3. Population dynamics, evolution

The dynamics of populations and evolution represent another application of the methods developed for the discussion of maps. In the following at first, a simplification is made in that the growth phenomena are treated as a smooth evolution along continuous degrees of freedom, even though it is clear that living entities are “quantized” and reproduce generation by generation. Consider first:

Binary cell division. At time t, assume that there is just one cell ($x_0 = 1$). In the second generation, after division, the number of cells is doubled (*n =2*). Hence, the (smooth) replication law is given by the differential equation (DEq)

$$\dot{x} = \frac{dx}{dt} = r \cdot x \quad \text{with} \quad \left(r = (n-1) = 1, 2, \ldots\right) \tag{31}$$

In principle, *r* can be a number between 0 and 1, if not all cells divide, divide at divide at different times, or even die. This will lead to an effective reproduction rate, as explained below.

DEq. 31 describes the gain in population and has the solution

$$x(t) = x(t = 0) \cdot e^{r \cdot t} \tag{32}$$

Of course, cells also die, at a rate

$$\dot{x} = \frac{dx}{dt} = -d \cdot x \tag{33}$$

which represents a loss to the population of cells. The average lifetime of a cell is $\tau = 1/d$. Combining gain and loss terms, one gets for the net rate of change in the population,

$$\dot{x} = \frac{dx}{dt} = \left(r - d\right) \cdot x = r_{eff} \cdot x \tag{34}$$

This section has largely been adapted from M. A. Nowak, *Evolutionary Dynamics*, Belknap Press of Harvard University Press, Cambridge, 2006

Obviously, r_{eff} is an effective reproduction rate corrected for the cell deaths. Therefore, if $r_{eff} > 0$, the population will grow exponentially, but with a reduced rate. If $r_{eff} < 0$, i.e., if more cells die than reproduce, the population will die out exponentially. For $r = d$, the cell population remains stable at the same number. Therefore, the quantity $r\text{-}d$ or r/d is the control parameter governing the evolution of the population of interest, its “fitness.”

Based on this simple math, one concludes that imposing the condition $r=d$ guarantees population stability. In practice, this is not possible, since both r and d are subject to fluctuations making rapid changes between $r/d > 1$ or $r/d < 1$. Nevertheless, actual populations do rarely grow or decay exponentially. Obviously, a description of a general population dynamics requires more complicated laws.

One obvious improvement to the law of Equ. 34 is based on the realization that r_{eff} itself is typically not a constant but depends on the size of the population $r_{eff} = r_{eff}(x)$. More specifically, it should depend (perhaps monotonically) on how close the actual population size is to the limiting size K (x = K), which is assumed to be a time-independent constant. Then, one has to consider a varying effective reproduction rate that is a function of the difference $(K\text{-}x)$, i.e., $r_{eff} = f(K\text{-}x)$ with $f(0)=0$. And, since reff has to decrease, one also has to require that $(d_{reff}/dx)_{x=K}<0$. For population sizes $x= K$, one can approximate the function r_{eff} to first order in a *Taylor* expansion around $x=K$. Then

$$\begin{aligned} r_{eff}(x) &= r_{eff}(x=K) - \left(\frac{dr_{eff}}{dx}\right)_K \cdot (K-x) + \ldots.. \\ &\approx \left(-\frac{dr_{eff}}{dx}\right)_K \cdot (K-x) = K \cdot \left(-\frac{dr_{eff}}{dx}\right)_K \left[1-\frac{x}{K}\right] \qquad (35) \\ &=: r \cdot \left[1-\frac{x}{K}\right] \end{aligned}$$

where r>0. The last line invokes a redefinition of the replication factor r. As x increases, the effective rate of reproduction slows down, which is plausible. Combining Equs. 34 and 35, one has, in essence, found a behavior of the population close to that described by the Logistic Map:

$$\dot{x} = \frac{dx}{dt} = r_{eff}(x) \cdot x = r \cdot x \cdot \left(1 - \frac{x}{K}\right) \qquad (36)$$

Here, r is no longer a function of x but a constant factor. Integrating this DE from initial condition $x=x_0$, one gets the solution

$$x(t) = \frac{K \cdot x_0}{K + x_0 \cdot \left(e^{r \cdot t} - 1\right)} \cdot e^{r \cdot t} \xrightarrow{t \to \infty} x^* = K \qquad (37)$$

With the asymptotic ($t \to \infty$) limit $x^* = K$. Interestingly, in equilibrium the population assumes its maximum possible size. Given the procedures discussed in Section 2.3, it is easy to show that this is, in fact, a stable state of the population.

Natural selection: Now the situation will be extended to include competition between two different subpopulations or species, A and B, with numbers x and y, respectively. Without any relations between them and no upper boundary, the subpopulations would each be described by the formalism just developed. Denoting with a and b the corresponding growth rates ("fitnesses") for A and B, the time dependent population sizes would both grow exponentially:

$$\begin{aligned} &\text{Species } A: \quad \dot{x} = a \cdot x \;\to x(t) = x(t) \cdot e^{a \cdot t} \\ &\text{Species } B: \quad \dot{y} = b \cdot y \;\to y(t) = y(0) \cdot e^{b \cdot t} \end{aligned} \qquad (38)$$

If one wants to make a comparison between the subpopulations, one can either consider the difference $\delta = (x(t) - y(t))$ or the ratio, $\rho(t) = x(t)/y(t)$. Obviously,

$$\rho(t) = \frac{x(t)}{y(t)} = \frac{x(0)}{y(0)} \cdot e^{(a-b) \cdot t} \qquad (39)$$

Therefore, if $b > a$, subpopulation B will outpace population A and grow indefinitely. In comparison to A, species B appears fitter will displace *A*. This is the principle of *selection or survival of the fittest.*

However, in nature there are always boundaries, to which both species would have to adjust, for example the total food supply that defines a maximum total population for the sum of *A* and *B*. How will the condition of a maximum total population

$$x + y = 1 \quad \rightarrow \quad \dot{x} + \dot{y} = 0 \tag{40}$$

normalized to unity, influence the competition between *A* and *B*? One would expect that the subpopulation with the greatest fitness will outgrow the competition but now only grow to the maximum possible size.

This can in fact easily be shown:
One suspects again that the reproduction rates (or fitnesses) of *A* and *B*, *a* and *b*, respectively, are no longer constants but are functions of the numbers (*x* and *y*) of *A* and *B* present at any given time. Therefore, instead of Equs.38, one may assume

$$\begin{aligned} \dot{x} &= a_{eff} \cdot x = \left[a - \phi(x,y)\right] \cdot x \\ \dot{y} &= b_{eff} \cdot y = \left[b - \phi(x,y)\right] \cdot y \end{aligned} \tag{41}$$

with a function that should depend (in the simplest case) linearly on either *x* or *y*. At the very beginning, when $x=y=0$, then one should also expect that $\phi(x,y)=0$. From Equ. 41 and the conservation of the total population (represented by Equ. 40), one has the condition for the function ϕ:

$$\begin{aligned} \dot{x} = \left[a - \phi(x,y)\right] \cdot x &= -\dot{y} = -\left[b - \phi(x,y)\right] \cdot y \\ ax - x\phi(x,y) &= -by + y\phi(x,y) \\ ax + by &= \underbrace{(x+y)}_{=1}\phi(x,y) = \phi(x,y) \end{aligned} \tag{42}$$

This determines the function

$$\boxed{\phi(x,y) = ax + by} \tag{43}$$

as the average fitness (reproduction rate) of the entire population. Upon insertion of the function into the rate equation for population *A,* one obtains

$$\begin{aligned}\dot{x} &= \left[a - ax - by\right]\cdot x = \left[a - ax - b(1-x)\right]\cdot x = \\ &= \left[a - b - (a-b)x\right]x = (a-b)\left[1-x\right]\cdot x\end{aligned} \tag{44}$$

and finally

$$\boxed{\begin{aligned}\dot{x} &= (a-b)\,x\cdot\left[1-x\right]\\ \dot{y} &= (b-a)\,y\cdot\left[1-y\right]\end{aligned}} \tag{45}$$

These functions are again reminiscent of the Logistic Map. What are the fixpoints of the DEq of Equ. 45? They are determined by the condition $\dot{x} = 0$, i.e., at $x=0$ ("extinction") and at $x=1$ ("take over") and correspondingly for y.

The situation described by Equs. 41 is a special case of

$$\begin{aligned}\dot{x} &= a\cdot x^{c} - \phi(x,y)\cdot x\\ \dot{y} &= b\cdot y^{c} - \phi(x,y)\cdot y\end{aligned} \tag{46}$$

with c=1. In that case, survival of the fittest always occurs, and the growth of the species with the highest fitness is exponential. However, for many species, two individuals have to meet to reproduce. The probability for two to meet increases with the second power of the population size (c=2)

$$\begin{aligned}\dot{x} &= a\cdot x^{2} - \phi(x,y)\cdot x\\ \dot{y} &= b\cdot y^{2} - \phi(x,y)\cdot y\end{aligned} \tag{47}$$

Here, the constancy of the size of the total population is guaranteed by the function

$$\phi(x,y) = ax^{2} + by^{2} \tag{48}$$

The resulting DEq is now much more complex than the Logistic Map and certainly highly non-linear. The general solution for any constant exponent *c* is given by,

$$\begin{aligned} \dot{x} &= f(x)\,x\cdot[1-x] \\ f(x) &= ax^{c-1} - b[1-x]^{c-1} \end{aligned} \tag{49}$$

which for *c=1* reduces to Equs. 45. Of course, there is the corresponding equation for rate in *y*. Because of the factors x and (1-x) in the rate equation (Equ. 49), fixpoints are again at $x_{1,2}=0,1$. This implies that for c=1, stable situations would result for homogeneous populations, either A or B, but no coexistence of both. The questions now are, whether the fixpoints are stable (attractors) or unstable (repellors). To decide that question, one can proceed and apply the criteria discussed in a previous section, i.e., calculate the derivatives of function (49) at the fixpoints.

However, for $c \neq 1$, there is one more fixpoint:

$$x_3 = \frac{1}{1+(a/b)^{1/(c-1)}} \tag{50}$$

i.e., there are three fixpoints altogether. It turns out that for different values of *c*, the behavior of the points x_1, x_2, and x_3 is different:

C	$\mathbf{x_1}$	$\mathbf{x_2}$	$\mathbf{x_3}$
<1	unstable	unstable	stable
=1	stable	stable	N/A
>1	stable	stable	unstable

This implies that for c>1, trajectories will be deflected away from x_3, either to x=0 or x=1. If *x (A)* > x_3, then *A* will outgrow *B* and displace it, and *vice versa*.

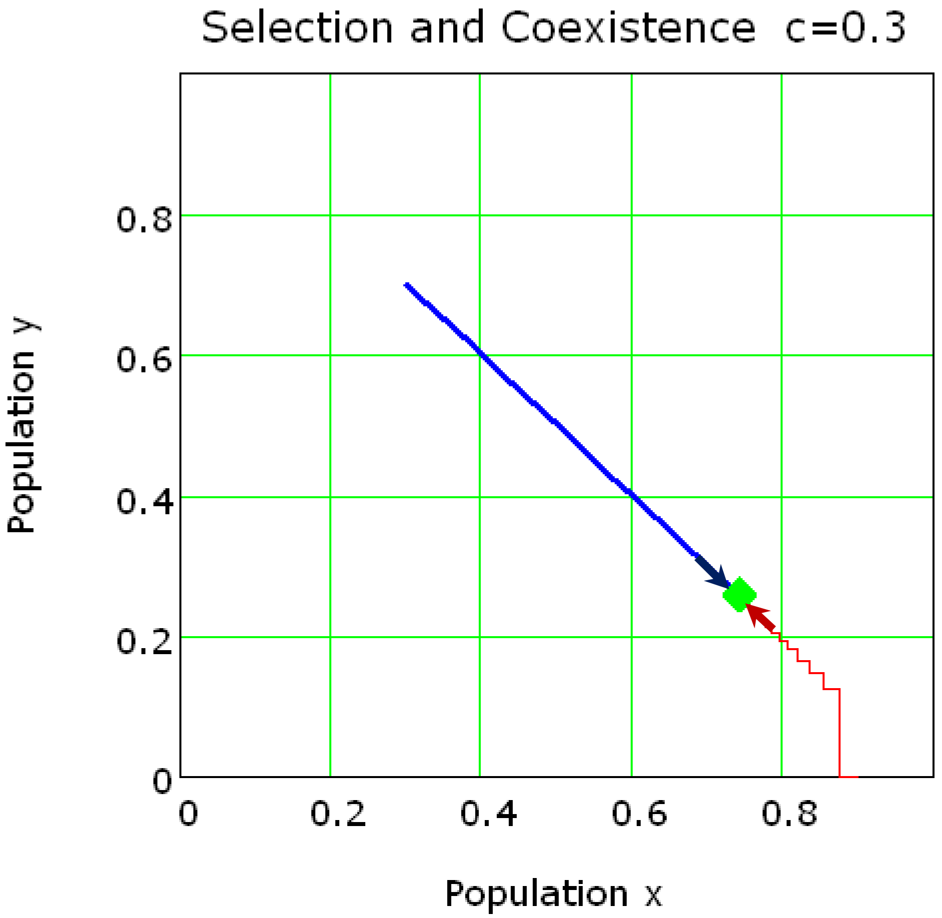

Figure 20: Population development for two initial conditions.

However, for *c<1,* the point 0 < x_3 < 1 is a stable fixpoint. This means that in this situation, populations *A* and *B* coexist with $x^*=x_3$ and $y^*=(1-x^*)$ denoting the stable population sizes for *A* and *B*, resp. These situations can be studied “experimentally” with the MathCad code PopDyn.mcd.

Mutation: Reproduction is not always perfect, sometimes *A* mutates to *B* and vice versa. For simplicity, consider only these two populations and fitness rates of $a=b=1$. Then also the mean fitness rate averaged over the entire population is unity: $\phi = 1$.

Assume the mutation rates m_A ($A\rightarrow B$) and m_B ($B\rightarrow A$), respectively. Then, one obtains from a consideration of gain and loss terms the rates of reproduction,

$$\begin{aligned}\dot{x} &= \left[1-m_A\right]\cdot x + m_B\cdot y - \phi\cdot x\\ \dot{y} &= m_A\cdot x + \left[1-m_B\right]\cdot y - \phi\cdot y\end{aligned}\tag{51}$$

and therefore,

$$\begin{aligned}\dot{x} &= \left[1-m_A\right]\cdot x + m_B\cdot\left(1-x\right) - x = m_B - \left[m_A+m_B\right]\cdot x\\ \dot{y} &= m_A\cdot\left(1-y\right) + \left[1-m_B\right]\cdot y - y = m_A - \left[m_A+m_B\right]\cdot y\end{aligned}\tag{52}$$

Again, because of the condition $x+y=1$, the two DEqs. (52) are linearly dependent and can be represented by just one,

$$\dot{x} = m_B - \left[m_A+m_B\right]\cdot x \tag{53}$$

Obviously, there is one fixpoint ($\dot{x}=0$) at

$$x_1 = \frac{m_B}{m_A+m_B} \qquad y_1 = \frac{m_A}{m_A+m_B} \tag{54}$$

This is a stable fixpoint, because the velocities for x-values close to x_1 drive the population always towards this point:

$$\begin{aligned}&\dot{x} = m_B - \left[m_A+m_B\right]\cdot x\\ &x < x_1 = \frac{m_B}{m_A+m_B} \rightarrow \dot{x} > m_B - \frac{m_B\cdot\left[m_A+m_B\right]}{m_A+m_B} = 0\\ &x > x_1 = \frac{m_B}{m_A+m_B} \rightarrow \dot{x} < m_B - \frac{m_B\cdot\left[m_A+m_B\right]}{m_A+m_B} = 0\end{aligned}\tag{55}$$

This implies that mutation provides a stable situation, where the two populations A and B coexist. It is plausible that the population with the smaller mutation rate is stable at a higher fraction of the total. Of course, if one mutation rate is much larger than the other, e.g., m_A » m_B˜ 0, then Equ. 53 degenerates to

$$\dot{x} = m_B - \left[m_A+m_B\right]\cdot x \approx -m_A\cdot x \tag{56}$$

This DEq. has an exponentially decaying solution x*(t)*. Population *A* becomes extinct, because there is a constant drain on it and no gain.

To generalize the formalism to many different (possible) populations, it is useful to use **vector algebra**. This is illustrated first with the above 2 populations A and B. Their time dependence can be considered in an **(***A***,** *B***)** or **(***x***,** *y***)** space. In this space, any partition of the total population can be represented by position vector $\vec{x}$. The corresponding rates of change make up the velocity vector $\dot{\vec{x}}$,

$$\vec{x} := \begin{pmatrix} y \\ x \end{pmatrix} \quad \text{and} \quad \dot{\vec{x}} := \begin{pmatrix} \dot{y} \\ \dot{x} \end{pmatrix} \tag{57}$$

Position and velocity vectors are connected by a 2x2 **Transition Matrix** *Q*

$$Q := \begin{pmatrix} m_A & [1 - m_B] \\ [1 - m_A] & m_B \end{pmatrix} \tag{58}$$

such that the velocity and position vectors are connected by the matrix equation

$$\boxed{\dot{\vec{x}} = Q \cdot \vec{x} - \phi \cdot \vec{x}} \tag{59}$$

Q is a **stochastic matrix**. It has no particular symmetry, except that the sum of the elements in each row is unity. Its elements are all in the interval [0, 1], as required for probabilities. The stochastic transition matrix distributes the original members of a population over all possible outcomes of the reproductions, such that the total population remains constant:

$$Q := (q_{ij}) \quad \text{with} \quad \sum_j q_{ij} = 1 \tag{60}$$

Now, for *n* different populations (x_i, $i=1,\dots,n$), the transition matrix is of the *n* x *n* type, and the corresponding matrix equation for the velocity vector is exactly equal to that of Equ. 59.

Finding the fixpoints is now equivalent to determining the eigenvalues of the stochastic matrix *Q,*

$$\vec{0} = Q \cdot \vec{x} - \phi \cdot \vec{x} \quad \rightarrow Q \cdot \vec{x} = \phi \cdot \vec{x} \tag{61}$$

Since the average fitness is again assumed to be equal to $\phi = 1$, the problem of finding a fixpoint, a possibly stable heterogeneous population, simplifies to

$$\left(Q - I\right) \cdot \vec{x} = \vec{0} \tag{62}$$

where *I* is the *n x n unit* matrix. This equation is equivalent to a set of *n* homogeneous equations with *n* unknown populations which are subject to one boundary condition (sum = 1). Hence, the system has non-trivial solutions for the populations $\{x_1, x_2, \ldots, x_n\}$.

Genetics: One of the early problems in understanding population dynamics was the fact that, without mutation, even after many generations the diversity of the population does not disappear. One does not end up with one genotype representing the completely homogeneous mix of all types that ever were present in a population. Today, it is known that genotypes do not "blend," but different parts of the genome are reshuffled and recombined in the offspring. The genetics is illustrated in the following simple example of reproduction of diploid organisms.

Consider sexual reproduction of diploid organisms possessing two matched sets of chromosomes (A_1 and A_2) in the cell nucleus, one set from each parent. The chromosomes are slight variants of the same genome and occur with probabilities

$$A_1 : p_1 \leq 1 \qquad A_2 : p_2 \leq 1 \qquad p_1 + p_2 = 1 \tag{63}$$

The chromosomes can combine to make 3 different genotypes, different types of individuals in the population, in each generational change (reproduction).

$$\{A_1, A_1\}: P_{11} \leq 1 \qquad \{A_1, A_2\}: P_{12} \leq 1 \qquad \{A_2, A_2\}: P_{22} \leq 1 \qquad (64)$$

with the corresponding different probabilities P_{ij}. The total probability has to add up to unity:

$$P_{11} + P_{12} + P_{22} = 1 \qquad (65)$$

Calculate the probabilities to find the two types of chromosomes in the population at some point in time (Generation X)

$$A_1 : p_1 = P_{11} + \frac{1}{2}P_{12} \qquad A_2 : p_2 = P_{22} + \frac{1}{2}P_{12} \qquad (66)$$

These expressions represent the fact that A_1 is contained twice in $\{A_1, A_1\}$ but only once in $\{A_1, A_2\}$, and similar for A_2.
Now assume that mating occurs randomly between genotypes. The random probabilities for producing the 3 genotypes in the following generation (Generation Y) are given by the relative probabilities (differentiated by a *prime*):

$$P'_{11} = p_1^2 \qquad P'_{12} = 2 \cdot p_1 \cdot p_2 \qquad P'_{22} = p_2^2 \qquad (67)$$

Obviously, the chromosome frequencies should not change, if none gets lost. That this is true in the formalism, one can see, e.g., by computing

$$p'_1 = P'_{11} + \frac{1}{2}P'_{12} = p_1^2 + p_1 \cdot p_2 = p_1 \cdot \left(p_1 + p_2\right) = p_1 \qquad (68)$$

and equivalent for $p'_2 = p_2$. Therefore, one also finds that the number of different genotypes in the Generation Y population has the same relation to the numbers of chromosomes than for Generation X:

$$P'_{11} = p_1^2 = p_1'^2 \qquad P'_{12} = 2 \cdot p_1 \cdot p_2 = 2 \cdot p'_1 \cdot p'_2 \qquad P'_{22} = p_2'^2 \qquad (69)$$