

Supplementary material to:

Chapter 9: General Solutions and Transformations – Models with Multiple Variables

From:

A Biologist's Guide to Mathematical Modeling in Ecology and Evolution

S. P. Otto and T. Day (2005)

Princeton University Press

Supplementary Material 9.1: Additive versus multiplicative interactions among loci

In the text, we used a perturbation analysis to approximate the linkage disequilibrium between two genes as:

$$\hat{D}_Q = \frac{(1-r)p_A(1-p_A)p_B(1-p_B)}{r}(w_1 - w_2 - w_3 + w_4) + O(\xi^2). \quad (\text{S9.1.1})$$

This approximation is accurate when $w_1 - w_2 - w_3 + w_4$ is large enough (of order ξ), but it becomes inaccurate when it is smaller (of order ξ^2). Here, we show how to obtain an approximation for the quasi-equilibrium value of the disequilibrium, \hat{D}_Q , that is accurate even when $w_1 - w_2 - w_3 + w_4$ is small.

To obtain a more accurate approximation, you might simply continue with the perturbation analysis and seek out the next higher-order term, \hat{D}_2 , in the approximation for the quasi-equilibrium, $\hat{D}_Q = \hat{D}_0 + \hat{D}_1 \xi + \hat{D}_2 \xi^2 + \dots$. This calculation turns out to be very messy. Yet we know that the answer is very nearly (S9.1.1) when the fitnesses are far from additive (i.e., $w_1 - w_2 - w_3 + w_4$ is large), and we only need a better approximation for those cases where the fitnesses are close to additive. Therefore, it is easiest to first re-write the fitness expressions in a form that allows us to specify the extent to which they deviate from additivity with a single (small) parameter.

One way to specify that the fitness regime is close to additive is by setting:

$$\begin{aligned}
 w_1 &= 1 + \tilde{s}_A \zeta + \tilde{s}_B \zeta + \tilde{\eta} \zeta^2 \\
 w_2 &= 1 + \tilde{s}_A \zeta \\
 w_3 &= 1 + \tilde{s}_B \zeta \\
 w_4 &= 1
 \end{aligned}
 \tag{S9.1.2}$$

where $\tilde{\eta} \zeta^2$ represents a very small departure from additive fitnesses. Using this fitness regime, along with $\hat{D}_Q = \hat{D}_0 + \hat{D}_1 \zeta + \hat{D}_2 \zeta^2 + \dots$ in (9.43), we can then repeat the perturbation analysis.

Setting $\zeta = 0$, we again find that $f(0) = 0$ only if $\hat{D}_0 = 0$. Furthermore, $\frac{df(0)}{d\zeta} = 0$ only if $\hat{D}_1 = 0$.

Finally, $\frac{d^2 f(0)}{d\zeta^2}$ is zero if:

$$\hat{D}_2 = \frac{(1-r)p_A(1-p_A)p_B(1-p_B)}{r} (\tilde{\eta} - \tilde{s}_A \tilde{s}_B).
 \tag{S9.1.3}$$

Using (S9.1.3), we can approximate the quasi-equilibrium level of disequilibrium as:

$$\begin{aligned}
 \hat{D}_Q &= \hat{D}_2 \zeta^2 + O(\zeta^3) \\
 &= \frac{(1-r)p_A(1-p_A)p_B(1-p_B)}{r} (\tilde{\eta} - \tilde{s}_A \tilde{s}_B) \zeta^2 + O(\zeta^3).
 \end{aligned}
 \tag{S9.1.4}$$

We can write this result in terms of the original fitnesses (S9.1.2) by noting that

$(\tilde{\eta} - \tilde{s}_A \tilde{s}_B) \zeta^2 = w_1 w_4 - w_2 w_3$. The quasi-equilibrium value of the disequilibrium is then:

$$\begin{aligned}
 \hat{D}_Q &= \hat{D}_2 \zeta^2 + O(\zeta^3) \\
 &= \frac{(1-r)p_A(1-p_A)p_B(1-p_B)}{r} (w_1 w_4 - w_2 w_3) + O(\zeta^3).
 \end{aligned}
 \tag{S9.1.5}$$

Thus, by making the explicit assumption that fitnesses were close to additive in (S9.1.2), we were able to refine our perturbation analysis to demonstrate that disequilibrium develops in a manner proportional to $w_1 w_4 - w_2 w_3$.

Equations (S9.1.1) and (S9.1.5) differ only in how they depend on fitness interactions. Equation (S9.1.5) suggests that the key determinant of the sign of linkage disequilibrium is the quantity $w_1w_4 - w_2w_3$, which measures the extent to which fitnesses depart from a *multiplicative* expectation. If alleles A_1 and B_1 have selection coefficients equal to s_A and s_B , respectively, a multiplicative fitness scheme can be written as:

$$\begin{aligned} w_1 &= (1 + s_A)(1 + s_B) \\ w_2 &= 1 + s_A \\ w_3 &= 1 + s_B \\ w_4 &= 1 \end{aligned} \quad (\text{S9.1.6})$$

(Check for yourself that $w_1w_4 - w_2w_3$ equals zero under a multiplicative fitness scheme.) The additive fitness scheme (9.46) and the multiplicative fitness scheme (S9.1.6) are very similar, except that w_1 is slightly higher, by an amount $s_A s_B$, in the multiplicative scheme. Assuming weak selection, this difference is very small, $O(\xi^2)$, and was ignored in the perturbation analysis leading to (9.45).

Aside from providing a more accurate approximation that can be used even when fitnesses are nearly additive, expression (S9.1.5) is also a more natural determinant of the sign of linkage disequilibrium. The reason for this comes directly from equation (9.40c), which describes the exact dynamics of linkage disequilibrium within a population. If the population begins in linkage equilibrium (i.e., if $D = 0$), then equation (9.40c) predicts the exact value of D in the next generation:

$$D(t+1) = (1-r)p_A(1-p_A)p_B(1-p_B) \frac{w_1w_4 - w_2w_3}{\bar{w}_h^2}, \quad (\text{S9.1.7})$$

Thus, without making any assumptions about the strength of selection, we see that the sign of the disequilibrium that develops depends on the fitness relationship $w_1w_4 - w_2w_3$, which measures the departure from a multiplicative fitness scheme.

Let's now consider what the quasi-equilibrium value of \hat{D}_Q (S9.1.5) means biologically. If $w_1w_4 > w_2w_3$, the fitnesses of A_1B_1 and A_2B_2 haploids are higher when multiplied together than

the fitnesses of A_1B_2 and A_2B_1 haploids; in this case, we say that the fitnesses exhibit *positive epistasis* as measured on a multiplicative scale (Figure S9.1.1). Equation (S9.1.5) tells us that positive epistasis in fitness generates positive disequilibrium, so that alleles A_1 and B_1 co-occur in the same haploid genotype more often than we would expect by chance, as do alleles A_2 and B_2 . Conversely, if $w_1w_4 < w_2w_3$, the fitnesses of A_1B_1 and A_2B_2 are lower when multiplied together than the fitnesses of A_1B_2 and A_2B_1 ; in this case, we say that the fitnesses exhibit *negative epistasis* on a multiplicative scale. Equation (S9.1.5) tells us that negative epistasis generates negative disequilibrium, causing alleles A_1 and B_2 (and A_2 and B_1) to co-occur in the same haploid genotype more often than we would expect by chance. In short, at quasi-equilibrium, the relatively more fit combinations of alleles become more frequent.

Figure S9.1.1: Fitness as a function of the number of mutant alleles carried by an individual, A_1 and B_1 . If each mutation affects fitness independently, there is no epistasis, measured on a multiplicative scale (thick curve). If fitness rises faster than multiplicatively, positive epistasis results (dashed curve), while if fitness rises slower than multiplicatively, negative epistasis results (dotted curve). To simplify the presentation, we have assumed that fitness depends on the number of mutations, n , according to the function, $e^{\alpha n + \beta n^2}$; epistasis as measured on a multiplicative scale then depends only on the sign of β (negative epistasis: $\beta < 0$; no epistasis: $\beta = 0$; positive epistasis: $\beta > 0$). (a) Mutant alleles are beneficial. (b) Mutant alleles are deleterious.

