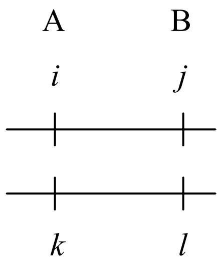
Ad Rice p. 47-51, following F.C Hoppensteadt (1976) A slow selection analyses of two locus, two allele traits. Theor. Pop. Biol. 9: 68-81.

Notation:

Zygote chromosomal pattern:



(relative) frequency: Q_{ijkl}

per capita birth rate: b assumed here to be the same for everybody (this assumption simplifies the formulas but is not essential)

per capita death rate: μ_{ijkl}

crossing over probability: r

Symmetries:

$$\mu_{ijkl} = \mu_{klij} \quad \text{(no parental effect)}$$
$$= \mu_{ilkj} \quad \text{(no cis-effect)}$$
$$\Rightarrow \quad = \mu_{kjil}$$

Consequence: The subspace $Q_{ijkl} = Q_{klij}$ is invariant. From now on it is assumed that $Q_{ijkl} = Q_{klij}$.

Notation: $x_{ij} = \sum_{kl} Q_{ijkl} = \sum_{kl} Q_{klij}$ (chromosome frequencies), $p_i = \sum_j x_{ij}$, $q_j = \sum_i x_{ij}$ (gene frequencies)

Basic balance law:

$$\frac{\mathrm{d}Q_{ijkl}}{\mathrm{d}t} = bx'_{ij}x'_{kl} - \left(\mu_{ijkl} + \underbrace{b - \bar{\mu}}\right)Q_{ijkl}\,, \qquad \bar{\mu} := \sum_{ijkl} \mu_{ijkl} Q_{ijkl}$$

↑ from the fact that the Q's are relative frequencies

gametes:

$$x'_{ij} = (1-r)\sum_{kl} Q_{ijkl} + r\sum_{kl} Q_{ilkj} \quad = x_{ij} - r\overbrace{\sum_{kl} \underbrace{\left(Q_{ijkl} - Q_{ilkj}\right)}_{D_{ijkl}}}^{D_{ij}}$$

Some useful relations:

$$\sum_j D_{ij} = 0\,, \qquad \sum_i D_{ij} = 0$$
$$D_{ijkl} = D_{klij}\,,$$
$$D_{ijkl} = -D_{ilkj} \qquad \Rightarrow \quad D_{ijkj} = 0$$
$$\Rightarrow \quad D_{ijkl} = -D_{kjil} \qquad \Rightarrow \quad D_{ijil} = 0$$

Example: two alleles per locus

$$D := Q_{1122} - Q_{1221}$$

(linkage disequilibrium in adults)

	D_{ijkl}	kl 11	12	21	22	D_{ij}
ij	11	0	0	0	D	D
	12	0	0	$-D$	0	$-D$
	21	0	$-D$	0	0	$-D$
	22	D	0	0	0	D

Notation:

$$\theta_{ij} := \mathrm{sign}(D_{ij})$$

$$D_{ij} = \theta_{ij} D$$

Two alleles per locus (from now on)

$$\frac{\mathrm{d}Q_{ijkl}}{\mathrm{d}t} = bx'_{ij}x'_{kl} - \left(\mu_{ijkl} + b - \bar{\mu}\right)Q_{ijkl}\,, \qquad \bar{\mu} := \sum_{ijkl}\mu_{ijkl}Q_{ijkl}\,, \qquad \mu^*_{ij} := \sum_{kl}\mu_{ijkl}Q_{ijkl}/x_{ij}\,,$$

$$x'_{ij} = x_{ij} - r\theta_{ij}D\,, \qquad D := Q_{1122} - Q_{1221}\,,$$

$$\frac{\mathrm{d}x_{ij}}{\mathrm{d}t} = bx'_{ij} - \left(\mu^*_{ij} + b - \bar{\mu}\right)x_{ij} = bx_{ij} - br\theta_{ij}D - \left(\mu^*_{ij} + b - \bar{\mu}\right)x_{ij} = -br\theta_{ij}D - \left(\mu^*_{ij} - \bar{\mu}\right)x_{ij}\,,$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\left(\mu^*_{i\mathrm{g}} - \bar{\mu}\right)p_i\,, \qquad \frac{\mathrm{d}q_j}{\mathrm{d}t} = -\left(\mu^*_{\mathrm{g}j} - \bar{\mu}\right)q_j\,,$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \frac{\mathrm{d}Q_{1122}}{\mathrm{d}t} - \frac{\mathrm{d}Q_{1221}}{\mathrm{d}t} = b\underbrace{\left(x'_{11}x'_{22} - x'_{12}x'_{21}\right)}_{D'} - \mu_{1122}\underbrace{\left(Q_{1122} - Q_{1221}\right)}_{D} - \left(b - \bar{\mu}\right)D$$

$$= b\left(D' - D\right) - \left(\mu_{1122} - \bar{\mu}\right)D$$

$$= b\left(D^0 - D\right) - brD - \left(\mu_{1122} - \bar{\mu}\right)D\,,$$

with

$$D^0 := x_{11}x_{22} - x_{12}x_{21} = \theta_{ij}\left(x_{ij} - p_iq_j\right). \qquad \text{[c.f. Rice formula (2.3)]}$$

(Use

$$D' = x_{11}x_{22} - r\left(x_{11}\theta_{22} + \theta_{11}x_{22}\right)D + r^2\theta_{11}\theta_{22}D^2$$
$$- x_{12}x_{21} + r\left(x_{12}\theta_{21} + \theta_{12}x_{21}\right)D - r^2\theta_{12}\theta_{21}D^2$$
$$= D^0 - rD.)$$

Interpretation: D' is the linkage disequilibrium in the newborn, D the linkage disequilibrium in the adult population. To interpret $D^0 = D' + rD$, consider the case where there is no selection so that $\mathrm{d}D/\mathrm{d}t = b\left(D' - D\right) = b\left(D^0 - D\right) - brD$. In that case, D changes through two processes, through a replacement of the initial adults by the new crop, and through recombination. The term $b\left(D' - D\right)$ still represents the combined effect of these processes. In $b\left(D^0 - D\right) - brD$ the contributions of two processes have been separated: D^0 is what would become of D if both selection and recombination were switched off.

Time scale separation

Convention: Here, and below, any functions called g_i satisfy *a priori* bounds, and any $K_i \in \mathbb{R}_+$.

Set $\mu_{ijkl} = \mu - \varepsilon v_{ijkl}$ (and potentially $r = f(\varepsilon)$), so that we can write

$$\frac{\mathrm{d}Q_{ijkl}}{\mathrm{d}t} = bx'_{ij}x'_{kl} + \left(-b + \varepsilon\left(v_{ijkl} - \bar{v}\right)\right)Q_{ijkl}$$

$$\frac{\mathrm{d}x_{ij}}{\mathrm{d}t} == -br\theta_{ij}D + \varepsilon\left(v^*_{ij} - \bar{v}\right)x_{ij}$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \varepsilon\left(v^*_{i\mathrm{g}} - \bar{v}\right)p_i\,, \qquad \frac{\mathrm{d}q_j}{\mathrm{d}t} = \varepsilon\left(v^*_{\mathrm{g}j} - \bar{v}\right)q_j\,, \qquad v^*_{i\mathrm{g}} = p_i^{-1}\sum_j v^*_{ij}x_{ij}\,,\ v^*_{\mathrm{g}j} = q_j^{-1}\sum_i v^*_{ij}x_{ij}$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = b\left(D^0 - D\right) - brD + \varepsilon\left(v_{1122} - \bar{v}\right)D\,.$$

Moreover, define the new variables $\tau = \varepsilon t$ (the time scale on which the selection effects become noticeable), and

$$q_{ijkl} = Q_{ijkl} - x'_{ij}x'_{kl}. \qquad d = D - D^0.$$

so that

$$\frac{\mathrm{d}q_{ijkl}}{\mathrm{d}t} = -bq_{ijkl} + \varepsilon\left(v^*_{ij} - \bar{v}\right)Q_{ijkl} - \frac{\mathrm{d}x'_{ij}}{\mathrm{d}t}x'_{kl} - x'_{ij}\frac{\mathrm{d}x'_{kl}}{\mathrm{d}t}, \qquad \frac{\mathrm{d}d}{\mathrm{d}t} = -bd + \left(-br + \varepsilon\left(v_{1122} - \bar{v}\right)\right)D - \frac{\mathrm{d}D^0}{\mathrm{d}t}.$$

In more detail,

$$\frac{\mathrm{d}q_{ijkl}}{\mathrm{d}t} = -bq_{ijkl} + \varepsilon\left[\left(v_{ijkl} - \bar{v}\right)Q_{ijkl} - \left(v^*_{ij} + v^*_{kl} - 2\bar{v}\right)x_{ij}x_{kl}\right] + \varepsilon\left[\left(v^*_{ij} - \bar{v}\right)x_{ij}\theta_{kl} + \left(v^*_{kl} - \bar{v}\right)\theta_{ij}x_{kl}\right]rD - 2b\theta_{ij}\theta_{kl}r^2D^2,$$

and, since

$$\begin{aligned}
\frac{\mathrm{d}D^0}{\mathrm{dt}} &= \frac{\mathrm{d}x_{11}}{\mathrm{d}\tau}x_{22} + x_{11}\frac{\mathrm{d}x_{22}}{\mathrm{d}\tau} - \frac{\mathrm{d}x_{12}}{\mathrm{d}\tau}x_{21} - x_{12}\frac{\mathrm{d}x_{21}}{\mathrm{d}\tau} \\
&= \varepsilon\left(v^*_{11} - \bar{v}\right)x_{11}x_{22} - br\theta_{11}Dx_{22} + \varepsilon\left(v^*_{22} - \bar{v}\right)x_{11}x_{22} - br\theta_{22}Dx_{11} \\
&\quad - \varepsilon\left(v^*_{12} - \bar{v}\right)x_{12}x_{21} + br\theta_{12}Dx_{21} - \varepsilon\left(v^*_{21} - \bar{v}\right)x_{12}x_{21} + br\theta_{21}Dx_{12} \\
&= \varepsilon\left(\left(v^*_{11} + v^*_{22}\right)x_{11}x_{22} - \left(v^*_{12} + v^*_{21}\right)x_{12}x_{21} - 2\bar{v}D^0\right) - brD \\
&= \varepsilon\left(\left(v^*_{12} + v^*_{21}\right)D^0 + \left(v^*_{11} + v^*_{22} - v^*_{12} - v^*_{21}\right)x_{11}x_{22} - 2\bar{v}D^0\right) - brD,
\end{aligned}$$

$$\frac{\mathrm{d}d}{\mathrm{d}t} = -bd + \varepsilon\left(-\left(v^*_{11} + v^*_{22}\right)x_{11}x_{22} + \left(v^*_{12} + v^*_{21}\right)x_{12}x_{21} + \left(v_{1122} + \bar{v}\right)D\right) = -bd + \varepsilon g_1(t).$$

Hence, for sufficiently large t: $\quad |d| < K_1\varepsilon$.

Case I: Set $r = \varepsilon\rho$ and let ε become small ("continous Hardy Weinberg")**.**

Then

$$\frac{\mathrm{d}q_{ijkl}}{\mathrm{d}t} = -bq_{ijkl} + \varepsilon g_1(t). \qquad \text{[c.f. Rice formula (2.16)]}$$

Hence, for sufficiently large t: $\quad \left|q_{ijkl}\right| < K_2\varepsilon$.

On the long time scale

$$\frac{\mathrm{d}x_{ij}}{\mathrm{d}\tau} = \left(v^{*0}_{ij} - \bar{v}^0\right)x_{ij} - b\rho\theta_{ij}D^0 + \varepsilon g_3(\tau)$$

$$\frac{\mathrm{d}D^0}{\mathrm{d}\tau} = \left(v^{*0}_{12} + v^{*0}_{21}\right)D^0 + \underbrace{\left(v^{*0}_{11} + v^{*0}_{22} - v^{*0}_{12} - v^{*0}_{21}\right)}_{E^0}x_{11}x_{22} - 2\bar{v}^0D^0 - b\rho D^0 + \varepsilon g_4(\tau)$$

c.f. Rice's formulas (2.19 & (2.20).

It follows that at equilibrium:

$$\hat{D} = \hat{D}^0 + \mathrm{O}(\varepsilon) = \frac{\hat{x}_{11}\hat{x}_{22}E^0}{2\bar{v}^0 - v^{*0}_{12} - v^{*0}_{21} + b\rho} + \mathrm{O}(\varepsilon). \qquad \text{[c.f Rice formule (2.21)]}$$

The conclusion is that in this case $\hat{D}$ will be close to 0 whenever the epistasis measure $E^0 = 0$, whereas $\hat{D}$ is O(1) when $E^0 \neq 0$. (Note that E^0 is not *a priori* fixed, but depends itself on the $\hat{x}_{ij}$.)

Case II: Let ε become small while keeping r constant.
Then

$$\frac{\mathrm{d}D^0}{\mathrm{d}t} = -brD^0 + \varepsilon g_5(\tau)$$

Hence, for sufficiently large t: $\left|D^0\right| < K_3\varepsilon$ and $|D| < K_4\varepsilon$.

Hence, for sufficiently large t:

$$\frac{\mathrm{d}q_{ijkl}}{\mathrm{d}t} = -bq_{ijkl} + \varepsilon g_6(t).$$

Since

$$x_{ij} = p_i q_j + \theta_{ij} D^0 \quad \text{[Rice page 42], and} \quad x'_{ij} = p_i q_j - r\theta_{ij} D \quad \text{[by definition]}$$

it follows that, for sufficiently large t: $\left|Q_{ijkl} - p_i q_j p_k q_l\right| < K_5\varepsilon$

and

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \varepsilon\left(v_{i\mathrm{g}}^{*00} - \bar{v}^{00}\right)p_i + \varepsilon^2 g_7(t), \qquad \frac{\mathrm{d}q_j}{\mathrm{d}t} = \varepsilon\left(v_{\mathrm{g}j}^{*00} - \bar{v}^{00}\right)q_j + \varepsilon^2 g_8(t),$$

where $v_{i\mathrm{g}}^{*00}$, $v_{\mathrm{g}j}^{*00}$ and $\bar{v}^{00}$ are calculated like $v_{i\mathrm{g}}^{*}$, $v_{\mathrm{g}j}^{*}$ and $\bar{v}$, but using $Q_{ijkl}^{00} = p_i q_j p_k q_l$ instead of Q_{ijkl}.
Hence, on the long time scale $\tau = \varepsilon t$

$$\frac{\mathrm{d}p_i}{\mathrm{d}\tau} = \left(v_{i\mathrm{g}}^{*00} - \bar{v}^{00}\right)p_i + \varepsilon g_7(t), \qquad \frac{\mathrm{d}q_j}{\mathrm{d}\tau} = \varepsilon\left(v_{\mathrm{g}j}^{*00} - \bar{v}^{00}\right)q_j + \varepsilon g_8(t),$$

and

$$\left|x_{ij} - p_i q_j\right| < K_6\varepsilon, \qquad \left|Q_{ijkl} - x_{ij}x_{kl}\right| < K_7\varepsilon.$$

Case III: Let $\varepsilon \ll r = \rho\varepsilon \ll 1$ and rescale $\tilde{D} = \rho D^0$.
Then

$$\frac{\mathrm{d}\tilde{D}}{\mathrm{d}\tau} = \left(v_{12}^{*0} + v_{21}^{*0}\right)\tilde{D} + \rho\left(v_{11}^{*0} + v_{22}^{*0} - v_{12}^{*0} - v_{21}^{*0}\right)x_{11}x_{22} - 2\bar{v}^0\tilde{D} - b\rho\tilde{D} + \varepsilon\rho g_4(\tau).$$

Define the intermediate time scale $\sigma = \rho\tau = rt$, i.e., σ is fast relative to τ, but slow relative to t. Then $x_{ij} = p_i q_j + \rho^{-1}\tilde{D}$ can be used to arrive at (use the equation on the first line to show that $\tilde{D}$ stays bounded before moving to the second line)

$$\begin{aligned}\frac{\mathrm{d}\tilde{D}}{\mathrm{d}\sigma} &= -b\tilde{D} + \left(v_{11}^{*0} + v_{22}^{*0} - v_{12}^{*0} - v_{21}^{*0}\right)x_{11}x_{22} + \rho^{-1}\left[\left(v_{12}^{*0} + v_{21}^{*0}\right)\tilde{D} - 2\bar{v}^0\tilde{D}\right] + \varepsilon g_4(\sigma) \\ &= -b\tilde{D} + \left(v_{11}^{*00} + v_{22}^{*00} - v_{12}^{*00} - v_{21}^{*00}\right)p_1q_1p_2q_2 + \rho^{-1}\left[\left(v_{12}^{*0} + v_{21}^{*0}\right)\tilde{D} - 2\bar{v}^0\tilde{D}\right] + \varepsilon g_4(\sigma) + \rho^{-1}g_9(\sigma)\end{aligned}$$

with $v_{ij}^{*00} = \sum_{kl} v_{ijkl}p_k q_l$ as before.

Let $|g_9| < K_9$ and $|g_4| < K_{10}$. Then $\frac{\mathrm{d}\tilde{D}}{\mathrm{d}\sigma} < 0$ when $b\tilde{D} - \left(v_{11}^{*00} + v_{22}^{*00} - v_{12}^{*00} - v_{21}^{*00}\right)p_1q_1p_2q_2 > \rho^{-1}K_4 + \varepsilon K_{10}$ and $\frac{\mathrm{d}\tilde{D}}{\mathrm{d}\sigma} > 0$ when $b\tilde{D} - \left(v_{11}^{*00} + v_{22}^{*00} - v_{12}^{*00} - v_{21}^{*00}\right)p_1q_1p_2q_2 < -\left(\rho^{-1}K_4 + \varepsilon K_{10}\right)$. Hence $\tilde{D}$ converges to

$$\left|\tilde{D} - \frac{\left(v_{11}^{*00} + v_{22}^{*00} - v_{12}^{*00} - v_{21}^{*00}\right)p_1q_1p_2q_2}{b}\right| < \frac{K_9\rho^{-1} + K_{10}\varepsilon}{b}.$$

Note: Although this expression seemingly differs from the one in formula (2.22) in Rice, they are actually the same, if one takes the error terms into account. To see this, remember that also Rice is considering the weak selection limit. Hence it is possible to write $w_{ijkl} = 1 + \varepsilon s_{ijkl}$ so that

$$\frac{w_{11}^* w_{22}^* - w_{12}^* w_{21}^*}{\bar{w}} = \frac{\left(1+\varepsilon s_{11}^*\right)\left(1+\varepsilon s_{22}^*\right) - \left(1+\varepsilon s_{12}^*\right)\left(1+\varepsilon s_{21}^*\right)}{1+\varepsilon \bar{s}}$$
$$= \varepsilon\left(s_{11}^* + s_{22}^* - s_{12}^* - s_{21}^*\right) + \mathrm{O}\left(\varepsilon^2\right) \qquad \square$$

On the long time scale $\tau = \varepsilon t$, similar to the situation in case II,

$$\frac{\mathrm{d}p_i}{\mathrm{d}\tau} = \left(v_{i\mathrm{g}}^{*00} - \bar{v}^{00}\right)p_i + \varepsilon g_7(\tau) + \rho^{-1} g_9(\tau), \qquad \frac{\mathrm{d}q_j}{\mathrm{d}\tau} = \left(v_{\mathrm{g}j}^{*00} - \bar{v}^{00}\right)q_j + \varepsilon g_8(\tau) + \rho^{-1} g_{10}(\tau).$$

General remarks

In cases II and III Fisher's Fundamental Theorem, as encountered in the one locus case, once again applies on the slow time scale.

These pages were mainly written to make sense of pages 47-51 of Rice. There he uses a continuous time model that cannot *prima facie* be interpreted in individual-based terms. The reason is that any continuous time model necessarily has overlapping generations. This means that although in the birth stream gene frequencies lie on the Hardy-Weinberg manifold, this is no longer the case for frequencies of individuals. Hence Rice's continuous time equations only make sense when seen as a description of the long-term effect of weak selection. However, the most immediate consideration of the weak selections case ends up as case II above. Rice's equations initially refer to case I, and further on, when he considers quasi-linkage equilibrium, to case III. However, he is never discusses the precise nature of these frameworks. In particular the necessity to rescale *D* can be as stumbling point for anyone trying to dig into the conceptual foundation of his equations.

In the models considered above it was possible to write down immediate easy error estimates. Often this is not possible. In that case it is necessary to fall back on the standard theorems from singular perturbation theory as found in, for example, F.C. Hoppensteadt (1993) Analysis and simulation of chaotic systems. Springer, New York, chapter 8.
In our case, applying these theorems would lead to the following statements, where each time only two time scales, to be called fast and slow, are considered relative to each other.
On the fast time scale, the solution of the original differential equation, for $\rho^{-1}, \varepsilon > 0$, will converge to the solution of the corresponding limit equation (the "fast equation"), with $\rho^{-1}, \varepsilon = 0$, uniformly on any bounded time interval.
If the fast equation has for each fixed value of the slow variables a unique globally attractive equilibrium, the solution of the full problem converges on the slow time scale to the solution of the corresponding limit equation (the "slow equation") at least uniformly on any time interval bounded away from zero and infinity, with as initial condition (for the slow time goes to zero) the equilibrium reached by the fast equation.
If the fast equation only has locally attracting equilibria the previous statement applies for all initial conditions in the domain of attraction of the equilibrium under consideration.
If the solution of the slow equation converges to an equilibrium point, the convergence of the solution of the full equation is uniform on time intervals that extend to infinity.
If the solution of the slow equation homes in on a limit cycle, then for $\varepsilon,\ \rho^{-1} \to 0$ the full forward orbit, starting from an initial condition within an ε- (ρ^{-1}-) distance of the equilibrium of the fast equation, converges to the forward orbit of the slow equation. (The difference with the previous case is that here time is left out, so that any larger effect of very small initial perturbations on the final phase do not mess up the picture.)
(Of course, all these results depend on sufficient smoothness conditions being fulfilled.)