Phase-plane analysis of multispecies population dynamics
5.10,6.2,6.3,6.6,6.7

The chemostat revisited
We have learnt that

\[
\begin{align*}
s' &= (s_0 - s) - \frac{sb}{1 + s}, \\
b' &= asb - b,
\end{align*}
\]

has two equilibrium points

\[
(s_0, 0), \quad \text{and} \quad \left(\frac{1}{a-1}, a(s_0 - \frac{1}{a-1})\right),
\]

of which the second requires \(a > 1\) to exist biologically. When \(a < 1\) the first is stable; it is unstable when the second exist, which in such a case is stable.

To get further insight we look at the isoclines, i.e., the curves where the trajectories are parallel to one of the axes. In this case the equation \(s' = 0\) corresponds to the curve \(b = (s_0 - s)\frac{1}{1+s}\) and the equation \(b' = 0\) to straightline: \(b = 0\) and \(s = \frac{1}{1+a}\).

In the graph below the isoclines are drawn in red and arrows are drawn to show the direction of the vector field. They all have unit length, which sometimes gives the wrong impression – e.g., all vectors on the curve are in the vertical direction. Also included are two example trajectories (blue).

Drawing sample arrows based mostly on signs of the components in order to indicate the direction of the trajectories is what a phase-plane analysis amounts to.

Modeling an epidemic
Introduction
The bubonic plague struck the small village of Eyam in England in the middle of the 1660th. Eyam, lying about 20 km away from Sheffield, had at the time 350 inhabitants. The epidemic was a side-show of the well-known epidemic in London 1664-1666 and started when a tailor received some clothes from London 1665 containing plague-infested fleas. The first Eyam victim was buried on the September 7 and during the next 9 months another 76 villages were buried without the epidemic really getting up to speed. Which it did during the summer of 1666:

<table>
<thead>
<tr>
<th></th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>19</td>
<td>56</td>
<td>77</td>
<td>24</td>
<td>14</td>
</tr>
</tbody>
</table>

Our problem is this: given the data at the end of August, can we say anything about what will happen to the village? We will assume that on June 1 there were 273 inhabitants, of which 2 were infected with the plague.

The model we will work with was published in 1927 by Kermack and McKendrick and provided a reasonably accurate description of the plague epidemic of Bombay in 1905-06.

The model equations
At each time point \(t\) we can divide the population of the village into three classes:

\[
S(t) = \text{the number susceptible to the disease},
\]

\[
l(t) = \text{the number with the disease},
\]

\[
R(t) = \text{the number removed, i.e. those that have had the disease, now being immune or dead}.
\]

Note that it is an assumption that once you have had the disease, you do not get it again. For the plague we will in fact assume that the disease has 100% mortality. Also note that

\[
S(t) + l(t) + R(t) = N
\]

where \(N\) is the initial village population.

Remark: We talk about the number of..., but actually what is relevant from the model perspective is the concentration of ... That should be kept in mind.

According to well-known mass-balance arguments the perhaps simplest model would be

\[
\begin{align*}
S'(t) &= -r l(t) S(t), \\
l'(t) &= r l(t) S(t) - \sigma I(t), \\
R'(t) &= \sigma I(t)
\end{align*}
\]

which assumes a process so fast that other causes of death are balanced by births and the village population is kept constant constant. The model constants have the following interpretation: \(r\) measures the degree of infectiveness of disease (how fast it spreads) and \(\sigma\) is related to how long an individual stays infective and can spread the disease.

Remark: The rate at which new infectives appear is according to the law of mass action well-known from chemistry. We can argue for it as follows. Assume all individuals mix well, and that during a short time interval \(h\) the probability is \(p\) that a particular infective transfers the disease to a particular susceptible. With perfect mixing and \(l\) infectives, the probability that a particular susceptible does not get infective is \((1-p)^l = e^{-al}\), where \(a = -ln(1-p)\). The total number of infectives at the end of the time interval is therefore \(e^{-al}S\), which means that the number of new infectives is \((1-e^{-al})S = alS = rlh\) if \(r = d'(0)\). The argument is the same in chemistry.

Remark: The opposite of a 100% lethal disease would be one that imposes no immunity at all, but returns the infectives to the susceptibles once disease-free. In that case we have \(l(t) + S(t) = N\) and get the single-equation model

\[
l'(t) = r l(t) (N - l(t)) - \sigma I(t),
\]

an equation we have encountered in a different setting.

Returning to the original model, we see that

\[
R(t) = \sigma \int_0^t l(s) ds,
\]
which is what we can measure. It also means that we can focus on the dynamics of \((S(t), I(t))\) to understand the disease.

**Example** If we use the data for June-August to determine the two unknown constants, it can be shown that the parameters that best fit available data are (time is months)

\[
r = 4.4226/273 \text{ and } \sigma = 2.4156.
\]

The model implications and fit to data is shown below.

Note that \(I/\sigma\) can be interpreted as the expected length of time an individual is infective, which in this case is about 12 days. We can use the model to compute both the predicted number of cases during September and October, which turns out to be 36 and 12, respectively. We can also predict that, when the epidemic is over, a total of 71 villages will have survived it. The true number was 83. Note that the model does not take into account that \(t\) probably is season-dependent.

As pointed out above, it is \(R(t)\) (or rather \(R'(t)\)) we observe (number of dead per time unit). It is possible to obtain a approximate expression for this along the following lines. We have, with \(\mu = 1/5\), that \(S'(t) = -\mu S(t)R'(t)\), which implies that

\[
S(t) = S(0) \exp(-\mu R(t)).
\]

This gives us

\[
R'(t) \approx \sigma (N - R(t) - S(0)(1 - \mu R(t) + \mu^2 R(t)^2/2)) = \\
\sigma (N - S(0)) + (\mu S(0) - 1)R(t) - (S(0)/2)(\mu R(t))^2,
\]

provided \(\mu R(t)\) is small. This equation can be solved, and after some calculations one find that \(R'(t)\) can be expressed in terms of a cosh-function. This function defines the epidemic curve of the disease outbreak.

**Analysing the SIR-model**

To analyse the model we first nondimensionalize it, so that \(S\) and \(I\) are measured as fractions of the total population (and with new time \(rNt\)). This give us

\[
\begin{align*}
S' &= -SI \\
I' &= (S - \rho)I \\
\rho &= \sigma/rN.
\end{align*}
\]

We see that \(S' < 0\) always, whereas \(I' > 0\) precisely when \(S > \rho\). Thus:

There will be no epidemic if \(S(0) < \rho\). If \(S(0) > \rho\) the function \(I(t)\) will have a global maximum when \(S(t) = \rho\) and decrease to zero thereafter.

**Remark** The number

\[
R_0 = \frac{rS(0)}{\sigma},
\]

is called the basic reproduction rate of the infection, and represents the number of secondary infections produced by a primary infection in a wholly susceptible population. If more than one secondary infection is produced from one primary infection, that is \(R_0 > 1\), an epidemic ensues.

**Example** In the Eyam-case we have \(\rho = 2.4156/4.4226 = 0.546\), which means we should have an outbreak. To prevent this, we need to increase \(\rho\) to about twice its size. The only factor we can influence is \(r\), which essentially means keeping people from mixing.

From the equations we easily deduce that \(dI = -dS + \rho \frac{dS}{S}\) which we can integrate to (since \(I(0) + S(0) = 1\))

\[
I(t) = 1 - S(t) + \rho \ln \frac{S(t)}{S(0)}.
\]

Plotted in the phase-plane the \((S, I)\)-curve looks like below: the red curve is for \(\rho = 0.546\) and the blue curve for \(\rho = 0.3\).

When \(S(t) \to 0\) we get \(I(t) \to -\infty\), which means that there is some time \(t_1\) such that \(I(t_1) = 0\) and \(S_{\infty} = S(t_1) > 0\). Thus

The epidemic ends before the reservoir of susceptibles is depleted – some individual will survive the epidemic without having been infected.

We can note that \(S_{\infty}\) satisfies the equation

\[
S_{\infty} - \rho \ln S_{\infty} = 1 - \rho \ln S(0).
\]

**Example** In Eyam-case this equation becomes \(S_{\infty} - 0.546 \ln S_{\infty} = 1 - 0.546 \ln 273\), which has the solution \(S_{\infty} = 0.259\). The number of survivors is 273 times this, namely 71. In fact, we can use the equation to estimate \(\rho\) also. In the Eyam-case this gives the estimate 0.59, similar to the one obtained from 3 month data.

**Consequences of model**

One consequence of the model is that we get an idea of what it takes to eradicate a disease. What we need is to ensure that \(S < \rho\), where \(S\) is the fraction of the population that is susceptible to the disease. For this we need to get an estimate of the contact number \(\rho\). For smallpox this was considered to be about 0.2 worldwide, which means that we need \(S < 0.2\). In other words, we need to vaccinate 80% of the population. That was doable, and a worldwide program started in 1958 and smallpox was formally declared eradicated in 1977.

For measles \(\rho = 0.06\), which means that we need to vaccinate 94% of the population. That is doable in a country like Sweden, but hardly world-wide.

**A modified model – the SIRS model**

With the model above the disease disappears when the epidemic is over. This does of course not happen in real life for various reasons: newborns added to susceptibles, regional aspects, animal reservoirs etc. For some diseases one reason may be that the immunity is not life-long. How will the situation change in such cases?

Assume the situation with a non-fatal disease and with a limited time spent being immune. After non-dimensionalisation the equations
then changes to

\[ \begin{align*}
S'(t) &= \kappa R(t) - I(t)S(t), \\
I'(t) &= I(t)S(t) - \rho I(t), \\
R'(t) &= \rho I(t) - \kappa R(t)
\end{align*} \]

(Again the total population is constant.) The equilibrium points for this is determined by

\[ \kappa R - IS = 0, \quad (S - \rho)I = 0, \quad R = \frac{\rho}{\kappa} I. \]

One solution is \( I = R = 0, S = 1 \), the other is (note that the sum should be one)

\[ S = \rho, \quad I = \frac{1 - \rho}{1 + \rho/\kappa}, \quad R = \frac{\rho}{\kappa} I, \]

which only exists when the threshold criterion is fulfilled. It is easy to show that it is always stable (first we reduce it to a system in \( S, I \) only – the trajectories spiral into the equilibrium).

**Lotka-Volterra equations**

In the middle of the 1920th, the italian biologist Umberto D’Ancona took an interest in how different fish populations interact. Among his data was the following table, which describes what percentage of the area. Rabbits eat lettuce and carrots and squirrels eat lettuce and nuts. Example

\[
\begin{array}{ccc}
1914 & 1915 & 1916 \\
1917 & 1918 & 1919 \\
1920 & 1921 & 1922 \\
1923 & & \\
11.9 & 21.4 & 22.1 \\
21.2 & 26.4 & 27.3 \\
16.0 & 15.9 & 14.8 \\
10.7 & & \\
\end{array}
\]

What puzzled D’Ancona about these data was that the relative increase in predators during the period of WWI, a time when fishing was reduced.

His good friend, the mathematician Vito Volterra, came up with a mathematical explanation 1926. With \( P \) being the concentration of predators and \( N \) the concentration of prey, he proposed the pair of differential equations

\[
\begin{align*}
N' &= rN - qNP \\
P' &= pNP - sP.
\end{align*}
\]

This system of equations is called the **Lotka-Volterra model**, because it was formulated also by Lotka a few years earlier for a hypothetical chemical reaction.

To analyse the system and get Volterra explanation we first nondimensionalize the system: let the new time be \( rt \) and introduce \( u = pN/s, \quad v = qP/r \). That gives us the system

\[
\begin{align*}
u' &= u(1 - v) \\
v' &= av(v - 1), \quad a = \frac{s}{r}.
\end{align*}
\]

We see that there are four equilibrium: (0, 0), (1, 0), (0, 1) and (1, 1). Of these only the last interest us, for obvious reasons. The system derivative is \( \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \), so the eigenvalues are \( \lambda = \pm i \), which means that we cannot determine if it is stable or not. This is actually for good reasons, because we can find an implicit equation for the solution by rewriting the equation as

\[
\frac{(1 - v)dv}{v} = a(u - 1)du, \quad a = \frac{s}{r}.
\]

which can be integrated to \( \ln v - v = a(u - \ln u) + C \). Exponentiating and we have the implicit equation

\[ ve^{-v} = Ku^{-a} e^{\alpha u}. \]

We can actually plot the trajectories by a trick, illustrated in the figure below.

It follows that the trajectories are closed curves around \((1, 1)\). This also means that the equilibrium is stable, but not asymptotically stable.

The catch reflects the average population over a year. By integrating \( du/u = 1 - v \) and using the periodicity we see that the integral over a period of \( u \) is the length of the period, so the average is one. The same is found for \( v \). But this is the equilibrium point, which corresponds to \( N = s/p, \quad P = r/q \) in the original system.

So far we have considered the ecological system without harvesting. If we assume that fishing removes a fraction \( \epsilon \) of each fish type, this is equivalent with replacing \( r \) with \( r - \epsilon \) and \( s \) with \( s + \epsilon \) in the original equations. The mean values are therefore \( (s + \epsilon)/p, (r - \epsilon)/q \). Thus we see that fishing diminishes the predator population and increases the preys. A diminished fishing level will lead to the result observed in the data.

**Competing species**

A different problem involving two populations of animals is the question of what happens when they compete for the same resources. For a single species we have the basic logistic equation

\[ N' = rN(1 - N/K), \]

where \( K \) is the carrying capacity for this species. If another species uses the same food resource, only a part of this is accessible to the first species. A simple and robust way of modeling this is to write

\[ N' = rN\left(1 - \frac{N + m}{K}\right), \]

where \( m \) is the amount of food resources occupied by the second species. In general we assume \( m = aM \), where \( a \) measures how much the second species affect the food availability for the first species. If the second species eat much more than the first species we would have \( a > 1 \). Assuming something similar for the effect of the first species on the second, we would produce the simple model

\[
\begin{align*}
N' &= rN\left(1 - \frac{N}{K} - \frac{a}{M} \right) \\
M' &= sM\left(1 - \frac{M}{T} - \beta \frac{N}{L} \right).
\end{align*}
\]

To nondimensionalize this system, introduce the new time \( rt \) and \( x = N/K, y = M/L, a_{12} = aL/K, a_{21} = \beta K/L \) to get the equations

\[
\begin{align*}
x' &= x(1 - x - a_{12}y) \\
y' &= \rho y(1 - y - a_{21}x), \quad \rho = s/r.
\end{align*}
\]

**Example** Assume that rabbits (x) and squirrels (y) live in the same area. Rabbits eat lettuce and carrots and squirrels eat lettuce and nuts. Since the lettuce the rabbits eat are part of the squirrels diet, they reduce the growth of the squirrel population. But since the squirrels also eat nuts, which the rabbits do not eat, the rabbits do not affect the
growth of squirrels as much as they do themselves. This means that $a_{12} < 1$.

If the nuts suddenly disappear, squirrels can only eat lettuce. In that situation the rabbits affect the growth for squirrels more than they affect their own growth, so we have $a_{21} > 1$.

To analyse the model, we first find the equilibrium which are up to four: $(0,0), (1,0), (0,1)$ and $(x^*, y^*)$. The last occurs when there is a positive solution to the equations

$$1 - x - a_{12}y = 0, \quad 1 - y - a_{21}x = 0$$

which occur when the two constants are on the same side of one. To investigate their stability, we first compute the system derivative:

$$\begin{pmatrix} 1 - 2x - a_{12}y & -a_{12}x \\ -\rho a_{21}y & \rho (1 - 2y - a_{21}x) \end{pmatrix}.$$  

Obviously the origin is an unstable equilibrium, with eigenvalues $1$ and $\rho$. For the point $(1,0)$ the eigenvalues are $-1$ and $\rho(1 - a_{21})$, which makes it stable if $a_{21} > 1$. The equilibrium $(0,1)$ is stable if $a_{12} > 1$.

Note that if both $a_1$ and $a_2$ are greater than one, both are stable. In that case the initial conditions will determine which equilibrium the system ends up in. In this case the fourth equilibrium is a saddle point, and its “stable” orbit will serve as a separatrix and divide the first quadrant into the domain attraction regions for the two equilibria.

Finally, at the fourth equilibrium we have

$$\begin{pmatrix} 1 - 2x^* - a_{12}y^* & -a_{12}x^* \\ -\rho a_{21}y^* & \rho (1 - 2y^* - a_{21}x^*) \end{pmatrix} = \begin{pmatrix} -x^* & -a_{12}x^* \\ -\rho a_{21}y^* & -\rho y^* \end{pmatrix},$$

which gives the characteristic polynomial

$$(\lambda + x^*)(\lambda + \rho y^*) - \rho a_{12}a_{21}x^*y^* = \lambda^2 + (x^* + \rho y^*)\lambda + \rho(1 - a_{12}a_{21})x^*y^*$$

This is stable precisely if $1 - a_{12}a_{21} > 0$, which means that both parameters must be less than one. In summary we see that the values of $a_{12}, a_{21}$ determine if one species will outcompete the other, or if they will be able to co-exist.

**Exercises**

**Exercise 1** Derive the epidemic SIR-model from the discrete time Reed-Frost model of lecture F3.

**Exercise 2** Determine the stability of the equilibria of the epidemic SIRS-model. Make a phase-plane analysis of the system.

**Exercise 3** Species may derive mutual benefit from their association; this type of interaction is known as mutualism. A possible model for this is

$$N'_1 = rN_1(1 - \frac{N_1}{K_1 + aN_2}), \quad N'_2 = rN_2(1 - \frac{N_2}{K_2 + bN_1}).$$

Explain how it captures mutualism and analyse it using phase-plane methods.

**Exercise 4** The following model has been suggested for the interactions between baleen whales and their main food source, krill, in the southern ocean:

$$x' = rx(1 - x/K) - axy, \quad y' = sy(1 - y/bx).$$

Explain this model and analyse it by phase-plane methods.

**Exercise 5** Estimating the species-competition parameters.

1. Show that

$$a_{12} = \frac{K_1(1 - N'_1)}{r_1 N_1 N_2}, \quad a_{21} = \frac{K_2(1 - N'_2)}{r_2 N_1 N_2}$$

and suggest how this observation can be used to estimate the parameters.