

1 Genetic background (simplified)

DNA

Deoxyribonucleic acid; the primary genetic material of a cell; a polymer consisting of four types of nucleotides: adenine (A), guanine (G), thymine (T), and cytosine (C); typically contains two polynucleotide chains in the form of double helix (with A-T and G-C pairings); DNA are the largest biologically active molecules known; three nucleotides code for an amino acid of which there are 20. A protein consists of a few hundred amino acids.

Gene

- Classically, a unit of inheritance transmitted from generation to generation by a gamete (sperm or egg), which controls a particular characteristic of an individual.
- In modern genetics, the basic unit of inheritance comprising a specific sequence of nucleotides on a DNA chain that has a specific function and occupy a specific position (**locus**) on a chromosome; alternative forms of a gene are known as **alleles**.

Table 1: Gene number

| | |
|--|----------------|
| Prokaryotes | 1,000-8,000 |
| Eukaryotes (except vertebrates) | 7,000-15,000 |
| Vertebrates | 25,000-100,000 |

(Prokaryotes are organisms in which the genetic material is not enclosed in a cell nucleus. Bacteria are prokaryotes. Eukaryotes are organisms consisting of cells in which the genetic material is contained within a distinct nucleus. All organisms except bacteria are eukaryotes. Vertebrates are animals that have backbones. Vertebrates include the fishes, amphibians, reptiles, birds and mammals.)

Chromosome is a single thread-like molecule of DNA surrounded by various proteins.

Haploid organisms have a single set of chromosomes (bacteria, algae, mosses, fungi).

Diploid organisms have a double set of chromosomes (higher plants and animals). Humans have 22 matched pairs of chromosomes plus one pair of sex chromosomes.

Gamete is a reproductive cell (sperm or egg) that fuses with another gamete to form a **zygote**. Gametes are haploid. Zygotes develop to whole organisms.

2 Asexual haploid populations

2.1 Basic equations

We consider a population of asexual haploid individuals. We assume that there is no gene exchange between different organisms. Individuals reproduce by division (examples: bacteria, algae, fungi), and each individual has a single parental organism. The population has discrete and non-overlapping generations (that is adults are replaced by their offspring at the end of each generation). Population size is very large (and, thus, stochastic effects can be neglected).

We start with a *simplified* deterministic model assuming that individuals are different with respect to a single locus with k alleles. In other words, there are k different genetic types (=genotypes) of individuals. Few words about notation. We will use different **bold** letters to represent different genes with subscripts to denote different alleles. Thus, in the model under consideration there are k alleles $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k$.

Let $n_i(t)$ be the number of alleles \mathbf{A}_i at the beginning of the generation t , and $N(t) = \sum_i n_i(t)$ be the (overall) population size. Both $n_i(t)$ and $N(t)$ are assumed to be sufficiently large. The life cycle of the population is illustrated in Figure 1.

Figure 1 will be here

We assume that individuals of different types differ with respect to *viability* v_i (defined as the probability that an \mathbf{A}_i offspring survives to reproductive age) and *fertility* f_i (defined as the average number of offspring of an \mathbf{A}_i adult). The number of alleles \mathbf{A}_i in the next generation can be found easily:

$$n_i(t+1) = v_i f_i n_i(t) \equiv w_i n_i(t), \quad (1)$$

where $w_i \equiv v_i f_i$ is the *fitness* of \mathbf{A}_i . Here, fitness gives the average number of surviving offspring. Note that if either $v_i = 0$ (that is allele \mathbf{A}_i is lethal) or $f_i = 0$ (that is \mathbf{A}_i is sterile), fitness $w_i = 0$.

Instead of absolute numbers n_i we will mainly use frequencies. The frequency of allele \mathbf{A}_i is

$$p_i(t) = n_i(t)/N(t). \quad (2)$$

Note that $\sum_i p_i = 1$.

Notation: we will use prime to denote the values of different variables in the next generation. For example: p'_i, n'_i, N' .

It is straightforward to write down the recurrent equations expressing the number of alleles \mathbf{A}_i and the population size N in the next generation as a function of their present values and parameters:

$$n'_i = w_i n_i, \quad (3a)$$

$$N' = \sum_i n'_i = \sum_i w_i n_i = \sum_i w_i \frac{n_i}{N} N = \left(\sum_i w_i p_i \right) N = \bar{w} N, \quad (3b)$$

where

$$\bar{w} = \sum_i w_i p_i, \quad (4)$$

is the *mean fitness* of the population. By dividing equation (3a) by (3b) one finds the recurrent equation for allele frequencies:

$$p'_i = n'_i/N' = \frac{w_i n_i}{\bar{w} N} = \frac{w_i}{\bar{w}} p_i. \quad (5)$$

The dynamics of the system under consideration can be completely described in terms of the overall population size and allele frequencies:

$$N' = \bar{w} N, \tag{6a}$$

$$p'_i = \frac{w_i}{\bar{w}} p_i. \tag{6b}$$

In general, fitness can depend on a number of factors (e.g. time, population size, population structure etc): $w_i = w_i(t, N, p)$, where p denotes the whole set of allele frequencies, $p = \{p_1, \dots, p_k\}$. Let us consider some special cases.

- *No selection* (= no genetic differences with respect to fitness). Note that there may be other genetic differences not considered here. If $w_i = w(N, t, p)$ for all i , then $\bar{w} = w(N, t, p)$ and equation (6) can be rewritten as

$$N' = w(N, t) N, \\ p'_i = p_i.$$

The second equation tells us that allele frequencies do not change (there is no evolution). The first equation does not depend on the genetic structure of the population. It represents a standard population ecology model for the changes in the size of an isolated population. For example, if $w = \text{const}$, one has a model of an exponential growth $N' = wN$, whereas if $w = a + bN$ where a and b are some constants, one has a model of a logistic growth $N' = (a + bN)N$.

- *No density dependence*. If fitness w does not depend on N , that is if $w_i = w(t, p)$, then the dynamics of allele frequencies are completely independent of N . Thus, one can analyze equation (6b) *independently* of equation (6a). Note that, in general, the dynamics of the population size will still depend on the genetic structure of the population: $N' = w(t, p)N$. In general, with no density dependence the population either explodes ($N \rightarrow \infty$ which is unrealistic) or goes extinct ($N \rightarrow 0$).
- *Population size, N , is regulated independently of the population structure, p* . Specifically, we assume that fitness w can be factorized into two terms: $w_i = w_b(N)w_{g,i}(t, p)$ one of which depends on the population size and the other depends on the genotype. Population regulation independent of the genetic structure may happen if, for example, each generation produces many offspring but there is only a limited amount of resource, or if selection and population regulation act at different stages of the life cycle (as in many butterflies where the population size is usually regulated at the larval stage whereas selection acts on adults). Under the above assumption about fitnesses the ratio

$$\frac{w_i}{\bar{w}} = \frac{w_b(N)w_{g,i}(t, p)}{\sum_j w_b(N)w_{g,j}(t, p)} = \frac{w_{g,i}(t, p)}{\sum_j w_{g,j}(t, p)}$$

does not depend on N . Thus, one can analyze equation (6b) *independently* of equation (6a).

Below in analyzing the dynamic equations for allele frequencies we will always imply that the population size is regulated independently of the genetic structure of the population in a way similar to the one just discussed.

Sometimes instead of using equations of type (6b), it is more convenient to consider changes between two subsequent generations:

$$\Delta p_i \equiv p'_i - p_i = \frac{w_i - \bar{w}}{\bar{w}} p_i. \tag{7}$$

Both equation (6b) and equation (7) tell us that the frequency of an allele increases or decreases ($\Delta p_i > 0$ or $\Delta p_i < 0$) depending on whether its fitness is larger or smaller than the average fitness ($w_i > \bar{w}$ or $w_i < \bar{w}$). Note also that multiplying all fitnesses by a constant does not change the ratio w_i/\bar{w} . Thus, one can normalize fitnesses.

2.1.1 Major points:

- population size regulation independent of the genetic structure is usually implied in population genetic models,

2.1.2 Constant fitnesses

Here, we assume that fitnesses are constant. Let \mathbf{A}_1 be the fittest allele in the population. That is $w_1 > w_i$ for all $i = 2, \dots, k$. Then using equation (6b)

$$\frac{p'_i}{p'_1} = \left(\frac{w_i}{\bar{w}} p_i \right) / \left(\frac{w_1}{\bar{w}} p_1 \right) = \frac{w_i}{w_1} \frac{p_i}{p_1} < \frac{p_i}{p_1}.$$

Thus, p_i/p_1 decreases in time for all $i = 2, \dots, k$. As time increases ($t \rightarrow \infty$), the frequencies of all alleles but the fittest approach zero ($p_i \rightarrow 0$), and the fittest allele becomes fixed ($p_1 \rightarrow 1$). Genetic variability will not be maintained in the population. One can say that *only the fittest survives*.

How long does it take for the fittest allele to take over? Let us consider two alleles: \mathbf{A}_i and \mathbf{A}_j . Then

$$p'_i/p'_j = (w_i/w_j)(p_i/p_j)$$

Let $r_t = p_i/p_j$ be the ratio of allele frequencies at the t -th generation. Then the above equation can be rewritten as

$$r_{t+1} = \frac{w_i}{w_j} r_t = \frac{w_i}{w_j} \frac{w_i}{w_j} r_{t-1} = \dots = \left(\frac{w_i}{w_j} \right)^t r_0.$$

For example, if initially the alleles are at equal frequencies ($r_0 = 1$) and the fitness of allele \mathbf{A}_i exceeds that of allele \mathbf{A}_j by 5% ($w_i/w_j = 1.05$), then in the 50-th generation the ratio of the frequencies will be 11.47, and will increase to 131.5 by generation 100.

Note that the following *linear* (in time t) relation is true:

$$\ln(r_t) = \ln(r_0) + t \ln\left(\frac{w_i}{w_j}\right).$$

This relationship can be (and actually has been) tested in experiments using linear regression (pp.215-216 in the Hartl-Clark book).

2.1.3 Major points:

In an asexual haploid population with constant fitnesses:

- genetic variation is not maintained,
- only the fittest allele survives,
- the mean fitness of the population increases monotonically at a rate proportional to the variance in fitness,
- the time scale for the fittest allele to take over is order $1/s$, where s is the difference in fitness between the fittest and the second fittest alleles.

ADD: graphs of allele frequencies changing in time (*Maple*)

Homework Problem: The change in the mean fitness and the time scale.

1. Show that the change in the mean fitness per generation, $\Delta\bar{w} = \bar{w}' - \bar{w}$, can be represented as

$$\Delta\bar{w} = \frac{V_g}{\bar{w}}, \tag{8}$$

where the genetic variance in fitness

$$V_g = \sum_i (w_i - \bar{w})^2 p_i.$$

What does eq.(8) tell us about the dynamics?

2. Show that the time for the fittest allele (with fitness w_1) to take over has order $1/(w_1 - w_2)$, where w_2 is the fitness of the second fittest allele. Hint: normalize fitnesses so that $w_2 = 1, w_1 = 1 + s$.

3 Selection in a single locus

We consider a (deterministic) model of a biological population where individuals are different with respect to a single locus with two alleles which we will denote **A** and **a**. Because in diploid organisms each individual has two genes at each locus (coming from maternal and paternal genotypes), there are three different genotypes: two *homozygotes* **AA** and **aa** and a heterozygote **Aa**. (We assume that heterozygotes **Aa** and **aA** are equivalent). Let x, y and z be the frequencies of genotypes **AA**, **Aa** and **aa** in the population ($x + y + z = 1$). The corresponding frequencies of alleles **A** and **a** are

$$\begin{aligned}p &= x + y/2, \\q &= z + y/2\end{aligned}$$

($p + q = 1$).

3.1 Random mating

We assume that reproduction is sexual, that individuals mate *randomly* with respect to the gene under consideration, and that generations are discrete and non-overlapping. We will consider two breeding schemes: (i) random union of gametes and (ii) random mating of individuals.

3.1.1 Random union of gametes

This method of modeling is applicable to organisms that release their gametes (sperm and eggs) into water, and to wind pollinated plants.

The probability that a sperm or egg carries **A** is p and the probability that a sperm or egg carries **a** is q . With random union of gametes the frequencies of genotypes **AA**, **Aa** and **aa** among offspring will be

$$\begin{aligned}x' &= p^2, \\y' &= pq + qp = 2pq, \\z' &= q^2,\end{aligned}$$

respectively. The frequencies of allele **A** and **a** among offsprings will be

$$\begin{aligned}p' &= x' + y'/2 = p^2 + pq = p, \\q' &= z' + y'/2 = q^2 + pq = q.\end{aligned}$$

Thus, allele frequencies do not change whereas genotype frequencies are in *Hardy-Weinberg proportions* in one generation:

$$\mathbf{AA} : \mathbf{Aa} : \mathbf{aa} = p^2 : 2pq : q^2.$$

Note that the population state can be described by a single variable, say p .

3.1.2 Random mating of individuals

Here we assume that individuals form mating pairs. If mating pairs are formed randomly then the proportion of mating pairs of type **AA** \times **AA** will be x^2 , that of type **AA** \times **aa** will be xz and so on. The table below shows the frequencies of all matings, and the frequencies of offspring produced from each mating.

Table 2: Mating types and offspring

| Mating types | | Frequency of mating | Offspring frequency | | |
|--------------|-----------|---------------------|---------------------|-----------|-----------|
| Female | Male | | AA | Aa | aa |
| AA | AA | x^2 | 1 | 0 | 0 |
| | Aa | xy | 1/2 | 1/2 | 0 |
| | aa | xz | 0 | 1 | 0 |
| Aa | AA | xy | 1/2 | 1/2 | 0 |
| | Aa | y^2 | 1/4 | 1/2 | 1/4 |
| | aa | yz | 0 | 1/2 | 1/2 |
| aa | AA | xz | 0 | 1 | 0 |
| | Aa | yz | 0 | 1/2 | 1/2 |
| | aa | z^2 | 0 | 0 | 1 |

To find the frequency of a genotype among the offspring one sums up the products of the frequency of mating with the corresponding offspring frequency:

$$\begin{aligned} x' &= x^2 \times 1 + 2xy \times 1/2 + y^2 \times 1/4 = (x + y/2)^2 = p^2, \\ y' &= 2xy \times 1/2 + 2xz \times 1/2 + y^2 \times 1/2 + 2yz \times 1/2 = 2(x + y/2)(z + y/2) = 2pq, \\ z' &= \dots = q^2. \end{aligned}$$

Thus, the genotype frequencies are in Hardy-Weinberg proportions after a single generation of random mating. It is easy to see that allele frequencies will not change.

Note that if an allele, say **a**, is rare, i.e., $q \ll 1$, then

$$y = 2pq = 2q(1 - q) \approx 2q.$$

That is, the frequency of heterozygotes is twice the frequency of the rare allele. This property can be used to estimate the frequency of rare alleles. Also, if the population is in Hardy-Weinberg proportions, then the following equality must be true: $y^2 = 4xz$. This property can be used to check for Hardy-Weinberg proportions.

3.1.3 More alleles

Let there be k different alleles: A_1, A_2, \dots, A_k . Now there are $k(k + 1)/2$ possible genotypes. Let $P_{ij} = \text{freq}(A_i A_j)$ be the genotype frequencies and $p_i = \sum_j P_{ij}$ be the allele frequencies. One can show that with random mating, allele frequencies do not change

$$p'_i = p_i,$$

and that Hardy-Weinberg proportions

$$P'_{ij} = 2p_i p_j$$

are attained in a single generation.

Note that if generations are overlapping or there is selfing, then Hardy-Weinberg proportions are attained gradually.

3.2 Viability selection

We consider a one-locus two-allele population with alleles **A** and **a**. There are three possible genotypes **AA**, **Aa** and **aa**. Let p and q and x, y and z be the corresponding allele and genotype frequencies. Immediately after reproduction, the genotype frequencies among offspring are in Hardy-Weinberg proportions: $x = p^2, y = 2pq, z = q^2$. We assume that genotypes are different with respect to *viability* defined as the probabilities of survival to the age of reproduction. Let w_{AA}, w_{Aa} and w_{aa} be the corresponding viabilities. Because viability selection is the only selection type incorporated in the model, we will also call these values *fitnesses*. The genotype frequencies after (viability) selection are

$$\begin{aligned}x' &= cw_{AA}x, \\y' &= cw_{Aa}y, \\z' &= cw_{aa}z,\end{aligned}$$

where c is a normalizing coefficient necessary to satisfy the condition $x' + y' + z' = 1$. This coefficient is $1/\bar{w}$ where

$$\bar{w} = w_{AA}x + w_{Aa}y + w_{aa}z = w_{AA}p^2 + w_{Aa}2pq + w_{aa}q^2,$$

is the mean fitness of the population. Let

$$\begin{aligned}w_A &= w_{AA}p + w_{Aa}q, \\w_a &= w_{Aa}p + w_{aa}q,\end{aligned}$$

be the *induced* fitnesses of alleles **A** and **a**. Note that the mean fitness of the population can be represented as $\bar{w} = w_A p + w_a q$. The allele frequency after selection is

$$p' = x' + y'/2 = \frac{w_{AA}x + w_{Aa}y/2}{\bar{w}} = \frac{w_{AA}p^2 + w_{Aa}pq}{\bar{w}} = \frac{w_A}{\bar{w}} p,$$

Because reproduction does not change the allele frequencies, the allele frequencies in the offspring will be p' as well. Thus, we have shown that under viability selection the dynamics of allele frequency are described by a recurrence equation

$$p' = \frac{w_A}{\bar{w}} p. \quad (9)$$

We can also consider the change in allele frequency between two subsequent generations, $\Delta p = p' - p$. One can show that the following equalities are true

$$\Delta p = \frac{w_A - \bar{w}}{\bar{w}} p \quad (10a)$$

$$= \frac{pq(w_A - w_a)}{\bar{w}} \quad (10b)$$

$$= pq \frac{w_{Aa} - w_{aa} + (w_{AA} - 2w_{Aa} + w_{aa})p}{\bar{w}} \quad (10c)$$

$$= \frac{pq}{2} \frac{d \ln \bar{w}}{dp} \quad (10d)$$

Note that equation (10d) describes a gradient-type dynamics. This equation tells us that allele frequency changes in such a way that the mean fitness always increases.

To obtain more detailed information about the allele frequency dynamics we will consider several cases. Let $w_{Aa} - w_{aa} + (w_{AA} - 2w_{Aa} + w_{aa})p \equiv g(p)$ so that equation (10c) can be rewritten as

$$\Delta p = \frac{p(1-p)g(p)}{\bar{w}}.$$

Note that $g(p)$ is a *linear* function of p .

1. $w_{aa} \leq w_{Aa} \leq w_{AA}$ (**A** is advantageous). Then

$$g(0) > 0, g(1) > 0 \Rightarrow g(p) > 0 \text{ for all } p \Rightarrow \Delta p > 0 \Rightarrow p \rightarrow 1,$$

that is allele **A** gets fixed.

2. $w_{aa} \geq w_{Aa} \geq w_{AA}$ (**a** is advantageous). Then

$$g(0) < 0, g(1) < 0 \Rightarrow g(p) < 0 \text{ for all } p \Rightarrow \Delta p < 0 \Rightarrow p \rightarrow 0,$$

that is allele **a** gets fixed.

3. $w_{aa} \geq w_{Aa} \leq w_{AA}$ (underdominance). Let

$$\hat{p} = \frac{w_{aa} - w_{Aa}}{w_{AA} - 2w_{Aa} + w_{aa}}$$

$$\Delta p = \frac{pq}{\bar{w}}(w_{AA} - 2w_{Aa} + w_{aa})(p - \hat{p}).$$

Then, if $p > \hat{p}$, then $\Delta p > 0$ and $p \rightarrow 1$. If $p < \hat{p}$, then $\Delta p < 0$ and $p \rightarrow 0$. Thus, the system evolves to a monomorphic state. Which allele is fixed depends on the initial conditions. Note that the state with $p = \hat{p}$ is an unstable equilibrium.

3. $w_{aa} \leq w_{Aa} \geq w_{AA}$ (overdominance). It is easy to see that

$$p \rightarrow \hat{p}$$

for all initial conditions. Thus, with overdominance genetic variability is maintained. Here, overdominance is actually both sufficient and necessary condition for the maintenance of genetic variation.

3.2.1 Major points

In randomly mating 1-locus 2-allele diploid populations under constant viability selection

- mean fitnesses does not decrease;
- the population always evolves to an equilibrium;
- overdominance is both necessary and sufficient for the maintenance of genetic variation
- with underdominance, both monomorphic equilibria are stable simultaneously.

4 Technical section

4.1 Differential approximation

Usually it is much easier to formulate a model using discrete time (and non-overlapping generations). But it is much easier to analyze a model in continuous time. Here we consider how one can justify using a differential approximation (and differential equations) for models described by difference equations.

We start with a *difference* equation

$$\Delta p_i = \frac{w_i - \bar{w}}{\bar{w}} p_i$$

describing the change in an allele frequency between two generations as a result of selection specified by fitnesses w_i . Let

$$w_i = 1 + \epsilon s_i,$$

where ϵ is a small positive parameter ($\ll 1$) and s_i are some coefficients. Note that the difference in fitness between any two alleles $w_i - w_j = \epsilon(s_i - s_j)$ and is order ϵ (i.e. very small). This is a **weak selection approximation** which is broadly used in modeling evolutionary processes. It is justified by a lot of experimental data showing that the strength of selection acting on individual alleles is generally small. [In general, we say that there is selection if individuals have different fitnesses.] Under the above approximation

$$\begin{aligned}\bar{w} &= \sum w_i p_i = \sum (p_i + \epsilon s_i p_i) = 1 + \epsilon \sum s_i p_i = 1 + \epsilon \bar{s} \\ w_i - \bar{w} &= \epsilon(s_i - \bar{s}).\end{aligned}$$

Thus, our basic equation can be rewritten as

$$\Delta p_i = \frac{\epsilon(s_i - \bar{s})}{1 + \epsilon \bar{s}} p_i \approx \epsilon(s_i - \bar{s}) p_i.$$

Finally, if we approximate a difference by a differential: $\frac{\Delta p_i}{1} \approx \frac{dp_i}{dt}$, and change time to $\tau = \epsilon t$, we end up with a *differential* equation

$$\frac{dp_i}{d\tau} = (s_i - \bar{s}) p_i.$$

One could get a similar equation by constructing a model for a population with overlapping generations.

4.1.1 Major points:

- a differential approximation of difference equations can be justified by invoking a weak selection assumption.

4.2 Stability analysis of systems of ordinary differential equations (ODE)

Exact solutions of even simple models are often unavailable. In such situations one has to use different *approximate* methods of which the analysis of stability of equilibria is very useful.

Let us consider a system of two ODE:

$$\dot{x} = f(x, y), \tag{11a}$$

$$\dot{y} = g(x, y). \tag{11b}$$

Definition. A point (x_0, y_0) is an equilibrium of (11) if $f(x_0, y_0) = g(x_0, y_0) = 0$. At equilibrium both $\dot{x} = \dot{y} = 0$ (no changes).

Existence of equilibria. To find equilibria, one has to solve a system of algebraic equations $f(x, y) = 0, g(x, y) = 0$.

4.2.1 Stability of equilibria

Linear systems. Before approaching the *nonlinear* system (11) let us first consider a system of two *linear* ODE:

$$\dot{x} = ax + by, \tag{12a}$$

$$\dot{y} = cx + dy. \tag{12b}$$

The general solution to (12) depends on the *eigenvalues* λ_1, λ_2 of stability matrix (Jacobian)

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

These eigenvalues satisfy to a *characteristic equation*

$$\lambda^2 - \lambda \operatorname{Tr}(S) + \operatorname{Det}(S) = 0, \tag{13}$$

where the *trace*, $\operatorname{Tr}(S)$, and the *determinant*, $\operatorname{Det}(S)$, of matrix S are

$$\operatorname{Tr}(S) = a + d, \operatorname{Det}(S) = ad - bc.$$

The solutions of the characteristic equation (13) are

$$\lambda = \frac{\operatorname{Tr}(S)}{2} \pm \sqrt{\left(\frac{\operatorname{Tr}(S)}{2}\right)^2 - \operatorname{Det}(S)}.$$

Note that $\operatorname{Tr}(S) = \lambda_1 + \lambda_2$, $\operatorname{Det}(S) = \lambda_1 \lambda_2$. If $\operatorname{Det}(S) \neq 0$, the point $(0, 0)$ is the only equilibrium of (12).

There are three different cases to consider

- (i) λ_1 and λ_2 are real and different. A general solution of (12) is

$$\begin{aligned} x(t) &= C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}, \\ y(t) &= C_3 e^{\lambda_1 t} + C_4 e^{\lambda_2 t}, \end{aligned}$$

where C_1, C_2, C_3, C_4 are coefficients that depend on initial conditions.

- (ii) λ_1 and λ_2 are real and equal ($\lambda_1 = \lambda_2 = \lambda$) A general solution of (12) is

$$\begin{aligned} x(t) &= C_1 e^{\lambda t} + C_2 t e^{\lambda t}, \\ y(t) &= C_3 e^{\lambda t} + C_4 t e^{\lambda t}. \end{aligned}$$

- (iii) λ_1 and λ_2 are complex ($\lambda = \alpha \mp i\beta$ where $i = \sqrt{-1}$). A general solution of (12) is

$$\begin{aligned}x(t) &= C_1 e^{\alpha t} \cos(\beta t) + C_2 e^{\alpha t} \sin(\beta t), \\y(t) &= C_3 e^{\alpha t} \cos(\beta t) + C_4 e^{\alpha t} \sin(\beta t).\end{aligned}$$

The intuitive *concept of stability*: the equilibrium is stable if the system returns to the equilibrium when slightly perturbed.

Case (i). Let $\lambda_1, \lambda_2 < 0$. Then both $x(t) \rightarrow 0, y(t) \rightarrow 0$ as $t \rightarrow \infty$. Thus, $(0, 0)$ is stable. Let $\lambda_1 > 0$ and/or $\lambda_2 > 0$. Then in general $|x(t)| \rightarrow \infty$ and/or $|y(t)| \rightarrow \infty$. Thus, $(0, 0)$ is unstable.

Case (ii). Let $\lambda < 0$. Then both $x(t) \rightarrow 0, y(t) \rightarrow 0$ as $t \rightarrow \infty$. Thus, $(0, 0)$ is stable. Let $\lambda > 0$. Then in general $|x(t)| \rightarrow \infty$ and $|y(t)| \rightarrow \infty$. Thus, $(0, 0)$ is unstable.

Case (iii). Let $\alpha < 0$. Then both $x(t) \rightarrow 0, y(t) \rightarrow 0$ as $t \rightarrow \infty$. Thus, $(0, 0)$ is stable. Let $\alpha > 0$. Then in general $|x(t)| \rightarrow \infty$ and $|y(t)| \rightarrow \infty$. Thus, $(0, 0)$ is unstable.

Summarizing: The equilibrium $(0, 0)$ of (12) is stable if both $Re(\lambda_1) < 0$ and $Re(\lambda_2) < 0$ and is unstable otherwise ($Re(\lambda)$ is the real part of λ).

Note that both eigenvalues of S have negative real parts if

$$Tr(S) < 0 \text{ and } Det(S) > 0. \tag{14}$$

If either of these inequalities is violated, $(0, 0)$ is unstable.

Nonlinear systems. Now let us return to the general system of two nonlinear ODE:

$$\dot{x} = f(x, y), \tag{15a}$$

$$\dot{y} = g(x, y). \tag{15b}$$

Let us consider an equilibrium (x_0, y_0) . One can expand functions f and g in a Taylor series in a neighborhood of this equilibrium:

$$f(x, y) \approx f(x_0, y_0) + \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + \text{smaller terms},$$

$$g(x, y) \approx g(x_0, y_0) + \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} + \text{smaller terms},$$

where all derivatives are evaluated at $x = x_0, y = y_0$.

Let $u = x - x_0$ and $v = y - y_0$ be the deviations of x and y from x_0 and y_0 , respectively. Making the variable change and neglecting second and higher order terms results in the nonlinear system of ODE (15) being approximated by a linear system of ODE

$$\dot{u} = \frac{\partial f}{\partial x} u + \frac{\partial f}{\partial y} v, \tag{16a}$$

$$\dot{v} = \frac{\partial g}{\partial x} u + \frac{\partial g}{\partial y} v \tag{16b}$$

where all derivatives are evaluated at $x = x_0, y = y_0$. The corresponding stability matrix (Jacobian) is

$$S = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix}.$$

where all derivatives are evaluated at $x = x_0, y = y_0$. Now we can use this matrix to deduce the behavior of the nonlinear system in a neighborhood of $x = x_0, y = y_0$.

General procedure:

- find an equilibrium
- compute the stability matrix (Jacobian) S at this equilibrium
- make conclusions about the eigenvalues of S
- make conclusions about conditions for stability/instability of the equilibrium

Table 3: Classification of equilibria

| eigenvalues | equilibrium |
|---|----------------|
| $\lambda_1 > 0, \lambda_2 < 0$ (or $\lambda_1 < 0, \lambda_2 > 0$) | saddle |
| $\lambda_1 < 0, \lambda_2 < 0$ | stable node |
| $\lambda_1 > 0, \lambda_2 > 0$ | unstable node |
| $\lambda = \alpha \mp \beta i, Re(\lambda) < 0$ | stable focus |
| $\lambda = \alpha \mp \beta i, Re(\lambda) > 0$ | unstable focus |

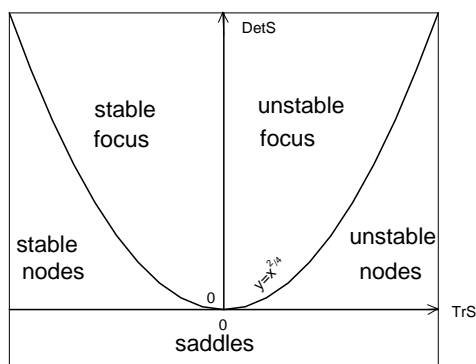


Figure 1: Conditions for stability given in terms of $Tr(S)$ and $Det(S)$.

4.2.2 Example

Let us consider a system of two ordinary differential equations

$$\begin{aligned}\dot{x} &= x(a - bx) - cxy, \\ \dot{y} &= dxy - (e + fy)y,\end{aligned}$$

where all coefficients are positive.

One can interpret x as the density of a prey population which grows logistically in the absence of predator (if $y = 0$, then $\dot{x} = x(a - bx)$), and y as the density of predator which experiences density-dependent mortality (with rate $-(e + fy)y$). Terms $-cxy$ and dxy stand for the rate at which predation decreases the prey population and increases the predator population, respectively.

The system can have up to three equilibria:

1. $x = 0, y = 0$. At this equilibrium, which exists always, both populations are extinct.
2. $x = a/b, y = 0$. At this equilibrium, which exists always, there is no predator whereas prey is at the carrying capacity.
3. $x = x^* \equiv (af + ce)/(bf + cd), y = y^* \equiv (ad - be)/(bf + cd)$. At this equilibrium, both populations are present. The equilibrium exists (that is, biologically meaningful) if $ad > be$.

Conditions for stability are as follows. The first equilibrium is never stable. (This is a saddle point with eigenvalues a and $-e$.) The second equilibrium is stable if $ad < be$ and is unstable otherwise. (The eigenvalues are $-a$ and $(ad - be)/b$ so that the equilibrium is a stable node if $ad < be$ and is a saddle point if $ad > be$.) The third equilibrium is stable if $ad > be$ that is whenever it exists. (The sign of the determinant of the stability matrix coincides with that of the difference $ad - be$, and if $ad > be$ then the trace is negative.)

Conditions for existence and stability of equilibria of this system can be summarized in a table:

| Equilibrium (x, y) | Conditions for existence | Conditions for stability |
|----------------------|--------------------------|--------------------------|
| $(0, 0)$ | always | never |
| $(a/b, 0)$ | always | $ad < be$ |
| (x^*, y^*) | $ad > be$ | $ad > be$ |

Thus, the general qualitative behavior is as follows: if $ad < be$, the predator dies out; if $ad > be$ both species coexist. (Note that conditions for stability do not depend on c and f . Other parameters affect conditions for stability in a specific way.)

4.3 Stability of equilibria of difference equations

Let us consider a *linear* difference equation

$$x_{t+1} = ax_t. \tag{17}$$

The general solution of (17) is $x_t = a^t x_0$ where x_0 is the initial value of x . Thus, asymptotically (i.e., as $t \rightarrow \infty$) $|x_t| \rightarrow 0$ if $|a| < 1$, and $|x_t| \rightarrow \infty$ if $|a| > 1$. The only equilibrium point of (17) is $x^* = 0$. If $|a| < 1$, this equilibrium is stable (after a perturbation x will return back to $x^* = 0$). If $|a| > 1$, this equilibrium is stable (after a perturbation, x will not return back to $x^* = 0$).

Next we consider a single *non-linear* difference equation:

$$x_{t+1} = f(x_t).$$

Let x^* be an equilibrium of this equation (meaning that $x^* = f(x^*)$). To determine local stability properties of x^* one approximates f in a small neighborhood of x^* by a linear function:

$$x_{t+1} \approx f(x^*) + \frac{df(x)}{dx} (x_t - x^*)$$

The resulting equation can be rewritten as

$$x_{t+1} - x^* = \frac{df(x)}{dx} (x_t - x^*)$$

or, by introducing $y_t = x_t - x^*$ (which is the deviation of x_t from the equilibrium x^*) as

$$y_{t+1} = \frac{df(x)}{dx} y_t.$$

The last equation has the same form as equation (17). Thus, x^* is stable to small perturbations, if $|df(x)/dx| < 1$ at x^* and is unstable otherwise.

In a similar way one can consider a system of difference equations. For example, let us consider a system of two difference equations

$$\begin{aligned} x_{t+1} &= f(x_t, y_t), \\ y_{t+1} &= g(x_t, y_t). \end{aligned}$$

Assume that this system has an equilibrium (x^*, y^*) :

$$x^* = f(x^*, y^*), y^* = g(x^*, y^*).$$

Consider the stability matrix (the Jacobian)

$$S = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix}$$

evaluated at the equilibrium (x^*, y^*) . One can show that this equilibrium is locally stable if the absolute values of all eigenvalues λ_i of the stability matrix S are smaller than 1: $|\lambda_i| < 1$. If at least for one eigenvalue $|\lambda_i| > 1$, the equilibrium is unstable.

5 Selection in a single locus (cont.)

5.1 Multiple alleles

We consider a diploid populations with k alleles A_1, A_2, \dots, A_k segregating at a locus of interest. Genotypes of individuals will be denoted as $A_i A_j$. Let p_1, p_2, \dots, p_k be the allele frequencies ($\sum p_i = 1$) and P_{ij} be the genotype frequencies which are in Hardy-Weinberg proportions in one generation: $P_{ij} = 2p_i p_j (i \neq j)$, $P_{ii} = p_i^2$. Let w_{ij} be fitnesses (viabilities) ($w_{ij} = w_{ji}$). We also define the induced fitness of allele A_i as

$$w_i = \sum_j w_{ij} p_j,$$

and the mean fitness of the population:

$$\bar{w} = \sum w_i p_i = \sum_{ij} w_{ij} p_i p_j.$$

After selection, genotype frequencies are

$$P'_{ij} = \frac{w_{ij}}{\bar{w}} P_{ij}.$$

Allele frequencies among surviving adults are

$$p'_i = \sum_j P'_{ij} = \frac{\sum_j w_{ij} P_{ij}}{\bar{w}} = \frac{\sum_j w_{ij} p_i p_j}{\bar{w}} = \frac{w_i p_i}{\bar{w}}. \quad (18)$$

Because reproduction does not change allele frequencies, p'_i is also the allele frequency in the next generation. The change in allele frequency between two generations is

$$\Delta p_i = \frac{w_i - \bar{w}}{\bar{w}} p_i. \quad (19)$$

The following summarizes some results known for the dynamical systems (18) and (19).

- There can be only one isolated polymorphic equilibrium with all $p_i^* > 0$. At this equilibrium $w_i = \bar{w}$ for all $i = 1, 2, \dots, k$. One can rewrite these conditions as $w_i = w_1$ for all $i = 2, 3, \dots, k$. Thus, there are $k - 1$ linear algebraic equations for $k - 1$ independent variables p_2^*, \dots, p_k^* . These equations can be solved using Cramer's rule. (Note that p_1^* can be found from the equality $p_1^* + p_2^* + \dots + p_k^* = 1$). For special values of w_{ij} there can be no admissible polymorphic solutions (with $0 < p_i^* < 1$), a single solution, or an infinite number of solutions.
- The overall number of possible equilibria of (18) and (19) is $2^k - 1$.
- A unique admissible polymorphic solution will be stable if and only if the matrix $W = \{w_{ij}\}$ has exactly one positive eigenvalue, and at least one negative eigenvalue (Kingman, 1961).
- If matrix W has j positive eigenvalues, at most $k - j + 1$ alleles will exist with positive frequencies at the equilibrium (Kingman, 1961).
- Three *necessary* (but not sufficient!) conditions for stability of the polymorphic equilibrium (Lewontin et al., 1978; Nagylaki, 1992):
 1. for each i there exists some j such that $w_{ii} < w_{ij}$,
 2. for all i, j , $w_{ij} > (w_{ii} + w_{jj})/2$,

3. “triangle inequality”: for all i, j there exists some k such that $w_{ij} < w_{ik} + w_{jk}$.

- Let D_i be the determinant of the $i \times i$ submatrix in the upper-left corner of W and $D_1 = w_{11} > 0$. Let $D_i \neq 0$ for all i . The completely polymorphic equilibrium is asymptotically stable if $(-1)^i D_i < 0$ for all i .
- If the completely polymorphic equilibrium is stable, no other equilibria are stable.
- In the two-allele case, overdominance is both necessary and sufficient for the maintenance of both alleles. In the k -allele case, pairwise overdominance ($w_{ij} > w_{ii}, w_{jj}$) is not necessary, and the total overdominance ($w_{ij} > w_{ii}$) is not sufficient for the maintenance of all k alleles (examples from Lewontin et al. 1978).
- Lewontin et al. (1978) performed large-scale numerical simulations using random fitness assignment and random initial conditions and measuring the proportion of runs resulting in stable polymorphic equilibria. They used 3 types of fitness assignment: totally random, random subject to pairwise heterosis (with $w_{ij} > w_{ii}, w_{jj}$ for all i, j) and random subject to total heterosis (with $w_{ij} > w_{kk}$ for all a, j, k). The following table summarizes their results:

Table. Proportion of runs resulting in stable polymorphic equilibria with k alleles present (100,000 runs in the case of totally random fitness assignment, 10,000 runs in case with pairwise and total heterosis).

| # alleles | totally random | pairwise heterosis | total heterosis |
|-----------|----------------|--------------------|-----------------|
| 2 | 0.33466 | 1.0000 | 1.0000 |
| 3 | 0.04237 | 0.5224 | 0.7120 |
| 4 | 0.00240 | 0.1259 | 0.3433 |
| 5 | 0.00006 | 0.0116 | 0.1041 |
| 6 | - | 0.0003 | 0.0137 |
| 7 | - | - | 0.0011 |
| 8 | - | - | - |

The overall conclusion of this analysis is that heterosis alone is not a mechanism for maintaining many alleles segregating at a locus.

The above results give an impression that details of the fitness matrix $\{W_{ij}\}$ are very important and that no generalizations are possible. However, some very general results with far-reaching interpretations do exist.

- The average fitness does not decrease

$$\bar{w}_{t+1} \geq \bar{w}_t.$$

$\Delta \bar{w} \equiv \bar{w}_{t+1} - \bar{w}_t = 0$ only at equilibrium. All equilibria are local maxima of \bar{w} . (Proof from Nagylaki, 1992.)

- Fisher’s fundamental theorem of natural selection (Fisher 1930)

$$\frac{d\bar{w}}{dt} = 2V_g,$$

where $V_g = \sum_i (w_i - \bar{w})^2 p_i$ is the additive genetic variance in fitness.

Sketch of the proof using a differential framework. We start with a system of differential equations

$$\dot{p}_i = (w_i - \bar{w})p_i, w_i = \sum_j w_{ij}p_j, \bar{w} = \sum_{ij} w_{ij}p_i p_j.$$

Then

$$\begin{aligned}
\frac{d\bar{w}}{dt} &= \sum_{ij} w_{ij}(\dot{p}_i p_j + p_i \dot{p}_j) \\
&= \sum_i \left(\sum_j w_{ij} p_j \right) \dot{p}_i + \sum_j \left(\sum_i w_{ij} p_i \right) \dot{p}_j \\
&= \sum_i w_i (w_i - \bar{w}) p_i + \sum_j w_j (w_j - \bar{w}) p_j \\
&= 2 \sum_i (w_i - \bar{w} + \bar{w})(w_i - \bar{w}) p_i = \sum_i (w_i - \bar{w})^2 p_i
\end{aligned}$$

- Svirezhev's optimality principle: natural selection operates in such a way that the path followed as a population changes from one state to another is the one that minimizes the total genetic variance over the path (Svirezhev 1972).

Svirezhev shows that the integral of the quantity

$$\sum_i \frac{\dot{p}_i^2}{p_i} + \sum_i p_i (w_i - \bar{w})^2$$

over the path from state 1 to state 2 \rightarrow min. This quantity is equivalent to the genetic variance in fitness V_g . Newton's laws can be derived from the principle of least action. Equations of gene-frequency and fitness change can be found from Svirezhev's optimality principle.

5.2 Proof that the mean fitness does not decrease

(Kingman 1961 given in Nagylaki, 1992, p.57-58)

We will need Jensen's inequality: if x takes values x_i with probabilities p_i and $\mu \geq 1$, then

$$\sum_i p_i x_i^\mu \geq \left(\sum_i p_i x_i \right)^\mu. \quad (20)$$

Proof:

$$\begin{aligned} \bar{w}' &= \sum_{ij} p'_i p'_j w_{ij} \\ &= \bar{w}^{-2} \sum_{ij} p_i w_i p_j w_j w_{ij} \end{aligned} \quad (21a)$$

$$= \bar{w}^{-2} \sum_{ijk} p_i p_j p_k w_{ij} w_{ik} \left(\frac{1}{2}\right) (w_j + w_k) \quad (21b)$$

$$\begin{aligned} &= \bar{w}^{-2} \sum_{ijk} p_i p_j p_k w_{ij} w_{ik} \left(\frac{1}{2}\right) (w_j + w_k) \\ &\geq \bar{w}^{-2} \sum_{ijk} p_i p_j p_k w_{ij} w_{ik} (w_j w_k)^{1/2} \end{aligned} \quad (21c)$$

$$\begin{aligned} &= \bar{w}^{-2} \sum_i p_i \left[\sum_j p_j w_{ij} (w_j)^{1/2} \right]^2 \\ &\geq \bar{w}^{-2} \left[\sum_i p_i \sum_j p_j w_{ij} (w_j)^{1/2} \right]^2 \end{aligned} \quad (21d)$$

$$= \bar{w}^{-2} \left(\sum_j p_j w_j^{3/2} \right)^2 \quad (21e)$$

$$\geq \bar{w}^{-2} \left[\left(\sum_j p_j w_j \right)^{3/2} \right]^2 \quad (21f)$$

$$= \bar{w}$$

The numbered equations in the above series come, respectively, from $p'_i = w_i p_i / \bar{w}$, $w_i = \sum_k p_k w_{ik}$, the elementary fact that $a + b \geq 2(ab)^{1/2}$, Jensen's inequality with $\mu = 2$, $w_j = \sum_i p_i w_{ij}$, and Jensen's inequality with $\mu = 3/2$.

5.3 Fertility selection

There are many different fitness components (viability, fertility, the ability to find a mate, the ability to complete fertilization etc). So far, we studied only one of many different fitness components - viability. Today, we will look at another important fitness component - fertility.

Let us consider a one-locus two-allele diploid population where individuals are different with respect to *fertility* rather than viability. We define fertility as the average number of offspring. In the most general case, fertility is a property of the mating pair. Therefore one needs to specify a 3×3 fertility matrix

| | | | |
|-----------|---------------|-----------|-----------|
| female | male genotype | | |
| genotype | AA | Aa | aa |
| AA | f_{11} | f_{12} | f_{13} |
| Aa | f_{21} | f_{22} | f_{23} |
| aa | f_{31} | f_{23} | f_{33} |

Let x, y, z be the frequencies of genotypes **AA**, **Aa** and **aa** in the population. To derive the dynamic equations for x, y, z one has to consider all possible mating times and resulting offspring.

Table 1. Mating types and offspring.

| Mating Types | | Frequency of mating | Fertility of mating | Offspring | | |
|--------------|-----------|---------------------|---------------------|-----------|-----------|-----------|
| Female | Male | | | AA | Aa | aa |
| AA | AA | x^2 | f_{11} | 1 | 0 | 0 |
| | Aa | xy | f_{12} | 1/2 | 1/2 | 0 |
| | aa | xz | f_{13} | 0 | 1 | 0 |
| Aa | AA | xy | f_{21} | 1/2 | 1/2 | 0 |
| | Aa | y^2 | f_{22} | 1/4 | 1/2 | 1/4 |
| | aa | yz | f_{23} | 0 | 1/2 | 1/2 |
| aa | AA | xz | f_{31} | 0 | 1 | 0 |
| | Aa | yz | f_{32} | 0 | 1/2 | 1/2 |
| | aa | z^2 | f_{33} | 0 | 0 | 1 |

Homework: derive general dynamic equations for x, y and z and find conditions for stability of the fixation equilibria: $x = 1, y = z = 0$ and $x = y = 0, z = 1$

In general, with fertility selection the population is not in Hardy-Weinberg proportions (thus, one needs to analyze genotype frequency dynamics rather than allele frequency dynamics), and the mean fitness (fertility) \bar{f} can decrease.

Three general fertility schemes have been studied in detail.

- Additive fertilities

$$f_{ij} = f_i + m_j,$$

that is the fertility of a pair is a sum of the fertilities of a male and a female. In this case, the dynamics are identical to that in a one-locus two-allele viability selection case with parameters $f_1 + m_1, f_2 + m_2$ and $f_3 + m_3$ playing the role of viabilities w_{AA}, w_{Aa} and w_{aa} . In particular, the population is in Hardy-Weinberg proportions in one generation, and the necessary and sufficient condition for the maintenance of genetic variation is overdominance ($f_2 + m_2 > f_1 + m_1, f_3 + m_3$).

- Multiplicative fertilities

$$f_{ij} = f_i m_j,$$

that is the fertility of a pair is a product of the fertilities of a male and a female. In this case the Hardy-Weinberg proportions are attained in one generation. If there is no sex differences ($f_i = m_i$), the model is equivalent to that of viability selection. However, if $f_i \neq m_i$, the dynamics are much more complicated. There can be three polymorphic equilibria of which two can be stable simultaneously. The monomorphic equilibria $x = 1$ and $z = 1$ are stable if $m_2/m_1 + f_2/f_1 \leq 2$ and $m_2/m_3 + f_2/f_3 \leq 2$, respectively.

Note that the case of multiplicative fertility selection is equivalent to the case of two-sex viability selection with fitnesses

| | AA | Aa | aa |
|---------|-------|-------|-------|
| females | f_1 | f_2 | f_3 |
| males | m_1 | m_2 | m_3 |

- Symmetric fertilities

$$\begin{pmatrix} \alpha & \beta & \gamma \\ \beta & \delta & \beta \\ \gamma & \beta & \alpha \end{pmatrix} \quad (22)$$

In this case, the productivity of a mating depends on the degree of heterozygosity of mating pairs. The dynamic equations are

$$\phi x' = \alpha x^2 + \beta xy + 1/4 \delta y^2, \quad (23a)$$

$$\phi y' = \beta xy + \beta yz + 2\gamma xz + 1/2 \delta y^2, \quad (23b)$$

$$\phi z' = \alpha z^2 + \beta zy + 1/4 \delta y^2, \quad (23c)$$

where ϕ is the mean fertility. This model allows for up to three symmetric polymorphic equilibria (with $x^* = z^*$) and two asymmetric polymorphic equilibria (with $x^* \neq z^*$)!

Let us consider an extreme case of self-incompatibility when $\alpha = \beta = 0$. The dynamic equations become

$$\phi x' = 1/4 \delta y^2, \quad (24a)$$

$$\phi y' = 2\gamma xz + 1/2 \delta y^2, \quad (24b)$$

$$\phi z' = 1/4 \delta y^2. \quad (24c)$$

Thus, $x' = z'$ after a single generation. Note that this implies that no fixation is possible. Let $u = 2x/y$ and $\epsilon = \gamma/\delta$. Then the dynamic equation for u is

$$u' = F(u) \equiv \frac{1}{1 + \epsilon u^2}.$$

The consideration of the graph of $F(u)$ shows that there is a single equilibrium (describing a polymorphic population) at which the first derivative of F is negative. If $\epsilon = 4$, this equilibrium is at $u = 1/2$; the corresponding first derivative of F is equal to -1 (see the Figure). If $\epsilon < 4$, the graph of $F(u)$ crosses the line $y = u$ above the point $(1/2, 1/2)$. If $\epsilon > 4$, the graph of $F(u)$ crosses the line $y = u$ below the point $(1/2, 1/2)$. In the former case, the first derivative $F'(u) > -1$ at the equilibrium meaning the equilibrium is stable. In the latter case, $F'(u) < -1$ at the equilibrium meaning the equilibrium is unstable.

One can show that in this case, there exists a *globally stable cycle* with period two! To do this, one has to solve an algebraic equation $u = F(F(u))$. The relevant roots are $u_{\mp} = (1 \mp \sqrt{1 - 4/\epsilon})/2$ corresponding to cyclic solutions $u_-, u_+, u_-, u_+, \dots$ and $u_+, u_-, u_+, u_-, \dots$.

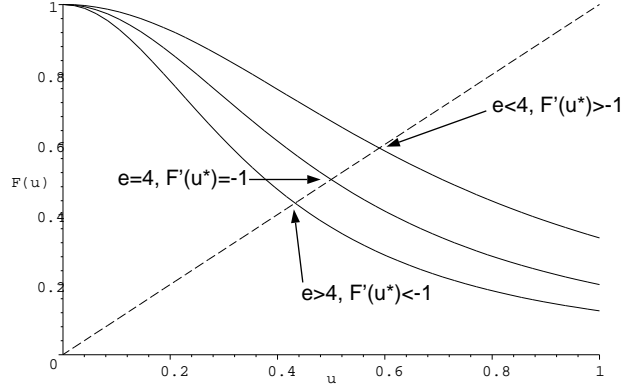


Figure 2: Function $F(u)$ for 3 different values of ϵ . The equilibria are at the intersection of $y = F(u)$ and $y = u$ (dashed line).

Note that one can interpret coefficients f_{ij} in the general fertility matrix as the probabilities of mating between genotypes i and j or as the viabilities of an offspring of mating between genotypes i and j . Thus, the framework (and results) developed for fertility selection can be directly applied to the cases of *sexual selection* and *assortative mating* as well as to the cases of *maternal* and *parental* selection.

Major points.

- In general, under fertility selection genotype frequencies are not in Hardy-Weinberg proportions, the mean fitness can decrease, there can be several simultaneously stable polymorphic equilibria, and cycling is possible.
- The additive fertility model and the multiplicative fertility model with no sex differences are mathematically equivalent to the model of constant viability selection.