}{2\pi} \int_{\mathbb{R}} \tilde{W}^*(k) \frac{d}{dk} [\tilde{p}(k)]\tilde{f}(k)dk \end{aligned}$$

where we have used the fact (i) that $\text{FT}[yp_0(y)] = i\tilde{p}'(k)$; (ii) FT transforms a convolution product into a simple product in reciprocal space and (iii) Parseval's theorem.

The same set of rules leads to

$$\begin{aligned} \iint_{\mathbb{R}^2} zp_0(y)W^*(z)f(z-y)dydz &= \int_{\mathbb{R}} W^*(z)z(p_0 * f)(z)dz \\ &= \frac{i}{2\pi} \int_{\mathbb{R}} \tilde{W}^*(k) \frac{d}{dk} [\tilde{p}(k)\tilde{f}(k)] dk \end{aligned}$$

Note that we can exchange the order of integration on y and z , write the first integral as a convolution product on functions of z and proceed to the second integral by using the FT on y . For R , we have

$$\begin{aligned} \iint_{\mathbb{R}^2} y p_0(y) W(z) f(z-y) dy dz &= \int_{\mathbb{R}} p_0^*(y) y (W^* f)(y) dy \\ &= \frac{i}{2\pi} \int_{\mathbb{R}} \tilde{p}_0^*(k) \frac{d}{dk} [\tilde{W}(k) \tilde{f}(k)] dk \end{aligned}$$

and for S we get

$$\begin{aligned} \iint_{\mathbb{R}^2} z p_0(y) W(z) f(z-y) dy dz &= \int_{\mathbb{R}} p_0^*(y) (z W * f)(y) dy \\ &= \frac{i}{2\pi} \int_{\mathbb{R}} \tilde{p}_0^*(k) \frac{d}{dk} [\tilde{W}(k)] \tilde{f}(k) dk \end{aligned}$$

The translation property was used in the derivation of the functional lags (eqs 19,20).

Finally, note that the FT of a Gaussian is a Gaussian:

$$\text{FT} \left[\frac{1}{\sqrt{2\pi s}} \exp\left(-x^2/(2s^2)\right) \right] = \exp(-s^2 k^2/2)$$

Computation of truncation selection with FT

Consider a normal breeding value and environmental factor distribution

$$p_0(y) = \mathcal{N}(0, \sigma_A; y)$$

$$f(x) = \mathcal{N}(0, \sigma_E; x)$$

where $\mathcal{N}(a, b; u) = (1/\sqrt{2\pi b}) \exp(-(u-a)^2/2b^2)$ to which we apply a truncation selection $W(z)$, where $W(z) = 1$ if $z_0 < z < z_1$ and 0 otherwise. The FT of these functions read:

$$\tilde{p}_0(k) = e^{-\sigma_A^2 k^2/2}$$

$$\tilde{f}(k) = e^{-\sigma_E^2 k^2/2}$$

$$\tilde{W}(k) = \frac{i}{q} (e^{-iqz_1} - e^{-iqz_0})$$

The mean fitness of the parental generation is

$$\begin{aligned} \bar{W}(z_1, z_0) &= \iint_{\mathbb{R}^2} p_0(y) W(z) f(z|y) dy dz \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{W}^*(k) \tilde{p}_0(k) \tilde{f}(k) dk \\ &= \frac{1}{2} \left(\text{erf}\left(\frac{z_1}{\sqrt{2}\Delta}\right) - \text{erf}\left(\frac{z_0}{\sqrt{2}\Delta}\right) \right) \end{aligned}$$

where $\Delta^2 = \sigma_A^2 + \sigma_E^2$ and $\text{erf}(u)$ is the error function. As

$$\frac{d}{dk} \left[e^{-s^2 q^2/2} \right] = -s^2 q e^{-s^2 q^2/2}$$

the computation of R and S , using their expression (13,14) is simple (the integrands are exact differentials) and leads to

$$R = \frac{1}{W} \frac{\sigma_A^2}{\sqrt{2\pi}\Delta} \left(e^{-z_0^2/2\Delta^2} - e^{-z_1^2/2\Delta^2} \right)$$

$$S = \frac{1}{W} \frac{\Delta}{\sqrt{2\pi}} \left(e^{-z_0^2/2\Delta^2} - e^{-z_1^2/2\Delta^2} \right)$$

and it is trivially verified that $R/S = \sigma_A^2/\Delta^2 = h^2$.

The same computations can be extended to the case where the breeding value distribution is a double Gaussian (Figure 2):

$$p_0(y) = \frac{1}{2}(\mathcal{N}(-m, s; y) + \mathcal{N}(m, s; y))$$

$$\tilde{p}_0(k) = \frac{1}{2}(e^{imk} + e^{-imk})e^{-s^2k^2/2}$$

The computations, although more cumbersome, involve the same level of technicality. For simplicity, we give the result only for extreme truncation $z_1 \rightarrow \infty$ (selecting all individuals with phenotype $> z_0$). Noting

$$\bar{W}(m) = \frac{1}{2} \left(1 - \operatorname{erf} \left(\frac{z_0 - m}{\sqrt{2\Delta}} \right) \right)$$

$$A(m) = \frac{s^2}{\sqrt{2\pi}\Delta} e^{-(z_0 - m)^2/2\Delta^2} + \frac{m}{2} \operatorname{erfc} \left(\frac{z_0 - m}{\sqrt{2\Delta}} \right)$$

$$B(m) = \frac{\Delta}{\sqrt{2\pi}} e^{-(z_0 - m)^2/2\Delta^2} + \frac{m}{2} \operatorname{erfc} \left(\frac{z_0 - m}{\sqrt{2\Delta}} \right)$$

Where $\Delta = s^2 + \sigma_E^2$. The selection and response function read:

$$R = \frac{A(m) + A(-m)}{W(m) + W(-m)}$$

$$S = \frac{B(m) + B(-m)}{W(m) + W(-m)}$$

The simple Gaussian case can be recovered from these expressions by $m = 0$ and $\sigma_A = S$.

Parent-offspring regression

The derivation of the breeder's equation sometimes uses the parent-offspring regression coefficient as an intermediate (Nagylaki 1992; Lynch and Walsh 1998). The linear regression between parent and offspring phenotype however is based on the same assumption of normal distribution of genotype and environmental factors.

The probability density of observing the phenotype z' in the offspring and z_a, z_b in the parents is

$$p(z'; z_a, z_b) = \iint_{\mathbb{R}^2} p(y_a) f(z_a | y_a) p(y_b) f(z_b | y_b) L(y_1 - (y_a + y_b)/2) f(z' | y_1) dy_a dy_b$$

and the conditional expectation of z' given z is

$$E(z' | z_a, z_b) = \int_{z' \in I} z' p(z', z) dz' / \int_{z' \in I} p(z', z) dz' = F(z_a, z_b)$$

It is not difficult to check that the function $F(z_a, z_b)$ is a linear function of its argument

$$F(z_a, z_b) = b(z_a + z_b)/2$$

if both the genotype and environment factors obey a normal distribution, in which case, the linearity coefficient is indeed $b = \sigma_A^2 / (\sigma_A^2 + \sigma_E^2)$. However, *even* if the parental generation follows a normal distribution, the *selected* parents do not (equation 7) and the use of parent-offspring regression poses even more of a problem than the direct method.

Segregation density function

Let $p_0(y)$ be the distribution of breeding value in the parental generation. In the absence of selection, after recombination-segregation, the distribution of breeding value in the progeny is

$$p_1(y) = \iint_{\mathbb{R}^2} p_0(y_a) p_0(y_b) L(y - (y_a + y_b)/2) dy_a dy_b \quad (28)$$

where the function $L(y)$ is the segregation density function capturing the inheritance process of the breeding value (Karlín 1979). $L(y)$ is a probability density function and in the absence of epistasis and dominance effect, its average is zero: $\int_{\mathbb{R}} y L(y) dy = 0$. In the infinitesimal model framework, $L(y)$ is a normal distribution of variance $\sigma_A^2/2$. However, any distribution probability $L(y)$ will lead to a stable, although not necessarily normal, probability distribution of breeding values after few round of reproduction. Let us set the origin of the breeding value at its average in the parental distribution, *i.e.* $\int_{\mathbb{R}} y p_0(y) dy = 0$. In Fourier space relation (28) is

$$\tilde{p}_1(k) = \tilde{p}_0^2(k/2)\tilde{L}(k)$$

and after n rounds of reproduction,

$$\tilde{p}_n(k) = \tilde{p}_0^{2^n}(k/2^n) \prod_{i=0}^{n-1} \tilde{L}^{2^i}(k/2^i)$$

As both $p_0(y)$ and $L(y)$ are probability distribution functions of zero mean, we have

$$\tilde{p}_0(0) = \tilde{L}(0) = 1$$

$$\tilde{p}'_0(0) = \tilde{L}'(0) = 0$$

and therefore

$$\tilde{p}''_n(0) = \frac{1}{2^n} \tilde{p}''_0(0) + \left(2 - \frac{1}{2^{n-1}}\right) \tilde{L}''(0)$$

Let $V = \int_{\mathbb{R}} y^2 L(y) dy$. We see then that

$$\text{Var}(Y_n) = \frac{1}{2^n} \text{Var}(Y_0) + \left(2 - \frac{1}{2^{n-1}}\right) V$$

So the variance of the breeding values converges fast to twice the variance of the segregation density function. The distribution function $p_n(y)$, however, converges to a normal distribution only if $L(y)$ is normal.