

# POPULATION GENETICS

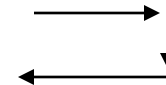
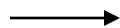
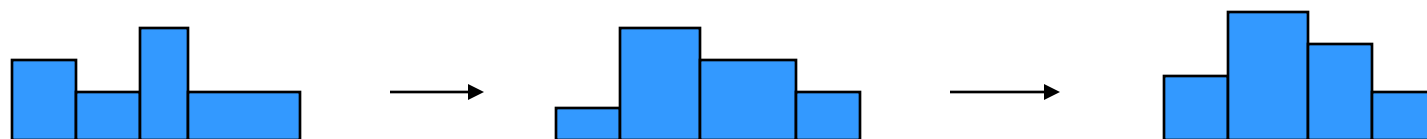
- JAN 2009

# POPULATION GENETICS

- The dynamics of genes in populations.
- Understanding the effects of population size, demography, mating system, viability and fertility differences, selection, recombination, epistasis, etc. on the genetic structure of populations.



# STOCHASTIC SYSTEMS.



# DYNAMICAL SYSTEMS.

- Suppose we are interested in a variable  $x$  through time,  $t$ . We write  $x(t)$  and express the change as a recurrence equation:- e.g.  
 $x(t+1) = f[x(t), x(t-1), \dots]$  ,  $t=0,1,\dots$
- Example:-  $x(t+1) = 2*x(t) + 6$ , then we have a linear system, and the solution of this is straightforward (see later) since  
 $(x(t+1) + 6) = 2 * (x(t) + 6)$ .

# DYNAMICAL SYSTEMS.

- $(x(t+1) + 6) = 2 * (x(t) + 6)$ , so  
 $(x(t) + 6) = 2^t * (x(0) + 6)$ .
- Thus if  $x(0) = 3$  then for  $t = 5$  we have  
 $(x(5) + 6) = 2^5 * (3 + 6) = 288$  so  
 $x(5) = 282$ .
- In fact  $x(.) = 3, 12, 30, 66, 138, 282$  while  
 $x(t)+6 = 9, 18, 36, 72, 144, 288$ .

# DYNAMICAL SYSTEMS.

- In general

if 
$$x(t+1) = \lambda * x(t) + b$$

(linear,  $\lambda$  not 1)

then 
$$(x(t+1) + c) = \lambda * (x(t) + c)$$

where  $c = b / (\lambda - 1)$  and  $\lambda$  is the rate (eigenvalue) at which the system grows.

# Dynamical Systems

- Linear System

$$x_1(t+1) = b_1 + a_{11}x_1(t) + \dots + a_{1n}x_n(t),$$

$$x_2(t+1) = b_2 + a_{21}x_1(t) + \dots + a_{2n}x_n(t),$$

.....

$$x_n(t+1) = b_n + a_{n1}x_1(t) + \dots + a_{nn}x_n(t),$$

which we write as

$$\mathbf{x}(t+1) = \mathbf{b} + \mathbf{A} \mathbf{x}(t) .$$



# DYNAMICAL SYSTEMS.

- Where  $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$   
and,

$$\mathbf{b} = (b_1, b_2, \dots, b_n)^T,$$

and  $\mathbf{A}$  is an  $n \times n$  matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

# DYNAMICAL SYSTEMS.

- Generally linear systems are

$$\mathbf{x}(t+1) = \mathbf{A} \mathbf{x}(t) + \mathbf{b}$$

(N.B. bold type indicates a vector/matrix)

where  $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$ ,  $\mathbf{A}$  is an  $n \times n$  matrix and  $\mathbf{b} = (b_1, b_2, \dots, b_n)^T$ .

# DYNAMICAL SYSTEMS

- Suppose we have (generically,  $\mathbf{b}$  eliminated)  $\mathbf{x}(t+1) = \mathbf{A} \mathbf{x}(t)$ .
- There are some special (column) vectors  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  and values  $\lambda_i$  such that  $\mathbf{A} \mathbf{u}_i = \lambda_i * \mathbf{u}_i$  for  $i=1, 2, \dots, n$ .
- Then if  $\mathbf{x}(t)=\mathbf{u}_i$  we have  $\mathbf{x}(t+1)=\mathbf{A}\mathbf{x}(t)=\mathbf{A}\mathbf{u}_i= \lambda_i\mathbf{u}_i$ .

# DYNAMICAL SYSTEMS

- Example.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad \text{then } \mathbf{u}_1 = (1, 1)^T \text{ with } \lambda_1 = 3$$

and  $\mathbf{u}_2 = (1, -1)^T$  with  $\lambda_2 = -1$ .

Now any vector  $\mathbf{x}(0)$  can be expressed in terms of  $\mathbf{u}_1$  and  $\mathbf{u}_2$  e.g.

$$\mathbf{x}(0) = (6, 2)^T = 4(1, 1)^T + 2(1, -1)^T$$

# DYNAMICAL SYSTEMS

- Thus if  $\mathbf{x}(0)=(6,2)^T$  then  $\mathbf{x}(1)=\dots$   
 $\mathbf{x}(1)=\mathbf{A}\mathbf{x}(0)=\mathbf{A}(4\mathbf{u}_1+2\mathbf{u}_2)=4\lambda_1\mathbf{u}_1+2\lambda_2\mathbf{u}_2=$   
 $12(1,1)^T-2(1,-1)^T=(10,14)^T,$   
which is what comes from the basic  
equation, and  
 $\mathbf{x}(2)=4\lambda_1^2\mathbf{u}_1+2\lambda_2^2\mathbf{u}_2=36\mathbf{u}_1+2\mathbf{u}_2$   
 $= (38,34)^T.$

# DYNAMICAL SYSTEMS

- Example.

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

so consider  $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$  where  $\mathbf{u} = (w_1, w_2)$  so

$$\lambda w_1 = w_1 + 2w_2 \text{ and } \lambda w_2 = 2w_1 + w_2$$

# DYNAMICAL SYSTEMS

- $\mathbf{u}=(w_1, w_2)$  so  $\lambda w_1=w_1+2w_2$  and  $\lambda w_2=2w_1+w_2$ ,  
 $(\lambda-1)w_1=2w_2$  and  $(\lambda-1)w_2=2w_1$  so  
 $(\lambda-1)/2=2/(\lambda-1)$  (equating  $w_2/w_1$ ) so  
 $(\lambda-1)^2=4$  and  $(\lambda-1)=+2$  or  $-2$ ,  $\lambda=3$  or  $-1$ ,  
and  $\mathbf{u}=(1,1)^T$  and  $\mathbf{u}=(1,-1)^T$  respectively.

Example (for class)

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$$



- $\lambda w_1 = 3w_1 + 2w_2$  &  $\lambda w_2 = 1w_1 + 4w_2$  so  
 $(\lambda - 3)w_1 = 2w_2$  &  $(\lambda - 4)w_2 = w_1$  so  $(\lambda - 3)(\lambda - 4) = 2$   
 $\lambda^2 - 7\lambda + 10 = 0 = (\lambda - 5)(\lambda - 2)$  so  $\lambda = 5$  &  $\lambda = 2$ .
- For  $\lambda = 5$   $2w_1 = 2w_2$  so  $w = (1, 1)$ .
- For  $\lambda = 2$   $-1w_1 = 2w_2$  so  $w = (2, -1)$

# Solution

- $\lambda w_1 = 3w_1 + 2w_2$  &  $\lambda w_2 = 1w_1 + 4w_2$  so  
 $(\lambda - 3)w_1 = 2w_2$  &  $(\lambda - 4)w_2 = w_1$  so  $(\lambda - 3)(\lambda - 4) = 2$   
 $\lambda^2 - 7\lambda + 10 = 0 = (\lambda - 5)(\lambda - 2)$  so  $\lambda = 5$  &  $\lambda = 2$ .
- For  $\lambda = 5$   $2w_1 = 2w_2$  so  $w = (1, 1)$ .
- For  $\lambda = 2$   $-1w_1 = 2w_2$  so  $w = (2, -1)$

# DYNAMICAL SYSTEMS

- NB In general, if

$$A = \begin{pmatrix} a & g \\ h & b \end{pmatrix} \quad \text{then the eigenvalues are the}$$

roots of  $\lambda^2 - (a+b)\lambda + (ab-gh)=0$ .

# Finding the Eigenvalues

- $\mathbf{A}\mathbf{u}=\lambda\mathbf{u}$  so  $au_1+gu_2=\lambda u_1$  &  $hu_1+bu_2=\lambda u_2$   
so  $(a-\lambda)u_1=-gu_2$  &  $hu_1=-(b-\lambda)u_2$ , and  
 $(a-\lambda)(b-\lambda)=gh$  (note this is  $|\mathbf{A}-\lambda\mathbf{I}|=0$  so  
quadratic for  $\lambda$ ).
- Example  $a=b=1$ ,  $g=h=2$ ,  $(1-\lambda)^2=4$  so  
 $(1-\lambda)=\pm 2$ ,  $\lambda=3$  or  $\lambda=-1$ .

# Finding the Eigenvalues

- In general since  $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$  then  $(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{0}$ , so  $\mathbf{u} = \mathbf{0}$  or  $|\mathbf{A} - \lambda\mathbf{I}| = 0$  (since if  $|\mathbf{A} - \lambda\mathbf{I}| \neq 0$  then  $(\mathbf{A} - \lambda\mathbf{I})^{-1}$  exists and  $\mathbf{u} = \mathbf{0}$  identically).
- The eigenvalues are the roots of the characteristic equation  $|\mathbf{A} - \lambda\mathbf{I}| = 0$ .

# Eigenvalues

- Note that we have  $A\mathbf{u}_i = \lambda_i\mathbf{u}_i$  for  $i=1, \dots, n$ .
- We also have  $\mathbf{v}_i A = \lambda_i\mathbf{v}_i$ .
- The  $\mathbf{v}$ 's are the left eigenvectors and the  $\mathbf{u}$ 's are the right eigenvectors.

# DYNAMICAL SYSTEMS

- We can use the  $\mathbf{u}_i$ 's to easily derive the behaviour of the system. Generically for any  $\mathbf{x}(0)$  we have

$$\mathbf{x}(0) = \sum \alpha_i \mathbf{u}_i, \text{ then}$$

$$\mathbf{x}(1) = A \mathbf{x}(0) = A \sum \alpha_i \mathbf{u}_i = \sum \alpha_i A \mathbf{u}_i$$

$$= \sum \alpha_i \lambda_i \mathbf{u}_i \text{ and repeating step we have}$$

$$\mathbf{x}(t) = \sum \alpha_i \lambda_i^t \mathbf{u}_i \longrightarrow \alpha_1 \lambda_1^t \mathbf{u}_1 + \text{negligible terms as } t \longrightarrow \infty \text{ for } \lambda_1 = \max \lambda_i.$$

# LESLIE MATRIX

- Example:- Leslie Matrix Model.
- $\mathbf{x}(t+1) = \mathbf{A} \mathbf{x}(t)$  where

$$A = \begin{pmatrix} 0 & 7 & 8 \\ .5 & 0 & 0 \\ 0 & .25 & 0 \end{pmatrix}$$



# LESLIE MATRIX

- Population with three age classes 0, 1 and 2.  
An individual of age 0 survives to age 1 with probability 0.5 but produces no offspring.  
An individual of age 1(2) survives to age 2(3) with probability 0.25(0.0) and produces 7(8) offspring.  
 $x(t) = (\text{no age 0, no age 1, no age 2})$  at time  $t$

# LESLIE MATRIX

- Now

number of age 0 at time  $t+1 =$

(number of age 1 at time  $t$ ) \* 7

+

(number of age 2 at time  $t$ ) \* 8,

number of age 1(2) at time  $t+1 =$

number of age 0(1) at time  $t$  \* 0.5(0.25)

# LESLIE MATRIX.

- First row is birth rates, others are survival rates.
- $\mathbf{x}(t) = (x_0(t), x_1(t), x_2(t))^T$  is the age-class frequency vector; thus
- $x_0(t+1) = 7 * x_1(t) + 8 * x_2(t),$
- $x_1(t+1) = 0.5 * x_0(t),$
- $x_2(t+1) = 0.25 * x_1(t).$

# LESLIE MATRIX

- Example:- Leslie Matrix Model.
- $\mathbf{x}(t+1) = A \mathbf{x}(t)$  where

$$A = \begin{pmatrix} 0 & 7 & 8 \\ .5 & 0 & 0 \\ 0 & .25 & 0 \end{pmatrix}$$

# Example

- Calculate the age-distribution when we start with equal frequencies in the three classes, i.e.  $(1,1,1)^T$ , for the next two generations.

# LESLIE MATRIX

- Example. Where  $\tau$  is empirical growth rate

$$x(0) = (1, 1, 1)^T \text{ then}$$

$$x(1) = (15, 0.5, 0.25)^T \quad \tau = 5.25$$

$$x(2) = (5.5, 7.5, .125)^T \quad \tau = 0.83$$

$$x(3) = (53.5, 2.75, 1.875)^T \quad \tau = 4.43$$

$$x(4) = (34.25, 26.75, 0.6875)^T \quad \tau = 1.17$$

$$x(5) = (192.75, 17.125, 6.6875)^T \quad \tau = 3.73$$

$$x(6) = (173.375, 86.375, 4.28125)^T \quad \tau = 1.22$$

# Eigenvalues

- $|A-\lambda I|=0=-\lambda^3+\lambda a_{12}a_{21}+a_{12}a_{13}a_{21}$   
 $= -\lambda^3+3.5\lambda+1$   
 $= -(\lambda-2)(\lambda^2+2\lambda+0.5)$

So eigenvalues  $\lambda=2$ ,  $\lambda=(-2 \pm \sqrt{2})/2$  and the largest of these is 2.

Find the eigenvector for  $\lambda=2$ .

# LESLIE MATRIX

- Then we have that  $A\mathbf{u} = \lambda * \mathbf{u}$  where  $\mathbf{u}=(1, 1/4, 1/32)^T$  and  $\lambda =2$  and two other such equations which we ignore (for now) since the  $\lambda$  values are smaller so the influence less.
- $\mathbf{u}$  specifies the stable age distribution and  $\lambda$  the population growth rate.



# DYNAMICAL SYSTEMS

- For Leslie matrix and  $\mathbf{x}(0) > \mathbf{0}$  we have that  $\mathbf{x}(t) \longrightarrow \text{“u”}$  as  $t \longrightarrow \text{inf.}$
- In general we will have that  $\mathbf{x}(t) \longrightarrow \text{“u”}$  as  $t \longrightarrow \text{inf.}$   $\mathbf{u}$  being the eigenvector corresponding to the dominant eigenvalue.

# DYNAMICAL SYSTEMS.

- Non-linear systems are much more complex. The system may settle down to a unique point, or cycle or show chaotic behaviour. The systems we meet will settle down to fixed points.
- Information about the stability of an equilibrium point in a non-linear system can be obtained by making a linear approximation near that point.

# CHAOS.The Shift Operator

- Consider decimal numbers less than 1
- If you want to represent fraction  $1/3$  in decimal then i.e.  
 $0.333333\dots\dots$ indefinitely.

# CHAOS. The Shift Operator.

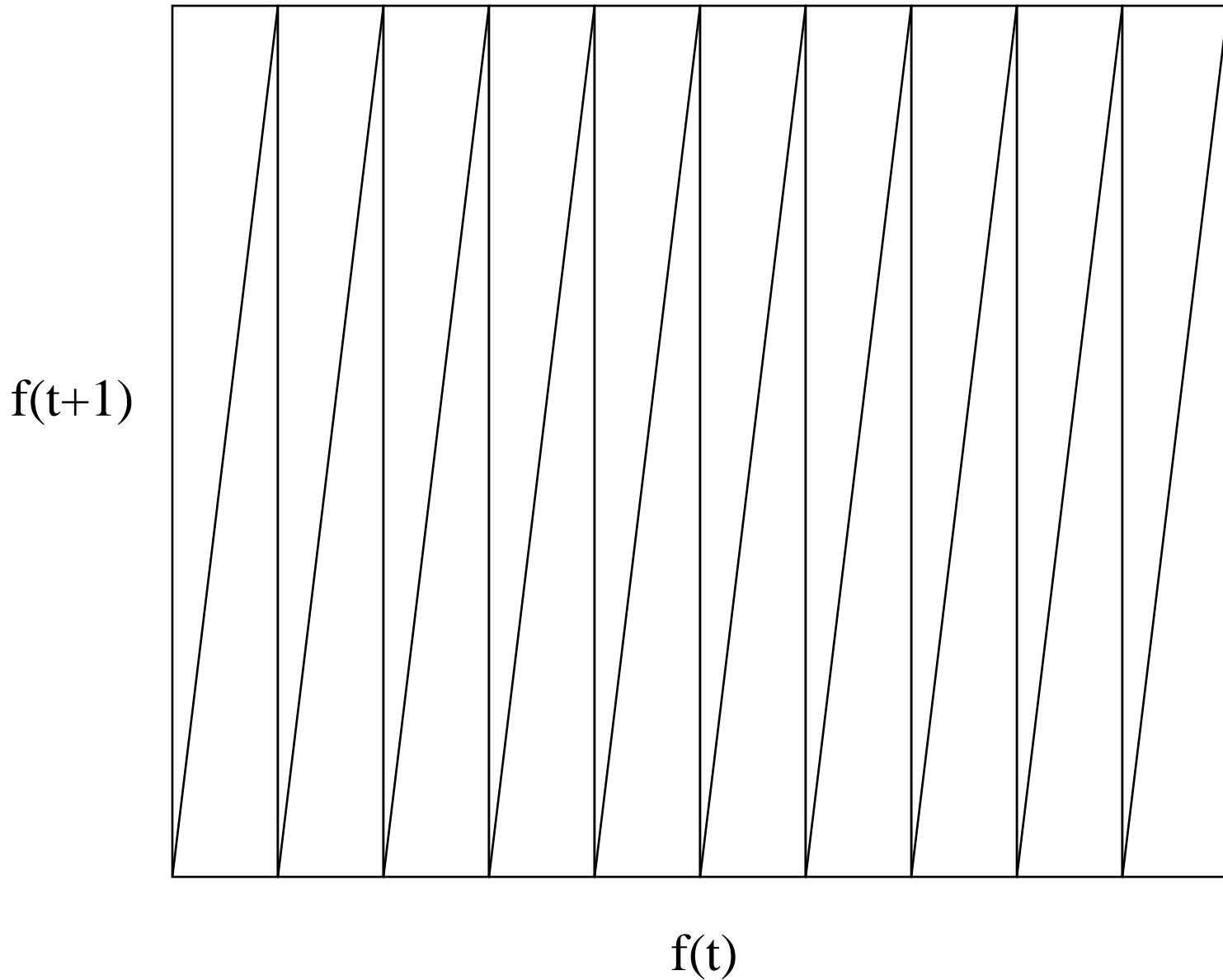
- Suppose  $D(x)$  = “decimal” part of  $x$  and

$$f(t+1) = D(10 * f(t))$$

i.e. just shift value to left and drop integer part

$$0 \leq f(0) \leq 1$$

# CHAOS. The Shift Operator



# CHAOS. The Shift Operator

- Examples.
- $f(0) = 0.5$   
 $f(t) = 0.0$  all  $t > 0$
- $f(0) = 0.(1)^{\text{inf}} = 1/9$   
 $f(t) = f(0)$  all  $t$  i.e. fixed point

# CHAOS. The Shift Operator

- Examples.
- $f(0) = 1/99 = 0.010101\dots\dots = 0.(01)^{\text{inf}}$  so  
 $f(1) = 0.(10)^{\text{inf}} \dots\dots = 10/99$  and so  
 $f(2) = 0.(01)^{\text{inf}} = 1/99 = f(0)$ .

A two cycle results whenever we have  
 $f(0) = 0.***\dots*****(ab)^{\text{inf}}$  where a n.e. b

# CHAOS. The Shift Operator

- There are three-cycles

$$f(0) = 1/999 = 0.(001)^{\text{inf}} \text{ and}$$

$$f(0) = 11/999 = 0.(011)^{\text{inf}} .$$

For any unequal a, b and c and

$$f(0) = \text{*****} \dots \text{*****} (abc\dots z)^{\text{inf}}$$





# CHAOS. The Shift Operator

- Clearly then certain points initiate periodic solutions, but others e.g.  $f(0) = 0.10110111011110\dots$  do not, this will give a chaotic solution.

# HARDY WEINBERG.

- Our first population genetic dynamic system. Suppose we have an infinite population of loci with no selective effects, with two alleles A and B, non-overlapping generations and with random mating.
- Suppose  $P(t)$ ,  $Q(t)$  and  $R(t)$  are the frequencies of AA, AB and BB in generation  $t$ .  $P(t) + Q(t) + R(t) = 1$ .

# HARDY WEINBERG

- Now we require an expression for  $P(t+1)$  etc.
- $$P(t+1) = P(t) * P(t) * 1 + 2 * P(t) * Q(t) * 0.5 + Q(t) * Q(t) * 0.25$$
$$= (P(t) + Q(t)/2)^2$$
$$= p(t)^2$$

where  $p(t)$  is the frequency of A at time  $t$ .

# HARDY WEINBERG

- Thus

$P(t+1) = p(t)^2$ , and similarly we find

- $Q(t+1) = 2 * p(t) * q(t)$ , and

$R(t+1) = q(t)^2$  where  $q(t) = 1 - p(t)$ .

- Then

$p(t+1) = P(t+1) + Q(t+1)/2 =$

$p(t) * (p(t) + q(t)) = p(t)$  i.e. allele frequency

does not change, so set  $p(t) = p$  and  $q(t) = q$ .

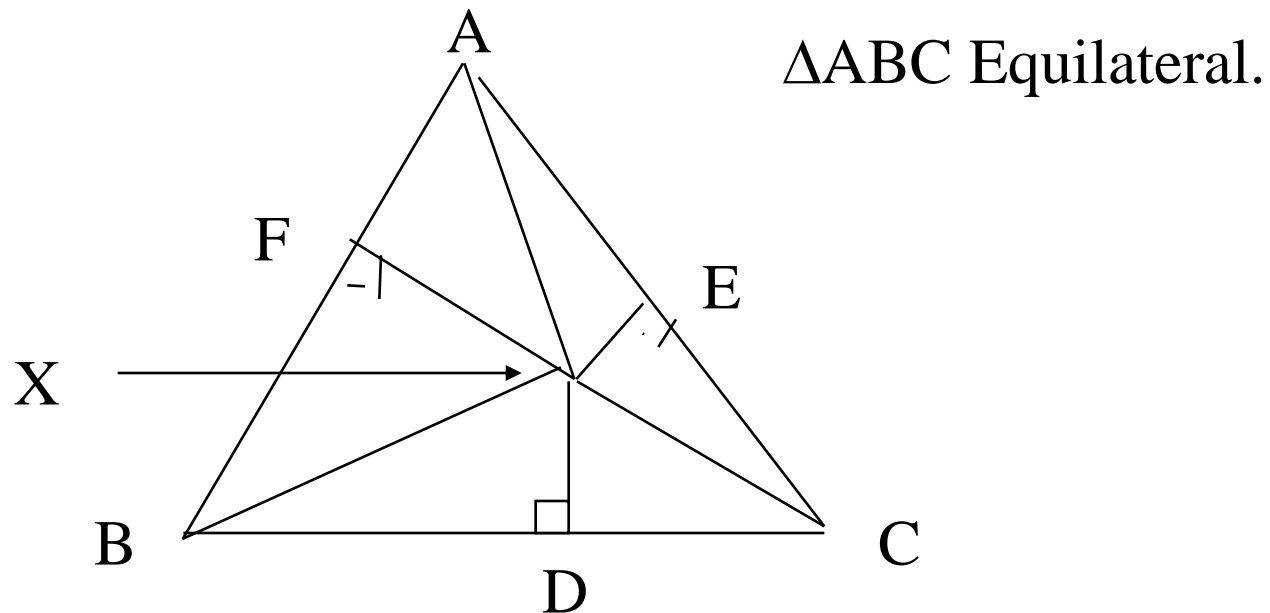
# HARDY WEINBERG

- Thus if we start with  $P(0)$ ,  $Q(0)$  and  $R(0)$ , then  $p(t) = p = P(0) + Q(0) / 2$  all  $t$ , and  $P(t) = p^2$ ,  $Q(t) = 2 * p * q$ ,  $R(t) = q^2$  for  $t = 1, 2, \dots$
- Thus Random Mating (in an infinite pop) is equivalent to Random Union of alleles.

# HARDY WEINBERG

- NB  $Q^2=4*P*R$

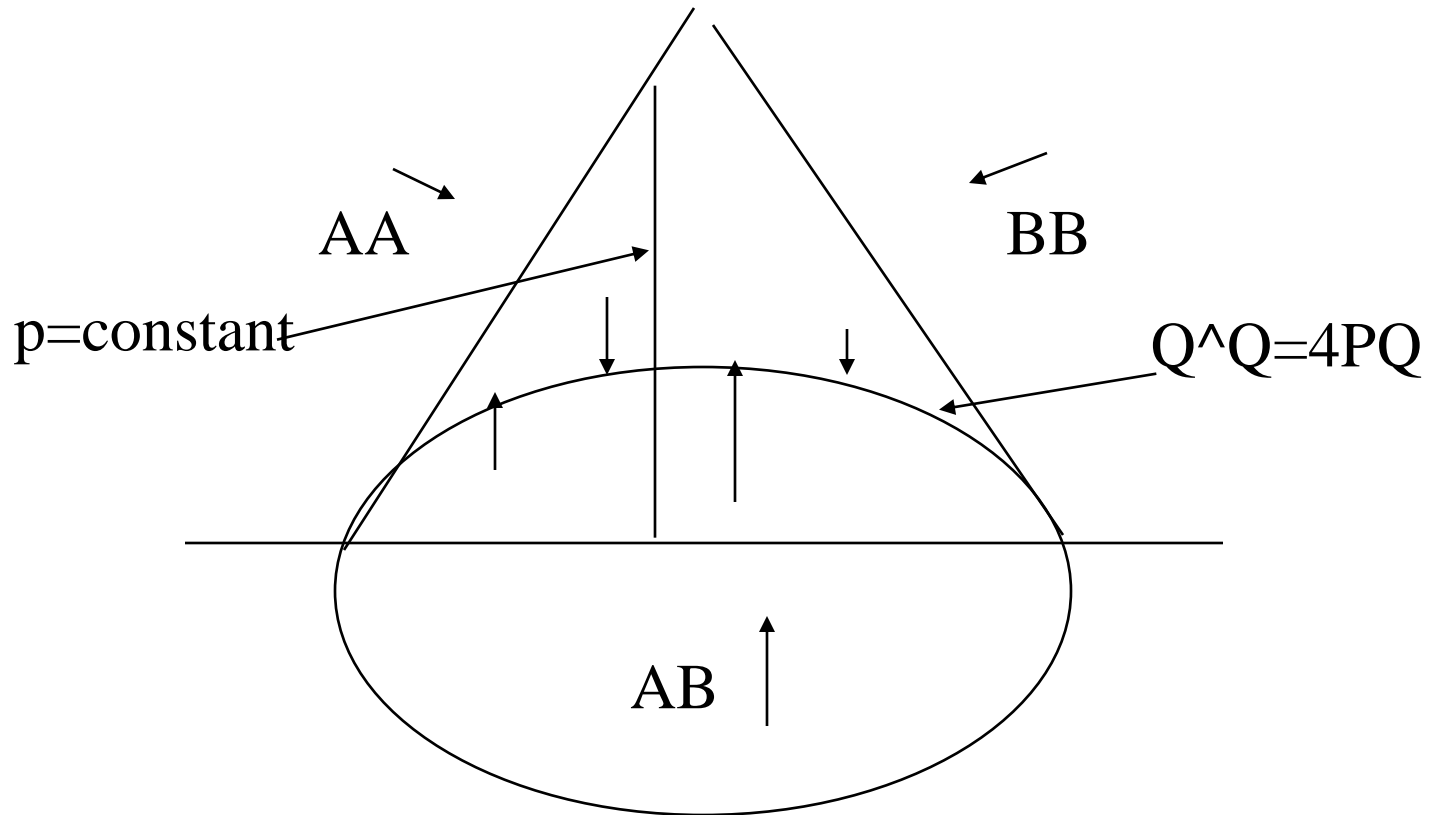
# Homogeneous Co-ordinates.



N.B.  $2\Delta = AB*FX + AC*EX + BC*DX = AB*(FX + EX + DX)$   
so for any point  $X$  in  $ABC$  (choose  $AB = 2/\text{sqrt}(3)$ )  $FX + EX + DX =$



# HARDY WEINBERG.



de Finetti diagram

# SEX-LINKED.

- Consider a locus which is completely sex-linked i.e. on the differential part of the X-chromosome, and has two alleles, a and b. Thus we have two types of X-chromosome  $X_a$  and  $X_b$ , and a single type of Y (it carries nothing at this locus). The genotypes are thus  $X_aX_a$ ,  $X_aX_b$  and  $X_bX_b$  females and  $X_aY$  and  $X_bY$  males.

# SEX-LINKED.

- Suppose we have chromosome frequencies  $p(t)$  and  $r(t)$  for the  $X_a$  in the females and males respectively, and genotype frequencies  $P(t)$ ,  $Q(t)$  and  $R(t)$  in females and  $r(t)$  and  $(1-r(t))$  in males.
- We have  $p(t+1) = (p(t) + r(t))/2$ , and  $r(t+1) = p(t)$ .

# SEX-LINKED.

- This is a linear system; writing  $\mathbf{v}(t) = (p(t), r(t))^T$  so  $\mathbf{v}(t+1) = A \mathbf{v}(t)$  where  $A = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}$

# SEX-LINKED.

- Finding the eigenvalues and eigenvectors.  
If  $Au = \lambda u$  where  $u=(z,1)^T$  (scale invariant)  
then  $(z+1)/2 = \lambda z$  and  $z = \lambda$ . Thus  
 $2\lambda^2 - \lambda - 1 = 0$  so  $\lambda_1 = +1$  and  $\lambda_2 = -1/2$ .

# EIGENVALUES & EIGENVECTORS

- We have matrix

$$\mathbf{A} = \begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}$$

and there are eigenvalues  $\lambda_1 = 1$  and  $\lambda_2 = -1/2$

# EIGENVALUES & EIGENVECTORS

- Suppose  $u_1 = (x, y)^T$  is eigenvector for  $\lambda_1=1$  so  
 $A u_1 = ((x+y)/2, x)^T = (x,y)^T$  so  
 $x = y$  so we take  $u_1 = (1,1)^T$  (any multiple would do).
- Suppose  $u_2 = (x, y)^T$  is eigenvector for  $\lambda_1=-1/2$   
so  $A u_2 = ((x+y)/2, x)^T = -(x,y)^T/2$  so  
 $x = -y/2$  so we take  $u_2 = (1,-2)^T$  (any multiple would do).

# Solving

- $v(0) = (p(0), r(0)) = (2p(0) + r(0))/3 (1, 1) + (p(0) - r(0))/3 (1, -2)$

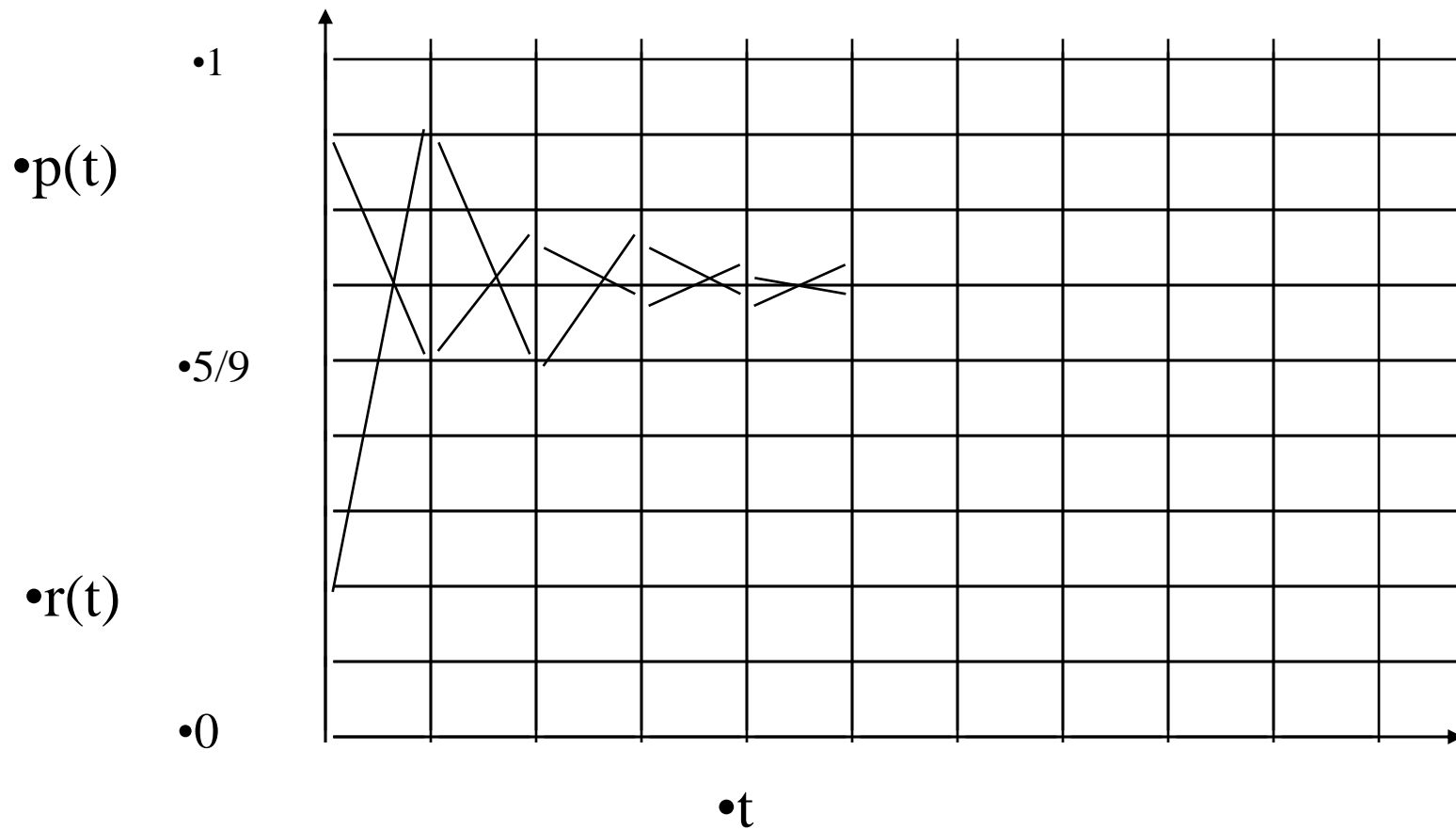
so

$$v(t) = (2p(0) + r(0))/3 1^t (1, 1) + (p(0) - r(0))/3 (-1/2)^t (1, -2)$$

so  $p(\infty) = r(\infty) = (2p(0) + r(0))/3$  and approach is oscillatory.



# SEX-LINKED.



# SEX-LINKED

- At equilibrium  $P = p^2$  ,  $Q = 2pq$ ,  $R = q^2$  in females and  $p$  and  $q$  in males.

## 2-LOCI

- There are two loci A with alleles  $A_1$  and  $A_2$ , and B with alleles  $B_1$  and  $B_2$ . Denote the frequency of the chromosome  $A_i B_j$  by  $P_{ij}(t)$  in the gametes of generation  $t$ . Assume random mating and infinite population and that the recombination rate is  $r$ .

## 2-LOCI

- $P_{11}(t)$  is frequency of  $A_1B_1$  in gametes.
- $$P_{11}(t+1) = P_{11}(t)*P_{11}(t) + 2*P_{11}(t)*P_{12}(t)/2 + 2*P_{11}(t)*P_{21}(t)/2 + 2*P_{11}(t)*P_{22}(t)*(1-r)/2 + 2*P_{12}(t)*P_{21}(t)*r/2 = P_{11}(t) + r*(P_{12}(t)*P_{21}(t) - P_{11}(t)*P_{22}(t)) = P_{11}(t) + r*D(t)$$

# 2-LOCI

- where

$D(t) = P_{12}(t) * P_{21}(t) - P_{11}(t) * P_{22}(t)$  is called the Linkage Disequilibrium (it is the cross-ratio of the 2\*2 Table).

## 2-LOCI

- Now since

$$\begin{aligned} P_{1.}(t+1) &= P_{11}(t+1) + P_{12}(t+1) \\ &= P_{1.}(t) = P_{11}(t) + P_{12}(t) \\ &= \text{frequency of A1, we have} \end{aligned}$$

$$P_{12}(t+1) = P_{12}(t) - r * D(t),$$

$$P_{21}(t+1) = P_{21}(t) - r * D(t), \text{ and}$$

$$P_{22}(t+1) = P_{22}(t) + r * D(t).$$

## 2-LOCI

- $$\begin{aligned}
 & D(t+1) \\
 &= P_{12}(t+1) * P_{21}(t+1) - P_{11}(t+1) * P_{22}(t+1) = \\
 & (P_{12}(t) - r * D(t)) * (P_{21}(t) - r * D(t)) \quad - \\
 & (P_{11}(t) + r * D(t)) * (P_{22}(t) + r * D(t)) \quad = \\
 & D(t) * (1 - r) \quad \longrightarrow 0 \text{ as } t \quad \longrightarrow \text{infinity.}
 \end{aligned}$$
- In the limit  $D=0$  so  $P_{12} * P_{21} = P_{11} * P_{22}$   
 so  $(P_{11} + P_{12}) * (P_{11} + P_{21}) = P_{11} + D = P_{11}$   
 and we have “independence” of loci.

## 2-LOCI

- Thus the frequencies at distinct loci are independent “in the limit” (with no epistatic effects).
- The rate of decay of the linkage disequilibrium  $D(t)$  is geometric with rate  $(1-r)$ , i.e. as  $r$  decreases rate of decay decreases. Close linkage then slower decay.



# SELECTION-AUTOSOMAL.

- Above was all with no differences in the fitness and/or fertility of the genotypes. Now consider an autosomal locus with alleles A and B and viabilities (survival rate from conception to maturity) of  $a$ ,  $h$  and  $b$  for the genotypes AA, AB and BB. Take  $P(t)$ ,  $Q(t)$  and  $R(t)$  as genotype and  $p(t)$  and  $q(t)$  as gene-frequencies at maturity.

# SELECTION-AUTOSOMAL.

- At **conception** we have frequencies  
 $p(t)*p(t)$ ,  $2*p(t)*q(t)$ ,  $q(t)*q(t)$ .
- At **maturity** (relative) frequencies  
 $a*p(t)*p(t)$ ,  $2*h*p(t)*q(t)$ ,  $b*q(t)*q(t)$ ,  
and  
 $V(t)=a*p(t)*p(t)+2*h*p(t)*q(t)+b*q(t)*q(t)$ ,  
is the mean viability (scale factor).

# SELECTION-AUTOSOMAL.

- Thus

$$p(t+1) = p(t) ( a*p(t) + h*q(t) ) / V(t),$$

and

$$q(t+1) = q(t) ( h*p(t) + b*q(t) ) / V(t).$$

- $$\frac{p(t+1)}{q(t+1)} = \frac{p(t) ( a*p(t) + h*q(t) )}{q(t) ( h*p(t) + b*q(t) )}$$

# SELECTION-AUTOSOMAL

- $$\frac{p(t+1)}{q(t+1)} = \frac{p(t) (a^*p(t) + h^*q(t))}{q(t) (h^*p(t) + b^*q(t))}$$

- Special cases.

(1) multiplicative viabilities  $a = \alpha^2$ ,

$h = \alpha\beta$  ,  $b = \beta^2$  so

$$\frac{p(t+1)}{q(t+1)} = \frac{p(t) \alpha ( \alpha^*p(t) + \beta^*q(t))}{q(t) \beta ( \alpha^*p(t) + \beta^*q(t))} = \frac{\alpha^*p(t)}{\beta^*q(t)}$$

# SELECTION-AUTOSOMAL

- Thus  $p(t)/q(t) \nearrow t$  if  $\alpha > \beta$ , and so  $p(t) \longrightarrow 1$  as  $t \longrightarrow \text{inf.}$

# SELECTION-AUTOSOMAL

- $$\frac{p(t+1)}{q(t+1)} = \frac{p(t) (a^*p(t) + h^*q(t))}{q(t) (h^*p(t) + b^*q(t))}$$
- (2) Recessive Lethal i.e.  $a = h = 1, b = 0$ ,  
so  $p(t+1)/q(t+1) = p(t)/q(t) * (1 / p(t))$   
 $\geq p(t)/q(t)$  with equality iff  $p(t)=1$ .  
Therefore  $p(t)$  increases steadily to 1.

# SELECTION-AUTOSOMAL

- $$\frac{p(t+1)}{q(t+1)} = \frac{p(t) (a * p(t) + h * q(t))}{q(t) (h * p(t) + b * q(t))}$$
- Taking  $f_t = p(t) / q(t)$  and dividing through by  $q(t) * q(t)$  we get  
 $f_{t+1} = f_t (a * f_t + h) / (g * f_t + b)$  where  $g (=h$   
here) adds generality.

# SELECTION-AUTOSOMAL.

- Denote fixed points by  $f(\text{inf}) = f$ , then  $f = f(a*f + h)/(g*f + b)$  so  $f = \text{“inf”}$ , 0 or  $(h - b) / (g - a)$ .
- The behaviour is determined by the parities of  $(h - b)$  and  $(g - a)$ ; we just look at the details of (i)  $h > b, g > a$  and describe the other three cases (ii)  $h < b, g < a$ , (iii)  $h < b, g > a$  and (iv)  $h > b, g < a$ .



$$df_{(t+1)}/df_t$$

$$\begin{aligned} df_{t+1} / df_t &= \{(2af_t + h)(gf_t + b) - f_t(af_t + h)g\} / \{(gf_t + b)^2\} \\ &= \{agf_t^2 + 2abf_t + bh\} / \{(gf_t + b)^2\} \geq 0 \end{aligned}$$

# SELECTION-AUTOSOMAL.

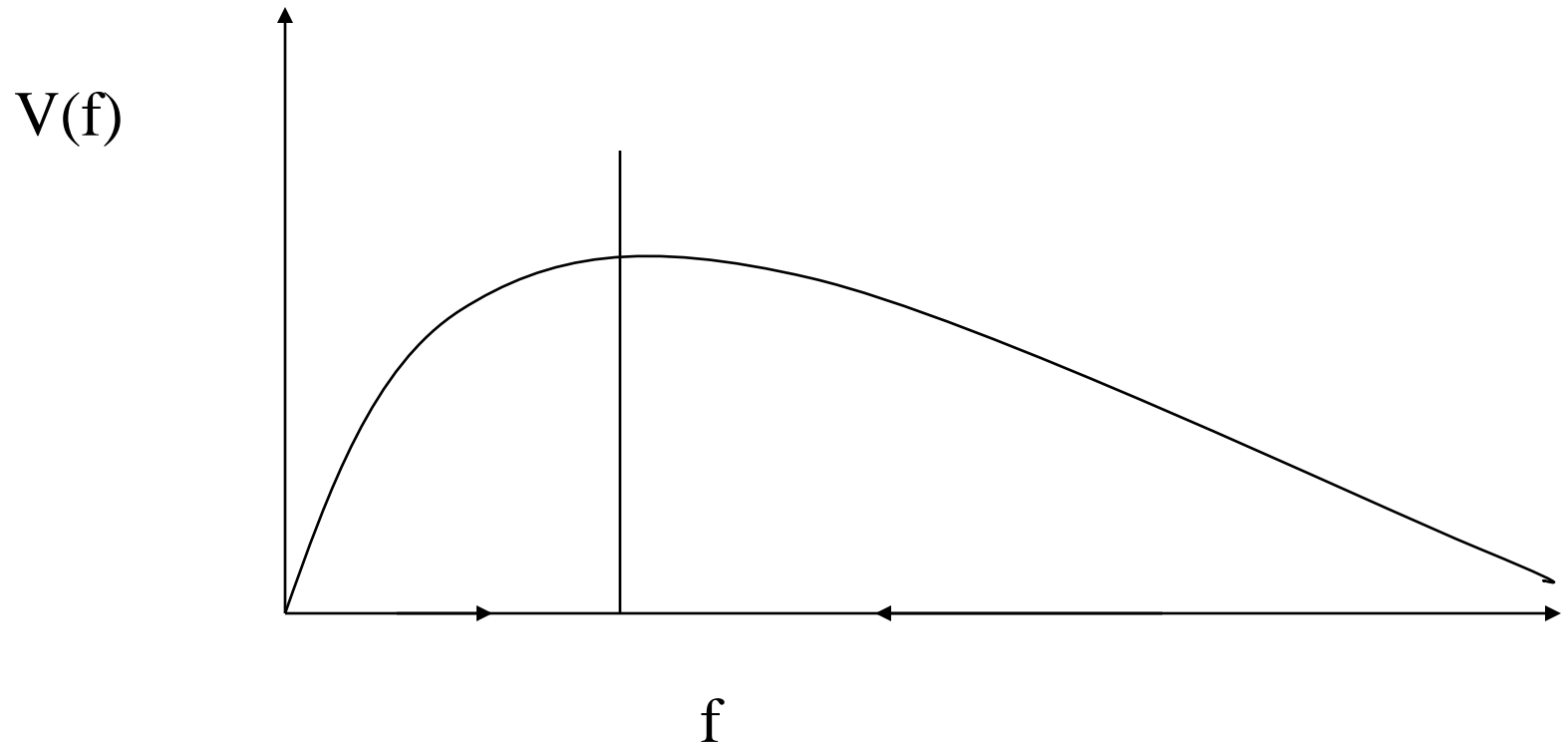
- $h > b$  ,  $g > a$  then take  $f = (h - b)/(g - a)$ .
- If  $f_t > f = (h - b)/(g - a)$  then  
 $f_t * (g - a) > (h - b)$  since  $g > a$  so  
 $g * f_t + b > a * f_t + h$  so  
 $(a * f_t + h) / (g * f_t + b) < 1$  so  $f_{t+1} < f_t$ .
- $df_{t+1}/df_t > 0$  and  $f_t > f$  implies  $f_{t+1} > f$ .

# SELECTION-AUTOSOMAL.

- Thus  $h > b$ ,  $g > a$  and  $f_t > f$  implies  $f_t > f_{t+1} > f$ .
- $h > b$ ,  $g > a$  and  $f_t < f$  implies  $f_t < f_{t+1} < f$ .
- Thus  $f_t \rightarrow f$  monotonically as  $t \rightarrow \infty$ .
- $V_t$  increases monotonically with  $t$ .

# SELECTION-AUTOSOMAL

- For  $h > a$ ,  $h > b$ ,



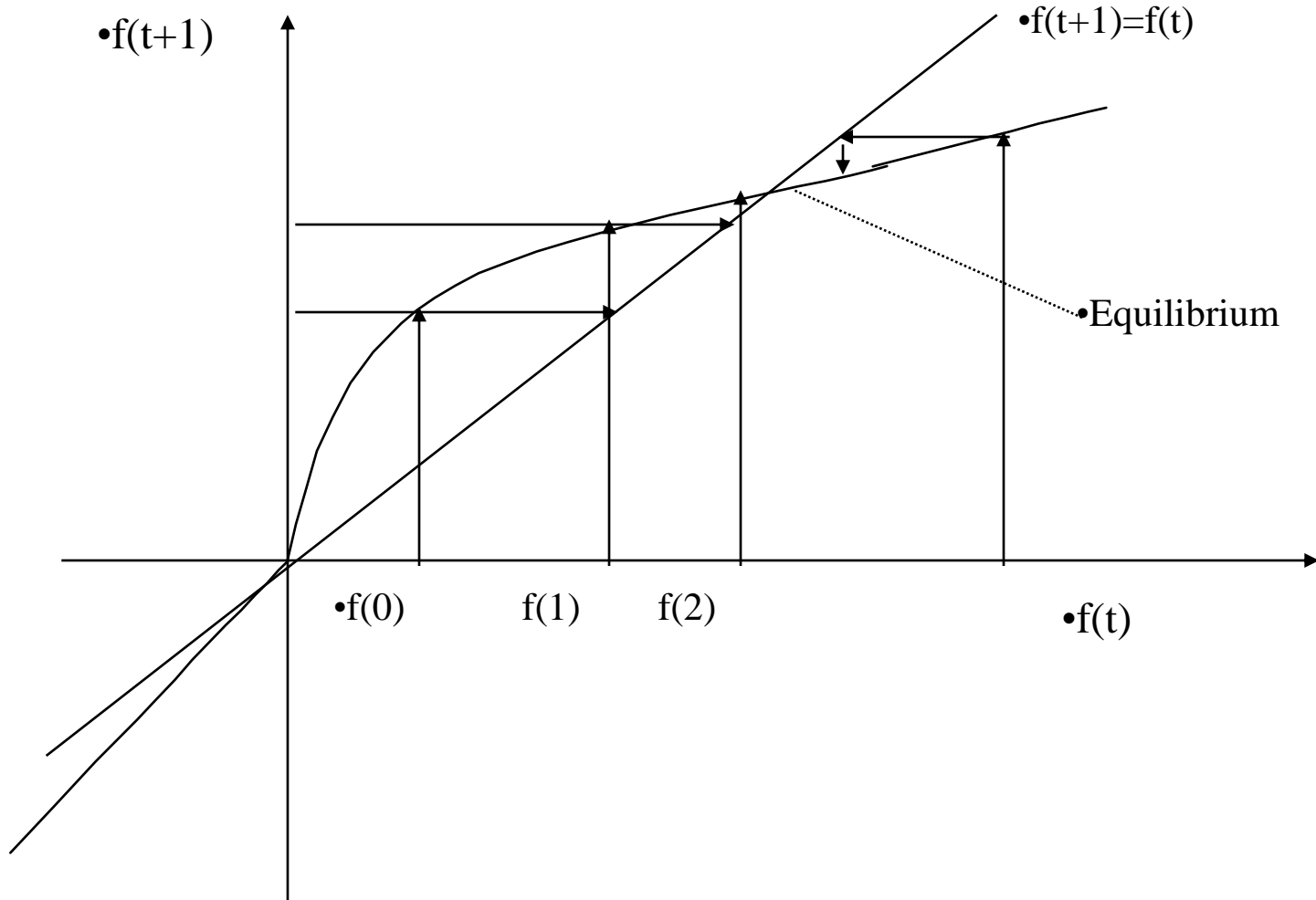
# SELECTION-AUTOSOMAL

- $h < b$  ,  $g < a$  so  $f = (h-b) / (g-a)$  is an internal unstable equilibrium. If  $f_0 < (h-b)/(g-a)$  then  $f_t \longrightarrow 0$  monotonically, and if  $f_0 > (h-b)/(g-a)$  then  $f_t \longrightarrow \text{inf}$ .
- $h > b$ ,  $a > g$  any  $f_0$  then  $f_t \longrightarrow \text{inf}$
- $h < b$ ,  $a < g$  any  $f_0$  then  $f_t \longrightarrow 0$ .

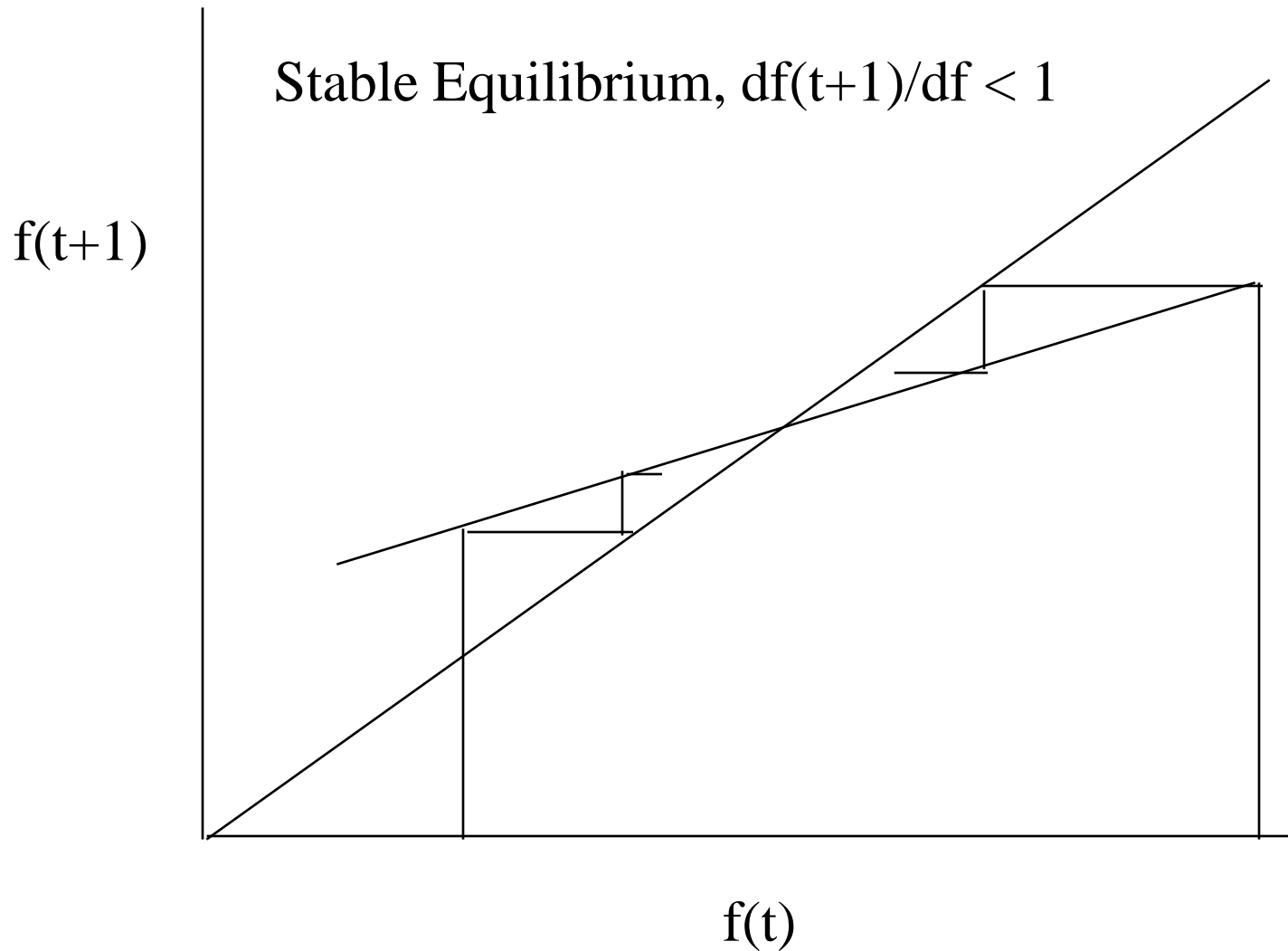
# POLYMORPHISM

- Thus an autosomal locus will remain polymorphic iff (if, and only if) the heterozygote is fitter than both homozygotes (heterosis).

# COBWEBBING.

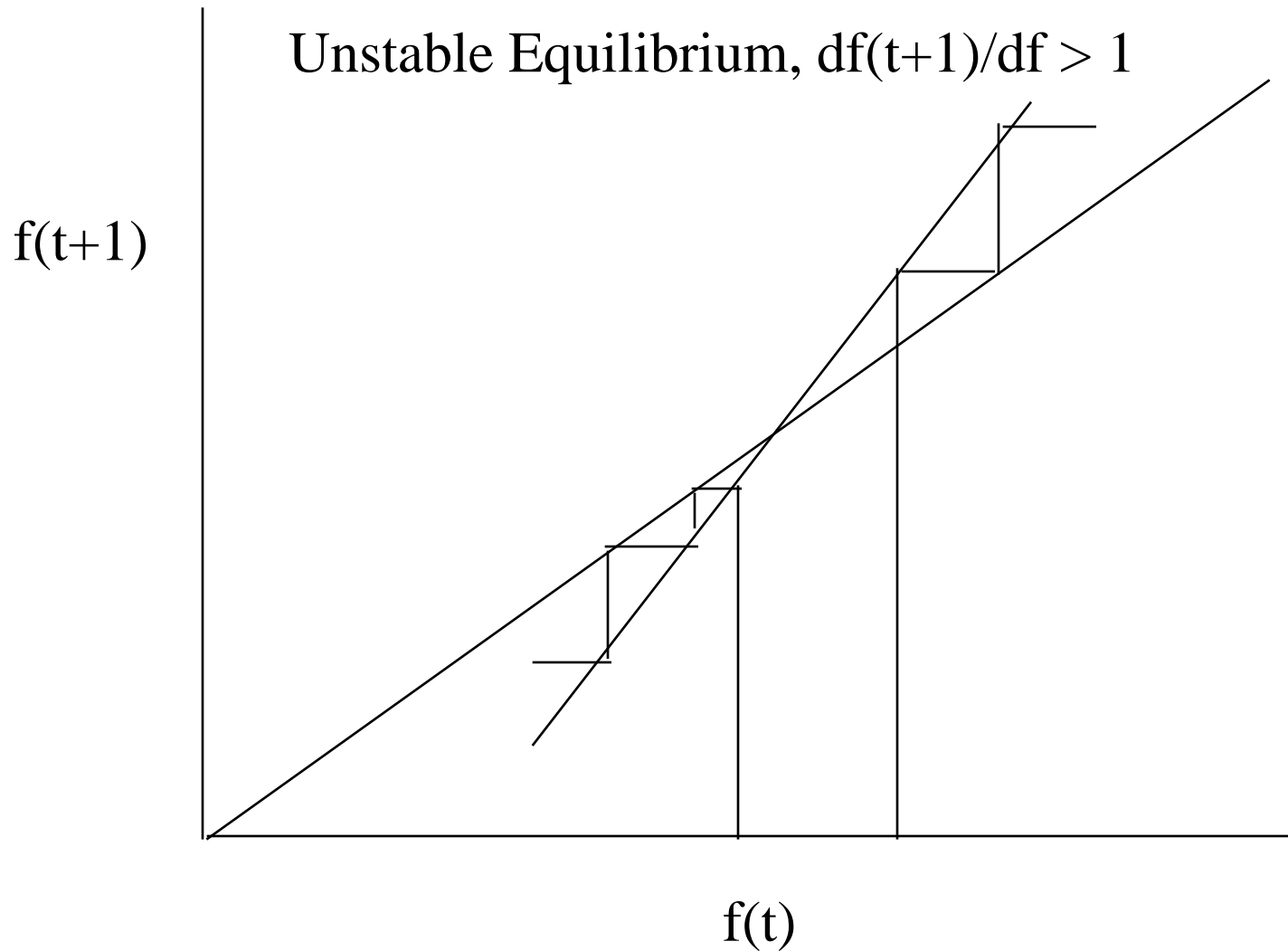


# COBWEBBING. Linear Approx.

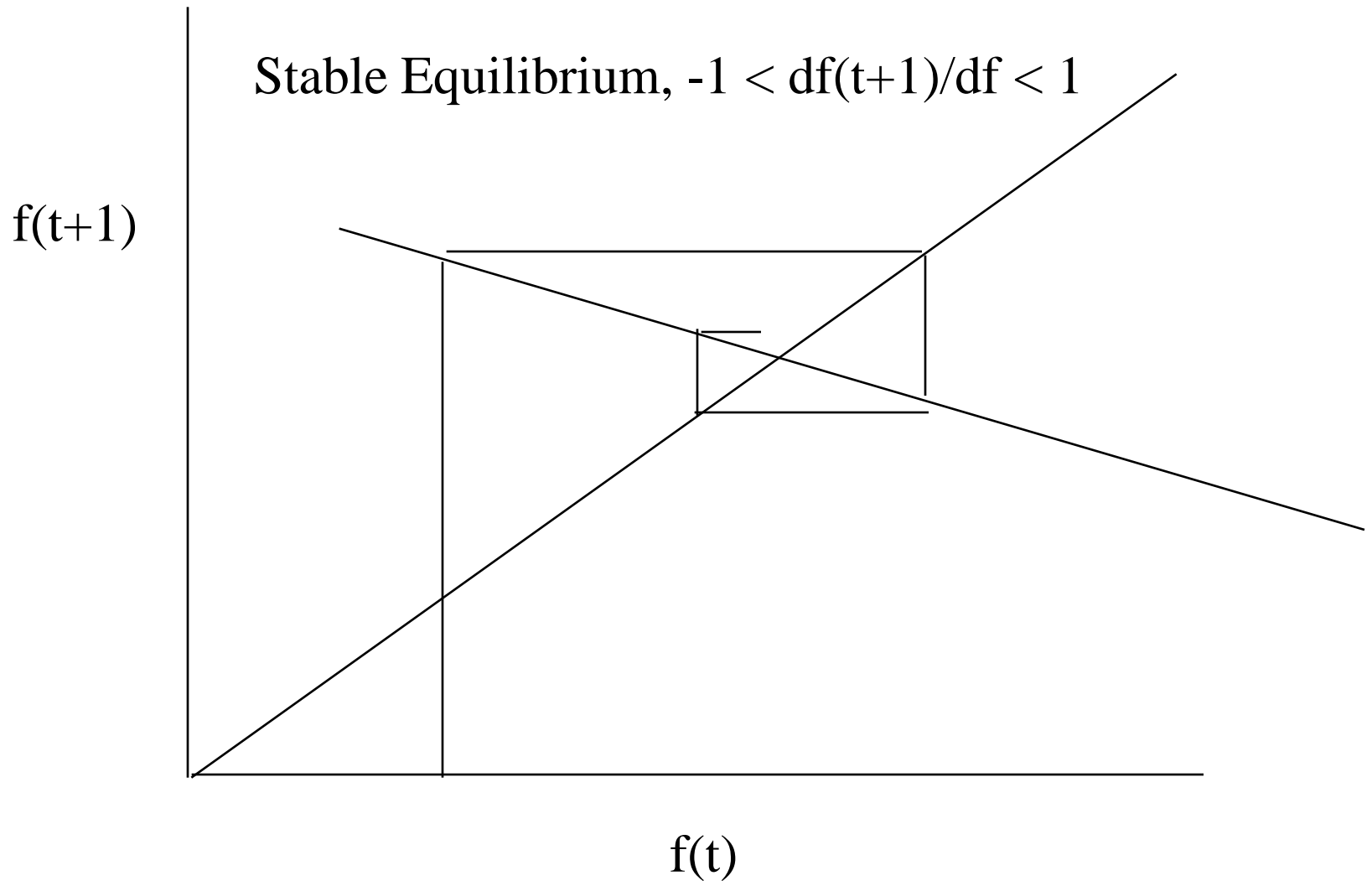




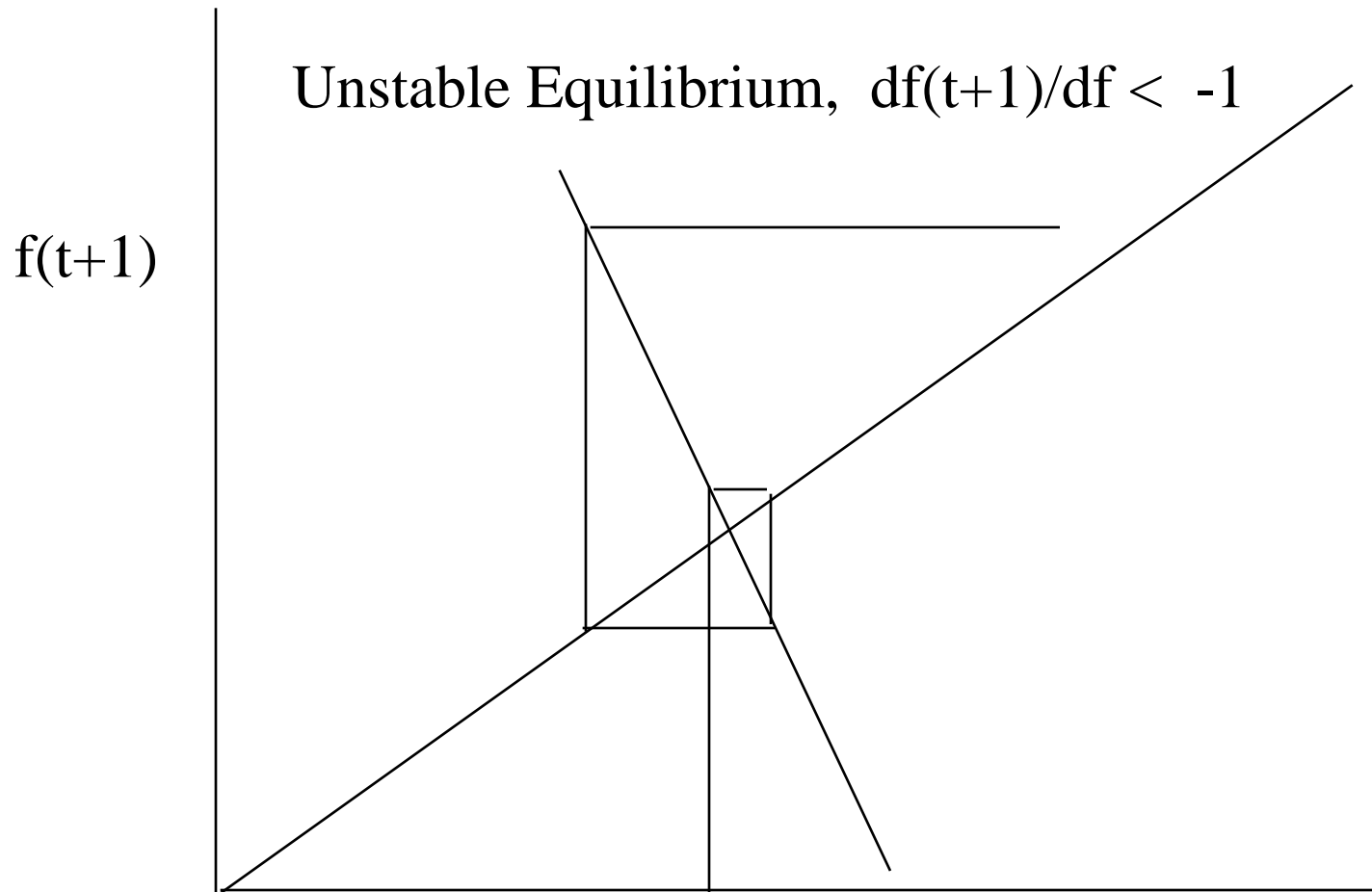
# COBWEBBING. Linear Approx.



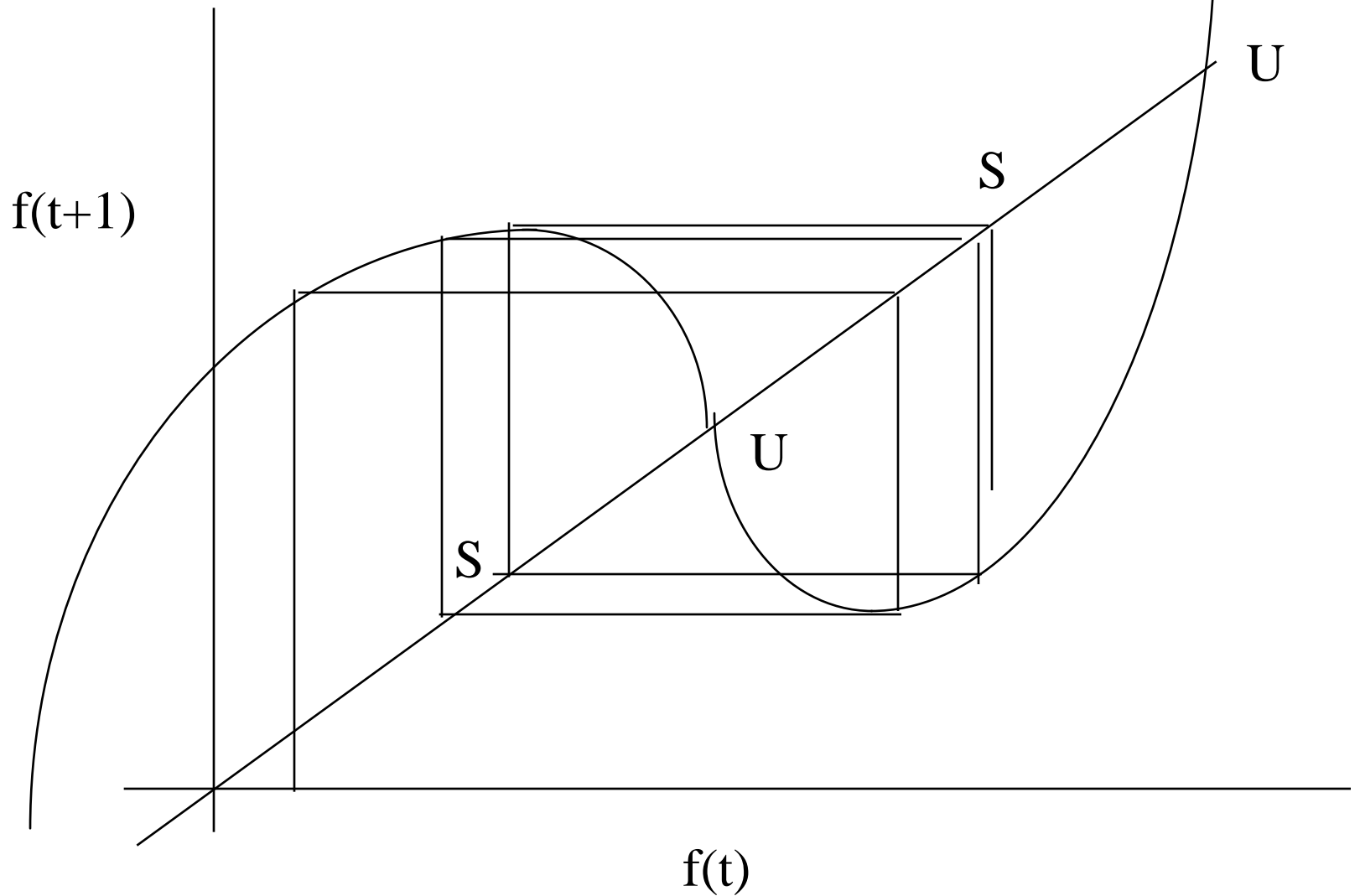
# COBWEBBING. Linear Approx



# COBWEBBING. Linear Approx



# COBWEBBING. 2-CYCLE.



# EXERCISE.

- Sketch the curves

$$f_{t+1} = f_t (a * f_t + h) / (g * f_t + b) \quad \text{for}$$

$$(1) a=b=1, g=h=2;$$

$$(2) a=b=2, g=h=1;$$

$$(3) a=1, b=3, g=h=2;$$

and check the behaviour of  $f_t$  with  $t$ .

# ADDING MUTATION.

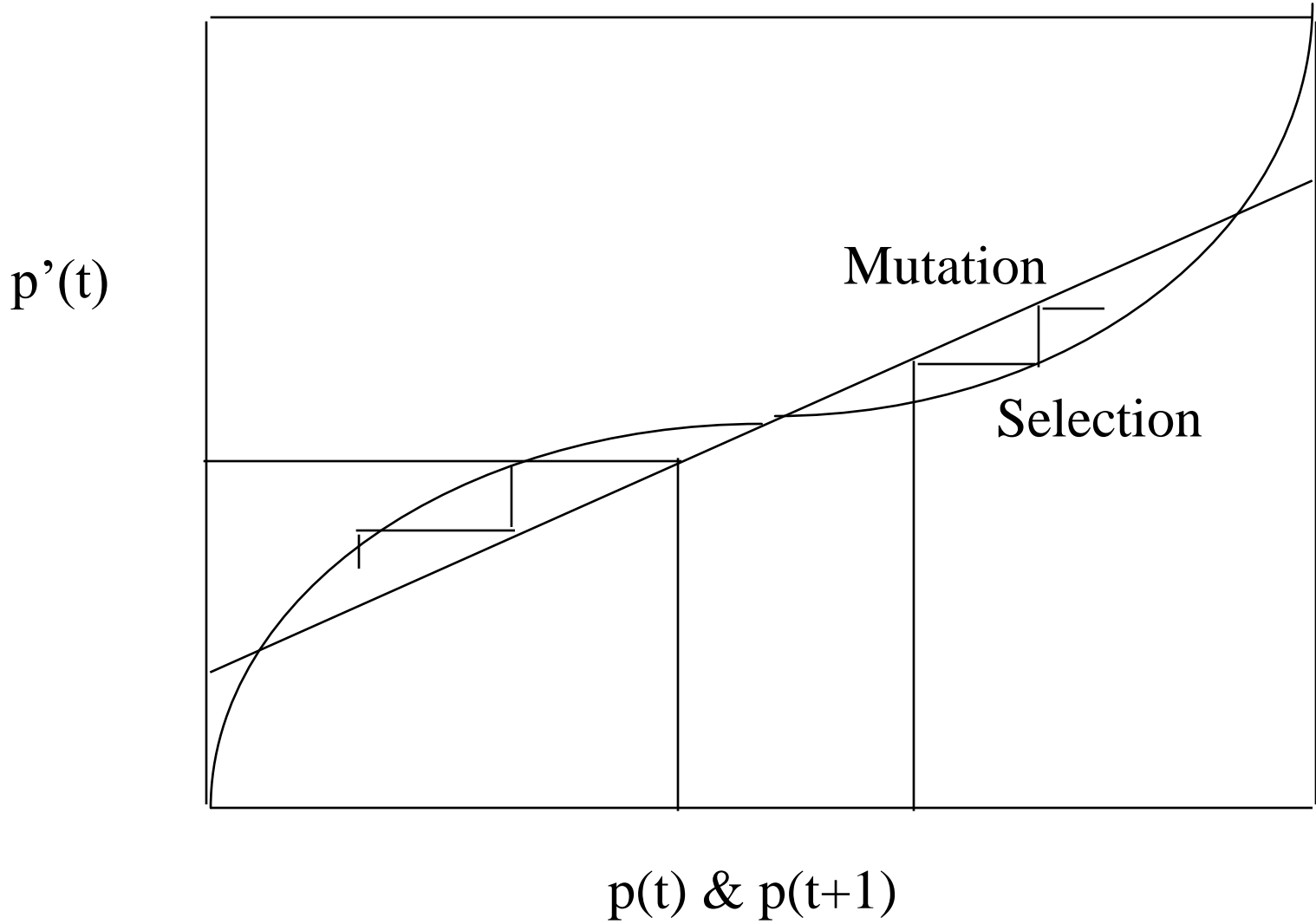
- Suppose mutation

$$p'(t) = p(t) * (1 - \alpha) + q(t) * \beta = h(p(t))$$

and selection  $p(t+1) = g(p'(t))$ , so

$$p(t) \longrightarrow p'(t) \longrightarrow p(t+1).$$

# ADDING MUTATION. Unstable



# Linearisation

- When there is a single variable then we can examine the stability of an equilibrium very easily. The linear approximation at the equilibrium  $f$  is just determined by the value of the derivative at  $f$ . The approximation is **(where the derivative is evaluated at  $f$ )**

$$f_{(t+1)} = f + df_{(t+1)} / df_t * (f_t - f)$$

$$(f_{(t+1)} - f) = \delta_{(t+1)} = \delta_t df_{(t+1)} / df$$



# Linearisation

- Then if the derivative at  $f$ ,  $\lambda$  in  $(-1,+1)$  then the system returns to  $f$  if perturbed a little, and moves away if  $\lambda > 1$  or  $\lambda < -1$ . The cases  $\lambda = 1$  and  $\lambda = -1$  require one to look at second derivatives.

# Autosomal locus

- Although we know how the system behaves when we have two alleles at an autosomal locus we use it here as an example of linearisation.

$$df_{(t+1)}/df_t$$

$$\dot{df}_{t+1} / df_t = \{(2af_t + h)(hf_t + b) - (af_t + h)g\} / \{(hf_t + b)^2\}$$

Now at f we have  $(af+h)=(hf+b)$  so

$$\lambda = \{(2af + h) - hf\} / \{hf + b\}$$

$$= \{(2a - h)(h - b) - h(h - a)\} / \{h(h - b) + b(h - a)\}$$

$$= \{ah + bh - 2ab\} / \{h^2 - ab\}$$

$$df_{(t+1)}/df_t$$

- Now if  $h > a, b$  then  
 $0 < (h-b)(h-a) = (h^2 - ab) - (ah + bh - 2ab)$  so  $0 < \lambda < 1$   
i.e.  $f$  is stable
- If  $h < a, b$  the  $\lambda > 1$  so  $f$  is unstable.

# End Points

- Sometimes it is useful to examine the end-points i.e. when  $f=0$  and “ $f=\text{inf}$ ”. We are then examining whether a mutation can invade the population. So at  $f=0$  where the population is all B's we examine the fate of an A.

# Autosomal

- If we take  $f$  near 0 the  $f(af+h)/(gf+b)$  is approximately  $fh/b$  (or the derivative at  $f=0$  is  $h/b$ ). Thus an A will invade if  $h>b$  (if  $h=b$  we need to look at second order terms).
- Similarly a B will invade a population of A's if  $h>a$ .
- On the basis of this we could say that  $h>a,b$  implies a protected polymorphism (though here we know much more anyway).

# OTHER CASES.

- Autosomal-locus  $n$  alleles, viabilities.
- Sex-linked locus 2 alleles.
- Differential selection in two sexes.
- Fertility differences.
- Multiple loci.

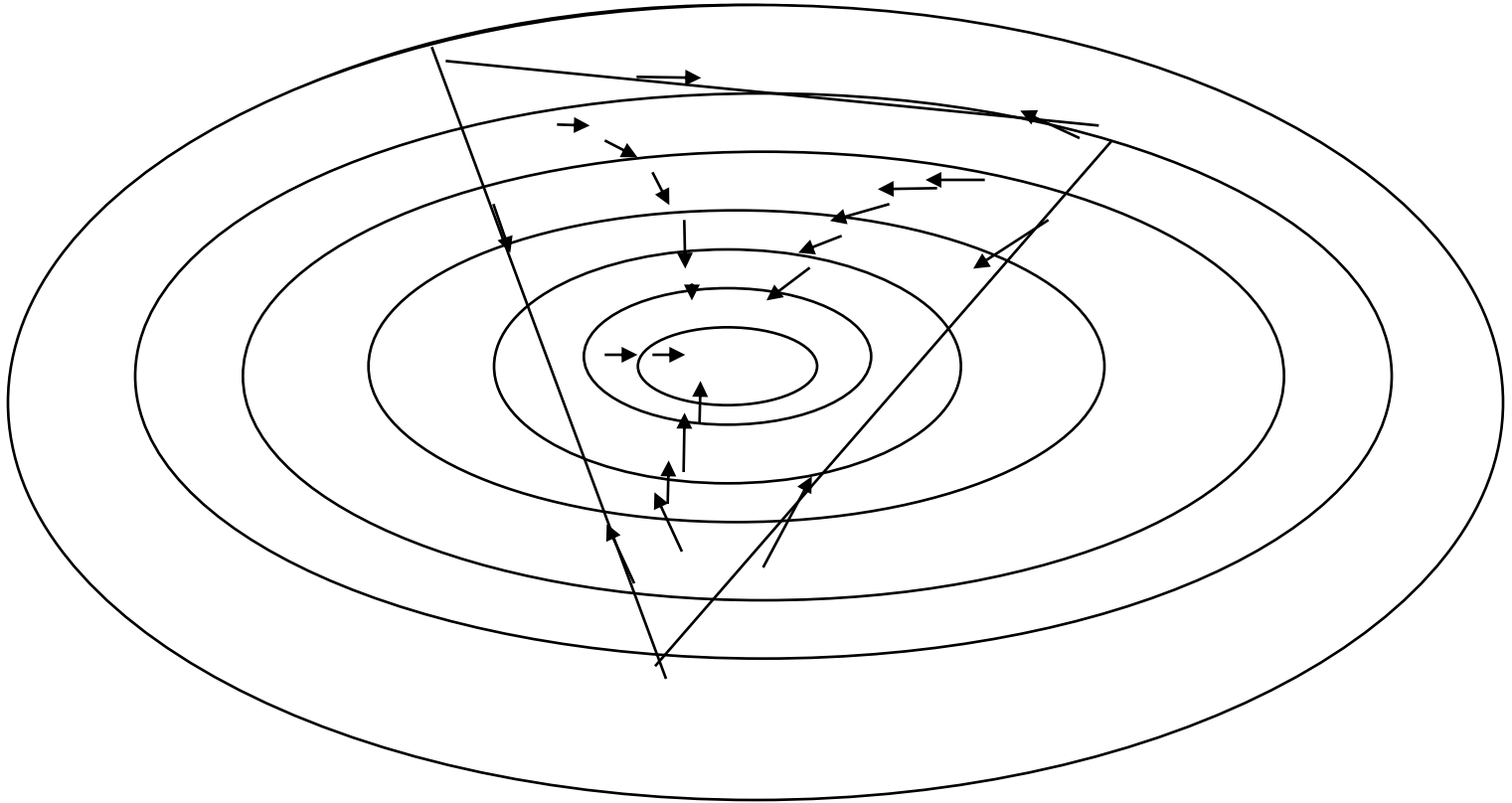
# AUTOSOMAL n-ALLELES.

- n alleles  $A_i$  for  $i=1,..n$  and viabilities  $w_{i,j}$  for  $A_i A_j$  then the mean viability  $V$  where  $V_t = \sum w_{i,j} p_i(t) p_j(t)$  and  $V_{t+1} \geq V_t$  with equality only if at a fixed point of system.



# AUTOSOMAL 3-ALLELES.

Contours of  $V$ .



# SEX-LINKED, VIABILITIES

- Suppose viabilities  $a$ ,  $h$  and  $b$  for  $X_a X_a$ ,  $X_a X_b$  and  $X_b X_b$  and  $u$  and  $v$  for  $X_a$  and  $X_b$  and suppose the frequencies of  $X_a$  in females is  $p(t)$  and in males  $r(t)$
- At conception (relative) frequencies are  $p(t)r(t)$ ,  $(p(t)s(t) + q(t)r(t))$ ,  $q(t)s(t)$  in females and  $p(t)$ ,  $q(t)$  in males.

# SEX-LINKED, VIABILITIES

- At **maturity** (relative) frequencies are  $a \cdot p(t)r(t)$ ,  $h \cdot (p(t)s(t) + q(t)r(t))$ ,  $b \cdot q(t)s(t)$  in females and  $u \cdot p(t)$ ,  $v \cdot q(t)$  in male with  
$$F(t) = a \cdot p(t)r(t) + h \cdot (p(t)s(t) + q(t)r(t)) + b \cdot q(t)s(t)$$
and  
$$M(t) = u \cdot p(t) + v \cdot q(t)$$

# SEX-LINKED, VIABILITIES

- Then

$$f_{t+1} = (2a * f_t * g_t + h * (f_t + g_t)) / (h * (f_t + g_t) + 2b)$$

$$\text{and } g_{t+1} = u * f_t / v.$$

# SEX-LINKED, VIABILITIES

- Now one can prove that the system converges.
- $f_t$  and  $g_t \longrightarrow (h^+ - b^+) / (h^+ - a^+)$  where  $a^+ = a * u$ ,  $b^+ = b * v$  and  $h^+ = (h * (u + v)) / 2$ , iff  $h^+ > a^+$  and  $h^+ > b^+$ .

# DIFFERENTIAL SELECTION IN THE SEXES.

- Infinite, random mating population, with differential viabilities.
- Suppose allele A has frequencies  $p_t$  and  $r_t$  in females and males at maturity in generation  $t$ ;  $q_t=1-p_t$  and  $s_t=1-r_t$ .
- At **conception** genotype frequencies are  $p_t * r_t$ ,  $p_t * s_t + q_t * r_t$ ,  $q_t * s_t$ .

# DIFFERENTIAL SELECTION IN THE SEXES.

- At **maturity** (relative) frequencies are  $a \cdot p(t)r(t)$ ,  $h \cdot (p(t)s(t) + q(t)r(t))$ ,  $b \cdot q(t)s(t)$  in females and  $c \cdot p(t)r(t)$ ,  $k \cdot (p(t)s(t) + q(t)r(t))$ ,  $d \cdot q(t)s(t)$  in male with  

$$L(t) = a \cdot p(t)r(t) + h \cdot (p(t)s(t) + q(t)r(t)) + b \cdot q(t)s(t)$$
and  

$$M(t) = c \cdot p(t)r(t) + k \cdot (p(t)s(t) + q(t)r(t)) + d \cdot q(t)s(t)$$

# DIFFERENTIAL SELECTION IN THE SEXES.

- Thus

$$\begin{aligned} p_{t+1}/q_{t+1} = f_{t+1} = \\ (2ap_t r_t + h(p_t s_t + q_t r_t)) / \\ (2bq_t s_t + h(p_t s_t + q_t r_t)) = \\ f_{t+1} = 2af_t m_t + h(f_t + m_t) / (2b + h(f_t + m_t)) \end{aligned}$$

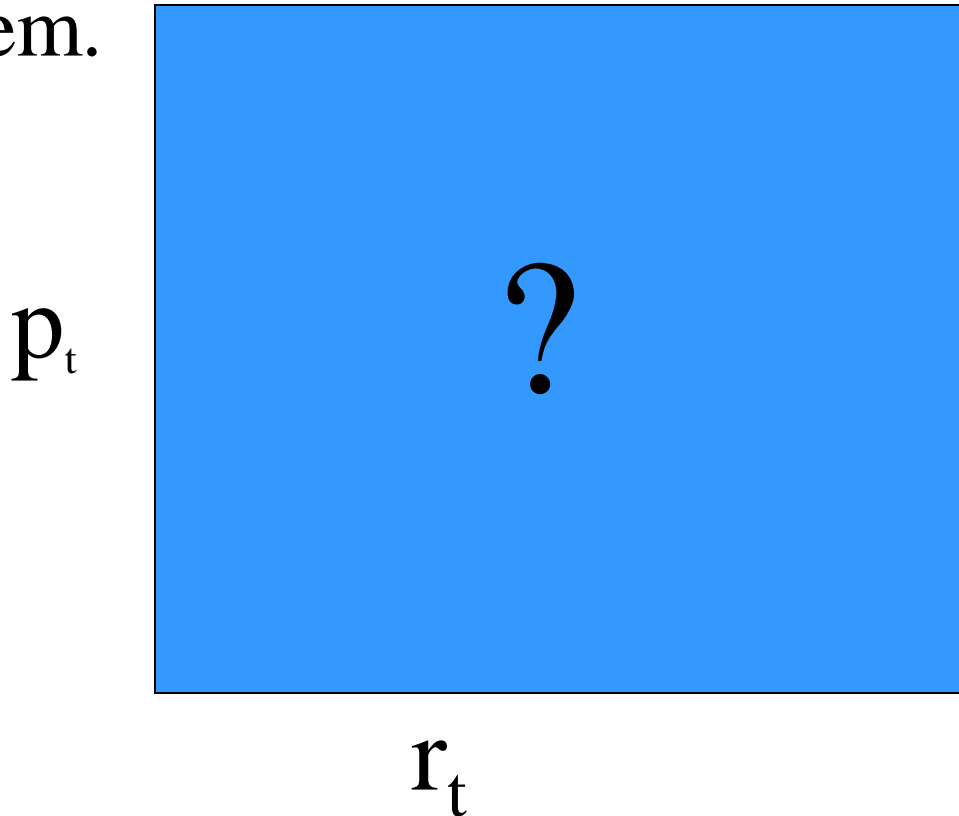
&

$$m_{t+1} = 2cf_t m_t + k(f_t + m_t) / (2d + k(f_t + m_t))$$



# DIFFERENTIAL SELECTION IN THE SEXES.

- This is a complex 2-dimensional dynamical system.



# End Point

- $f_{t+1} = \frac{2af_t m_t + h(f_t + m_t)}{2b + h(f_t + m_t)}$   
&

$$m_{t+1} = \frac{2cf_t m_t + k(f_t + m_t)}{2d + k(f_t + m_t)}$$

If  $f_t$  and  $m_t$  are small then we have  
approximately

$$f_{t+1} = (f_t + m_t) * h / 2b \text{ and } m_{t+1} = (f_t + m_t) * k / 2d$$

# End Point

- Thus 
$$\begin{pmatrix} f_{t+1} \\ m_{t+1} \end{pmatrix} = \begin{pmatrix} h/2b & h/2b \\ k/2d & k/2d \end{pmatrix} \begin{pmatrix} f_t \\ m_t \end{pmatrix}$$

and the eigenvalues of the matrix are

0 and  $\frac{h}{2b} + \frac{k}{2d}$  and so the A will invade

a population of B's if the latter  $> 1$  (heterosis!)

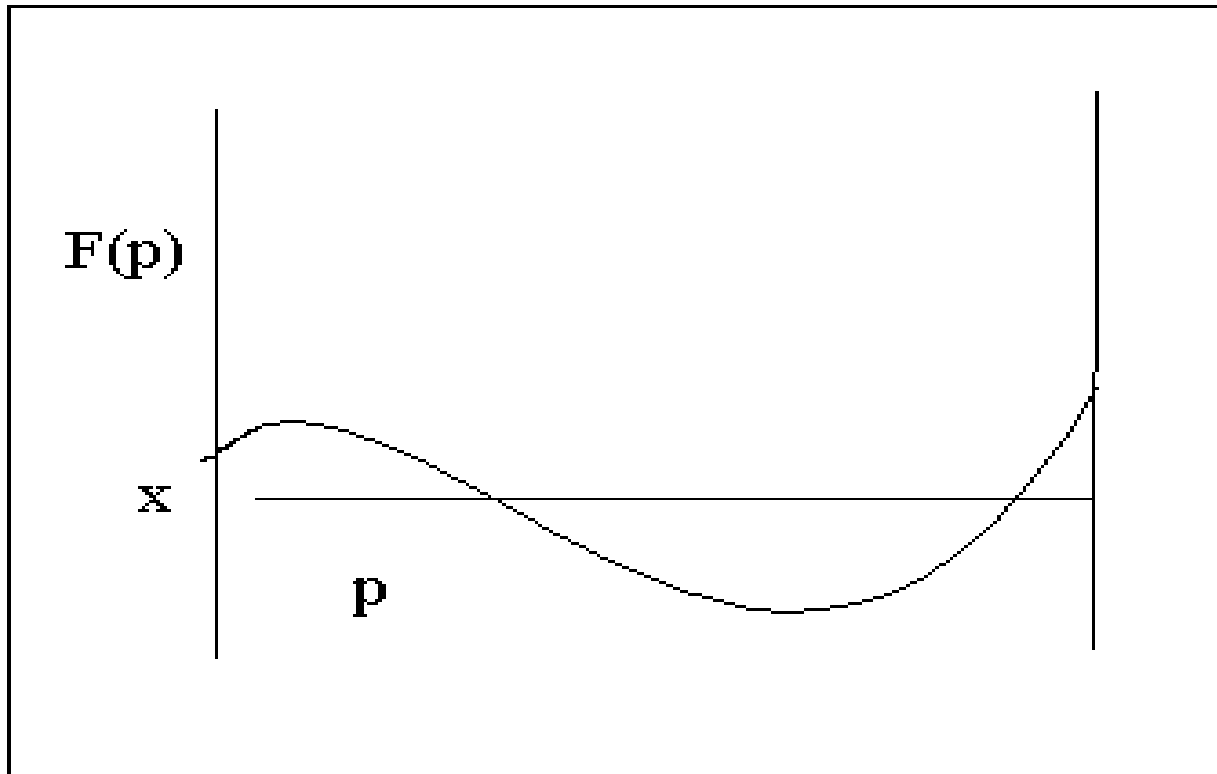
# DIFFERENTIAL SELECTION IN THE SEXES.

- If we had same selection in the sexes (AA viability  $a$ , Aa  $h$  and aa  $b$ ) then we would have a stable, polymorphism if, and only if,  $h > a, b$  with convergence assured provided start has both A and a present.

# DIFFERENTIAL SELECTION IN THE SEXES

- Equilibrium when  $p_{t+1} = p_t = p$  and  $r_{t+1} = r_t = r$  we obtain  $p=0$ ,  $p=1$  or any solution of a specific cubic for  $p$  [i.e.  $F(p)=(u \cdot p^3 + v \cdot p^2 + w \cdot p + x)=0$ ]. Thus we could have 2, 3, 4 or 5 equilibria, with up to 3 being internal (i.e. both alleles coexisting)

# DIFFERENTIAL SELECTION IN THE SEXES



# DIFFERENTIAL SELECTION IN THE SEXES

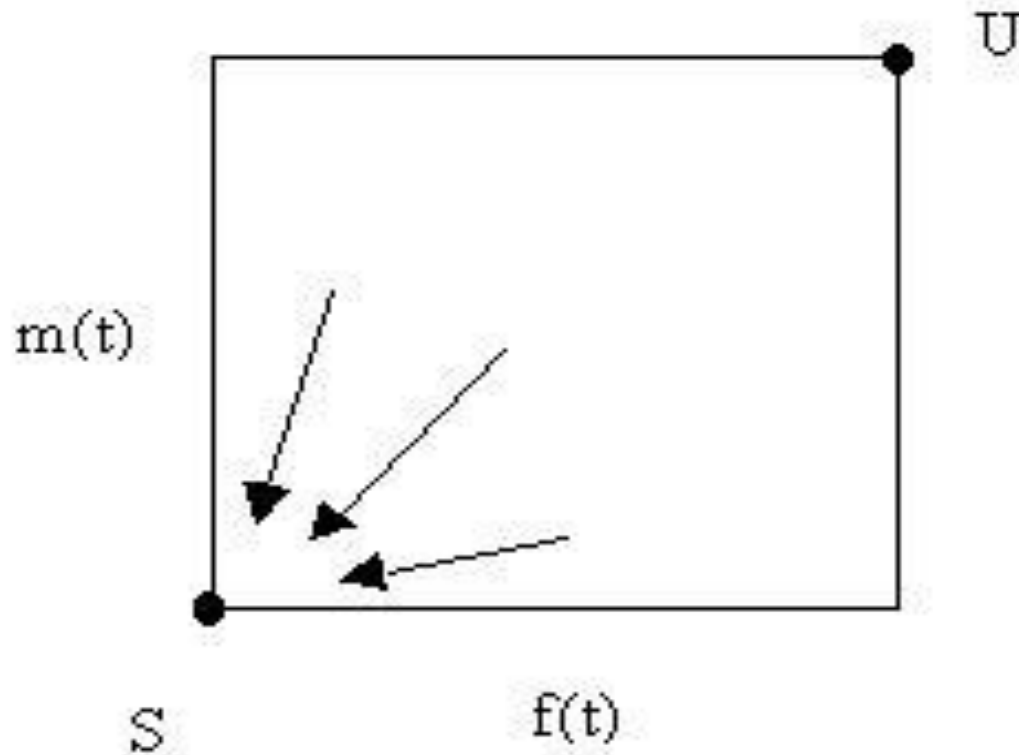
- $f_{t+1} = 2af_t m_t + h(f_t + m_t)/(2b + h(f_t + m_t))$   
 $m_{t+1} = 2cf_t m_t + k(f_t + m_t)/(2d + k(f_t + m_t))$   
and thus if  $f_t$  and  $m_t$  small we have  
 $f_{t+1} = h^*(f_t + m_t)/2b + \text{second order terms}$   
 $m_{t+1} = k^*(f_t + m_t)/2d + \text{second order terms}$   
so  $(0,0)$  is unstable if  $h/2b + k/2d > 1$   
a heterozygous advantage condition.

# DIFFERENTIAL SELECTION IN THE SEXES.

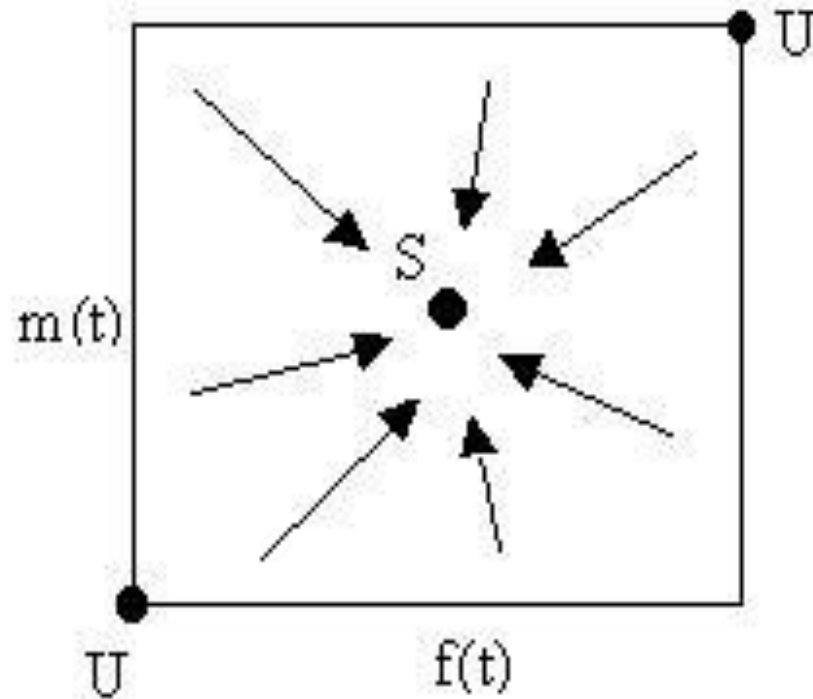
- It can be proved (with some difficulty) that irrespective of the starting position the system converges to an equilibrium (i.e. no cycles or chaos) Selgrade and Ziehe (1987).
- This enables us to move easily from the cubic to the dynamic (qualitatively).



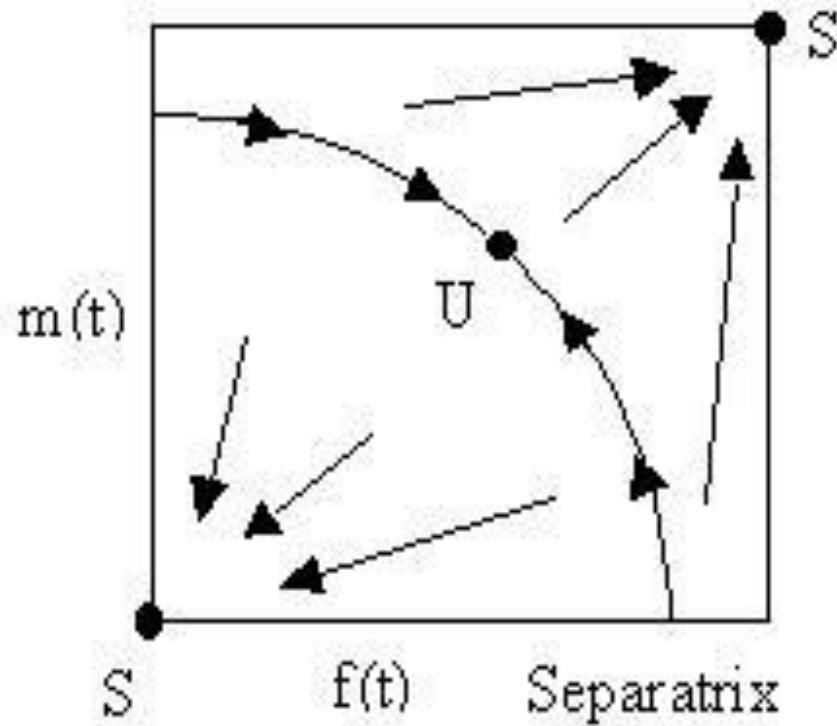
# DIFFERENTIAL SELECTION IN THE SEXES



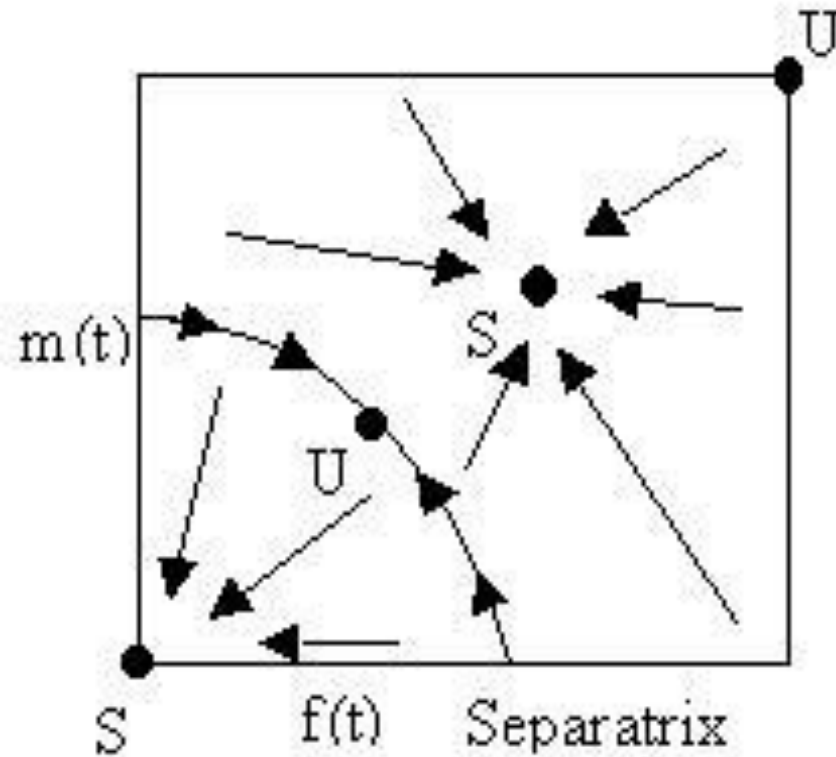
# DIFFERENTIAL SELECTION IN THE SEXES



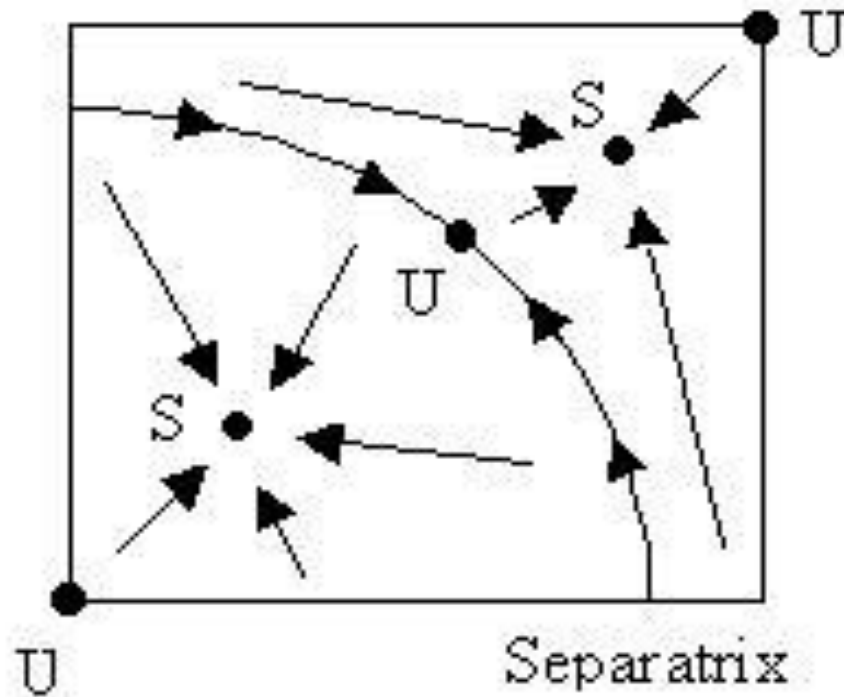
# DIFFERENTIAL SELECTION IN THE SEXES



# DIFFERENTIAL SELECTION IN THE SEXES



# DIFFERENTIAL SELECTION IN THE SEXES



# DIFFERENTIAL SELECTION IN THE SEXES

