Biomathematics

Marco Di Francesco

January 7, 2022

Contents

Chapter 1

Continuous population models for single species

1.1 The Malthus' model

Let us consider a population $N(t) \geq 0$ (of animals, cells, bacteria, etc.) depending on time $t \geq 0$. The basic approach of continuous models in population dynamics requires information on "how the population changes in time", more precisely

$$\frac{dN(t)}{dt} = \text{change of the number of individuals per unit time.}$$

It is reasonable to write the right-hand side above in terms of a *rate* of change, namely

$$\frac{dN(t)}{dt} = \text{rate of change of the number of individuals per unit time } \times \; N(t).$$

The effects contributing to the above rate of change can be grouped into *growth* mechanisms and *decrease* mechanisms. The former usually account of new births, the latter of deaths and/or migration phenomena. In the classical Malthus model [1] (dating back to 1798), the rate is assumed to be *constant*. Hence, we write

$$\frac{dN(t)}{dt} = (b - d)N(t), \tag{1.1.1}$$

where $b > 0$ is the *birth rate* and $d > 0$ is the *death rate*. The linear differential equation (1.1.1) is explicitly solvable as

$$N(t) = N(0)e^{(b-d)t}\,.$$

Clearly, if $b > d$ the population grows indefinitely as $t \to +\infty$, whereas it dies out to zero if $b < d$.

1.2 The Logistic model

The indefinite growth for $b > d$ makes (1.1.1) somewhat unrealistic, as this is not what happens in real life. The finiteness of the available resources for the population makes a slower growth rate more likely when the population is very large. This makes a dependency on $N(t)$ of the rate of change of the population more realistic. A first example in this direction is provided by the *logistic model* introduced by Verhulst in 1838 [2], in which a *carrying capacity* $K > 0$ is introduced as a parameter accounting for the available sustainable resources for the population $N(t)$. The models reads

$$\frac{dN(t)}{dt} = rN(t)\left(1 - \frac{N(t)}{K}\right), \tag{1.2.1}$$

where $r > 0$ is the growth rate of the population for small values of $N(t)$. The rate of change of the population is now given by

$$r\left(1 - \frac{N(t)}{K}\right)$$

which depends on N and is positive for $N < K$ and negative for $N > K$. In particular, if the population is higher than the carrying capacity the growth is not sustainable and the population decreases.

The differential equation (1.2.1) can be easily solved by separation of the variables:

$$\begin{aligned}
&\frac{dN}{N\left(1 - \frac{N}{K}\right)} = rdt \\
&\int_{N(0)}^{N(t)} \frac{dN}{N\left(1 - \frac{N}{K}\right)} = rt \\
&\int_{N(0)/K}^{N(t)/K} \frac{dM}{M\left(1 - M\right)} = rt \\
&\log\left(\frac{M}{|1 - M|}\right)\Big|_{N(0)/K}^{N(t)/K} = rt,
\end{aligned}$$

which leads, after few lines of calculations, to

$$\frac{N(t)\left|1 - \frac{N(0)}{K}\right|}{N(0)\left|1 - \frac{N(t)}{K}\right|} = e^{rt}\,.$$

Since the Cauchy problem for (1.2.1) admits a unique global solution (exercise!) and since $N(t) \equiv K$ is a stationary solution, the sign of $1 - \frac{N}{K}$ is preserved in time. Therefore, we can remove the absolute values and get

$$N(t) = \frac{KN(0)e^{rt}}{K - N(0) + N(0)e^{rt}}\,.$$

In this simple model we were able to find an explicit solution, which allows to detect properties of the solution which can be interpreted with respect to the population growth problem we are considering. For small times and for $N(0) < K$, $N(t)$ behaves like

$$N(t) \approxeq N(0)e^{rt}$$

which is the solution to the Malthus equation with growth rate r. For larger times, we observe that $N(t) \to K$ as $t \to +\infty$. The constant K plays somehow the role of a sustainable size of the population given the conditions of the surrounding environment. Notice that if $N_0 > k$ the population size is decreasing in time for all $t \geq 0$. This represents a situation in which, due to compensating effects of overcrowding, births are less than deaths and migration effect.

1.3 General population models for one species

Depending on the specific situation we may be dealing with, we may consider a population to be governed by

$$\frac{dN(t)}{dt} = f(N(t))\,, \tag{1.3.1}$$

where $[0,+\infty) \ni N \mapsto f(N) \in \mathbb{R}$ is a general nonlinear function of N. We assume for simplicity that f is $C^1(\mathbb{R})$. Equilibria N^* are solutions to $f(N^*) = 0$. From a simple linearised stability analysis, we perturb N^* as

$$N(t) = N^* + n(t)\,,$$

with $|n(t)|$ small, and we compute

$$\frac{dn}{dt} = f(N^* + n) \sim f(N^*) + nf'(N^*) + \text{higher horder terms}\,.$$

It is well known that, for small values of n, the solution $n(t)$ is well approximated by the solution to the linear problem

$$n(t) = n(0)e^{f'(N^*)t},$$

which yields, for the solution $N(t)$ to (1.3.1),

$$N(t) \approxeq N^* + (N(0) - N^*)e^{f'(N^*)t}\,.$$

Hence, the steady state N^* is asymptotically stable if $f'(N^*) < 0$, whereas it is unstable is $f'(N^*) > 0$. A more general situation with more than one equilibrium can be considered. As a simple exercise, assume there exists three stationary solutions $0 < N_1 < N_2 < N_3$. Assume

$$f(N)\begin{cases} > 0 & \text{if } N \in [0, N_1) \\ < 0 & \text{if } N \in (N_1, N_2) \\ > 0 & \text{if } N \in (N_2, N_3) \\ < 0 & \text{if } N > N_3\,. \end{cases}$$

By choosing $N(0) \in [0, N_1)$, the unique solution $N(t)$ to (1.3.1) exists at least for short times. Moreover, since $f(N_1) > 0$, $f(N(t)) < 0$ for the local-in-time existence time interval due to the preservation of the sign for continuous functions, so $N(t)$ will be increasing for short times. By a simple continuation principle for differential equations, $N(t)$ cannot cease to exist in a finite time, because it cannot "touch the boundary of the domain of f" (that is $\mathbb{R}$) and it cannot "blow-up" because otherwise it should "touch" the line $N = N_1$, but this would imply more than one solution to the Cauchy problem for (1.3.1) with initial condition N_1 at the time in which N "touches" N_1. Hence, $N(t)$ exists globally and is always increasing, since it remains in the interval $[0, N_1)$. Now, monotone functions have limits at $+\infty$, and so does $N(t)$. Said limit cannot be $< N_1$, because otherwise the time derivative of N would have a non-zero limit at infinity. Clearly, said limit can only be positive, but this implies $N(t)$ would "stay above a strictly increasing straight line", which contradicts the fact that N must stay below N_1 for all times. Hence, the only possible limit for $N(t)$ at $+\infty$ is N_1. By a similar argument, one can prove that a solution with $N(0) \in (N_1, N_2)$ is always decreasing in time and tends to N_1 as $t \to +\infty$. Hence, N_1 is *globally stable* for initial data in $[0, N_2)$. A similar arguments shows that N_3 is globally stable for initial data in $[N_2, +\infty)$. The advantage of the global analysis performed in this example is that we can predict the asymptotic behavior of $N(t)$ not just for small perturbation of the equilibria but for arbitrary initial conditions.

1.4 A model with predation: the "spruce budworm"

The stability or instability of a steady state is not the only relevant information one can get out of a mathematical model. When the model exhibits dependency on several parameters, one can see how those parameters affect the qualitative behavior of the model. In this example, $N(t)$ is the population of the "spruce budworm", an insect which can defoliate trees, so the control of its population is of practical interest. The model is nothing but a logistic equation plus a *predation term*:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) - p(N)\,, \tag{1.4.1}$$

where

$$p(N) = \frac{BN^2}{A^2 + N^2}\,,$$

which models the fact that predation (e.g. by birds) typically saturates for large N. In (1.4.1), r, K, A, B are four positive parameters. We already know the meaning of r and K. Here B is the "saturated" predation rate, that is the predation rate attained for large population size. As for A, a simple exercise shows that $p(N)$ has an inflection point at $N = \frac{A}{\sqrt{3}}$, which is also the point at which the slope of p reaches its maximum. Therefore, A is a parameter denoting a "switch" mechanism, a threshold value in which the effect of the predation term changes drastically, from very low to a significant one.

As we will do very often, we will now "rescale" the model (1.4.1) with the goal of *reducing the number of parameters*. In fact, very often the interpretation of the results involves "groups" of parameters. There are many possible ways to rescale a model. Typically, the new variables are *dimensionless*. In this case, we decide to compare N with the aforementioned switch parameter A, which has the same dimension as N. So we set

$$u = \frac{N}{A}\,.$$

Quite naturally, while trying to "remove" all constants from the predation terms, we end up with a multiplier A/B of the time derivative. Therefore, we may define a new time variable

$$\tau = \frac{B}{A}t$$

and $u(\tau) = N(t)$, which gives

$$\frac{du}{d\tau} = \frac{rA}{B}u\left(1 - \frac{Au}{K}\right) - \frac{u^2}{1+u^2}\,.$$

This suggests the introduction of two parameters

$$\rho = \frac{rA}{B}\,, \qquad q = \frac{K}{A}\,,$$

which make (1.4.1) re-writable in the scaled version

$$\frac{du}{d\tau} = u\left[\rho\left(1 - \frac{u}{q}\right) - \frac{u}{1+u^2}\right]\,. \tag{1.4.2}$$

Similarly to the all the models studied so far, $u = 0$ is a steady state. In order to detect other possible steady states, let us set

$$g(u) = \rho \left(1 - \frac{u}{q}\right) , \qquad h(u) = \frac{u}{1 + u^2} .$$

A nontrivial steady state u^* is found solving the nonlinear equation $g(u^*) = h(u^*)$. On the positive half-line $u \geq 0$, $g(u)$ is a decreasing straight line with $g(0) = \rho$ and with $g(q) = 0$. We now compute

$$h'(u) = \frac{1 - u^2}{(1 + u^2)^2}$$

which gives that h is increasing on $[0, 1)$, it has a global maximum $h(1) = 1/2$, and is decreasing on $u > 1$. By drawing the diagrams of h and g it is immediately seen that, with a fixed $q > 0$, there is a range $\rho[0, R_1]$ in which $h(u) = g(u)$ for only one $u = u_1 \in (0, 1)$. Then, there is a critical value $\rho = R_1$ in which $h(u)$ is tangent to $g(u)$ at some point $u = u_2 > 1$. In a suitable range $\rho \in (R_1, R_2)$ g and h have three intersection points $u_1 < u_2 < u_3$, which will become two at $\rho = R_2$ and just one $u = u_3$ for $\rho > R_2$. A simple analysis shows that the steady states 0 and u_2 are unstable, whereas u_1 and u_3 are stable.

The two parameters ρ and q have therefore a drastic impact on the dynamics. If ρ is very small compared to q, which is the case for example of a very small value of the predation switch parameter A, the only stable steady state is $u = u_2$, which is where the population will settle for large times regardless the initial condition. On the other hand, a very large value of A implies the population will settle to the larger steady state u_3.

1.5 Logistic model with delay

In the models treated so far, the population is assumed to react instantaneously to the surrounding environment, since the rate of change is computed at the same time at which it affects the growth or decrease of the population itself. In practice, this reaction to the surrounding environment “may take some time”. In many cases this reaction time is negligible. In other situations it is a significant intrinsic property of the model.

We shall consider the logistic model as an example. Assuming the population is affected by the rate of change with a delay $T > 0$, the model looks like

$$\frac{dN(t)}{dt} = rN(t) \left(1 - \frac{N(t - T)}{K}\right) . \tag{1.5.1}$$

Before discussing (1.5.1), we remark that a reasonable variant of said model is

$$\frac{dN(t)}{dt} = rN(t) \left(1 - \frac{1}{K} \int_{-\infty}^{t} W(t - s)N(s)ds\right) . \tag{1.5.2}$$

In (1.5.2), $W(\tau)$ is a *weight function* having a maximum at $\tau = T$, as an example one can consider $W(\tau) = G(\tau - T)$ where G is a standard Gaussian distribution. Note that the integral in (1.5.2) combines values for several times s, with the highest weight on $s = t - T$. The integral only considers $s \leq t$. Therefore, the signal inferred by the rate of change of the population propagates with many possible delay times, distributed in a non-homogeneous way among the population, the most frequent delay time being T.

Getting back to (1.5.1), we observe that for small times $t < T$ the rate of change is computed on negative times, namely $1 - \frac{1}{K}N(t-T)$ with $t - T < 0$. Now, with the classical approach of imposing *just* and initial condition $N(0) = N_0$, the model (1.5.1) may be not well posed. In fact, a whole set of values $N(\tau)$ with $\tau \in [-T, 0]$ needs to be prescribed in order to have the model well posed. The initial condition is not a number anymore, it is a function. This is just to remark that (1.5.1) brings along a complexity that is much higher that the one of a single ordinary differential equation encountered so far. We need indeed a "functional setting" for it.

Let us now try to guess the behavior of the solution to (1.5.1) via a simple heuristic argument. Let us assume that at some time $t = t_1$ the population $N(t_1)$ achieves the equilibrium value K and, for $t < t_1$ we have $N(t) < K$. By computing the time derivative of $N(t)$ at time t_1 we see that, since $N(t_1 - T) < K$, $dN(t_1)/dt$ is actually *positive*, so the population continues to grow. In particular, it does not settle at K as it happens for instance in the case without delay. For $t > t_1$, there will be a time at which the time derivative becomes zero. That time is actually $t = t_1 + T$, as in this case $1 - N(t-T)/K = 1 - N(t_1+T-T)/K = 1 - N(t_1)/K = 1 - 1 = 0$. For $t > t_1 + T$ we have a negative time derivative, since in this case $N(t - T) > K$. This situation persists until N possibly touches the value K with a negative slope. Let us call that time t_2. For $t > t_2$ the rate of changes computes N at times in which N is larger than K, which means N keeps on decreasing. At time $t = t_2 + T$ the rate of change is zero again and N starts growing again after it, until eventually it gets back to K. This heuristic argument suggests an *oscillatory* behavior, or periodic behavior. By assuming that all time intervals $[t_1, t_1 + T]$, $[t_1 + T, t_2]$, $t_2, t_2 + T]$, above have the same size, we expect a period of oscillations of the order of $4T$, where T is the delay time.

Let us now try a more rigorous approach to the resolution. First of all, we may rescale the model (1.5.1) by setting

$$N^* = \frac{N}{K}, \qquad t^* = rt, \qquad T^* = rT.$$

By dropping the $*$ symbols for simplicity, we get the rescaled model

$$\frac{dN(t)}{dt} = N(t)\left(1 - N(t-T)\right). \tag{1.5.3}$$

The equation (1.5.3) has two constant steady states: $N(t) \equiv 0$ and $N(t) \equiv 1$. By performing a standard linearisation around the zero steady state, we immediately see the that term $N(t)N(t-T)$ is of higher order, therefore the linearised equation is

$$\frac{dN(t)}{dt} = N(t)$$

which clearly implies that zero is an unstable steady state. Let us now perturb the steady state $N(t) \equiv 1$. We set

$$N(t) = 1 + n(t)$$

and substitute in (1.5.3) to obtain

$$\frac{dn(t)}{dt} = -(1 + n(t))n(t-T).$$

By linearising, we obtain

$$\frac{dn(t)}{dt} = -n(t-T). \tag{1.5.4}$$

We look for solutions of the form

$$n(t) = ce^{\lambda t}\,, \qquad \lambda \in \mathbb{C}\,.$$

Substituting said expression into (1.5.4) we get, assuming $c \neq 0$,

$$\lambda = -e^{-\lambda T}\,. \tag{1.5.5}$$

Assuming without restriction $\lambda \neq 0$ (as $\lambda = 0$ does not solve the equation), setting $z = 1/\lambda$ we get

$$1 + ze^{-\frac{T}{z}} = 0\,.$$

The complex function on the above left-hand side is holomorphic on $z \neq 0$ and has an essential singularity ar $z = 0$. Therefore, Picard's theorem implies it has countably many zeroes. Hence, (1.5.5) has countably many solutions. We set $\lambda = \mu + i\omega$, with $\mu \in \mathbb{R}$ and $\omega \in [0, 2\pi)$. Stability arises if $\mu > 0$ for all solutions to (1.5.5), whereas instability would be in place if $\mu < 0$ for at least one of them. Now, first of all we want to show that all solutions to (1.5.5) has real part that is bounded from above. To see this, assume that a subsequence λ_k of solutions to (1.5.5) satisfies $|\lambda_k| \to +\infty$. As

$$|\lambda_k| = e^{-\mu_k T} \to +\infty$$

we have $-\mu_k T \to +\infty$, which implies $\mu_k \to -\infty$, and hence the statement. Therefore, all eigenvalues have real part μ bounded above by some constant μ_0.

We first explore the case $\omega = 0$ in (1.5.5), which becomes in this case

$$\mu = -e^{-\mu T}\,, \qquad \mu \in \mathbb{R}\,.$$

Now, the above has clearly no solutions $\mu \geq 0$. It has, in fact, negative solutions only for small enough T. Indeed, a solution μ is characterised by being a zero of $g(\mu) = \mu + e^{-\mu T}$. We compute $g'(\mu) = 1 - Te^{-\mu T}$, which implies g has a global minimum at $\mu = \frac{\log T}{T}$, which equals $f((\log T)/T) = (1 + \log T)/T$, and this value is negative for $T < 1/e$. A negative global minimum for g implies g has zeroes. Hence, no eigenvalues arise for $\omega = 0$ in the unstable range.

Let us now investigate the case $\omega \neq 0$. Equating real and imaginary part of (1.5.5), we get

$$\begin{cases} \mu = -e^{-\mu T}\cos(\omega T) \\ \omega = e^{-\mu T}\sin(\omega T) \end{cases}$$

Assuming for the moment that ω is small, so that $\sin(\omega T)$ is not zero, we recover from the second equation above

$$e^{-\mu T} = \frac{\omega}{\sin(\omega T)}\,,$$

which substituted into the first equation gives

$$\mu T = -\omega T \frac{\cos(\omega T)}{\sin(\omega T)}\,.$$

Setting $\alpha = \omega T$, the function $[0, +\infty)\alpha \mapsto \alpha \cot \alpha$ decreases from the value 1, achieved as $\alpha \searrow 0$, towards the value 0 achieved at $\alpha = \pi/2$. Hence, solutions to (1.5.5) with $\mu > 0$ are not possible as long as $\omega T < \pi/2$. At $\omega T = \pi/2$, $\mu = 0$ is a solution with $\omega = e^{-0T}\sin(\pi/2) = 1$. As $\omega T = \pi/2$, we obtain that the "first" solution λ to (1.5.5) with $\mu = 0$ occurs for $T = \pi/2$. At this value, by scaling back to the original vabiable, the solution to the rescaled model (1.5.3) is well approximated by the solution to the linearised model $K(1 + ce^{irt})$. This solution is purely oscillatory. It has period $2\pi/r$. On the other hand, the delay T at which such solution exists is, in the original variable, equal to $\frac{\pi}{2r}$. Therefore, we obtain the heuristics prediction of the period of oscillations being equal to four times the delay time T.

1.6 A model in ecology

In this section we consider a simple model in ecology as an example of how to get useful information from the solution of a model not just from the stability of the solutions but rather from "structural" properties which may lead to information that are of practical use.

The context is that of the "harvesting" of a given animal population. The model is similar to the one we considered for the spruce budworm, with the exception that we shall consider a simpler predation term, in a way to model it via an "external" input representing the *effort* performed in the harvesting. We consider

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) - EN\,, \tag{1.6.1}$$

in which the constant $E > 0$ represents the harvesting *relative effort*, so that the quantity $Y(t) = EN(t)$ represents the *total yield* per unit time, at a given time $t \geq 0$. We would like to "control" the constant E in order to optimise the harvesting strategy, and we will do that by just looking at the equilibria. It is easily seen that (1.6.1) has two steady states: $N_1(t) \equiv 0$ and

$$N_2(t) \equiv K\left(1 - \frac{E}{r}\right)\,,$$

and the latter is relevant if and only if $E < r$. The right-hand side of (1.6.1) is a parabola, touching the x-axis at 0 and N_2. If $E < r$ then 0 is unstable and N_2 is stable (Exercise!). On the other hand, if $E > r$ then the only steady state zero is stable. In fact, one can prove that in the latter case zero attracts *all* solutions (for arbitrary initial datum) of (1.6.1). To see this, write

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K} - \frac{E}{r}\right) \leq -\alpha rN\,, \qquad \alpha := \frac{E}{r} - 1 > 0\,,$$

and a simple comparison principle implies

$$N(t) \leq N(0)e^{-\alpha t}$$

which proves the assertion. Clearly, choosing $E > r$ is not a good strategy as this implies that the population dies out for large times.

In case $E < r$, the (globally!) stable equilibrium $N_h = K\left(1 - \frac{E}{r}\right)$ implies a (constant) yield per unit time

$$Y_h = N_h E = EK\left(1 - \frac{E}{r}\right)\,,$$

which can be seen as a very practical formula to determine the total yield per unit time as a function of the yield effort. A simple calculus exercise implies that the above function has a global maximum at $E = r/2$ given by

$$Y_M = \frac{rK}{4}\,.$$

The corresponding steady state reads

$$N_M = K\left(1 - \frac{r}{2r}\right) = \frac{K}{2}\,.$$

We now try to get some useful information (from both practical and "ecological" point of view) out of these results in the case $E < r$. First of all, linearising around a general equilibrium N_h leads to

$$\frac{d(N(t) - N_h)}{dt} \approxeq \left(r - E - \frac{2rN_h}{K}\right)(N(t) - N_h) = -(r - E)(N(t) - N_h)$$

which gives

$$N(t) \approxeq N_h + \text{constant}\, e^{-(r-E)t}\,.$$

The quantity $r - E$ has dimension of $1/t$, and its reciprocal $1/(r - E)$ is the order of magnitude of the "recovery time" of the population to adjust to equilibrium from small perturbations. Let us call this "recovery time"

$$T_R(E) = \frac{1}{r - E}\,.$$

Clearly, T_R is a function of the relative effort E. When $E = 0$ (no effort, that is no harvesting), we have $T_R = T_R(0) = 1/r$. We introduce the "normalised recovery time"

$$\frac{T_R(E)}{T_R(0)} = \frac{r}{r - E} = \frac{1}{1 - \frac{E}{r}}\,.$$

At the "optimal" equilibrium $N_M = K/2$, in which $E = r/2$, we get

$$\frac{T_R(E)}{T_R(0)} = 2\,.$$

The recovery time (and it normalised variant) is clearly very important for practical uses. On the other hand, computing it as a function of the relative effort may be not easy. This is why we try now to compute it as a function of the total yield Y_h above, which is a function of E. By solving $Y = Y_h(E)$ with respect to E we get

$$\frac{E}{r} = \frac{1}{2}\left(1 \pm \sqrt{1 - \frac{Y}{Y_M}}\right),$$

which gives

$$\frac{T_R(E)}{T_R(0)} = \frac{2}{1 \mp \sqrt{1 - \frac{Y}{Y_M}}}\,.$$

By plotting this curve as a multi-valued function of Y/Y_M, the latter being confined in the interval $[0, 1]$, we see that it has two branches L_1, L_+. The branch L_+ grows from the value

1 attained at $Y/Y_M = 0$ to the value 2 attained at $Y/Y_M = 1$ with infinite slope. Then the branch L_- starts from $(1, 2)$ and approaches the line $Y/Y_M = 0$ with $\frac{T_R(E)}{T_R(0)} \to +\infty$.

Let us provide some comment. This model is totally deterministic, which clearly makes is more like a toy model to be improved in more specific situations. In any case, it is clear that a more refined model must include probabilistic effects. Therefore, one useful information here is that, by considering an increasing effort in the L^+ branch (E small and Y small), “we don’t want to get too close to the critical point $Y/Y_M = 1$”, because by some stochastic effect we may otherwise fall into the L^- branch, which may result into a catastrophic behavior with the recovery time blowing up to infinity with obvious negative repercussions.

Exercise: consider the model

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) - Y_0\,, \tag{1.6.2}$$

in which the total yield per unit time is constant, independent of the population size. Analyse (1.6.2) similarly as a above and deduce that it is definitely less convenient with respect to a population size depending yield as in (1.6.1).

Chapter 2

Continuous population models with many species

In this chapter we consider more than one species, and we focus in particular on the case of two species. The main point is that an interaction of a given species with another one may affect the evolution of both. This is why we expect many species modelling to improve the description of the population behavior. We shall consider three main families of interactions:

- Predator-prey interaction: one of the two species (predator) benefits form the interaction with the other species (prey), whereas the latter is negatively affected by the predator.
- Competition interaction: both species are negatively affected by the presence of the opposite species, for example because the use the same resources.
- Mutualism (or symbiosis) interaction: both species benefit from the interaction with the opposite species.

2.1 The Lotka-Volterra model

A classical model first formulated by Lotka [**?**] and in parallel by Volterra [**?**] for population dynamics purposes reads as

$$\begin{cases} \frac{dN}{dt} = N(a - bP) \\ \frac{dP}{dt} = P(cN - d)\,. \end{cases} \tag{2.1.1}$$

Here N is the prey population and P is the predator population. The parameters a, b, c, d are all positive. In the first equation of (2.1.1), the rate of change of the population of the prey N has is assumed to be constant and positive (Malthusian) in absence of P, whereas it will decrease as P increases, because it is assumed that the prey grows without being affected by the predator. Similarly, in the second equation the predator is assumed to have a decreasing population with constant negative rate if there is no prey around, whereas the presence of the prey provides a positive term in the rate of change of P.

As done before, we now rescale the model to reduce the number of parameters. We set

$$\tau = at\,, \quad \alpha = \frac{d}{a}\,, \quad u = \frac{c}{d}N\,, \quad v = \frac{b}{a}P\,.$$

System (2.1.1) becomes

$$\begin{cases} \frac{du}{d\tau} = u(1-v) \\ \frac{dv}{d\tau} = \alpha v(u-1)\,. \end{cases} \tag{2.1.2}$$

Hence, we are left with just one parameter $\alpha > 0$. Model (2.1.2) has two equilibria $(0,0)$ and $(1,1)$. We can analyse the trajectories in detail by dividing the second equation by the first

$$\frac{dv}{du} = \alpha \frac{v(u-1)}{u(1-v)}$$

and separating the variables to obtain

$$\frac{1-v}{v} dv = \alpha \frac{u-1}{u} du\,.$$

Integrating on both sides we get

$$\alpha(u - \log u) + v - \log v = \text{constant} =: H\,. \tag{2.1.3}$$

As $u - \log u \geq 1$ for all $u \geq 0$ (exercise) and the minimum is achieved at $u = 1$ $(1 - \log 1 = 1)$, we obtain that the orbits of system (2.1.2) are provided by the curves (2.1.3) with $H \geq 1 + \alpha$. These curves are closed and bounded. They confine wider and wider regions as H increases. They all contain the equilibrium point $(1,1)$. They never touch the $u = 0$ or the $v = 0$ axis, since otherwise we would get an infinite term on the left-hand side of (2.1.3). Since, of course, none of these trajectories touch the equilibrium point, it is easily seen that one loop in a trajectory is performed in a finite time, and the solution is *periodic in time*.

We provide later on a heuristic computation to estimate the period of the oscillations, but before that we would like to comment on a key feature of this model. The fact that the two species cover a periodic orbit in time is not per se unrealistic. Indeed, essentially the periodicity in time is the feature this model was introduced to represent. However, this model is *conservative*, which means that the oscillatory orbit is prescribed by the initial condition via

$$H = \alpha(u(0) - \log u(0)) + v(0) - \log v(0)\,.$$

If we perturb the initial condition through an arbitrary direction on the u, v-plane we almost certainly move to a separate trajectory, which makes the model somewhat "structurally unstable". What we would like to achieve (see the next section) is a situation in which, for a fairly general class of initial conditions, all solutions get closer and closer to a fixed oscillatory profile.

In order to estimate the period of oscillations, let us study the Jacobian matrix of the system (2.1.2). Let us set

$$f(u,v) = u(1-v)\,, \qquad g(u,v) = \alpha v(u-1)\,.$$

We have

$$\frac{D(f,g)}{D(u,v)} : A = \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{pmatrix} = \begin{pmatrix} 1-v & -u \\ \alpha v & \alpha(u-1) \end{pmatrix}.$$

We immediately see, as a remark, that $(0,0)$ is a saddle point (and therefore it is unstable) as in that case the eigenvalues are 1 and $-\alpha$. If we consider the steady state $(1,1)$ we get

$$A = \begin{pmatrix} 0 & -1 \\ \alpha & 0 \end{pmatrix},$$

which has $\pm i\sqrt{\alpha}$ as eigenvalues. Therefore, purely oscillatory eigenfunctions arise in this case with period of oscillations $2\pi/\sqrt{\alpha}$, which in the original variables reads $2\pi\sqrt{a/d}$. In particular, an increase of the linear growth of the prey and a decrease of the predator death rate produce an extension of the period of oscillations. Quite surprisingly (to a certain extent), the predation constants play no role in the period of oscillations.

2.2 A complex version of Lotka-Volterra

Assume we have k species of prey and k species of predator, with k a given positive integer. A natural generalisation of (2.1.1) is the system of $2k$ differential equations

$$\begin{cases} \frac{dN_i}{dt} = N_i \left[a_i - \sum_{j=1}^{k} b_{ij} P_j \right] & i = 1, \ldots, k \\ \frac{dP_i}{dt} = P_i \left[-d_i - \sum_{j=1}^{k} c_{ij} N_j \right] & i = 1, \ldots, k \,. \end{cases} \tag{2.2.1}$$

Let us write the above system in a more compact form. We set

$$\mathbf{N} = (N_1, \ldots, N_k) \in [0, +\infty)^k \,, \qquad \mathbf{P} = (P_1, \ldots, P_k) \in [0, +\infty)^k \,,$$

and introduce the column vectors

$$\mathbf{a} = (a_1, \ldots, a_k) \in [0, +\infty)^k \,, \quad \mathbf{d} = (d_1, \ldots, d_k) \in [0, +\infty)^k$$

and the matrices

$$\mathbf{B} = (b_{ij})_{i,j=1}^{k} \,, \qquad \mathbf{C} = (c_{ij})_{i,j=1}^{k} \,.$$

Hence, (2.2.1) becomes

$$\begin{cases} \frac{d\mathbf{N}}{dt} = \mathbf{N}(t)^T \left[\mathbf{a} - \mathbf{B} \cdot \mathbf{P}(t) \right] \\ \frac{d\mathbf{P}}{dt} = \mathbf{P}(t)^T \left[-\mathbf{d} - \mathbf{C} \cdot \mathbf{N}(t) \right] \,. \end{cases} \tag{2.2.2}$$

System (2.2.2) has the trivial equilibrium $(\mathbf{N}^*, \mathbf{P}^*) = (0, 0)$ as well as (possible) nontrivial equilibria arising as solutions $(\mathbf{N}^*, \mathbf{P}^*)$ to the linear system

$$\begin{cases} \mathbf{B}\mathbf{P}^* = \mathbf{a} \\ \mathbf{C}\mathbf{N}^* = \mathbf{d} \,. \end{cases}$$

The trivial steady state is clearly unstable as it has as eigenvalues all the entries of the vector $\mathbf{a}$ and the opposite of all the entries of the vector $\mathbf{d}$ (saddle point). In case of nontrivial steady state $(\mathbf{N}^*, \mathbf{P}^*)$, we perturb

$$\mathbf{N} = \mathbf{N}^* + \mathbf{u} \,, \qquad \mathbf{P} = \mathbf{P}^* + \mathbf{v} \,,$$

and consider the linearised system

$$\begin{cases} \frac{d\mathbf{u}}{dt} \approxeq -\mathbf{N}^* \cdot \mathbf{B}\mathbf{v} \\ \frac{d\mathbf{v}}{dt} \approxeq \mathbf{P}^* \cdot \mathbf{C}\mathbf{u} \,. \end{cases}$$

It is easily seen that the Jacobian matrix of this system has *zero trace*, which means that the sum of the eigenvalues is zero. Hence, for a given eigenvalue λ, $-\lambda$ is also an eigenvalue. Now, in case λ is imaginary as in the $k = 1$ case this simply means that the conjugate comples of λ is an eigenvalue and the solution is purely oscillatory. Clearly, the larger k the more likely the Jacobian matrix will have an eigenvalues with nontrivial real part, and this implies in this case the existence of an eigenvalue with positive real part. This heuristic argument shows that with this extent of complexity we should expect *instability*.

2.3 Realistic version of the predator-prey system

We now try to improve the previous version of the 2×2 version of Lotka-Volterra, in a way to remove the structural instability discussed above. We shall introduce three main variants making the model more reasonable and more realistic:

- We replace the Malthus growth of the prey by a logistic growth.
- We prescribe a carrying capacity of the predator which is proportional to the prey population size.
- We require that the relative predation rate in the N equation decreases to zero for large N.

This leads to

$$\begin{cases} \frac{dN}{dt} = N\left[r\left(1-\frac{N}{K}\right) - \frac{kP}{N+D}\right] \\ \frac{dP}{dt} = sP\left(1-\frac{hP}{N}\right) , \end{cases} \tag{2.3.1}$$

where $r, s > 0$ are linear growth rate of the prey and the predator respectively, $K > 0$ is the carrying capacity of the prey, $k, h, D > 0$ are parameters referring to the predation terms. We rescale the above system as follows:

$$\begin{aligned} u &= \frac{N}{K}, \quad v = \frac{hP}{K}, \quad \tau = rt \\ a &= \frac{k}{hr}, \quad b = \frac{s}{r}, \quad d = \frac{D}{K}. \end{aligned}$$

Then system (2.3.1) becomes

$$\begin{cases} \frac{du}{dt} = f(u,v) := u\left[1 - u - \frac{au}{u+d}\right] \\ \frac{dv}{dt} = g(u,v) := bv\left(1-\frac{v}{u}\right) . \end{cases} \tag{2.3.2}$$

Let us just consider nontrivial equilibria $u, v > 0$. From $g(u,v) = 0$ and $u \neq 0$ we obtain $u = v$, which substituted into $f(u,v) = 0$ gives

$$1 - u - \frac{au}{u+d} = 0 . \tag{2.3.3}$$

A simple computation leads to the two solutions

$$u = \frac{1}{2}\left[-(a+d-1) \pm \sqrt{(a+d-1)^2 + 4d}\right] ,$$

in which only one is meaningful (the other one is negative), namely

$$u^* = \frac{1}{2}\left[-(a+d-1) + \sqrt{(a+d-1)^2 + 4d}\right] .$$

We now perturb the steady state

$$u(\tau) = u^* + x(\tau) , \qquad v(\tau) = u^* + y(\tau) ,$$

and linearise (2.3.2) around it to get

$$\begin{pmatrix} \frac{dx}{d\tau} \\ \frac{dy}{d\tau} \end{pmatrix} = A \cdot \begin{pmatrix} x(\tau) \\ y(\tau) \end{pmatrix} , \qquad A = \left. \begin{pmatrix} 1 - 2u - \frac{avd}{(u+d)^2} & -\frac{au}{u+d} \\ b\frac{v^2}{u^2} & b - 2b\frac{v}{u} \end{pmatrix} \right|_{(u,v)=(u^*,u^*)} .$$

The top left term at the equilibrium can be rewritten, using (2.3.3) with $u = u^*$, as

$$\begin{aligned} 1 - 2u^* - \frac{au^*d}{(u^*+d)^2} &= -u^* + \frac{au^*}{u^*+d} - \frac{au^*d}{(u^*+d)^2} \\ = u^* \left[-1 + \frac{1}{(u^*+d)^2} \left(au^* + ad - ad \right) \right] &= u^* \left(-1 + \frac{au^*}{(u^*+d)^2} \right) . \end{aligned}$$

Other simple computations for the other terms provide the matrix

$$A = \begin{pmatrix} u^* \left(\frac{au^*}{(u^*+d)^2} - 1 \right) & -\frac{au^*}{u^*+d} \\ b & -b \end{pmatrix} .$$

With the well known formula for the eigenvalues λ

$$\lambda^2 - \mathrm{tr}(A)\lambda + \det(A) = 0$$

we get that stability is achieved provided

$$\mathrm{tr}(A) < 0 \qquad \text{and} \qquad \det(A) > 0 .$$

The latter reads

$$0 < bu^* \left[1 - \frac{au^*}{(u^*+d)^2} + \frac{a}{u^*+d} \right] = \frac{bu^*}{(u^*+d)^2} \left[(u^*+d)^2 + ad \right] > 0 ,$$

which is always satisfied. Hence, stability is equivalent to

$$u^* \left(\frac{au^*}{(u^*+d)^2} - 1 \right) < b ,$$

which, upon substituting the expression of u^* and $1 - u^* = \frac{au^*}{u^+ + d}$ becomes

$$\begin{aligned} b > u^* \left(\frac{au^*}{(u^*+d)^2} - 1 \right) &= \frac{u^*}{u^*+d} (1 - 2u^* - d) \\ = \frac{1}{a}(1-u^*)(1-2u^*-d) &= \frac{1}{a} \left[a + d + 1 - \sqrt{(1-a-d)^2 + 4d} \right] \left[a - \sqrt{(1-a-d)^2 + 4d} \right] . \end{aligned} \tag{2.3.4}$$

Both squared brackets above are non-increasing functions of $d > 0$ (exercise!). Therefore, a simple sufficient condition for the local stability of (u^*, u^*) is (2.3.4) with $d = 0$, namely

$$b > \frac{1}{a}(a + 1 - |1-a|)(a - |1-a|)$$

which reads

$$b > \begin{cases} 2a - 1 & \text{if } a \in [0,1] \\ 1/a & \text{if } a > 1 . \end{cases}$$

Recall that in the original variables $a = \frac{k}{hr}$ and $b = \frac{s}{r}$. Hence, stability of the steady state is ensured, for example, when the predator has a much higher linear growth rate than the prey, or when the predation factor k in the prey equation is very small. The latter is interpreted, for example, by the fact that with a very small predation term the prey has time to converge to a steady state in time to prevent instabilities due to the predation term. We shall not focus on the nature of the stability here (it may be a node or a spiral).

We will instead ask ourselves *what happens in the unstable range?*. We are going to determine the behavior of our system in this case by invoking *Poincaré-Bendixson Theorem*, which states as follows with the notation of our system: assume (u^*, v^*) is a fully unstable steady state (unstable node or spiral) included in an *invariant domain* I, then the system admits a stable *limit cycle*. The latter means that the system is asymptotically close to a periodic trajectory for large times, and said trajectory is the same for all initial data. This is a big improvement compared to the classical Lotka-Volterra model, it is indeed the structural stability property we were looking for. We now observe that (u^*, u^*) is indeed fully unstable in case (2.3.4) is not satisfied, since the determinant of A is always positive.

To find a confined set it is essential and always informative to draw the null clines of the system, that is, the curves in the phase plane where $du/d\tau = 0$ and $dv/d\tau = 0$. From (2.3.2) these are the curves $f(u,v) = 0$ and $g(u,v) = 0$ which are illustrated in Figure 2.1. The sign of the vector components of $(f(u,v), g(u,v))$ indicate the direction of the vector $(du/d\tau, dv/d\tau)$ and hence the direction of the (u,v) trajectory. So if $f > 0$ in a domain, $du/d\tau > 0$ and u is thus increasing there. On DE, EA, AB and BC, the trajectories clearly point inwards because of the signs of $f(u,v)$ and $g(u,v)$ on them. It can be shown that a line DC exists such that on it $\mathbf{n} \cdot (du/d\tau, dv/d\tau) < 0$; that is, $\mathbf{n} \cdot (f(u,v), g(u,v)) < 0$ where $\mathbf{n}$ is the unit vector perpendicular to DC. For instance, take the straight line $v = \alpha u$ with $\alpha > 1$ on a suitable u-interval according to Figure 2.1. On that line $f = f(u, \alpha u) = u\left(1 - u + \frac{\alpha a u}{u+d}\right)$ and $g = g(u, \alpha u) = -\alpha b(\alpha - 1)u < 0$. In order to point inwards we need $g/f > \alpha$, which means $g - \alpha f < 0$ on the set $f < 0$. It is left as an exercise to prove that such a choice of α is possible.

2.4 Competition model

In a competition model, both species N_1 and N_2 "penalise" the growth of the opposite species. This results in a system like

$$\begin{aligned} \frac{dN_1}{dt} &= r_1 N_1 \left[1 - \frac{N_1}{K_1} - \frac{b_{12} N_2}{K_1}\right] \\ \frac{dN_2}{dt} &= r_2 N_2 \left[1 - \frac{N_2}{K_2} - \frac{b_{21} N_1}{K_2}\right] . \end{aligned} \tag{2.4.1}$$

As above, we perform a scaling

$$\begin{aligned} u_1 &= \frac{N_1}{K_1} \qquad u_2 = \frac{N_2}{K_2} \qquad \tau = r_1 t \\ \rho &= \frac{r_1}{r_2} \qquad a_{12} = b_{12}\frac{K_2}{K_1} \qquad a_{21} = b_{21}\frac{K_1}{K_2} , \end{aligned}$$

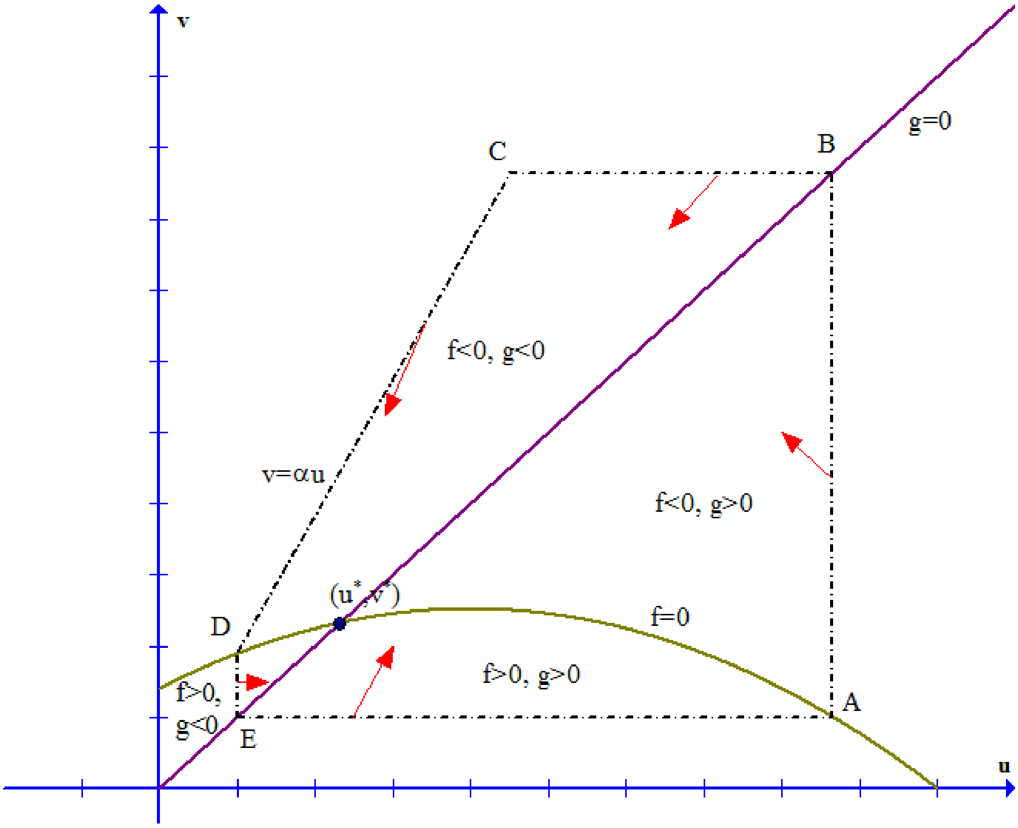

Figure 2.1: Null clines $f(u,v) = 0$, $g(u,v) = 0$ for the system (2.3.2); note the signs of f and g on either side of their null clines. $ABCDEA$ is the boundary of the invariant compact set about (u^*, v^*) on which the trajectories all point inwards.

which, upon a lengthy computation, leads to the rescaled system

$$\begin{aligned}\frac{du_1}{d\tau} &= u_1\left[1 - u_1 - a_{12}u_2\right] := f_1(u_1,u_2)\\ \frac{du_2}{d\tau} &= \rho u_2\left[1 - u_2 - a_{21}u_1\right] := f_2(u_1,u_2)\,. \end{aligned} \tag{2.4.2}$$

It is immediately seen that the three steady states $(0,0)$, $(1,0)$, and $(0,1)$ always exist. In order to see if there is a (fourth) steady state in which both components are positive, we need to solve the system

$$1 - u_1 - a_{12}u_2 = 0\,, \qquad 1 - u_2 - a_{21}u_1 = 0\,.$$

which leads to

$$u_1 = u_1^* = \frac{1 - a_{12}}{1 - a_{12}a_{21}}\,, \qquad u_2 = u_2^* = \frac{1 - a_{21}}{1 - a_{12}a_{21}}\,.$$

It is immediately seen that there are two cases in which such a steady state makes sense (only positive values for both u_1 and u_2 are allowed): the first case (case a) is $a_{12} < 1$ and $a_{21} < 1$, the second case (case b) is $a_{12} > 1$ and $a_{21} > 1$. The case c, with $a_{12} < 1$ and $a_{21} > 1$, and in case d, with $a_{12} > 1$ and $a_{21} < 1$, the fourth steady state doesn't exist.

Let us now analyse the stability of those states. The linearisation of system (2.4.2) leads to the Jacobian matrix

$$A = \begin{pmatrix} 1 - 2u_1 - a_{12}u_2 & -a_{12}u_1 \\ -\rho a_{21}u_2 & \rho(1 - 2u_2 - a_{21}u_1) \end{pmatrix}\,.$$

Substituting $(u_1,u_2) = (0,0)$ we easily obtain the diagonal matrix with entries 1 and ρ, so we get an unstable node. As for $(u_1,u_2) = (1,0)$, we obtain the two eigenvalues $\lambda_1 = -1$, $\lambda_2 = \rho(1 - a_{21})$, which is a stable node if $a_{21} > 1$ and a saddle point if $a_{21} < 1$. In case $(u_1,u_2) = (0,1)$ we obtain the two eigenvalues $\lambda_1 = 1 - a_{12}$ and $\lambda_2 = -\rho$, which is a stable node if $a_{12} > 1$ and a saddle point if $a_{12} < 1$.

Let us now substitute the fourth steady state $(u_1,u_2) = (u_1^*,u_2^*)$. The Jacobian matrix becomes

$$A = (1 - a_{12}a_{21})^{-1}\begin{pmatrix} a_{12} - 1 & a_{12}(a_{12} - 1) \\ \rho a_{21}(a_{21} - 1) & \rho(a_{21} - 1) \end{pmatrix}\,,$$

which has the two eigenvalues

$$\begin{aligned}\lambda_\pm = {}& \left[2\left(1 - a_{12}a_{21}\right)\right]^{-1}\Big[\rho(a_{21} - 1) + a_{12} - 1 \\ & \pm\sqrt{\left(\rho(a_{21} - 1) + a_{12} - 1\right)^2 + 4\rho(a_{12} - 1)(a_{21} - 1)(a_{12}a_{21} - 1)}\Big]\,. \end{aligned} \tag{2.4.3}$$

We compute the term in the squared root

$$\begin{aligned}&\left(\rho(a_{21} - 1) + a_{12} - 1\right)^2 + 4\rho(a_{12} - 1)(a_{21} - 1)(a_{12}a_{21} - 1)\\ &= \rho^2(a_{21} - 1)^2 + (a_{12} - 1)^2 + 2\rho(a_{21} - 1)(a_{12} - 1) - 4\rho(a_{12} - 1)(a_{21} - 1) + 4\rho a_{12}a_{21}(a_{12} - 1)(a_{21} - 1)\\ &= \left[\rho(a_{21} - 1) - (a_{12} - 1)\right]^2 + 4\rho a_{12}a_{21}(a_{12} - 1)(a_{21} - 1) > 0\,,\end{aligned}$$

therefore the two eigenvalues λ_- and λ_+ are real in both cases a and b (the only ones in which the fourth steady state exists).

In case a ($a_{12} < 1$ and $a_{21} < 1$), the expression (2.4.3) clearly shows that both eigenvalues are negative. In this case (u_1^*, u_2^*) is a stable node, whereas both $(1,0)$ and $(0,1)$ are saddle points. As for the latter, it is clear that the "stable" direction for $(1,0)$ ($(0,1)$ respectively) is the one with $u_2 = 0$ ($u_1 = 0$ respectively). Therefore, unless the initial condition is such that one of the two species size is zero, all trajectories move away from these two steady states. The steady state (u_1^*, u_2^*) attracts all trajectories. The explicit formula for (u_1^*, u_2^*) shows that both values u_1^* and u_2^* are less than one. The interpretation behind that is that the two species tend to adjust asymptotically to a value strictly less than their respective carrying capacities. This is possible because the two predation constants a_{12} and a_{21} are not too big (case a), so this is a case of *mild (or soft) competition*, a situation in which both species survive in the large time range.

In case b ($a_{12} > 1$ and $a_{21} > 1$), the expression (2.4.3) shows that the two eigenvalues have opposite sign, and hence we have a saddle point. This is the same case in which the two other non-trivial steady states $(1,0)$ and $(0,1)$ are fully stable. In this case, one could prove the existence of a *separatrix curve* in $\{u_1 > 0\,,\ u_2 > 0\}$, which splits the first quadrant into two regions, I and II, with $(1,0)$ belonging to I and $(0,1)$ belonging to II. Each region is the basin of attraction of the steady states it contains. This is the case of *aggressive competition*, in which the two predation constants are high, and only one species survives depending on the initial condition (with the only exception of an initial condition exactly on the separatrix). This is an example of application of the *principle of competitive exclusion* in evolution theory.

In case c (d respectively) $(1,0)$ is stable (unstable respectively) and $(0,1)$ is unstable (stable respectively) and there are no more steady states. This is another example of principle of competitive exclusion. However, while in case b the final outcome is determined by the initial condition, in cases c and d the final outcome is determined by the structural properties of the system.

As an exercise, one can see that this system as well as an invariant region $\{0 \leq u_1 \leq U_1,\ 0 \leq u_2 \leq U_2\}$ for suitably large constants U_1, U_2.

2.5 Mutualism or symbiosis

In a mutualism model, both species gain from the presence of the opposite species. A very simply way to model this is

$$\begin{aligned} \frac{dN_1}{dt} &= r_1 N_1 + a_1 N_1 N_2 \\ \frac{dN_2}{dt} &= r_2 N_2 + a_2 N_2 N_1, \end{aligned} \tag{2.5.1}$$

with $r_1, r_2, a_1, a_2 > 0$. Only one steady state occurs here, $(0,0)$ which is clearly unstable. Indeed, the velocity vector points strictly away from $(0,0)$ unless we start from the origin itself. More precisely

$$\frac{dN_1}{dt}(t) \geq r_1 N_1(t)\,, \qquad \frac{dN_2}{dt}(t) \geq r_2 N_2(t)\,,$$

which proves by a simple comparison principle that

$$N_1(t) \geq N_1(0) e^{r_1 t} \qquad N_2(t) \geq N_2(0) e^{r_2 t}\,,$$

and therefore the norm of $(N_1(t), N_2(t))$ diverges unless the initial datum is the origin. Hence, in this model both populations grow at infinity if both are initially positive, therefore this model is not that interesting.

A better version is the following one, in which the growth of both species is modelled via a logistic growth rate:

$$\begin{aligned}\frac{dN_1}{dt} &= r_1 N_1 \left(1 - \frac{N_1}{K_1} + b_{12}\frac{N_2}{K_1}\right)\\ \frac{dN_2}{dt} &= r_2 N_2 \left(1 - \frac{N_2}{K_2} + b_{21}\frac{N_1}{K_2}\right).\end{aligned} \tag{2.5.2}$$

By performing a similar scaling as for the competition model we get

$$\begin{aligned}\frac{du_1}{d\tau} &= u_1\left(1 - u_1 + a_{12}u_2\right) =: f_1(u_1, u_2)\\ \frac{du_2}{d\tau} &= u_2\left(1 - u_2 + a_{21}u_1\right) =: f_2(u_1, u_2).\end{aligned} \tag{2.5.3}$$

The pairs $(0, 0)$, $(1, 0)$, and $(0, 1)$ are steady states of (2.5.3). To detect possible further steady states, let us solve the system

$$1 - u_1 + a_{12}u_2 = 0\,, \qquad 1 - u_2 + a_{21}u_1 = 0$$

and get

$$u_1 = \frac{1 + a_{12}}{1 - a_{12}a_{21}}\,, \qquad u_2 = \frac{1 + a_{21}}{1 - a_{12}a_{21}}\,.$$

Clearly, the above steady state only makes sense if $a_{12}a_{21} < 1$, which is a case of *mild* symbiosis. The linearised matrix around a steady state is

$$A = \begin{pmatrix} 1 - 2u_1 + a_{12}u_2 & a_{12}u_1 \\ \rho a_{21}u_2 & \rho(1 - 2u_2 + a_{21}u_1) \end{pmatrix}$$

which shows that the three steady states $(0, 0)$, $(0, 1)$ and $(1, 0)$ are unstable. An easy computation shows that, when it exists, the fourth steady state is stable. Hence, in case of mild symbiosis the two species will settle to a finite value. On the other hand, when the fourth steady state does not exists there are no stable steady states and the two species grow indefinitely.

Chapter 3

Dynamics of Infectious Diseases: Epidemic Models and AIDS

The history of epidemics is an ever fascinating area. Today this subject has enormous impact on society due to Covid-19. The modelling of the latter is producing a very dense and complex literature. In this chapter we will consider example of classical epidemics models. We will then consider a more complex model for AIDS, which was probably the first epidemics for which more complex models were developed.

3.1 Simple Epidemic Models and Practical Applications

In the classical models we consider here the total population is taken to be constant. If a small group of infected individuals is introduced into a large population, a basic problem is to describe the spread of the infection within the population as a function of time. Of course this depends on a variety of circumstances, including the actual disease involved, but as a first attempt at modeling directly transmitted diseases we make some not unreasonable general assumptions. Consider a disease which, after recovery, confers immunity which, if lethal, includes deaths: dead individuals are still counted. Suppose the disease is such that the population can be divided into three distinct classes: the susceptibles, S, who can catch the disease; the infectives, I , who have the disease and can transmit it; and the removed class, R, namely, those who have either had the disease, or are recovered, immune or isolated until recovered. The progress of individuals is schematically represented by

$$S \to I \to R.$$

Such models are often called SIR models. The number of classes depends on the disease. SI models, for example, have only susceptible and infected classes while SEIR models have a suceptible class, S, a class in which the disease is latent, E, an infectious class, I, and a recovered or dead class, R. The assumptions made about the transmission of the infection and incubation period are crucial in any model; these are reflected in the terms in the equations and the parameters. With $S(t)$, $I(t)$ and $R(t)$ as the number of individuals in each class we assume here that: (i) The gain in the infective class is at a rate proportional to the number of infectives and susceptibles, that is, rSI, where $r > 0$ is a constant parameter. The susceptibles are lost at the same rate. (ii) The rate of removal of infectives to the removed class is proportional to the number of infectives, that is, aI where $a > 0$ is a constant; $1/a$ is a measure of the time spent in the infectious state. (iii) The incubation period is short enough to be negligible; that is, a susceptible who contracts the disease is infective right away.

We now consider the various classes as uniformly mixed; that is, every pair of individuals has equal probability of coming into contact with one another. This is a major assumption and in many situations does not hold as in most sexually transmitted diseases (STD's). The model mechanism based on the above assumptions is then

$$\frac{dS}{dt} = -rSI \tag{3.1.1}$$

$$\frac{dI}{dt} = rSI - aI \tag{3.1.2}$$

$$\frac{dR}{dt} = aI, \tag{3.1.3}$$

where $r > 0$ is the infection rate and $a > 0$ the removal rate of infectives. This is the classic Kermack–McKendrick (1927) model. We are, of course, only interested in nonnegative solutions for S, I and R. This is a basic model but, even so, we can make some highly relevant general comments about epidemics and, in fact, adequately describe some specific epidemics with such a model.

The constant population size is built into the system (3.1.1)–(3.1.3) since, on adding the equations,

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \quad \Rightarrow \quad S(t) + I(t) + R(t) = N, \tag{3.1.4}$$

where N is the total size of the population. Thus, S, I and R are all bounded above by N. The mathematical formulation of the epidemic problem is completed given initial conditions such as

$$S(0) = S_0 > 0, \qquad I(0) = I_0 > 0, \qquad R(0) \geq 0. \tag{3.1.5}$$

A key question in any epidemic situation is, given r, a, S_0 and the initial number of infectives I_0, whether the infection will spread or not, and if it does how it develops with time, and crucially when it will start to decline. From (3.1.2),

$$\left[\frac{dI}{dt}\right]_{t=0} = I_0(rS_0 - a) \qquad \begin{cases} > 0 & \text{if} \;\; S_0 > \rho \\ < 0 & \text{if} \;\; S_0 < \rho \end{cases}, \qquad \rho = \frac{a}{r}. \tag{3.1.6}$$

Since, from (3.1.1), $dS/dt \leq 0$, $S \leq S_0$ we have, if $S_0 < a/r$,

$$\frac{dI}{dt} = I(rS - a) \leq I(rS_0 - a) \leq 0, \qquad \text{for all} \;\; t \geq 0, \tag{3.1.7}$$

in which case $I_0 > I(t) \to 0$ as $t \to +\infty$ and so the infection dies out; that is, no epidemic can occur. On the other hand if $S_0 > a/r$ then $I(t)$ initially increases and we have an epidemic. The term 'epidemic' means that $I(t) > I_0$ for some $t > 0$; see Figure 3.1. We thus have a threshold phenomenon. If $S_0 > S_c = a/r$ there is an epidemic while if $S_0 < S_c$ there is not. The critical parameter $\rho = a/r$ is sometimes called the *relative removal rate* and its reciprocal $\sigma = r/a$ the infection's *contact rate.*

We write

$$R_0 = \frac{rS_0}{a},$$

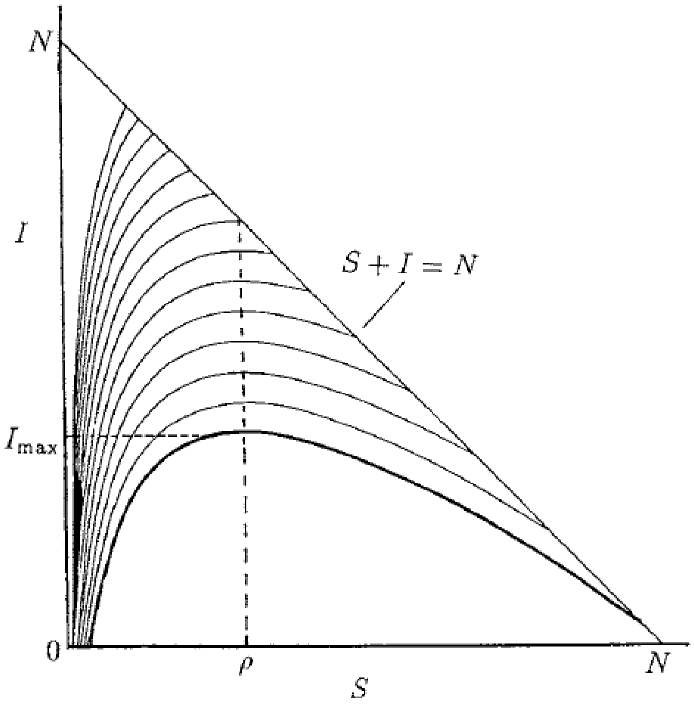

Figure 3.1: Phase trajectories in the susceptibles (S)–infectives (I) phase plane for the SIR model epidemic system (3.1.1)–(3.1.3). The curves are determined by the initial conditions $I(0) = I_0$ and $S(0) = S_0$. With $R(0) = 0$, all trajectories start on the line $S + I = N$ and remain within the triangle since $0 < S + I < N$ for all time. An epidemic situation formally exists if $I(t) > I_0$ for any time $t > 0$; this always occurs if $S_0 > \rho = a/r$ and $I_0 > 0$

where R_0 is the *basic reproduction rate* of the infection, that is, the number of secondary infections produced by one primary infection in a wholly susceptible population. Here $1/a$ is the average infectious period. If more than one secondary infection is produced from one primary infection, that is, $R_0 > 1$, clearly an epidemic ensues. The whole question of thresholds in epidemics is obviously important. The definition and derivation or computation of the basic reproduction rate is crucial and can be quite complicated.

We can derive some other useful analytical results from this simple model. From (3.1.1) and (3.1.2)

$$\frac{dI}{dS} = -\frac{(rS-a)I}{rSI} = -1 + \frac{\rho}{S}, \qquad (I \neq 0).$$

The singularities all lie on the $I = 0$ axis. Integrating the last equation gives the (I, S) phase plane trajectories as

$$I + S - \rho \log S = \text{constant} = I_0 + S_0 - \rho \log S_0, \tag{3.1.8}$$

where we have used the initial conditions (3.1.5). The phase trajectories are sketched in Figure 3.1. Note that with (3.1.5), all initial values S_0 and I_0 satisfy $I_0 + S_0 = N$ since $R(0) = 0$ (it is reasonable to choose as initial time that in which there are no removed individuals, since the epidemics still has to start spreading!) and so for $t > 0$, $0 \leq S + I < N$.

If an epidemic exists we would like to know how severe it will be. From (3.1.7) the maximum I, I_{max} , occurs at $S = \rho$ where $dI/dt = 0$. From (3.1.8), with $S = \rho$,

$$\begin{aligned} I_{max} &= \rho \log \rho - \rho + I_0 + S_0 - \rho \log S_0 \\ &= I_0 + (S_0 - \rho) + \rho \log\left(\frac{\rho}{S_0}\right) \\ &= N - \rho + \rho \log\left(\frac{\rho}{S_0}\right). \end{aligned} \tag{3.1.9}$$

For any initial values I_0 and $S_0 > \rho$, the phase trajectory starts with $S > \rho$ and we see that I increases from I_0 and hence an epidemic ensues. It may not necessarily be a severe epidemic as is the case if I_0 is close to I_{max}. It is also clear from Figure 3.1 that if $S_0 < \rho$ then I decreases from I_0 and no epidemic occurs.

Since the axis $I = 0$ is a line of singularities, on all trajectories $I \to 0$ as $t \to +\infty$. From (3.1.1), S decreases since $dS/dt < 0$ for $S \neq 0$, $I \neq 0$. From (3.1.1) and (3.1.3),

$$\frac{dS}{dR} = -\frac{S}{\rho}$$

$$\Rightarrow \qquad S = S_0 e^{-R/\rho} \geq S_0 e^{-N/\rho} > 0 \tag{3.1.10}$$

$$\Rightarrow \qquad 0 < S(+\infty) \leq N. \tag{3.1.11}$$

In fact from Figure (3.1.1), $0 < S(+\infty) < \rho$. Since $I(+\infty) = 0$, (3.1.4) implies that $R(+\infty) = N - S(+\infty)$. Thus, from (3.1.10),

$$S(+\infty) = S_0 e^{-R(+\infty)/\rho} = S_0 e^{-\frac{N-S(+\infty)}{\rho}}$$

and so $S(+\infty)$ is the positive root $0 < z < \rho$ of the transcendental equation

$$S_0 e^{-\frac{N-z}{\rho}} = z. \tag{3.1.12}$$

We then get the total number of susceptibles who catch the disease in the course of the epidemic as

$$I_{total} = I_0 + S_0 - S(+\infty), \tag{3.1.13}$$

where $S(+\infty)$ is the positive solution z of (3.1.12). An important implication of this analysis, namely, that $I(t) \to 0$ and $S(t) \to S(+\infty) > 0$, is that the disease dies out from a lack of infectives and not from a lack of susceptibles.

The threshold result for an epidemic is directly related to the relative removal rate, ρ: if $S_0 > \rho$ an epidemic ensues whereas it does not if $S_0 < \rho$. For a given disease, the relative removal rate varies with the community and hence determines whether an epidemic may occur in one community and not in another. The number of susceptibles S_0 also plays a major role, of course. For example, if the density of susceptibles is high and the removal rate, a, of infectives is low (through ignorance, lack of medical care, inadequate isolation and so on) then an epidemic is likely to occur. Expression (3.1.9) gives the maximum number of infectives while (3.1.13) gives the total number who get the infection in terms of $\rho = a/r$, I_0, S_0 and N.

In most epidemics it is difficult to determine how many new infectives there are each day since only those that are removed, for medical aid or whatever, can be counted. Public Health records generally give the number of infectives per day, week or month. So, to apply the model to actual epidemic situations, in general we need to know the number removed per unit time, namely, dR/dt, as a function of time.

From (3.1.10), (3.1.4) and (3.1.3) we get an equation for R alone; namely,

$$\frac{dR}{dt} = aI = a(N - S - R) = a\left(N - S - S_0 e^{-R/\rho}\right), \qquad R(0) = 0, \tag{3.1.14}$$

which can only be solved analytically in a parametric way: the solution in this form however is not particularly convenient. Of course, if we know a, r, S_0 and N it is a simple matter to compute the solution numerically. Usually we do not know all the parameters and so we have to carry out a best fit procedure assuming, of course, the epidemic is reasonably described by such a model. In practice, however, it is often the case that if the epidemic is not large, R/ρ is small (at least $R/\rho < 1$). Following Kermack and McKendrick (1927) we can then approximate (3.1.14) by

$$\frac{dR}{dt} = aI = a(N - S - R) = a\left[N - S_0 + \left(\frac{S_0}{\rho} - 1\right)R - \frac{S_0 R^2}{2\rho^2}\right].$$

We leave the solution of the above equation as an exercise.

3.2 Modelling Venereal Diseases

The incidence of sexually transmitted diseases (STDs), such as gonorrhea (Neisseria gonorrhoeae), chlamydia, syphilis and, of course, AIDS, is a major health problem in both developed and developing countries.

In this section we present a simple classical epidemic model which incorporates some of the basic elements in the heterosexual spread of venereal diseases. We have in mind such diseases as gonorrhea; AIDS we discuss separately later in the chapter.

Since the incubation period for venereal diseases is usually quite short when compared to the infectious period, we use an extension of the simple epidemic model in the previous section. We divide the promiscuous male population into susceptibles, S, infectives, I , and a removed class, R; the similar female groups we denote by S^*, I^* and R^*. If we do not include any transition from the removed class to the susceptible group, the infection dynamics is schematically

$$\begin{tikzcd} S \arrow[r] & I \arrow[r] \arrow[dl] & R \\ S^* \arrow[r] & I^* \arrow[r] \arrow[ul] & R^* \end{tikzcd} \tag{3.2.1}$$

Here I^* infects S and I infects S^*.

As we noted above, the contraction of gonorrhea does not confer immunity and so an individual removed for treatment becomes susceptible again after recovery. In this case a better dynamics flow diagram for gonorrhea is

$$\begin{tikzcd} & R \arrow[dl] \\ S \arrow[r] & I \arrow[u] \arrow[dl] \\ S^* \arrow[r] & I^* \arrow[ul] \arrow[d] \\ & R^* . \arrow[ul] \end{tikzcd} \tag{3.2.2}$$

An even simpler version involving only susceptibles and infectives is

$$\begin{tikzcd} S \arrow[r, shift right] & I \arrow[l, shift right] \arrow[dl] \\ S^* \arrow[r, shift right] & I^* \arrow[l, shift right] \arrow[ul] \end{tikzcd} \tag{3.2.3}$$

which, by way of illustration, we now analyze. It is a *criss-cross SI* model.

We take the total number of males and females to be constant and equal to N and N^* respectively. Then, for (3.2.3),

$$S(t) + I(t) = N, \qquad S^*(t) + I^*(t) = N^*. \tag{3.2.4}$$

As before we now take the rate of decrease of male susceptibles to be proportional to the male susceptibles times the infectious female population with a similar form for the female rate. We assume that once infectives have recovered they rejoin the susceptible class. A model for (3.2.3) is then (3.2.4) together with

$$\begin{aligned} \frac{dS}{dt} &= -rSI^* + aI, & \frac{dS^*}{dt} &= -r^*S^*I + a^*I^* \\ \frac{dI}{dt} &= rSI^* - aI, & \frac{dI^*}{dt} &= r^*S^*I - a^*I^*, \end{aligned} \tag{3.2.5}$$

where r, a, r^* and a^* are positive parameters. We are interested in the progress of the disease given initial conditions

$$S(0) = S_0, \quad I(0) = I_0, \quad S^*(0) = S_0^*, \quad I^*(0) = I_0^*. \tag{3.2.6}$$

Although (3.2.5) is a 4th–order system, with (3.2.4) it reduces to a 2nd–order system in either S and S^* or I and I^*. In the latter case we get

$$\frac{dI}{dt} = rI^*(N - I) - aI, \qquad \frac{dI^*}{dt} = r^*I(N^* - I^*) - a^*I^*, \tag{3.2.7}$$

which can be analyzed in the (I, I^*) phase plane in the standard way.

The equilibrium points, that is, the steady states of (3.2.7), are $I = 0 = I^*$ and

$$I_s = \frac{NN^* - \rho\rho^*}{\rho + N^*}, \quad I_s^* = \frac{NN^* - \rho\rho^*}{\rho^* + N}, \qquad \rho = \frac{a}{r}, \quad \rho^* = \frac{a^*}{r^*}. \tag{3.2.8}$$

Thus nonzero positive steady state levels of the infective populations exist only if $NN^*/\rho\rho^* > 1$: this is the threshold condition somewhat analogous to that found in the previous section.

With the experience gained from the previous chapters, we now expect that, if the positive steady state exists then the zero steady state is unstable. This is indeed the case. The eigenvalues λ for the linearization of (3.2.7) about $I = 0 = I^*$ are given by

$$\begin{vmatrix} -a - \lambda & rN \\ r^*N^* & -a^* - \lambda \end{vmatrix} = 0$$

$$\Rightarrow \quad 2\lambda = -(a + a^*) \pm \left[(a + a^*)^2 + 4aa^*\left(\frac{NN^*}{\rho\rho^*} - 1\right)\right]^{1/2}.$$

So, if the threshold condition $NN^*/\rho\rho^* > 1$ holds, $\lambda_1 < 0 < \lambda_2$ and the origin is a saddle point in the (I, I^*) phase plane. If the threshold condition is not satisfied, that is, $NN^*/\rho\rho^* < 1$, then the origin is stable since both $\lambda < 0$. In this case positive I_s and I_s^* do not exist.

If I_s and I_s^* exist, meaning in the context here that they are positive, then linearizing (3.2.7) about it, the eigenvalues λ satisfy

$$\begin{vmatrix} -a - rI_s^* - \lambda & rN - rI_s \\ r^*N^* - r^*I_s^* & -a^* - r^*I_s - \lambda \end{vmatrix} = 0,$$

that is,

$$\lambda^* + \lambda[a + a^* + rT_S^* + r^*I_s] + [a^*rI_s^* + ar^*I_s + rr^*(I^*N + IN^*) + aa^* - rr^*NN^*] = 0,$$

the solutions of which have $Re\lambda < 0$ and so the positive steady state (I_s, I_s^*) in (3.2.8) is stable.

The threshold condition for a nonzero steady state infected population is $NN^*/\rho\rho^* > 1$. We can interpret each term as follows. If every male is susceptible then rN/a is the average number of males contacted by a female infective during her infectious period; a reciprocal interpretation holds for r^*N^*/a^*. These quantities, rN/a and r^*N^*/a^*, are the maximal male and female contact rates respectively.

Although parameter values for contacts during an infectious stage are notoriously unreliable from individual questionnaires, what is abundantly clear from the statistics since 1950 is that an epidemic has occurred in a large number of countries and so $NN^*/\rho\rho^* > 1$. From data given by a male and a female infective, in the U.S.A. in 1973, regarding the number of contacts during a period of their infectious state, figures of maximal contact rates of $N/\rho \approx 0.98$ and $N^*/\rho^* \approx 1.15$ were calculated for the male and female respectively which give $NN^*/\rho\rho^* \approx 1.127$.

3.3 AIDS: Modelling the Transmission Dynamics of the Human Immunodeficiency Virus (HIV)

3.3.1 Human Immunodeficiency Virus (HIV). Background

The human immunodeficiency virus, HIV, leads to acquired immune deficiency syndrome, AIDS. HIV is a retrovirus and like most of the viruses in this family of viruses, the Retroviridae, only replicates in dividing cells.

Infection by the virus HIV–1, the most common variety, has many highly complex characteristics, most of which are still not understood. The fact that the disease progression can last more than 10 years from the first day of infection is just one of them. Another is that while most viral infections can be eliminated by an immune response, HIV is only briefly controlled by it. HIV primarily infects a class of white blood cells or lymphocytes, called CD4 T–cells, but also infects other cells such as dendritic cells. The virus has a high affinity for a receptor present on the cell surface of each of these cells which guides the virus to their location in vivo. When the CD4 T–cell count, normally around $1000/\mu L$, decreases to $200/\mu L$ or below, a patient is characterized as having AIDS.

The reason for the fall in the T–cell count is unknown. T–cells are normally replenished very quickly in the body, so the infection may affect the source of new T–cells or the life span of preexisting ones. Although HIV can kill cells that it infects, only a small fraction of CD4 T–cells are infected at any given time. Because of the central role of CD4 T–cells in immune regulation, their depletion has widespread deleterious effects on the functioning of the immune system as a whole and this is what leads to AIDS.

Most models describing the evolution of AIDS are deterministic, with the attempt to reflect the dynamic changes in mean cell numbers. Such models are more applicable to later stages of the process when the population is large. These models typically consider the dynamics of the CD4 cells, latently infected cells and virus populations as well as the effects of drug therapy.

Because of the ethics, among other things, of doing experiments on humans, fundamental information has been lacking about the dynamics of HIV infection. For example, since the disease takes an average of 10 years to develop it was widely thought that the components of the disease process would also be slow. A combination of mathematical modeling and experiment has shown this is not the case by showing that there are a number of different timescales in HIV infection, from minutes to hours and days to months. The current understanding of the rapidity of HIV infection has totally changed the manner in which HIV is treated in patients and has had a major impact in extending peoples' lives.

Figure 3.2 shows a typical course of HIV infection. Immediately after infection the amount of virus detected in the blood, V, increases rapidly. After a few weeks to months the symptoms disappear and the virus concentration falls to a lower level. An immune response to the virus occurs and antibodies against the virus can be detected in the blood. A test, now highly refined, to detect these antibodies determines if a person has been exposed to HIV. If the antibodies are detected, a person is said to be HIV–positive.

The level the virus falls to after the initial infection has been called the set–point. The viral concentration then seems to remain at a quasi-steady state level during which the concentration of CD4 T–cells measured in blood slowly declines. This period in which the virus concentration stays relatively constant but in which the T–cell count slowly falls is typically a period in which the infected person has no disease symptoms.

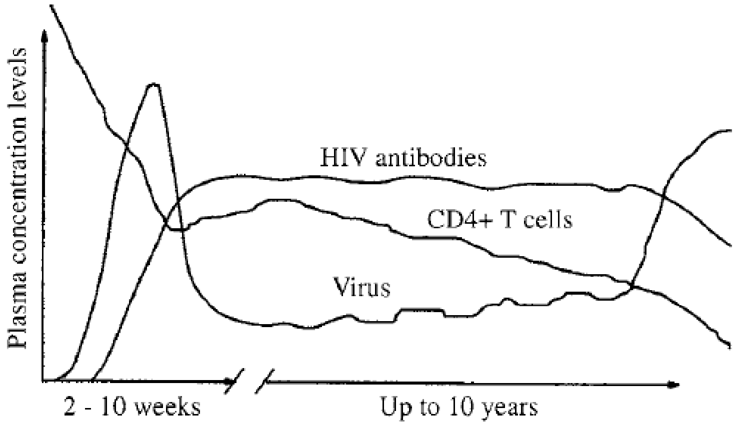

Figure 3.2: Schematic time course of a typical HIV infection in an infected adult. The viral concentration, the level of antibodies and the CD4 T-cells are sketched as a function of time. The early peak corresponds to the primary infection which leads to a period of latency. Note the typical gradual decline in the level of CD4 T–cells over the years. Eventually the symptoms of full–blown AIDS start to appear. (Taken from the Book of J. D. Murray)

A key question then is what is going on during this asymptomatic period. Many believed that the virus was simply quiescent or latent during this period, as seen in other viral diseases, such as herpes. One method of determining whether or not the virus is active is to perturb the host–virus system during the asymptomatic period. In the mid– 1990's work started on new antiretroviral drugs, the protease inhibitors. With their introduction it became possible to perturb the host–virus system during the asymptomatic period. In 1994, David Ho (Aaron Diamond AIDS Research Center) ran an experiment which examined the response of 20 patients infected with HIV to the protease inhibitor, ritonavir. The results were dramatic. The amount of virus measured in blood plasma fell rapidly once the drug was given. Alan Perelson (Los Alamos National Laboratory) and his colleagues then developed a model system which was applied to the patient data and estimations of crucial parameters were obtained. The work is reported in Ho et al. (1995).

Before discussing a model which includes protease inhibitor treatment, we first describe an early model by Anderson et al. (1986) for pedagogical reasons since it is a common way of constructing an epidemic model using a flow chart. It is much less specific and less directly related to current HIV thinking than the one we discuss below in relation to the data and qualitative behavior of the virus as shown in Figure 3.2.

3.3.2 Basic Epidemic Model for HIV Infection in a Homosexual Population

Here we are interested in the development of an AIDS epidemic in a homosexual population. Let us assume there is a constant immigration rate B of susceptible males into a population of size $N(t)$. Let $X(t)$, $Y(t)$, $A(t)$ and $Z(t)$ denote respectively the number of susceptibles, infectious males, AIDS patients and the number of HIV–positive or seropositive men who are noninfectious. We assume susceptibles die naturally at a rate μ; if there were no AIDS, the steady state population would then be $N^* = B/\mu$. We assume AIDS patients die at a rate d: $1/d$ is of the order of months to years, more often the latter. Figure 3.3 is a flow diagram of the disease on which we base our model.

As in previous models we consider uniform mixing. A reasonable first model system, based on the flow diagram in Figure 3.3, is then

$$\frac{dX}{dt} = B - \mu X - \lambda c X, \qquad \lambda = \frac{\beta Y}{N}, \tag{3.3.1}$$

$$\frac{dY}{dt} = \lambda c X - (v + \mu) Y, \tag{3.3.2}$$

$$\frac{dA}{dt} = pvY - (d + \mu) A, \tag{3.3.3}$$

$$\frac{dZ}{dt} = (1 - p) vY - \mu Z, \tag{3.3.4}$$

$$N(t) = X(t) + Y(t) + A(t) + Z(t). \tag{3.3.5}$$

Here B is the recruitment rate of susceptibles, μ is the natural (non–AIDS–related) death rate, λ is the probability of acquiring infection from a randomly chosen partner ($\lambda = \beta Y/N$ where β is the transmission probability), c is the number of sexual partners, d is the AIDS–related death rate, p is the proportion of HIV–positives who are infectious and v is the rate of conversion from infection to AIDS here taken to be constant. $1/v$, equal to D say, is then the average incubation time of the disease. (Actually λ here is more appropriately $\beta Y/(X+Y+Z)$ but A is considered small in comparison with N.) Note that in this model the total population

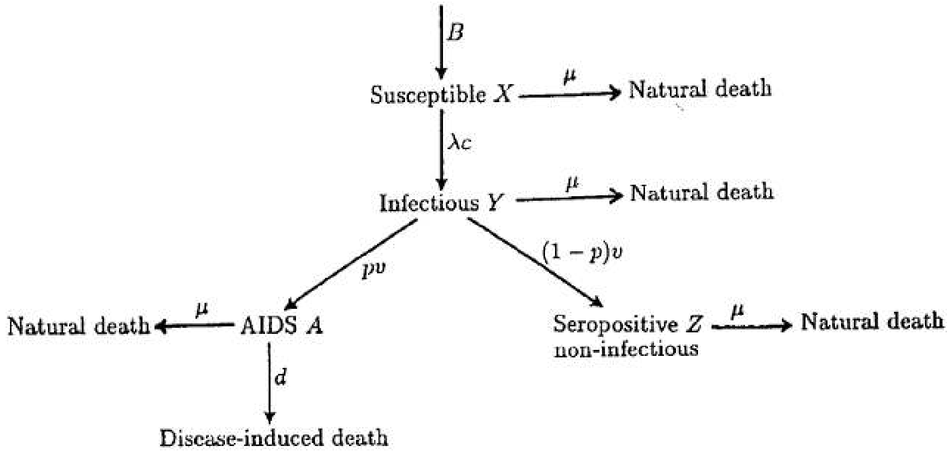

Figure 3.3: The flow diagram of the disease as modelled by the system (3.3.1)–(3.3.5). B represents the recruitment of susceptibles into the homosexual community. The rate of transferral from the susceptible to the infectious class is λc, where λ is the probability of acquiring infection from a randomly chosen partner and c is the number of sexual partners. A proportion of the infectious class is assumed to become noninfectious with the rest developing AIDS. Natural (non–AIDS induced) death is also included in the model. Parameters are defined in the text.

$N(t)$ is not constant, as was the case in the epidemic models in the previous sections. If we add equations (3.3.1)–(3.3.5) we get

$$\frac{dN}{dt} = B - \mu N - dA. \tag{3.3.6}$$

An epidemic ensues if the basic reproductive rate $R_0 > 1$: that is, the number of secondary infections which arise from a primary infection is greater than 1. In (3.3.5) if, at $t = 0$, an infected individual is introduced into an otherwise infection–free population of susceptibles, we have initially $X \approx N$ and so near $t = 0$,

$$\frac{dY}{dt} \approx (\beta c - v - \mu)Y \approx v(R_0 - 1)Y \tag{3.3.7}$$

since the average incubation time, $1/v$, from infection to development of the disease, is very much shorter than the average life expectancy, $1/\mu$, of a susceptible; that is, $v \ll \mu$. Thus the approximate threshold condition for an epidemic to start is, from the last equation,

$$R_0 \approx \frac{\beta c}{v} > 1. \tag{3.3.8}$$

Here the basic reproductive rate R_0 is given in terms of the number of sexual partners c, the transmission probability β and the average incubation time of the disease $1/v$.

When an epidemic starts, the system (3.3.1)–(3.3.5) evolves to a steady state given by

$$\begin{aligned} X^* &= \frac{(v+\mu)N^*}{c\beta}, \qquad Y^* = \frac{(d+\mu)(B-\mu N^*)}{pvd} \\ Z^* &= \frac{(1-p)(d+\mu)(B-\mu N^*)}{pd\mu}, \qquad A^* = \frac{B-\mu N^*}{d} \\ N^* &= \frac{B\beta[\mu(v+d+\mu)+vd(1-p)]}{[v+\mu][b(d+\mu)-pv]}. \end{aligned} \tag{3.3.9}$$

If we linearize about this steady state it can be shown that (X, Y, Z, A) tends to (X^*, Y^*, Z^*, A^*) in a damped oscillatory manner with a period of oscillation given in terms of the model parameters; the method to obtain this is exactly the same as described in the previous chapters but the algebra is messy. We omit the details.

3.3.3 HIV: Modeling Combination Drug Therapy

We start this subsection with the simple, but experimentally based model, proposed by Perelson et al. (1996). We then develop a more complex nonlinear model which includes treatment for HIV infection with a protease inhibitor and a reverse transcription inhibitor such as AZT.

The Ho et al. (1995) model was a simple linear first-order equation which accounted for viral production and viral decline; namely,

$$\frac{dV}{dt} = P - cV, \tag{3.3.10}$$

where P represented a source of viral peptides and c was the viral clearance rate. While many factors play a role in the clearance of viral peptides such as immune cells, fluid flow and absorption into other cells, c did not distinguish between them. After introduction of the protease inhibitor (the specific type of drug used on the patients) it was assumed that the drug would be completely effective, or in other words, the drug would block all viral production after being introduced. Hence $P = 0$, and we are left with the simple equation

$$\frac{dV}{dt} = -cV \qquad \Rightarrow \qquad V(t) = V_0 e^{-ct}, \tag{3.3.11}$$

where V_0 is measured as the mean viral concentration in the plasma before treatment. Plotting lnV against t and using linear regression to determine the slope gave an estimate for c and hence for the half–life of the virus in the plasma; namely, $t_{1/2} = ln2/c$. The mean for the half–life was $t_{1/2} = 2.1 \pm 0.4$ days. The experimentalists then assumed that the patients were in a quasi–steady state before treatment: that is, the levels of viral load measured in the plasma remained fairly constant. With this assumption, and knowing the value for c and the initial viral concentration, V_0, they were able to compute the viral production before therapy by solving $P = cV$. While these results were minimal estimates, based on the assumption of a perfect drug (with no delays), they still provided an estimate of over 1 billion viral particles being produced daily. This important result was contrary to the belief that the viral dynamics during this latent period was close to dormant. It is an excellent example where even simple, mathematically trivial, models can be of immense help in extracting crucial information from patient data.

Due to these results many more models have been developed to study the HIV. In the rest of this section we examine one model which looks at combination drug therapy.

Protease inhibitors are drugs which target the protease enzymes in the cell and cause newly produced viruses to be noninfectious. To date there is no single drug (nor even a combination of them) which completely kills the HIV infection because of the ability of the virus to mutate into a drug resistant form. It takes time, however, for a new form to evolve. The idea behind combination drug treatment is when the virus is presented with two quite different antiviral drugs the time it takes for a mutiple–drug resistant strain to emerge is much longer than if the virus had to contend with only one toxic drug. The use of multiple drug treatments, such as protease inhibitors together with AZT, has already had a major effect (in the developed world) in significantly slowing down the progression from HIV infection to full–blown AIDS. It has not, however, effected a cure for the disease.

We consider each drug to be less then perfect, which thus allows for viral mutation to a resistant form if administered independently. Let n_p be a measure of the effectiveness of a protease inhibitor or combination of protease inhibitors in blocking production of infectious virions so this will affect the viral dynamics directly and the T–cells indirectly. Other commonly used drugs are reverse transcriptase inhibitors, of which AZT is perhaps the best known. After the development of the protease inhibitors, a combination, or cocktail, therapy which included multiple drugs was prescribed. For instance, patients would take a combination of three drugs made of up of a protease inhibitor and two reverse transcriptase inhibitors. This combination was dramatic initially in reducing the number of viral peptides detectable in the patient and it was thought that this might be the cure for the AIDS virus. Unfortunately, with a virus as complex as the HIV it was only a matter of time before the emergence of resistant viruses. While the combination treatment is still showing promise for prolonging the lives of infected patients, it is too early (2001) to say whether or not the virus is even permanently controlled, far less cured.

We develop (just for illustrative purposes, we don't develop any mathematical tool here) a four–species model which includes an equation for uninfected T–cells, T, productively infected

T–cells, T^* (not all infected T–cells produce the virus), infectious viruses, V_I and noninfectious viruses, V_{NI}. The model consists of the following equations which we motivate in turn below.

$$\begin{aligned}
\frac{dT}{dt} &= s + pT\left(1 - \frac{T}{T_{max}}\right) - d_T T - kV_I T,\\
\frac{dT^*}{dt} &= (1 - n_{rt})kV_I T - \delta T^*\\
\frac{dV_I}{dt} &= (1 - n_p)N\delta T^* - cV_I \qquad (3.3.12)\\
\frac{dV_{NI}}{dt} &= n_p N\delta T^* - cV_{NI}.
\end{aligned}$$

In the T–cell equation we consider the cells to be destroyed proportional to the number of infected viruses and cells with clearance parameter k. In the absence of infection there is a nonzero steady state, T_{s1}, so we have a quadratic polynomial in T for the uninfected T–cell dynamics: s, p, T_{max}, d_T and k are positive constants. The specific form of the T–cell kinetics, namely, with a logistic form plus another source (s) and a clearance term ($-d_T T$), is because of the form of T–cell recovery after therapy as indicated by patient data. With the reverse transcriptase (RT) drug like AZT, the RT–inhibitor acts on the source term for productively infected T–cells with $0 \leq n_{rt} \leq 1$ the measure of its efficacy; if $n_{rt} = 1$ it is completely effective and prevents all production of infected T–cells while if $n_{rt} = 0$ it implies no RT–inhibitor is given. In the T^* equation the effect of the RT–inhibitor is to reduce the production of the infected cells. These cells also have a natural death with a rate parameter, δ. The protease inhibitor acts on the source of the virus and so appears in the V_I equation with n_p a measure of its efficacy. The specific appearance in the equations for the effects of the drugs is due to the cellular mechanisms of each drug and the stage at which they aim to target during infection.

When a drug is completely effective we set $n_p = 1$ or $n_{rt} = 1$. In the infected virus V_I equation there is a factor N which is the bursting parameter for the viral production after lysis (essentially the breaking up, or death, of the cell due to its penetration by the infected virus and subsequent generation of a large number of viruses); it is of the order of 480 virions/cell (a virion is a complete virus with all its coating, proteins and so on). The infected viruses are considered to die naturally at a rate c. Finally the noninfectious viruses are produced with a rate dependent on the protease drug and we assume they die off at the same rate as the infected ones. This model lets us explore the effect of the drugs on the HIV by varying, in particular, the parameters n_{rt} and n_p. For example, if $n_p = 0$ we are using only the reverse transcriptase, or RT–inhibitors.

Chapter 4

Introduction to space-time depending models

The models covered in the first half of the biomathematics course describe the evolution of time-dependent quantities such as a total population of an animal species or the total concentration of a chemical substance. Although many of them are complex enough to describe the interplay between initial parameters and final outcome (i. e. asymptotic behavior, stability, and so on), they are somewhat useless in those situations in which one expects the described quantity (or quantities) to behave an a *inhomogeneous* way, that is, to behave differently depending on the *position* of the individuals. In this case, the dependence of a *space* variable x should be introduced. Depending on the context, x may be an element of $\mathbb{R}^n$, $n = 1, 2, 3$. For instance, in a model for the concentration of humans in a crowded tunnel, one can (with good approximation) consider the position x to be one-dimensional. In cell biology, very often $x \in \mathbb{R}^2$. More generally, x is a three-dimensional vector.

We shall therefore start dealing with models where the population of a certain species is not homogeneous in the reference environment. Therefore, it has to be modeled via a density function which depends on a *space* variable x besides the time variable t. Consequently, the differential models we shall recover in this context are in terms of *partial differential equations* or systems. Due to the extreme complexity of the area of partial differential equations and systems, it is not worthwhile to devote a whole section to partial differential equations (PDE's), whereas we shall recall the needed mathematical tools depending on the different contexts.

4.1 Transport and reaction effects

The main distinction we emphasize between all the possible space-dependent effects considered here is the following classification:

- **Transport effects**. When the total amount of individuals in an arbitrary region $A \subset \mathbb{R}^n$ changes only because of the flow through the boundary ∂A, and there is no birth-death mechanism in the evolution, we say that individuals are subject to a transport effect. This definition is reminiscent of the fact that the distribution of individuals at a certain position can change only because of their *movement*.
- **Reaction effects**. When total number of individuals in $A \subset \mathbb{R}^n$ changes also in

view of birth-death mechanisms, which produces a possible change in time of the total population, we talk about reaction effects. The terminology comes from chemistry, and it is reminiscent of the creation or destruction of a chemical substance after a reaction process.

We now introduce the basic equation to describe a *transport effect*, namely the *continuity equation*. Such equation has been already introduced in the classical mechanics and electrodynamics courses, in a non-rigorous form. Typically, the main assumption behind the formulation of a continuity equation is the so-called *continuum assumption*, i.e. we assume that the population under study (particles, bacteria, animals) can be well approximated by a a continuum quantity, in which individuals are not distinguishable. Such an assumption is well justified in many contexts of classical physics such as fluid mechanics and electromagnetism. On the other hand, the *individual-based* nature of models in biology and social sciences make this assumption less trivial to accept. This is why we choose here to make a first, partially formal, justification of the continuity equation as derived from a simple model of moving particles, or moving agents.

We start by considering a very simple case, namely that of an individual with position $x(t)$ moving on a region in space, say in R^n. Here, $t \geq 0$ is the time variable. We assume that we now the velocity of x, and that such a velocity depends on the position in which x is located and on the time t. Clearly, we have

$$\dot{x}(t) = v(t, x(t)). \tag{4.1.1}$$

We assume that v has the properties that ensure existence and uniqueness for (4.1.1) with a given initial condition $x(0) = x_0 \in \mathbb{R}^d$. Now we define an object which is known from distribution theory, namely the Dirac's delta measure δ_x. For a given test function $\phi \in C_c^\infty(\mathbb{R}^d)$, we have

$$\langle \delta_x, \phi \rangle = \phi(x).$$

We shall use simple properties of distributions such as linearity, definition of derivatives[1], multiplication by a smooth function[2]. We set

$$\mu(t) = \delta_{x(t)}.$$

Due to the time dependency, the action of the above distribution has to be defined on a space-time space. More precisely, given $\phi \in C_c^\infty((0, +\infty) \times \mathbb{R}^d)$, we have

$$\langle \mu(t), \phi \rangle = \int_0^{+\infty} \phi(t, x(t)) dt.$$

Lemma 4.1.1 *For all $T \geq 0$ and for all $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$, we have*

$$\langle \mu_t + \mathrm{div}(\mu v), \phi \rangle = 0,$$

i.e., the measure $\mu(t)$ satisfies the continuity equation

$$\partial_t \mu + \mathrm{div}(\mu v) = 0,$$

in the sense of distributions.

---

[1] `https://en.wikipedia.org/wiki/Distribution_(mathematics)#Derivatives_of_distributions`

[2] `https://en.wikipedia.org/wiki/Distribution_(mathematics)#Multiplication_by_a_smooth_function`

Proof.

Compute, due to (4.1.1),

$$\begin{aligned}
&\langle \mu_t + \operatorname{div}(\mu v), \phi\rangle = -\langle \mu, \phi_t\rangle - \langle \mu, v\phi_x\rangle \\
&\quad = -\int_0^T \left(\phi_t(t,x(t)) + v(t,x(t))\phi_x(t,x(t))\right) dt = -\int_0^T \left(\phi_t(t,x(t)) + \dot{x}(t)\phi_x(t,x(t))\right) dt \\
&\quad = -\int_0^T \frac{d}{dt}\phi(t,x(t))dt = \phi(0,x(0)) - \phi(T,x(T)) = 0,
\end{aligned}$$

where the last step is due to the fact that ϕ is supported in a compact subset of the open strip $0 < t < T$.

The previous computation can be easily generalized to the case of N particles with mass $m_1, \dots, m_N > 0$, located on $x_1(t), \dots, x_N(t)$. We consider the distribution

$$\mu(t) = \sum_{i=1}^{N} m_i \delta_{x_i(t)},$$

and assume that

$$\dot{x}_i(t) = v(t, x_i(t)),$$

which means that the individuals are moving according to a velocity field v, which is known. We leave as an exercise to prove that μ satisfies the continuity equation $\mu_t + \operatorname{div}(\mu v) = 0$ in the sense of distributions.

The above approach assumes that one can describe transport phenomena by *counting* the number of individuals (with their masses) on a given region. Now, in many phenomena it is not convenient to work with set of particles and distributions, but rather with *densities.* So, we assume that our set of individuals form a *continuum* in space, namely particles are infinitesimal, and they cannot be distinguished, but rather appear similar to a fluid.

If we replace then $\mu(t)$ by a density $\rho(t,x)$, we obtain the classical continuity equation

$$\partial_t \rho + \operatorname{div}(\rho v) = 0.$$

It is natural to assume that the quantity

$$\int_{\mathbb{R}^n} \rho(0,x)dx,$$

expressing the total amount of particles at the initial time, is finite. Now, if we assume $\rho(0,\cdot)$ to be *compactly supported* and that particles travel with *finite speed* v at all times, we easily obtain that the support of $\rho(t,\cdot)$ is compact at all times. Therefore

$$\frac{d}{dt}\int_{\mathbb{R}^n} \rho(t,x)dx = \int_{\mathbb{R}^n} \rho_t(t,x)dx = -\int_{\mathbb{R}^n} \operatorname{div}(\rho v)dx,$$

and by the divergence theorem the above quantity is equal to the surface integral

$$\int_{\partial B_t} \rho(t,x)v(t,x)d\sigma_x dt,$$

where B_t is the support of ρ at time t. Hence, such a quantity is equal to zero, and we get

$$\frac{d}{dt}\int_{\mathbb{R}^n}\rho(t,x)dx = 0,$$

which implies

$$\int_{\mathbb{R}^n}\rho(t,x)dx = \int_{\mathbb{R}^n}\rho(0,x)dx$$

for all times. Such a property is called *conservation of the total mass*, and is typical of transport phenomena.

Particularly interesting is the quantity ρv in the continuity equation. It is called the *flux*, and is sometimes denoted with $J = \rho v$. The flux has the dimension of a mass divided by a volume (which is the dimension of a density) times a velocity, which translates into mass per time times area of a surface. So, physically speaking, the flux on a given point x describes the amount of mass passing through an infinitesimal surface around x. The flux is a vector quantity, directed towards the velocity v. Given a smooth surface S, one can define the flux of ρ on S as

$$\int_{\partial S} J(x)\cdot \nu(x)d\sigma_x,$$

where $\nu(x)$ is the unit normal of S on the point x. If S is a closed surface, e.g. the boundary of a bounded domain in $\mathbb{R}^n$, such a quantity describes *the amount of mass that is leaving the domain* per unit time.

The continuity equation only expresses transport of individuals. It is not taking into account of loss or gain of the total number of individuals. This is the reason behind the conservation of the total mass. In order to deal with *reaction phenomena*, in which the total mass may change, one has to consider the local rate of change of the total mass. We define

$$f(t,x,\rho(x))$$

as the growth rate of the population described by the density $\rho(t,x)$. This means that on an arbitrary region $A\subset\mathbb{R}^n$ we have

$$\frac{d}{dt}\int_A \rho(t,x)dx = \int_A f(t,x,\rho(x))dx - \int_{\partial A}\rho v\cdot \nu d\sigma_x,$$

in which we have also considered the outgoing flux, i.e. the amount of mass leaving the region due to the transport phenomenon. In the most general case, f depends on the position x, on the time t, and typically in nonlinear phenomena also on the density itself. The above identity gives

$$\int_A \rho_t(t,x)dx + \int_{\partial A}\rho v\cdot\nu d\sigma_x = \int_A f(t,x,\rho(x))dx.$$

Once again, the divergence theorem implies

$$\int_A \rho_t(t,x)dx + \int_A \operatorname{div}(\rho v)dx = \int_A f(t,x,\rho(x))dx.$$

The above identity can be rephrased as

$$\int_A \left[\rho_t + \operatorname{div}(\rho v) - f(t,x,\rho)\right]dx = 0.$$

Since A is an arbitrary (measurable) set, one gets from a classical result in measure theory, the reaction-transport equation

$$\rho_t(t,x) + \operatorname{div}(\rho(t,x)v(t,x)) = f(t,x,\rho) \tag{4.1.2}$$

Chapter 5

Diffusion equations

The most classical phenomenon in biology accounting for a pure transport of individuals is *diffusion.*

In an assemblage of particles, for example, cells, bacteria, chemicals, animals and so on, each particle usually moves around in a random way. The particles spread out as a result of this irregular individual particle's motion. When this microscopic irregular movement results in some macroscopic or gross regular motion of the group we can think of it as a *diffusion* process. Of course there may be interaction between particles, for example, or the environment may give some bias in which case the gross movement is not simple diffusion. To get the macroscopic behavior from a knowledge of the individual microscopic behavior is much too hard so we derive a continuum model equation for the *global* behavior in terms of a particle *density* or *concentration.*

It is instructive to start with a random process which we look at probabilistically in an elementary way, and then derive a deterministic model. Before doing that, we first provide another classical interpretation of diffusion in terms of the celebrated *Fick's law.*

5.1 Diffusion–reaction processes and Fick's law

Diffusion mechanism models the movement of many individuals in an environment or media. The individuals can be very small such as basic particles in physics, bacteria, molecules, or cells, or very large objects such as animals, plants, or certain kind of events like epidemics, or tumors. The particles reside in a region, which we call Ω and we assume that Ω is open subset of $\mathbb{R}^n$ with $n \geq 1$. In particular, we are interested in the cases of $n = 1, 2$ and 3, but most material here are true regardless of the dimensions of the space. The main mathematical variable we consider here is the density function of the particles $\rho(t, x)$, where t is the time, and $x \in \Omega$ is the location. The dimension of the population density usually is number of particles or organisms per unit area (if $n = 2$) or unit volume (if $n = 3$).

The question we are interested now is how the function $\rho(t, x)$ changes as time t evolves, and as the location x varies. It is a natural phenomenon that a substance goes *from high density regions to low density* regions. The 'high to low' principle means that the flux always points to the most rapid decreasing direction of $\rho(t, x)$, which is the negative gradient of $\rho(t, x)$. This principle is called *Fick's law*, and it can be represented as

$$J(t, x) = -d(x)\nabla_x \rho(t, x) \tag{5.1.1}$$

where J is the flux of ρ, $d(x) \geq 0$ is called *diffusion coefficient* at x.

On the other hand, the number of particles at any point may change because of other reasons like birth, death, hunting, or chemical reactions. We assume that the rate of change of the density function due to these reasons is $f(t, x, P)$, which we usually call the *reaction rate*. Now we derive a differential equation using the balanced law.

In terms of the reaction-transport equation (4.1.2), here no reaction occurs, so $f = 0$, and J is given by Fick's law. The diffusion coefficient $d(x)$ is not a constant in general since the environment is usually heterogeneous. So, in general we get

$$\frac{\partial \rho(t,x)}{\partial t} = \operatorname{div}(d(x)\nabla\rho(t,x)). \tag{5.1.2}$$

When the region of the diffusion is approximately homogeneous, we can assume that $d(x) \equiv d$, then (5.1.2) can be simplified to

$$\frac{\partial \rho(t,x)}{\partial t} = d\Delta\rho(t,x), \tag{5.1.3}$$

where Δ is the Laplacian operator. In classical mathematical physics, the equation $\partial_t T = \Delta T$ is called *heat equation*, where T is the temperature function. In our context, (5.1.3) is called *linear diffusion equation.*

5.2 Simple Random Walk and Derivation of the Diffusion Equation

In an assemblage of particles, for example, cells, bacteria, chemicals, animals and so on, each particle usually moves around in a random way. The particles spread out as a result of this irregular individual particle's motion. When this microscopic irregular movement results in some macroscopic or gross regular motion of the group we can think of it as a *diffusion process.* Of course there may be interaction between particles, for example, or the environment may give some bias in which case the gross movement is not simple diffusion. To get the macroscopic behavior from a knowledge of the individual microscopic behavior is much too hard so we derive a continuum model equation for the global behavior in terms of a particle density or concentration. It is instructive to start with a random process which we look at probabilistically in an elementary way, and then derive a deterministic model.

For simplicity we consider initially only one–dimensional motion and the simplest random walk process. The generalization to higher dimensions is then intuitively clear from the one–dimensional equation.

Suppose a particle moves randomly backward and forward along a line in fixed steps Δx that are taken in a fixed time Δt. If the motion is unbiased then it is equally probable that the particle takes a step to the right or left. After time $N\Delta t$ the particle can be anywhere from $-N\Delta x$ to $N\Delta x$ if we take the starting point of the particle as the origin. The spatial distribution is clearly not going to be uniform if we release a group of particles about $x = 0$ since the probability of a particle reaching $x = N\Delta x$ after N steps is very small compared with that for x nearer $x = 0$.

We want the probability $p(m, n)$ that a particle reaches a point m space steps to the right (that is, to $x = m\Delta x$) after n time–steps (that is, after a time $n\Delta t$). Let us suppose that to reach $m\Delta x$ it has moved a steps to the right and b to the left. Then

$$m = a - b, \qquad a + b = n \qquad \Rightarrow a = \frac{n+m}{2}, \qquad b = n - a = \frac{n-m}{2}.$$

The number of possible paths that a particle can reach this point $x = m\Delta x$ is

$$\frac{n!}{a!b!} = \frac{n!}{a!(n-a)!} = C_a^n,$$

where C_a^n is the binomial coefficient. The total number of possible n–step paths is 2^n and so the probability $p(m, n)$ (the favorable possibilities/total possibilities) is

$$p(m,n) = \frac{1}{2^n}\frac{n!}{a!(n-a)!}, \qquad a = \frac{m+n}{2}, \tag{5.2.1}$$

notice that $n + m$ is always even. Note that

$$\sum_{m=-n}^{n} p(m,n) = 1,$$

as it must since the sum of all probabilities must equal 1. It is clear mathematically since

$$\sum_{m=-n}^{n} p(m,n) = \sum_{m=-n}^{n} C_a^n \left(\frac{1}{2}\right)^a \left(\frac{1}{2}\right)^{n-a} = \left(\frac{1}{2} + \frac{1}{2}\right)^n = 1,$$

$p(m, n)$ is the *binomial distribution.*

If we now let n be large so that $n \pm m$ are also large we have, asymptotically,

$$n! \sim (2\pi n)^{1/2} n^n e^{-n}, \qquad n \gg 1,$$

which is Stirling's formula. Using that in (5.2.1) we get

$$p(m,n) \sim \frac{1}{(2\pi)^{1/2}} \left(\frac{n}{a(n-a)}\right)^{1/2} \left(\frac{n}{2a}\right)^a \left(\frac{n}{2(n-a)}\right)^{n-a}.$$

From $a = (n + m)/2$ we infer

$$p(m,n) \sim \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{n}{n^2 - m^2}\right)^{1/2} \left(\frac{n}{n+m}\right)^{\frac{n+m}{2}} \left(\frac{n}{n-m}\right)^{\frac{n-m}{2}}.$$

We now assume that the ratio m/n tends to zero as asymptotically, which corresponds to have the number or space-steps grow more slowly than the number of time-steps. This is a key assumption which characterizes a diffusion process, as we shall see later on. As $n \to +\infty$ we clearly see that

$$\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{n}{n^2 - m^2}\right)^{1/2} \sim \left(\frac{2}{\pi n}\right)^{1/2}.$$

Moreover, the remaining factors in $p(m,n)$ are approximated by the second order Taylor approximation of the map $x \mapsto \log(1+x)$ around $x=0$ as follows

$$\begin{aligned}
&\log\left(\frac{n}{n+m}\right)^{\frac{n+m}{2}}\left(\frac{n}{n-m}\right)^{\frac{n-m}{2}}\\
&=\left(\frac{n+m}{2}\right)\log\left(1-\frac{m}{n+m}\right)+\left(\frac{n-m}{2}\right)\log\left(1+\frac{m}{n-m}\right)\\
&=-\frac{m}{2}-\frac{m^2}{4(n+m)}+\frac{m}{2}-\frac{m^2}{4(n-m)}+o(1/n)=-\frac{m^2}{2n}+o(1/n).
\end{aligned}$$

Therefore,

$$p(m,n)\sim\left(\frac{2}{\pi n}\right)^{1/2}e^{-\frac{m^2}{2n}}, \tag{5.2.2}$$

which is a *Gaussian probability distribution.*

Now set

$$m\Delta x = x, \qquad n\Delta t = t,$$

where x and t are the continuous space and time variables. If we anticipate letting $m\to+\infty$, $n\to+\infty$, $\Delta t\to 0$, $\Delta x\to 0$ so that x and t are finite, then it is not appropriate to have $p(m,n)$ as the quantity of interest since this probability must tend to zero: the number of points on the line tends to $+\infty$ as $\Delta x\to 0$. The relevant dependent variable is more appropriately $u=p/(2\Delta x)$: $2u\Delta x$ is the probability of finding a particle in the interval $(x,x+\Delta x)$ at time t. From (5.2.2) with $m=x/\Delta x$, $n=t/\Delta t$,

$$u(x,t)=\lim_{\Delta x\to 0,\Delta t\to 0}\frac{p\left(\frac{x}{\Delta x},\frac{t}{\Delta t}\right)}{2\Delta x}=\lim_{\Delta x\to 0,\Delta t\to 0}\left\{\frac{\Delta t}{2\pi t(\Delta x)^2}\right\}^{1/2}e^{-\frac{x^2}{2t}\frac{\Delta t}{(\Delta x)^2}}.$$

If we assume

$$\lim_{\Delta x\to 0,\Delta t\to 0}\frac{(\Delta x)^2}{2\Delta t}=D>0, \tag{5.2.3}$$

the last equation gives

$$u(x,t)=\left(\frac{1}{4\pi Dt}\right)^{1/2}e^{-\frac{x^2}{4Dt}}. \tag{5.2.4}$$

Please notice that (5.2.3) is consistent with the assumption $m/n\to 0$ above. Here, D is the *diffusion coefficient* or *diffusivity* of the particles; note that it has dimensions (length)^2/(time). It is a measure of how efficiently the particles disperse from a high to a low density. For example, in blood, haemoglobin molecules have a diffusion coefficient of the order of $10.7 cm^2 sec^{-1}$ while that for oxygen in blood is of the order of $10.5 cm^2 sec^{-1}$.

Let us now relate this result to the classical approach to diffusion, namely, Fickian diffusion. In the simplest case this gives the linear heat equation

$$\frac{\partial c}{\partial t}=D\frac{\partial^2 c}{\partial x^2} \tag{5.2.5}$$

satisfied by the concentration of the species c.

If we release an amount Q of particles per unit area at $x=0$ at $t=0$, that is,

$$c(x,0)=Q\delta_0(x)$$

where $\delta_0(x)$ is the Dirac delta function, then the solution of (5.2.5) is

$$c(x,t) = \frac{Q}{(4\pi Dt)^{1/2}} e^{-x^2/(4Dt)}, \qquad t > 0$$

which, with $Q = 1$, is the same result as (5.2.4), obtained from a random walk approach when the step and time sizes are small compared with x and t.

5.3 The gaussian distribution

In this section we prove the last assertion of the previous section, namely we aim to find a solution to the diffusion equation

$$c_t = Dc_{xx}, \qquad x \in \mathbb{R}, \qquad t > 0, \tag{5.3.1}$$

of *self–similar* form

$$c(x,t) = \frac{C}{\sqrt{t}} U\left(\xi\right), \qquad \xi = \frac{x}{\sqrt{t}}, \qquad C > 0. \tag{5.3.2}$$

In order to do that, let us plug the above expression (5.3.2) into the equation (5.3.1). After simple computations, assuming $t > 0$ we recover

$$U_{\xi\xi} + \frac{1}{2D}(\xi U)_\xi = 0,$$

which is equivalent to

$$\left(U\left(\log U + \frac{\xi^2}{4D}\right)_\xi\right)\xi = 0.$$

A solution can be found imposing that the term in the inner bracket is zero, namely

$$\log\frac{U(\xi)}{C} = -\frac{\xi^2}{4D},$$

which yields

$$U(\xi) = Ce^{-\frac{\xi^2}{4D}}.$$

In the original variables, this formula gives the solution

$$c(x,t) = G(x,t) = \frac{Q}{(4Dt\pi)}^{1/2} e^{-\frac{x^2}{4Dt}},$$

where $Q > 0$ is the initial total mass of u. G is called *Gaussian distribution*, or Gaussian solution to the diffusion equation.

It is an easy exercise to prove that, as $t \to 0$, $G(x,t) \to 0$ if $x \neq 0$. In case $x = 0$ we clearly have $G(0,t) \to +\infty$ as $t \to 0$. On the other hand, the mass of G is constant in time, so that the initial datum of G must have positive mass Q. This is not possible in case the initial datum of G is an L^1 function ($G(\cdot, 0)$ should have zero mass because it is zero almost everywhere). Function theory is not suitable to give an interpretation of such phenomenon and one has to deal with distribution theory, which is not among the purposes of this course.

In fact the initial datum of G is the *Dirac delta distribution* δ_0, which we will consider to be defined as the 'function' which satisfies

$$\int_{-\infty}^{+\infty} \delta_0(x)dx = 1, \qquad \delta_0(x) = 0 \ \text{ for all } \ x \neq 0.$$

The idea behind the concept of a delta distribution is to model a situation in which all the particles (individuals, cells...) are all packed together at one point (or very close to one point), and the concentration is zero elsewhere.

The Gaussian solution models exactly the fact that particles initially concentrated at one point start spreading as $t > 0$ towards regions with zero concentration, according to Fick's law. The maximum value of the concentration, initially $+\infty$ achieved at $x = 0$ (the maximum can be achieved at any other point just by translation, please notice that the equation (5.3.1) is invariant after translation with respect to x), decreases as t increases. It is always achieved at $x = 0$ and it is given by $\frac{Q}{(4Dt\pi)}^{1/2}$. Clearly, the concentration of particles at any other point than zero *increases*, and the profile of the distribution takes the shape of a bell with fatter and fatter tails as time increases, see figure 5.1.

It is straightforward to notice that at a fixed time t, the profile of G has fatter tails when the diffusion coefficient D is larger. This corresponds to the intuition that a larger diffusion coefficient implies a faster *spreading* process of the particles.

5.4 Smoothing and decay properties of the diffusion operator

The linear diffusion operator in a reaction diffusion equation (or system) accounts for the spreading of the individuals far from each other due to Fick's law. It is a sort of repulsive effect. In case only diffusion is present in the model, supposing for instance that the model is set up in a container $\Omega \subset \mathbb{R}^n$ with Neumann boundary conditions, we expect that all individuals tend more and more to distribute homogenously, eventually reaching a constant distribution state. This is due to the fact that the initial flux of individuals from regions with high density toward regions with low density is soon compensated by the movement of the individuals, so that the flux intensity decreases more and more until it stabilized to a constant values. As long as we are dealing with linear phenomena (as it is the case for the Heat equation), one can use the method of separation of variables in order to see this fact from the mathematical point of view. However, we shall prove this stabilization phenomenon by a *energy method*, and keep the method of separation of variables for a later case. In order to simplify the coverage, we consider the following one–dimensional problem

$$\begin{cases} u_t = Du_{xx} & x \in [0, L], \ t \geq 0 \\ u(0, x) = u_0(x) & x \in [0, L] \\ u_x(t, 0) = u_x(t, L) = 0 & t \geq 0. \end{cases} \tag{5.4.1}$$

Clearly, all constants $u \equiv C$ are solutions of the heat equation $u_t = Du_{xx}$ with homogeneous Neumann boundary conditions $u_x(t, 0) = u_x(t, L) = 0$. However, only *one* constant solution is feasible as a candidate to represent the *large time behavior* of the solution of the transient problem (5.4.1), that is

$$u^\infty(t, x) \equiv \frac{1}{L} \int_0^L u_0(x)dx = \frac{1}{L} \int_0^L u(t, x)dx =: \frac{M}{L}$$

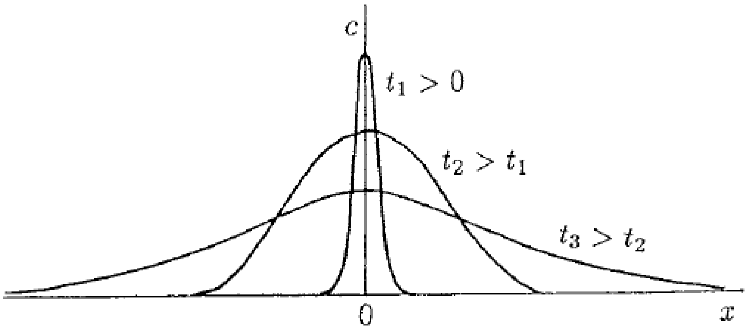

Figure 5.1: Schematic particle concentration distribution arising from Q particles released at $x = 0$ at $t = 0$ and diffusing according to the diffusion equation (5.3.1).

where the equality in the middle is a trivial consequence of the Neumann boundary conditions; it expresses the fact that the total amount of individuals remains unchanged in time.

The energy method consists in evaluating the evolution of a certain functional related to the solution $u(t)$. This functional should contain enough information in order to infer qualitative properties of the solution itself. A typical choice in case of linear problems if the L^2 norm of the solution $u(t)$

$$\int_0^L u(t,x)^2 dx.$$

Actually, since we aim to prove that $u(t)$ behaves like u^∞ as time goes to infinity, a smarter choice of the energy is

$$E(t) := \int_0^L \left(u(t,x) - u^\infty\right)^2 dx,$$

so that $E(t)$ being small as $t \to +\infty$ implies that $u(t)$ is approaching u^∞ in some sense. As for the interpretation of the use of the L^2 norm, we refer to basic textbooks of Functional Analysis. We just remark here that, in the present context, $E(t) \to 0$ implies $u(t,x) \to u^\infty$ for almost all $x \in [0,L]$ as $t \to +\infty$.

Let us then compute

$$\frac{d}{dt}E(t) = 2\int_0^L u u_t dx = 2D\int_0^L u u_{xx} dx = -2D\int_0^L u_x^2 dx$$

and the last step is due to integration by parts and the Neumann boundary conditions. The above computation already gives an interesting result: the energy $E(t)$ is non increasing. Actually we can do better and use the following version of the Poincaré inequality, the proof of which can be found for instance in the PDE book by L. C. Evans.

Theorem 5.4.1 (Poincaré inequality) *Let $f : [0,L] \to \mathbb{R}$ such that*

$$\int_0^L (f'(x))^2 dx < +\infty, \qquad \int_0^L f(x) dx = 0.$$

Then, there exists a constant $C(L)$ depending only on L such that

$$\int_0^L f(x)^2 dx \leq C(L)\int_0^L (f'(x))^2 dx.$$

Due to the above inequality, the energy estimate becomes

$$\frac{d}{dt}E(t) \leq -\frac{2D}{C(L)}E(t)$$

and therefore the following exponential decay trivially follows

$$E(t) \leq E(0)e^{-\frac{2D}{C(L)}t}$$

and in particular one has

$$\lim_{t\to+\infty}\int_0^L \left(u(t,x) - u^\infty\right)^2 dx = 0.$$

Notice that the fact that u and u^∞ have the same mass is crucial in order to apply Poincaré inequality.

The exponential convergence to equilibrium states is not the only 'nice' property of the heat equation. Among the others, an important one is given by the so–called *smoothing effect*, a property which can be resumed as follows: no matter how singular the initial datum u_0 is (discontinuous, unbounded, etc...), the solution $u(t)$ belongs to C^∞ both in space and in time at any positive time t. In order to see this, let us recover the exact formula for u by *the method of separation of the variables*. Let us choose u of the form

$$u(t,x) = f(x)g(t)$$

for certain functions f and g. Substituting the above ansatz into the equation in (5.4.1) yields

$$f(x)g'(t) = Df''(x)g(t)$$

and therefore

$$\frac{g'(t)}{g(t)} = D\frac{f''(x)}{f(x)} = D\lambda \in \mathbb{R}$$

for a certain constant λ, because of the fact that the two sides in the first inequality above depend only on t and x respectively. We assume that both f and g never vanish. Due to the boundary conditions we have $f'(0) = f'(L) = 0$, therefore we have to solve the limiting problem for all possible $\lambda \in \mathbb{R}$

$$\begin{cases} f'' = \lambda f & x \in [0,L] \\ f'(0) = f'(L) = 0. \end{cases}$$

By imposing the boundary conditions one easily infer that non trivial solutions only exist if $\lambda \leq 0$. In particular, in case $\lambda = 0$ all constants $f(x) \equiv C$ are solutions. In case of a negative $\lambda = -\omega^2$ we have

$$f(x) = A\cos(\omega x) + B\sin(\omega x).$$

Again, the boundary condition $f'(0) = 0$ implies $B = 0$ whereas $f'(L) = -A\omega\sin(\omega x) = 0$ implies $\omega L = k\pi$ for an integer k. Therefore, all values $\lambda = -\left(\frac{k\pi}{L}\right)^2$ allow for nontrivial solutions of the form $f_k(x) = A_k\cos\left(\frac{k\pi}{L}x\right)$. Solving the equation for g implies

$$g'(t) = -D\left(\frac{k\pi}{L}\right)^2 g(t)$$

and hence we have a family of solutions indexed by the integer k

$$g_k(t) = E_k e^{-D\left(\frac{k\pi}{L}\right)^2 t}.$$

Due to the superposition principle for linear PDEs, any combination of $f_k g_k$ is still a solution of the PDE in (5.4.1) endowed with the boundary conditions. Therefore, in order to achieve the unique solution $u(t)$ we express u in the following series expression

$$u(t,x) = \sum_{k=0}^{+\infty} A_k e^{-D\left(\frac{k\pi}{L}\right)^2 t}\cos\left(\frac{k\pi}{L}x\right). \tag{5.4.2}$$

The coefficients A_k are easily obtained by matching $u|_{t=0}$ with the initial datum u_0

$$u_0(x) = \sum_{k=0}^{+\infty} A_k\cos\left(\frac{k\pi}{L}x\right),$$

which implies that A_k are the Fourier's coefficients of u_0. Let us assume that $\int |u_0(x)|dx < +\infty$. Then, standard results on the Fourier's series imply that all the coefficients A_k are uniformly bounded. It is then clear that we can differentiate with respect to x and t infinitely many times in the formula (5.4.2) and this proves the smoothing effect for the heat equation. Note that the only assumption on u_0 is $\int |u_0(x)|dx < +\infty$, which possibly allows for discontinuities and unboundedness in a set of zero Lebesgue measure.

Exercise 5.4.2 Use formula (5.4.2) to prove that the solution u converges toward its initial average $\frac{1}{L}\int u_0(x)dx$ as $t \to +\infty$.

One can easily prove that a similar phenomenon occurs on unbounded domains as well, in any dimensions. For instance, consider the Cauchy problem on $\mathbb{R}^n$

$$\begin{cases} u_t = D\Delta u & x \in \mathbb{R}^n,\ t \geq 0 \\ u(0,x) = u_0(x) \geq 0 & x \in \mathbb{R}^n. \end{cases} \tag{5.4.3}$$

It is well known (cf. classical textbooks in PDEs) that the solution $u(x,t)$ can be represented via convolution with the Gaussian solution, namely

$$u(x,t) = \int_{\mathbb{R}^n} G(x-y,t)u_0(y)dy.$$

Now, suppose as an example that u models the density of ants on a two dimensional 'large' domain such as the floor or a very large room. Suppose that a group of ants is initially concentrated on a small tile $T \subset \mathbb{R}^2$. This means that u_0 is supported on T. Let us now consider a point x_0 arbitrarily far away form the tile T, and consider a time $t_0 > 0$ arbitrarily close to zero. A direct evaluation of the convolution integral above shows that $u(x_0,t_0) > 0$, which means that ants have been so fast in moving that some of them could reach x in an arbitrarily small time. With such an argument, one can prove that ants can move with arbitrarily large (and therefore infinite, somewhat) speed. Such phenomenon is somehow non realistic. One possible way to improve the model in this sense is described in the following section.

5.5 Nonlinear diffusion

Diffusion models form a reasonable basis for studying insect and animal dispersal and invasion; one extension of the classical diffusion model which is of particular relevance to insect dispersal is when there is an increase in diffusion due to population pressure. One such model has the diffusion coefficient, or rather the flux J, depending on the population density n such that the diffusivity D increases with n; that is,

$$J = -D(n)\nabla n, \qquad D'(n) > 0.$$

A typical form for $D(n)$ is $D_0(n/n_0)^m$, where $m > 0$ and D_0 and n_0 are positive constants. The dispersal equation for n without any growth term is then

$$\frac{\partial n}{\partial t} = D_0 \mathrm{div}\left[\left(\frac{n}{n_0}\right)^m \nabla n\right].$$

In one dimension
$$\frac{\partial n}{\partial t} = D_0 \frac{\partial}{\partial x}\left[\left(\frac{n}{n_0}\right)^m \frac{\partial n}{\partial x}\right].$$
After scaling the space and the time variable suitably, one can recover the following dimensionless form
$$\frac{\partial u}{\partial t} = \Delta u^\gamma, \qquad \gamma > 1. \tag{5.5.1}$$
Now, for such an equation a solution of self similar form can be recovered in a similar way as for the linear diffusion equation. We therefore look for a solution of the form

$$u(x,t) = t^{-\beta}U(xt^{-\alpha}), \qquad x \in \mathbb{R}, \qquad t > 0$$

to the one–dimensional nonlinear diffusion equation (also called Porous Medium Equation)

$$u_t = (u^\gamma)_{xx}.$$

First of all, the conservation of the total mass easily implies

$$\int_{\mathbb{R}} u(x,t)dx = t^{-\beta}\int_{\mathbb{R}} U(xt^{-\alpha})dx = t^{-\beta+\alpha}\int_{\mathbb{R}} U(\xi)d\xi$$

and therefore we have $\beta = \alpha$.

In the new variables U and $\xi = xt^{-\alpha}$ we obtain

$$\alpha t^{-\alpha-1}(\xi U)_\xi + t^{-\alpha(m+2)}(U^m)_{\xi\xi} = 0,$$

which requires the choice
$$\alpha = \frac{1}{m+1}$$
in order to get rid of the time variable. The above can be written as

$$\partial_x\left(U\partial_x\left(\frac{m}{m-1}U^{m-1} + \frac{\xi^2}{2(m+1)}\right)\right) = 0.$$

In the domain where $U > 0$, we can impose

$$U^{m-1}(\xi) = \left(C - \frac{m-1}{2m(m+1)}\xi^2\right).$$

Clearly, the above gives problems in case U becomes negative. Since $U = 0$ on an interval solves the equation for U, we can introduce the following solution

$$U(\xi)\left[C - \frac{m-1}{2m(m+1)}\xi^2\right]_+^{\frac{1}{m-1}},$$

which in the original variables reads

$$u(x,t) = B(x,t) = t^{\frac{1}{m+1}}\left[C - \frac{m-1}{2m(m+1)}\left(\frac{x}{\sqrt{t}}\right)^2\right]_+^{\frac{1}{m-1}}$$

which is called *Barenblatt solution.* Such a solution has compact support on the interval

$$-t^{\frac{1}{m+1}}\sqrt{\frac{2mC(m+1)}{m-1}} \leq x \leq t^{\frac{1}{m+1}}\sqrt{\frac{2mC(m+1)}{m-1}}$$

which grows as t grows. Moreover, it can be easily proven that $B(\cdot,0)$ is a multiple of the Dirac delta distribution, exactly as it is the case for the Gaussian solution of the linear diffusion equation.

Analyzing the differences with the linear diffusion case, we clearly see that the Barenblatt solution is not smooth, opposite to the Gaussian solution to the linear diffusion equation, which is C^∞. More precisely, the Barenblatt solution has possible lack of smoothness at the boundary of its support, where it eventually features discontinuities in a space derivative of a certain order depending on the exponent m. At any other points, B is C^∞.

A key difference between G and B is that the support of G becomes unbounded immediately after $t = 0$ (infinite speed of propagation) whereas the support of B is *compact for all times*, it travels with *finite speed.* This fact depends on the density dependent diffusivity $D(n)$ being equal to zero as $n = 0$. Such an assumptions corresponds to state that when the concentration of particles is very low, particles are subject to a much lower diffusivity (in other context we could say ‘pressure’) than in situations with higher concentration.

Later on in this course we shall devote ourselves to the existence theory for the nonlinear diffusion equation for general initial conditions.

Chapter 6

Reaction–diffusion models for one single species

In this chapter we start to consider models combining the diffusion effect described in the previous chapter with *reaction* effects, typically corresponding with the evolutionary mechanism described in part 1. At this first stage, we start considering only scalar equations.

6.1 Diffusive Malthus equation and critical patch size

Let us now apply the method of separation of the variables to the simplest reaction diffusion model we can imagine, namely the one–dimensional linear *Diffusive–Malthus equation*

$$u_t = Du_{xx} + au, \qquad D > 0, \quad a \in \mathbb{R}, \tag{6.1.1}$$

posed on the interval $x \in [0, L]$ with initial datum $u(0, x) = u_0(x)$ and endowed with the homogeneous Dirichlet boundary conditions $u(t, 0) = u(t, L) = 0$. We are therefore modeling a context of a population with a linear exponential growth and with diffusion, the two phenomena being respectively quantified by the two parameters a and D. The homogeneous Dirichlet boundary conditions are usually referred to as *hostile boundary* conditions, since they model the fact that individuals are somehow pushed away from the boundary.

In order to solve (6.1.1), let us set $u(t, x) = f(x)g(t)$. We then obtain

$$Df''(x)g(t) + af(x)g(t) = f(x)g'(t), \qquad D\frac{f''}{f} + a = \frac{g'}{g} = D\lambda, \quad \lambda \in \mathbb{R},$$

which suggests looking for solutions to the two ordinary differential equations

$$\begin{aligned} &Df''(x) = \left(\lambda - \frac{a}{D}\right) f(x), \qquad f(0) = f(L) = 0 \\ &g'(t) = \lambda g(t). \end{aligned}$$

Similarly to the previous calculations, we have $\lambda - a/D < 0$ (the case $\lambda - a/D = 0$ this time only gives the zero solution due to the Dirichlet boundary data), therefore $\lambda - a/D = -\omega^2$, $\omega \in \mathbb{R}$ and this corresponds to

$$f(x) = A\sin(\omega x) + B\cos(\omega x).$$

The boundary condition $f(0) = 0$ implies $B = 0$, whereas $f(L) = 0$ implies $\omega L = k\pi$. Therefore, once again we have a sequence of (eigen)values $\lambda_k = a - \left(\frac{k\pi}{L}\right)^2$ corresponding to the sequence of (eigen)solutions

$$f_k(x) = A_k \sin\left(\frac{k\pi}{L}x\right).$$

The equation for g then turns into

$$g'(t) = \left[\frac{a}{D} - \left(\frac{k\pi}{L}\right)^2\right] g(t)$$

which yields

$$g_k(t) = C_k e^{\left[\frac{a}{D} - \left(\frac{k\pi}{L}\right)^2\right] t}.$$

By the superposition principle, similarly to what we saw in the previous section, we get

$$u(t,x) = \sum_{k=1}^{+\infty} A_k e^{\left[\frac{a}{D} - \left(\frac{k\pi}{L}\right)^2\right] t} \sin\left(\frac{k\pi}{L}x\right) \tag{6.1.2}$$

where A_k are the Fourier's coefficients of the initial datum u_0.

The above formula (6.1.2) clearly shows that if $a < 0$ the zero solution is asymptotically stable, since all the terms in the series decay to zero exponentially fast as $t \to +\infty$. We know that this is also the case when diffusion is absent, as we have seen at the very beginning of the chapter **??**. We shall show that the presence of diffusion *improves* the range of a's for which we have stability.

Let us then consider the case $a > 0$: we clearly see from the first term in the above series that, if

$$L < \sqrt{\frac{D}{a}}\pi, \tag{6.1.3}$$

then the solution $u(t,x)$ decays to zero as $t \to +\infty$ provided the usual standard conditions are required on u_0 (see the previous section). The critical value in (6.1.3) is called *Critical Patch Size*. The interpretation of the condition (6.1.3) is the following. When the size of the interval is not enough, the individuals reach the boundary soon enough to be subject to the hostile boundary conditions (which 'kill' the concentration) rather than to the reaction term. Of course, another possible interpretation of (6.1.3) can be given in terms of the relative size of D and a. More precisely, when the ratio D/a is high enough, diffusion is the dominant effect. Now, it is easy to check (exercise!) that $u \equiv 0$ is the only stationary solution to (6.1.1) with homogeneous Dirichlet boundary conditions. Moreover, this solution is asymptotically stable when there's no reaction, namely when $a = 0$ (easy exercise! Use the representation formula (6.1.2)). Therefore, we expect that $u \equiv 0$ is the asymptotic limit as $t \to +\infty$ of the particle concentration when diffusion dominates the process. In fact, this is the case.

In case (6.1.3) is not satisfied, then the behavior of u depends on the initial datum u_0. This is clear form the representation formula (6.1.2). For instance, if the first Fourier's coefficient of u_0, A_1, is not zero, then the exponential factor $e^{\left(\frac{a}{D} - \frac{\pi^2}{L^2}\right)t}$ appears in the solution formula. In this case the solution $u(t,x)$ grows to $+\infty$ as t goes to $+\infty$ when (6.1.3) is violated. On the other hand, the initial datum could have a certain number of zero Fourier's coefficients, say the first $m > 1$ coefficients, $A_1 = \cdots = A_m = 0$. In this case the condition yielding decay to zero of u would be

$$L < (m+1)\sqrt{\frac{D}{a}}\pi \tag{6.1.4}$$

since the first exponential factor appearing in (6.1.3) would be $e^{\left(\frac{a}{D}-\frac{((m+1)\pi)^2}{L^2}\right)t}$. Clearly, (6.1.4) is weaker than (6.1.3), which means that the critical patch size is higher: particles need a much higher reaction rate a in order to grow exponentially due to the Malthus term. The interpretation of that is given in term of the Fourier series of u_0

$$u_0(x) = \sum_{k\geq 0} A_k \sin\left(\frac{k\pi}{L}x\right).$$

The higher the order k is, the higher is the value of the gradient of u_0 in absolute value, which means the more diffusion is dominant (remember Fick's law!). Therefore, even with a larger interval, particles are subject to such a higher diffusion effect that we still don't see the growth due to the reaction as t becomes large.

6.2 Local existence of solutions to the reaction diffusion equation

We now focus on the nonlinear reaction-diffusion equation

$$u_t = Du_{xx} + g(u), \tag{6.2.1}$$

for a general nonlinear function $g \in C^1([0,+\infty))$ with the assumption $g(0) = 0$. The presence of a nonlinear term makes more unlikely that an explicit formula for such an equation with given initial condition holds. On the other hand, we can make use of the so-called *Duhamel principle.*

Lemma 6.2.1 *Let* $G(t,x) = (4D\pi t)^{1/2}e^{-x^2/4Dt}$ *be the one-dimensional heat kernel, and let* $a(t,x)$ *be a continuous function on* $[0,+\infty)\times\mathbb{R}$ *and* $u_0 \in L^1 \cap L^\infty(\mathbb{R})$. *Set*

$$u(t,x) = \int_{\mathbb{R}} G(t,x-y)u_0(y)dy + \int_0^t\int_{\mathbb{R}} G(t-s,x-y)a(s,y)dsdy.$$

Then, $u \in C^\infty_{t,x}$ *and* u *solved the Cauchy problem*

$$\begin{cases} u_t = Du_{xx} + a(t,x) \\ u(0,x) = u_0(x). \end{cases} \tag{6.2.2}$$

Proof. We only give sketch of the proof, and leave the details as an exercise. At $t = 0$, the initial condition is satisfied because $G(0,\cdot)$ i the Dirac delta distribution, and hence $\int_{\mathbb{R}} G(t,x-y)u_0(y)dy = G(t) * u_0(x) = u_0(x)$. Then, the differentiation of u with respect to t yields

$$u_t(t,x) = \int_{\mathbb{R}} G_t(t,x-y)u_0(y)dy + \int_{\mathbb{R}} G(0,x-y)a(t,y)dy + \int_0^t\int_{\mathbb{R}} G_t(t-s,x-y)a(s,y)dyds.$$

Now, since G satisfies the heat equation, and since $G(0,\cdot) = \delta_0$, we have

$$u_t(t,x) = D\partial_{xx}\int_{\mathbb{R}} G(t,x-y)u_0(y)dy + a(t,x) + D\partial_{xx}\int_0^t\int_{\mathbb{R}} G(t-s,x-y)a(s,y)dyds = Du_{xx} + a.$$

Now, our goal is to use the above Lemma to solve the nonlinear problem (6.2.1), with the idea of substituting $g(\rho)$ to a in (6.2.2). Clearly, the formula provided by the above Lemma will not be an explicit formula. Indeed, we get

$$u(t,x) = \int_{\mathbb{R}} G(t,x-y)u_0(y)dy + \int_0^t\int_{\mathbb{R}} G(t-s,x-y)g(u(s,y))dsdy. \tag{6.2.3}$$

Formula (6.2.3) expresses a mild formulation of the problem (6.2.1) with initial condition $u(0,x) = u_0(x)$. It is not an explicit formula, but the result in the previous Lemma shows (upon minor modifications that are left as an exercise) that if u satisfies (6.2.3), then u solves the Cauchy problem for (6.2.1) with initial condition u_0.

Hence, in order to solve such a problem we have to solve (6.2.3). Such an equation can be seen as a fixed point condition for an operator on a Banach space. For fixed $T > 0$, we set

$$|||u|||_T = \sup_{0\leq t\leq T}\left[\|u(t,\cdot)\|_{L^1(\mathbb{R})} + \|u(t,\cdot)\|_{L^\infty(\mathbb{R})}\right].$$

Then, for another fixed constant $R > 0$ we set

$$X_{T,R} = \{f : [0,T]\times\mathbb{R}\to\mathbb{R};\ |||f|||_T \leq R\}.$$

We then define the linear operator $A : X_{T,R} \to X_{T,R}$

$$(Au)(t,x) = \int_{\mathbb{R}} G(t,x-y)u_0(y)dy + \int_0^t\int_{\mathbb{R}} G(t-s,x-y)g(u(s,y))dsdy.$$

We first prove that A is well posed as an operator from $X_{T,R}$ to itself. Given $u \in_{T,R}$, set $M = \|u_0\|_{L^1(\mathbb{R})} + \|u_0\|_{L^\infty(\mathbb{R})}$. We use the Young inequality for convolutions $\|f * g\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q}$ with $1 + 1/r = 1/p + 1/q$ and recall that $\|G(t,\cdot)\|_{L^1(\mathbb{R})} = 1$. We get

$$\begin{aligned}&\|Au(t,\cdot)\|_{L^1(\mathbb{R})} \leq \|G(t,\cdot) * u_0\|_{L^1(\mathbb{R})} + \int_0^t \|G(t-s,\cdot) * g(u(s,\cdot))\|_{L^1(\mathbb{R})}ds\\ &\leq M + \int_0^t \|g'\|_{L^\infty([0,R])}\|u(s,\cdot)\|_{L^1(\mathbb{R})}ds.\end{aligned}$$

Here we have use the mean value theorem and $g(0) = 0$ to justify $g(u(s,y)) = g'(\bar{u}(s,y))u(s,y)$ for some $\bar{u}(s,y) \in [0,R]$, and then we easily see that $g'(\bar{u}(s,y)) \leq \|g'\|_{L^\infty([0,R])}$. Such a term is a continuous function of R on $[0,+\infty)$. Similarly, we get

$$\|Au(t,\cdot)\|_{L^\infty(\mathbb{R})} \leq M + \int_0^t \|g'\|_{L^\infty([0,R])}\|u(s,\cdot)\|_{L^\infty(\mathbb{R})}ds.$$

The above computations imply

$$|||Au|||_T \leq 2M + 2T\|g'\|_{L^\infty([0,R])}R.$$

It is easy to see that for R large enough and for T small enough we can enforce

$$2M + 2T\|g'\|_{L^\infty([0,R])}R \leq R,$$

which is equivalent to

$$T \leq \frac{R-2M}{\|g'\|_{L^\infty([0,R])}R}.$$

With such restrictions on T and R, we have that A is well defined from $X_{T,R}$ to itself. We now want to prove that for a suitable choice of T, A is a contraction on $X_{T,R}$. For given $u_1, u_2 \in X_{T,R}$, let us compute

$$\|Au_1(t,\cdot) - Au_2(t,\cdot)\|_{L^1(\mathbb{R})} \leq \|g'\|_{L^\infty([0,R])}\int_0^t \|G(t-s,\cdot)\|_{L^1(\mathbb{R})}\|u_1(s,\cdot)-u_2(s,\cdot)\|_{L^1(\mathbb{R})}ds,$$

and a similar computation holds for the L^∞ difference $\|u_1(t,\cdot)-u_2(t,\cdot)\|_{L^\infty(\mathbb{R})}$, and this shows that for $T < \frac{1}{2}\|g'\|_{L^\infty([0,R])}$ the map A is a contraction. Since the space $X_{T,R}$ is a closed subset of a complete metric space, we can invoke Banach fixed point theorem to show that A has a unique fixed point.

As we stated above, a similar proof to that of the above Lemma shows that u is a classical solution to $u_t = Du_{xx} + g(u)$ on a time interval $[0,T]$. Since u is $X_{T,R}$, we have $\sup_{t\in[0,T]}\left(\|u(t)\|_{L^1(\mathbb{R})} + \|u(t)\|_{L^\infty(\mathbb{R})}\right) \leq R$. We now prove that u is nonnegative on $[0,T]$. Consider the negative part function

$$u_- = \max\{0,-u\},$$

and its approximation

$$\eta_n(u) = \begin{cases} \frac{-1}{2n} - u & \text{if } u \leq -1/n \\ n\frac{u^2}{2} & \text{if } -1/n \leq u \leq 0 \\ 0 & \text{if } u \geq 0. \end{cases}$$

It is easy to see that $\eta_n(u) \to u_-$ as $n \to +\infty$ for all $u \in \mathbb{R}$. Moreover η_n is non increasing and convex. we estimate

$$\frac{d}{dt}\int_{\mathbb{R}} \eta_\delta(u(t,x))dx = \int_{\mathbb{R}} \eta_n'(u)u_t dx = \int_{\mathbb{R}} \eta_n'(u)(Du_{xx} - g(u))dx.$$

Now, neglecting the boundary conditions (we omit the details, but essentially the boundary terms are vanishing due to the fact that u decays at infinity) we can integrate by parts and get

$$\frac{d}{dt}\int_{\mathbb{R}} \eta_\delta(u(t,x))dx = -D\int_{\mathbb{R}} \eta_n''(u)u_x^2 dx + \int_{\mathbb{R}} \eta_n'(u)g(u)dx \leq \int_{\mathbb{R}} \eta_n'(u)g(u)dx.$$

Now, since u is bounded on $[0,T]$ and since $g(0)=0$, we can find a constant $C>0$ such that

$$|g(u)| \leq C|u|,$$

for all $t \in [0,T]$. Hence,

$$\frac{d}{dt}\int_{\mathbb{R}} \eta_\delta(u(t,x))dx \leq C\int_{u\leq 0} \eta_n'(u)|u|dx = C\int_{u\leq 0} |\eta_n'(u)|u_- dx.$$

Integrating on $[0,t]$ with $0 \le t \le T$, we get

$$\int_{\mathbb{R}} \eta_\delta(u(t,x))dx \le \int_{\mathbb{R}} \eta_\delta(u_0(x))dx - C\int_0^t \int_{u\le 0} \eta_n'(u)u_- dx.$$

Now we let $n \to +\infty$ to get

$$\int_{\mathbb{R}} (u(t,x))_- dx \le \int_{\mathbb{R}} (u_0(x))_- dx + C\int_0^t \int_{\mathbb{R}} (u(t,x))_- dx,$$

where we have used $-\eta_n'(u)u_- \to u_-$ as $n \to +\infty$. Now, if we assume that u_0 is nonnegative (as it is the case, since the initial condition expresses a density), we have $(u_0)_- = 0$, thus Gronwall inequality implies that $\int_{\mathbb{R}}(u(t,x))_- dx = 0$ for all $t \in [0,T]$. Since u_- is a nonnegative function, this implies that $u_- = 0$ on $[0,T]\times\mathbb{R}$, which implies that u is never negative on $[0,T]\times\mathbb{R}$.

As we could see above, we need to choose T small in order to find a solution u to our problem. Hence, we have proven a *local existence and uniqueness* result. On the other hand, the following example shows that in general one cannot expect a global in time solution for any g.

Example 6.2.2 *Let us consider the initial boundary value problem*

$$\begin{cases} u_t = Du_{xx} + u^2(t,x) & x \in [0,L]\ t \ge 0 \\ u(0,x) = u_0(x) & x \in [0,L] \\ u_x(t,0) = u_x(t,L) = 0 & t \ge 0. \end{cases}$$

Assume the initial condition is constant, i. e. $u_0(x) = \bar{u}$*. Then, let us solve the ODE*

$$\begin{cases} U_t = U^2 & t \ge 0 \\ U(0) = \bar{u}. \end{cases}$$

We immediately get the solution

$$U(t) = \frac{\bar{u}}{1-\bar{u}t},$$

which shows that $u(t,x) = U(t)$ *solves the above initial boundary value problem. Now, it is immediately seen that* u *blows up as* $t \to 1/\bar{u}$*, so we cannot expect global existence for all times in general.*

On the other hand, there are reasonable conditions on g which ensure global existence. Let us estimate the L^2 norm of solutions to the equation (6.2.1) on the whole real line. Neglecting the boundary conditions (same considerations as above in the computation of the positivity of u), we get

$$\begin{aligned} \frac{d}{dt}\int_{\mathbb{R}} u^2(t,x)dx &= -\int_{\mathbb{R}} u_x^2(t,x)dx + \int_R g(u)udx \\ &\le \|g'\|_{L^\infty(\mathbb{R})}\int_{\mathbb{R}} u^2. \end{aligned}$$

The above computations show that if either $g(u) \le 0$ for all u or g is globally Lipschitz, then the L^2 norm of $u(t)$ if finite at all times, and therefore no blow up occurs in the L^2 norm.

Exercise 6.2.3 *Show that global existence in* L^2 *holds for* (6.2.1) *with the logistic growth reaction term* $g(u) = u(1-u)$.

6.3 Asymptotic stability of constant states for general nonlinear reaction diffusion equations

In this section we tackle the problem of the asymptotic stability of a constant (homogeneous) state for the unknown variable u in a quite general context, namely we consider the *nonlinear reaction diffusion equation*

$$u_t = Du_{xx} + g(u), \tag{6.3.1}$$

where g is a general nonlinear function modeling a reaction process, $D > 0$ is the diffusivity. Our goal is to prove that under certain conditions on g, a homogeneous stationary state is *asimptotically stable*. Clearly, by homogeneous stationary state we mean a solution $u(x,t) \equiv u^\infty$ with $g(u^\infty) = 0$.

Going back to the scalar ODE model

$$u' = g(u)$$

studied at the beginning of the course, we recall that a general condition for the stability of u^∞ was

$$g'(u^\infty) < 0. \tag{6.3.2}$$

Therefore, hinted by the results in Section 6.1, we expect condition (6.3.2) to be sufficient for stability as we know that diffusion contributes to reinforce stability.

6.3.1 The case with Neumann condition

We work on a bounded interval $x \in [0,L]$ with homogenous Neumann boundary conditions

$$u_x(0,t) = u_x(L,t) = 0 \tag{6.3.3}$$

and with an initial datum

$$u(x,0) = u_0(x). \tag{6.3.4}$$

Of course, the presence of diffusion makes the problem of proving stability under the effect of diffusion much more complicated, since it is a matter of stability in a *functional* sense, i. e. the solution $u(x,t)$ is seen as a curve $t \mapsto u(\cdot,t)$ in a certain functional space (in which the independent variable is x). In order to clarify this point, we introduce the functional spaces we shall need in this section. The space $L^2([0,L])$ is defined as the space of functions f on the interval $[0,L]$ such that $\int_0^L f(x)^2 dx < +\infty$. Such a definition may be not completely well posed mathematically, for instance in case the function f is not continuous. We send the reader to a Functional Analysis textbook for a more clear explanation. To our purposes, we can assume that all involved functions are continuous. The L^2 norm of f is defined then as

$$\|f\|_{L^2} = \left(\int_0^L f(x)^2 dx\right)^{1/2}.$$

We shall need also the functional spaces $H^s([0,L])$, $s \in \mathbb{N}$, namely, $f \in H^s([0,L])$ if and only if f and its derivatives up to order s are all elements of L^2, or equivalently

$$\int_0^L f(x)^2 dx + \int_0^L f'(x)^2 dx + \ldots + \int_0^L D^s f(x)^2 dx < +\infty.$$

We recall the definition

$$\|f\|_{H^s}^2 = \int_0^L f(x)^2 dx + \int_0^L f'(x)^2 dx + \ldots + \int_0^L D^s f(x)^2 dx < +\infty.$$

The reason why we have introduced all the above functional norms is that we want to use the *energy method* to prove stability. This is probably the only way to develop a general theory for a general g, more specific cases can be dealt with depending on the particular form of g. In order to prove that the perturbation of the stationary state is small for large times, we denote

$$u = u^\infty + v$$

and use the Taylor expansion of g around u^∞ in the equation for the perturbation v:

$$\begin{cases} v_t = Dv_{xx} + g(u^\infty + v) = Dv_{xx} + g'(u^\infty)v + h(v) \\ v(x,0) = u_0(x) - u^\infty \\ v_x(0,t) = v_x(L,t) = 0, \end{cases} \tag{6.3.5}$$

where h is a continuous function satisfying $h(v)/v \to 0$ as $v \to 0$. Let us compute the evolution of the squared L^2–norm of v. Here we use integration by parts and the boundary conditions in the computation of the term involving the diffusion.

$$\begin{aligned} \frac{d}{dt}\int_0^L v^2 dx &= 2\int_0^L v v_t dx = 2\int_0^L v(Dv_{xx} + g'(u^\infty)v + h(v))dx \\ &= -2D\int_0^L v_x^2 dx - 2D\left(v v_x\right)_0^L + 2g'(u^\infty)\int_0^L v^2 dx + 2\int_0^L h(v) v dx \\ &= -2D\int_0^L v_x^2 dx + 2g'(u^\infty)\int_0^L v^2 dx + 2\int_0^L \frac{h(v)}{v} v^2 dx, \end{aligned}$$

where we have used that $g'(u^\infty) < 0$ and the zero boundary conditions on v. Now, since $h(v)/v \to 0$ as $v \to 0$, we can write

$$\sup_{x\in[0,L], t\geq 0} \left| \frac{h(v(x,t))}{v(x,t)} \right| \to 0 \quad \text{as} \quad \|v(t)\|_{L^\infty} \to 0,$$

where the L^∞–norm of v is given by $\|v(t)\|_{L^\infty} = \sup_{x\in[0,L]} |v(x,t)|$. Therefore we can estimate the last term in the previous estimate as follows:

$$2\int_0^L \frac{h(v)}{v} v^2 dx \leq H(\|v(t)\|_{L^\infty}) \int_0^L v^2 dx,$$

for some continuous function H such that $H(x) \to 0$ ad $x \to 0$. We can assume that such a function H is *increasing*. For instance, once can define H as follows:

$$H(v) := v + \sup\left\{ \frac{|h(z)|}{z},\ 0 < z \leq v \right\}, \quad \text{as } v > 0,$$

and extend H to $\mathbb{R}$ as an even function. It is an easy exercise to prove that $H \geq 0$, that $H(v) \to 0$ as $v \to 0$ and that

$$|H(v_1)| \leq H(|v_2|) \quad \text{if} \quad |v_1| \leq |v_2|.$$

Now the monotonicity of H in the above estimate implies

$$\frac{d}{dt}\int_0^L v^2 dx \leq -2\left[-g'(u^\infty) - H(\|v(t)\|_{L^\infty})\right]\int_0^L v^2 dx.$$

It is clear from the above estimate that, in case we would be able to say a priory that $\|v(t)\|_{L^\infty}$ is smaller than a certain threshold, then we could prove exponential decay as in the case of the linear diffusion equation provided $g'(u^\infty) < 0$. Of course we cannot get to such a conclusion so far. We need some more machinery. The strategy is to estimate the L^∞–norm of v, $\|v(t)\|_{L^\infty}$, by the L^2 norm of its space derivative. To this purpose we shall use the following special case of the *Sobolev inequality.*

Lemma 6.3.1 (Sobolev inequality) *Let $f \in H^1([0,L])$. Then there exists $C > 0$ independent of f such that*

$$\|f\|_{L^\infty([0,L])} \leq C\|f\|_{H^1([0,L])}, \tag{6.3.6}$$

Proof. Suppose for simplicity that f is continuous (the general case can be proven by an approximation argument which goes beyond the purposes of this course, see for instance the PDE book by Evans). Then, define the average

$$\bar{f} = \frac{1}{L}\int_0^L f(x)dx.$$

The mean value theorem implies that there is a point $c \in [0,L]$ such that $f(c) = \bar{f}$. Let us use the fundamental theorem of integral calculus as follows: for $x \in [0,L]$ we have

$$f^2(x) = f^2(c) + \int_c^x \frac{d}{dy}(f^2(y))dy$$

which implies due to Cauchy-Schwarz inequality

$$|f^2(x)| = |f^2(c)| + 2\left|\int_{x_0}^x f(y)f'(y)dy\right| \leq 2\int_0^L |f(y)||f'(y)|dy$$
$$\leq |f^2(c)| + \int_0^L f^2(y)dy + \int_0^L (f')^2(y)dy = \|f\|^2_{H^1([0,L])}.$$

Now, Cauchy-Schwarz inequality also implies

$$|f^2(c)| = \left|\frac{1}{L}\int_0^L f(x)dx\right|^2 \leq \frac{1}{L}\int_0^L f^2(y)dy,$$

so the desired inequality (6.3.6) easily follows.

Then, we can use (6.3.6) plus the fact that H is increasing to get

$$H(\|v(t)\|_{L^\infty}) \leq H(C\|v_x\|_{H^1})$$

and therefore

$$\frac{d}{dt}\int_0^L v^2 dx \leq -2\left[-g'(u^\infty) - H(C\|v(t)\|_{H^1})\right]\int_0^L v^2 dx. \tag{6.3.7}$$

Now, in order to 'close' such an estimate, we need to estimate the L^2 norm of v as well. To this aim, we write the equation satisfied by v_x:

$$(v_x)_t = D(v_x)_{xx} + g(u^\infty + v)_x = D(v_x)_{xx} + g'(u^\infty + v)v_x = D(v_x)_{xx} + g'(u^\infty)v_x + l(v)v_x,$$

where we have used the Taylor expansion of g' around u^∞. Here $l(v)$ is a continuous function of v such that $l(v) \to 0$ as $v \to 0$. Let us compute, by using the same strategy as before for the diffusion term, namely integration by parts and the boundary conditions,

$$\begin{aligned}
&\frac{d}{dt}\int_0^L v_x^2 dx = 2\int_0^L v_x v_{xt} dx = 2\int_0^L v_x(D(v_x)_{xx} + g'(u^\infty)v_x + l(v)v_x)dx \\
&= -2D\int_0^L v_{xx}^2 dx + 2g'(u^\infty)\int_0^L v_x^2 dx + 2\int_0^L l(v)v_x^2 dx \\
&\leq -2D\int_0^L v_{xx}^2 dx + 2\left(g'(u^\infty) + L(\|v\|_{L^\infty})\right)\int_0^L v_x^2 dx,
\end{aligned}$$

where, as before, L is an increasing and continuous function such that $L(0) = 0$. We apply once again the Sobolev inequality to get

$$\frac{d}{dt}\int_0^L v_x^2 dx \leq -2\left(-g'(u^\infty) - L(C\|v\|_{H^1})\right)\int_0^L v_x^2 dx. \tag{6.3.8}$$

We take the sum of (6.3.7) and (6.3.8) and we obtain

$$\frac{d}{dt}\|v\|_{H^1}^2 \leq -2\left(-g'(u^\infty)\right)\|v\|_{H^1}^2 + \alpha(\|v\|_{H^1}^2)\|v\|_{H^1}^2, \tag{6.3.9}$$

where α is a continuous function such that $\alpha(0) = 0$ obtained by combining M and N in (6.3.7) and (6.3.8) respectively. It is now clear that we shall be able to achieve a satisfactory sufficient condition for stability by requiring

$$g'(u^\infty) < 0. \tag{6.3.10}$$

Let us therefore assume (6.3.10) holds and let $E = |g'(u^\infty)|$. We have then

$$\frac{d}{dt}\|v\|_{H^1}^2 \leq -2E\|v\|_{H^1}^2 + \alpha(\|v\|_{H^1}^2)\|v\|_{H^1}^2.$$

Now, by the comparison principle for ODE's, we can estimate $\|v(t)\|_{H^1}^2$ from above with the solution of the following ODE

$$Y' = -2EY + \alpha(Y)Y, \qquad Y(0) = \|v(0)\|_{H^1}^2.$$

Performing a qualitative study of the above ODE as in the first chapter of this notes, due to the fact that $\alpha(0) = 0$ we can easily see that there exists a fan $A := \{|Y| \leq \delta\}$, for some $\delta > 0$, such that the solutions in A all point towards the axis $Y = 0$, since they have a negative derivative above the axis and a positive derivative below it. Therefore, *if* $|Y(0)| \leq \delta$, then $Y(t) \to 0$ as $t \to +\infty$ and by comparison principle as well $\|v(t)\|_{H^1}^2 \to 0$ as $t \to +\infty$. By using once again the Sobolev inequality, we obtain

$$\|v(t)\|_{L^\infty([0,L])} \to 0 \qquad \text{as} \qquad t \to +\infty,$$

which proves that the perturbation from the equilibrium state converges to zero provided we have the following additional condition:

$$\|v(0)\|_{H^1} \qquad \text{small enough.}$$

The above condition says that the L^2 norm of the perturbation and the L^2 norm of its space derivative should be less than a small enough constant in order to get asymptotic stability. Therefore, we see that even in the case of nonlinear reaction, due to linear diffusion, we get asymptotic stability of homogeneous states under a similar condition for the reaction compared to the ODE case.

Now, one can wonder whether or not such an estimate is sharp, i. e. whether or not it can be improved. The answer is *no*. Take a constant initial condition $\bar{u}$. Assume the solution u depends only on t, say $u(t,x) = U(t)$. We get

$$U_t = g(U),$$

so U solves an ODE. The stability results on ODEs show that $U(t) - u_\infty$ converges to zero after small (constant in space) perturbations of the steady state u_∞. Moreover, it is easy to see that the rate of convergence is provided by the constant $g'(u_\infty)$, and no contribution is played by the diffusion term (the population is already homogeneous!).

In fact, this is not surprising. The linear diffusion term drives the population towards a homogenized state, which is only determined by the fact that the total mass M is preserved, while the reaction leads to the stable steady state. If reaction wouldn't be there, the diffusion operator would lead to the homogeneous state M/L, which is not necessarily equal to u_∞. So, the two terms lead to two different states. Now, when combined together, the two effects interact in the following way: the mass is no longer preserved, so the diffusion operator is only leading towards a generic constant state; it is the reaction term that 'decides' what this state should be, namely the stable steady state u_∞.

6.3.2 The case with Dirichlet conditions

We now consider homogeneous Dirichlet boundary conditions

$$u(0,t) = u(L,t) = 0, \tag{6.3.11}$$

which model a hostile boundary condition as in the linear case. In this case, suggested by the linear diffusion case, we expect that the diffusion operator drives the solution towards the zero state. On the other hand, the reaction term can cooperate to this process only if the steady state u_∞ is the zero state. If one wants to consider a general steady state $u_\infty > 0$, the correct boundary conditions are

$$u(0,t) = u(L,t) = u_\infty. \tag{6.3.12}$$

In this case, the perturbation

$$v = u - u_\infty$$

will be zero at $x = 0$ and $x = L$. Hence, v will satisfy

$$\begin{cases} v_t = Dv_{xx} + g(u^\infty + v) = Dv_{xx} + g'(u^\infty)v + h(v) \\ v(x,0) = u_0(x) - u^\infty \\ v(0,t) = v(L,t) = 0, \end{cases} \tag{6.3.13}$$

A similar energy estimate as in the previous case implies

$$\frac{d}{dt}\int_0^L v^2 dx = -2D\int_0^L v_x^2 dx + \left(2g'(u_\infty) + 2H(\|v\|_\infty)\right)\int_0^L v^2 dx,$$

where H are defined as above.

Theorem 6.3.2 (Poincaré inequality, Dirichlet version) *Let $f : [0, L] \to \mathbb{R}$ such that*

$$\int_0^L (f'(x))^2 dx < +\infty, \qquad f(0) = 0.$$

Then, there exists a constant $C_D(L)$ depending only on L such that

$$\int_0^L f(x)^2 dx \le C_D(L)\int_0^L (f'(x))^2 dx.$$

Proof. The proof works essentially similarly to the Sobolev inequality. It is left as an exercise.

The above inequality implies

$$\frac{d}{dt}\int_0^L v^2 dx \le \left(-\frac{2D}{C_D(L)} + 2g'(u_\infty) + 2H(\|v\|_\infty)\right)\int_0^L v^2 dx.$$

We now have to use the same strategy used in the Neumann case to close the above estimate. We estimate the L^2 norm of v_x as before:

$$\frac{d}{dt}\int_0^L v_x^2 dx \le -2D\int_0^L v_{xx}^2 dx + 2\left(g'(u^\infty) + L(\|v\|_{L^\infty})\right)\int_0^L v_x^2 dx.$$

Now, since $\int_0^L v_x dx = v(L) - v(0) = 0$, we can use the usual Poincaré inequality with zero average constraint. We get

$$\frac{d}{dt}\int_0^L v_x^2 dx \le \left(-\frac{2D}{C(L)} + 2g'(u^\infty) + L(\|v\|_{L^\infty})\right)\int_0^L v_x^2 dx.$$

Hence, it is an easy exercise to generalize the computation we did in the case of Neumann conditions to obtain the following statement: if

$$g'(u_\infty) < 2D\left(\frac{1}{C_D(L)} + \frac{1}{C(L)}\right),$$

then the stationary state u_∞ is asymptotically stable under small perturbations.

We have therefore obtained a similar result to the one on the linear case, namely the range of possible reaction rates $g'(u_\infty)$ which yield stability can slightly exceed the range of positive real numbers provided it is not too big compared to the diffusion constant.

6.4 Travelling waves

There is a vast number of phenomena in biology in which a key element or precursor to a developmental process seems to be the appearance of a *travelling wave* of chemical concentration, mechanical deformation, electrical signal and so on. Looking at almost any film of a developing embryo it is hard not to be struck by the number of *wavelike* events that appear after fertilization. Mechanical waves are perhaps the most obvious. There are, for example, both chemical and mechanical waves which propagate on the surface of many vertebrate eggs. In the case of the egg of the fish Medaka a calcium (Ca++) wave sweeps over the surface; it emanates from the point of sperm entry. Chemical concentration waves are visually dramatic examples. From the analysis on insect dispersal we can also expect wave phenomena in that area, and in interacting population models where spatial effects are important. Another example, related to interacting populations, is the progressing wave of an epidemic, of which the rabies epizootic currently spreading across Europe is a dramatic and disturbing example. The movement of microorganisms moving into a food source, chemotactically directed, is another. The slime mould Dictyostelium discoideum is a particularly widely studied example of chemotaxis; we discuss this phenomenon later.

The point to be emphasized is the widespread existence of wave phenomena in the biomedical sciences which necessitates a study of traveling waves in depth and of the modeling and analysis involved. This section (and another section in the next chapter) deal with various aspects of wave behavior where diffusion plays a crucial role.

We must first decide what we mean by a traveling wave. Customarily a traveling wave is taken to be a *wave which travels without change of shape*, and this will be our understanding here. So, if a solution $u(x,t)$ represents a traveling wave, the shape of the solution will be the same for all time and the speed of propagation of this shape is a constant, which we denote by c. If we look at this wave in a traveling frame moving at speed c it will appear stationary. A mathematical way of saying this is that if the solution u can be represented as

$$u(x,t) = u(x-ct) = u(z), \qquad z = x - ct, \tag{6.4.1}$$

then $u(x,t)$ is a traveling wave, and it moves at constant speed c in the positive x–direction. Clearly if $x - ct$ is constant, so is u. It also means the coordinate system moves with speed c. A wave which moves in the negative x–direction is of the form $u(x+ct)$ with positive c. The wavespeed c generally has to be determined. The dependent variable z is sometimes called the *wave variable*. When we look for traveling wave solutions of an equation or system of equations in x and t in the form (6.4.1), we have

$$\frac{\partial u}{\partial t} = -c\frac{du}{dz}$$
$$\frac{\partial u}{\partial x} = \frac{du}{dz},$$

So partial differential equations in x and t become ordinary differential equations in z. To be physically realistic $u(z)$ has to be bounded for all z and nonnegative with the quantities with which we are concerned, such as chemicals, populations, bacteria and cells.

Let us first point out that without reaction there can be no traveling waves. To see this, consider a solution of the form (6.4.1) to the equation

$$u_t = Du_{xx}.$$

Then we have

$$D\frac{d^2u}{dz^2} + c\frac{du}{dz} = 0$$

which implies

$$u(x,t) = A + Be^{-\frac{c}{D}(x-ct)}$$

for two constants A and B. Since u has to be bounded for all x and t, B must be zero since the exponential becomes unbounded as $x - ct \to -\infty$. $u(z) = A$, a constant, is not a wave solution. In marked contrast the parabolic reaction diffusion equation

$$u_t = Du_{xx} + f(u) \tag{6.4.2}$$

can exhibit traveling wave solutions, depending on the form of the reaction/interaction term $f(u)$. This solution behavior was a major factor in starting the whole mathematical field of reaction diffusion theory.

6.5 An example of travelling wave: the Fisher–Kolmogoroff equation

The classic simplest case of a nonlinear reaction diffusion equation (6.4.2) is

$$\frac{\partial n}{\partial t} = kn\left(1 - \frac{n}{K}\right) + D\frac{\partial^2 u}{\partial x^2}, \tag{6.5.1}$$

where k, D and K are positive parameters. It was suggested by Fisher (1937) as a deterministic version of a stochastic model for the spatial spread of a favored gene in a population. It is also the natural extension of the logistic growth population model discussed when the population disperses via linear diffusion. In this sense, k represent the classical linear growth rate and K the carrying capacity.

As a first step we can consider a first dimensionless form of (6.5.1)

$$\frac{\partial u}{\partial t} = ku(1-u) + D\frac{\partial^2 u}{\partial x^2}, \tag{6.5.2}$$

simply obtained by putting $u = n/K$.

This equation and its traveling wave solutions have been widely studied, as has been the more general form with an appropriate class of functions $f(u)$ replacing $ku(1-u)$. We discuss this model equation in some detail, not because in itself it has such wide applicability but because it is the prototype equation which admits traveling wavefront solutions. It is also a convenient equation from which to develop many of the standard techniques for analyzing single species models with diffusive dispersal.

Although (6.5.2) is now referred to as the *Fisher–Kolmogoroff equation*, the discovery, investigation and analysis of traveling waves in chemical reactions was first reported by Luther (1906).

It is convenient at the outset to rescale (6.5.2) by writing

$$t^* = kt, \qquad x^* = x\left(\frac{k}{D}\right)^{1/2} \tag{6.5.3}$$

and, omitting the asterisks for notational simplicity, (6.5.2) becomes

$$\frac{\partial u}{\partial t} = u(1-u) + \frac{\partial^2 u}{\partial x^2}. \tag{6.5.4}$$

In the spatially homogeneous situation the steady states are $u = 0$ and $u = 1$, which are respectively unstable and stable. This suggests that we should look for traveling wavefront solutions to (6.5.4) for which $0 \leq u \leq 1$; negative u has no physical meaning with what we have in mind for such models.

If a traveling wave solution exists it can be written in the form (6.4.1), say

$$u(x,t) = U(z), \qquad z = x - ct, \tag{6.5.5}$$

where c is the wavespeed. We use $U(z)$ rather than $u(z)$ to avoid any nomenclature confusion. Since (6.5.4) is invariant if $x \to -x$, c may be negative or positive. To be specific we assume $c \geq 0$. Substituting this travelling waveform into (6.5.4), $U(z)$ satisfies

$$U'' + cU' + U(1-U) = 0, \tag{6.5.6}$$

where primes denote differentiation with respect to z. A typical wavefront solution is where U at one end, say, as $z \to -\infty$, is at one steady state and as $z \to +\infty$ is at the other. So here we have an eigenvalue problem to determine the value, or values, of c such that a nonnegative solution U of (6.5.6) exists which satisfies

$$\lim_{z\to-\infty} U(z) = 0, \qquad \lim_{z\to+\infty} U(z) = 1. \tag{6.5.7}$$

At this stage we do not address the problem of how such a traveling wave solution might evolve from the partial differential equation (6.5.4) with given initial conditions $u(x,0)$; this point is quite interesting, but it will not covered here (see the book by J.D. Murray).

We study (6.5.6) for U in the (U,V) phase plane where

$$U' = V, \qquad V' = -cV - U(1-U), \tag{6.5.8}$$

which gives the phase plane trajectories as solutions of

$$\frac{dV}{dU} = \frac{-cV - U(1-U)}{V}. \tag{6.5.9}$$

This has two singular points for (U,V), namely, $(0,0)$ and $(1,0)$: these are the steady states of course. A linear stability analysis shows that the eigenvalues λ for the singular points are

$$\begin{aligned}
(0,0): \quad \lambda_\pm &= \frac{1}{2}\left[-c \pm (c^2-4)^{1/2}\right] \quad \Rightarrow \quad \begin{cases} \text{stable node} & \text{if } c^2 > 4 \\ \text{stable spiral} & \text{if } c^2 < 4 \end{cases} \\
(1,0): \quad \lambda_\pm &= \frac{1}{2}\left[-c \pm (c^2+4)^{1/2}\right] \quad \Rightarrow \quad \text{saddle point.}
\end{aligned} \tag{6.5.10}$$

Figure 6.1 (a) illustrates the phase plane trajectories. If $c \geq c_{min} = 2$ we see from (6.5.10) that the origin is a stable node, the case when $c = c_{min}$ giving a degenerate node. If $c^2 < 4$ it is a stable spiral; that is, in the vicinity of the origin U oscillates. By continuity arguments, or simply by heuristic reasoning from the phase plane sketch of the trajectories in Figure 6.1

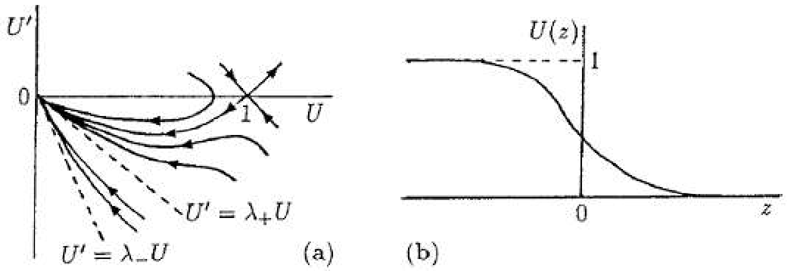

Figure 6.1: (a) Phase plane trajectories for equation (6.5.6) for the traveling wavefront solution: here $c^2 > 4$. (b) Traveling wavefront solution for the Fisher–Kolmogoroff equation (6.5.4): the wave velocity $c \geq 2$.

(a), there is a trajectory from $(1, 0)$ to $(0, 0)$ lying entirely in the quadrant $U \geq 0$, $U' \leq 0$ with $0 \leq U \leq 1$ for all wavespeeds $c \geq c_{min} = 2$. In terms of the original dimensional equation (6.5.2), the range of wavespeeds satisfies

$$c \geq c_{min} = 2(kD)^{1/2}. \tag{6.5.11}$$

Figure 6.1 (b) is a sketch of a typical travelling wave solution. There are travelling wave solutions for $c < 2$ but they are physically unrealistic since $U < 0$, for some z, because in this case U spirals around the origin. In these, $U \to 0$ at the leading edge with decreasing oscillations about $U = 0$.

Chapter 7

Solution and asymptotic behaviour of the nonlinear-diffusion equation

7.1 Solution to the Cauchy problem

We devote this chapter to show how to solve the nonlinear diffusion equation on the whole space $\mathbb{R}^d$. More precisely, we consider the Cauchy problem

$$\begin{cases} \partial_t \rho = \Delta \rho^m \\ \rho(x,0) = \rho_0(x), \end{cases} \tag{7.1.1}$$

in which ρ_0 is taken in $L_c^\infty(\mathbb{R}^d)$ and nonnegative.

A technique that recurs often in nonlinear PDEs to solve a complicated problem is to approximate the problem with a simpler one. In this case, the problem with (7.1.1) is that the solution may feature a singularity in the gradient similar to the one featured by the Barenblatt solution. Such a singularity appears near the state $\rho = 0$ in the case of the Barenblatt solution, so very likely such a problem would not occur if ρ would be detached from zero. On the other hand, the only diffusion equation that we know how to solve so far is the linear diffusion equation, therefore we need to approach the solution to (7.1.1) by several steps.

7.1.1 Linear diffusion equation with non-degenerate coefficients

We start by solving the following problem. Let $D = D(x) \geq d > 0$ a positive diffusion coefficient that depends on $x \in \mathbb{R}^d$. For a fixed $R > 0$, consider the Cauchy-Neumann problem

$$\begin{cases} \partial_t \rho = \operatorname{div}(D(x,t)\nabla\rho) & x \in B_R(0),\ t \geq 0 \\ \rho(x,0) = \rho_0(x) & x \in B_R(0), \\ \partial_\nu \rho(x,t) = 0 & x \in \partial B_R(0),\ t \geq 0, \end{cases} \tag{7.1.2}$$

with $\rho_0(x) \geq r > 0$. The solution to such a problem can be easily found using a well known method by Galerkin, see the Advanced Analysis course. Indeed, the operator on the right-hand-side is uniformly parabolic due to $D \geq r > 0$.

7.1.2 Nonlinear nondegenerate diffusion equation

We now move to the problem

$$\begin{cases} \partial_t \rho = \operatorname{div}(D(\rho)\nabla\rho) & x \in B_R(0),\ t \geq 0 \\ \rho(x,0) = \rho_0(x) & x \in B_R(0), \\ \partial_\nu \rho(x,t) = 0 & x \in \partial B_R(0),\ t \geq 0, \end{cases} \tag{7.1.3}$$

with $D = D(\rho) \geq d > 0$, $D(\rho) \leq \bar{D}$, and $\rho_0(x) \geq r > 0$. We construct the solution via an iterative argument. Fix $T \geq 0$. Given $\bar{\rho} \in L^2(B_R(0) \times [0,T])$, we consider the semi-linear problem

$$\begin{cases} \partial_t \rho = \operatorname{div}(D(\bar{\rho})\nabla\rho) & x \in B_R(0),\ t \geq 0 \\ \rho(x,0) = \rho_0(x) & x \in B_R(0), \\ \partial_\nu \rho(x,t) = 0 & x \in \partial B_R(0),\ t \geq 0, \end{cases} \tag{7.1.4}$$

the solution of which exists in view of the previous case. We set $A[\bar{\rho}] = \rho$ as the solution to (7.1.4). We perform the following estimates. Multiply the equation in (7.1.4) by ρ, integrate on $B_R(0)$. Integration by parts implies

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\int_{|x|\leq R} \rho(x,t)^2 dx &= \int_{|x|\leq R} \rho \operatorname{div}(D(\bar{\rho}(x,t))\nabla\rho(x,t))dx \\ = -\int_{|x|\leq R} D(\bar{\rho})|\nabla\rho(x,t)|^2 dx &\leq -d\int_{|x|\leq R} |\nabla\rho(x,t)|^2 dx, \end{aligned}$$

and integrating on t we get

$$\frac{1}{2}\int_{|x|\leq R} \rho(x,t)^2 dx + d\int_0^t \int_{|x|\leq R} |\nabla\rho(x,s)|^2 dxds = \frac{1}{2}\int_{|x|\leq R} \rho_0(x)^2 dx. \tag{7.1.5}$$

The above estimate allows to estimate ρ in the $L^2([0,T];\ H^1(B_R(0)))$. In particular, if we assume that

$$\|\bar{\rho}\|_{L^2([0,T]\times B_R(0))} \leq M,$$

then we easily obtain

$$\int_0^T \int_{B_R(0)} \rho^2(x,t)dxdt \leq T\|\rho_0\|^2_{L^2(B_R(0))}.$$

Therefore, if we choose $M \geq \sqrt{T}\|\rho_0\|_{L^2(B_R(0))}$ we get that the map A is well posed from the set

$$\mathcal{B}_M = \left\{ \rho \in L^2([0,T] \times B_R(0))\ :\ \int_0^T \|\rho(t,\cdot)\|^2_{L^2(B_R(0))} dt \leq M \right\}$$

into itself.

We also need an estimate of the time derivative. However, it will suffice to estimate ρ_t in some negative Sobolev norm. We recall that, given an L^2 function h, the H^{-1} norm of h is found by solving

$$\Delta\psi = h$$

with zero homogeneous Dirichlet conditions, and by considering

$$\|h\|^2_{H^{-1}} = \int |\nabla\psi|^2 dx.$$

In the case of ρ_t, we have to solve

$$\Delta\varphi = \rho_t,$$

and we get

$$\int_0^T \int |\nabla\varphi(x,t)|^2 dxdt = -\int_0^T \int \varphi\Delta\varphi dxdt = -\int_0^T \int \varphi\rho_t dxdt = \int_0^T \int \nabla\varphi \cdot D(\bar{\rho})\nabla\rho dxdt,$$

and by a simple weighted Young inequality we get

$$\begin{aligned} &\int_0^T \int |\nabla\varphi(x,t)|^2 dxdt \\ &\leq D\int_0^T \int |\nabla\rho(x,t)||\nabla\varphi(x,t)|dx \leq \frac{1}{2}\int_0^T \int |\nabla\varphi(x,t)|^2 dxdt + \frac{D^2}{2}\int_0^T \int |\nabla\rho(x,t)|^2 dxdt, \end{aligned}$$

which gives

$$\int_0^T \int |\nabla\varphi(x,t)|^2 dxdt \leq D^2 \int_0^T \int |\nabla\rho(x,t)|^2 dxdt,$$

and the last term is finite in view of (7.1.5). We now invoke the so called Aubin-Lions lemma.

Theorem 7.1.1 (Aubin-Lions lemma) *Let T be a positive number and Ω be a bounded open set in $\mathbb{R}^d$. Then, the set*

$$\{\rho : [0,T] \times \Omega \to \mathbb{R} \;:\; \rho \in L^2([0,T]\,;\; H_1(\Omega)) \text{ and } \rho_t \in L^2([0,T]\,;\; H^{-1}(\Omega))\}$$

is relatively compact in $L^2([0,T] \times \Omega)$.

The above theorem implies that the above defined map A defined on $\mathcal{B}_M$ mapping $\bar{\rho}$ into ρ is pre-compact. We then invoke another important theorem:

Theorem 7.1.2 (Schauder's fixed point theorem) *Every compact map defined on a closed, bounded, and convex domain of a Banach space in itself has a fixed point.*

Clearly, $\mathcal{B}_M$ is a closed, bounded, and convex subset of a Banach space. Schauder's fixed point theorem then implies that there exists an element $\rho \in L^2([0,T] \times B_R(0))$ such that $A\rho = \rho$, which is a classical solution to the problem (7.1.3).

We need to ensure that ρ stays nonnegative if ρ_0 is nonnegative. This is something which can be proven in exactly the same way we did for reaction-diffusion equations in the previous sections, and is left as an exercise. Finally, we need to show that the minimum of ρ gets never smaller than the minimum of the initial condition. To see that, let us multiply our equation for ρ

$$\partial_t\rho = \operatorname{div}(D(\rho)\nabla\rho),$$

by a suitable approximation of $(\rho - c)_-$ with $c = \inf_{x\in B_R(0)} \rho_0(x)$ and proceed as above. The details are left as an exercise.

7.1.3 Nonlinear degenerate diffusion equations

We now tackle the main problem of this section, namely the Cauchy problem (7.1.1). Our approximation technique works as follows. For a given initial condition $\bar{\rho} \in L^1 \cap L^\infty(\mathbb{R}^d)$ and a given $R > 0$, we consider

$$\bar{\rho}_R = \mathbf{1}_{B_R(0)}(x)(\bar{\rho}(x) + \frac{1}{R}).$$

The PDE in (7.1.1) can be written as

$$\partial_t \rho = \operatorname{div}(D(\rho)\nabla\rho)\,, \qquad D(\rho) = m\rho^{m-1}.$$

We now consider the following approximation of $D(\rho)$ for all $R > 0$:

$$D_R(\rho) = m \min\left\{(\max\{\rho, 1/R\})^{m-1}, R^{m-1}\right\}.$$

For all $R > 0$, the problem

$$\begin{cases} \partial_t \rho = \operatorname{div}(D_R(\rho)\nabla\rho) \\ \rho(x,0) = \bar{\rho}_R, \\ \partial_\nu \rho(x,t) = 0, \qquad\qquad x \in \partial B_R(0) \end{cases} \tag{7.1.6}$$

has a solution ρ_R in view of the result in the previous subsection. This is due to the fact that $D_R(\rho)$ is uniformly positive and uniformly bounded. Moreover, the results in the previous subsection imply that this solution ρ_R satisfies $\rho_R(x,t) \geq 1/R$. We can also provide an L^∞ estimate of the solution ρ_R. Let $\lambda = \|\bar{\rho}\|_{L^\infty}$. Consider a suitable monotone, convex regularization $\eta_\delta(\rho)$ of the function $(\rho - \lambda - 1/R)_+$. We get:

$$\frac{d}{dt}\int_{B_R(0)} \eta_\delta(\rho_R)dx = \int_{B_R(0)} \eta'(\rho_R)\operatorname{div}(m\rho_R^{m-1}\nabla\rho_R)dx = -m\int_{B_R(0)} \eta''(\rho_R)\rho_R^{m-1}|\nabla\rho|^2 dx \leq 0,$$

which implies, upon sending $\delta \searrow 0$, that

$$\int_{B_R(0)} (\rho_R(x,t) - \lambda - 1/R)_+ dx \leq \int_{B_R(0)} (\bar{\rho}_R(x) - \lambda - 1/R)_+ dx = 0,$$

and this implies that $\rho_R(x,t) \leq \lambda + 1/R$ almost everywhere on $B_R(0)$. The above estimates show that

$$D_R(\rho_R) = D(\rho_R) := m\rho_R^{m-1}.$$

We now want to prove that ρ_R has a limit (up to a subsequence) as $R \to +\infty$ and that such a limit is a weak solution to (7.1.1). To perform this task, let us integrate by parts and compute

$$\begin{aligned} \frac{d}{dt}\int_{B_R(0)} \rho_R^2(x,t)dx &= 2\int_{B_R(0)} \rho\rho_{R,t}dx = -2\int_{B_R(0)} D(\rho_R)|\rho_R(x,t)|^2 dx \\ &= -2m\int_{B_R(0)} \rho_R^{m-1}|\nabla\rho_R|^2 dx. \end{aligned}$$

A simple computation then shows

$$\frac{d}{dt}\int_{B_R(0)} \rho_R^2(x,t)dx = -\frac{8m}{(m+1)^2}\int_{B_R(0)} |\nabla\rho^{\frac{m+1}{2}}|^2 dx.$$

Integration on $t \in [0,T]$ then gives

$$\int_0^T \int_{B_R(0)} |\nabla \rho^{\frac{m+1}{2}}|^2 dx dt \leq C,$$

for some C depending on $\|\rho_0\|_{L^2}$.

Now, it turns out that the Aubin-Lions strategy does not work here. In order to estimate the time derivative, we have to first compute (we omit the index R)

$$\begin{aligned}
&\frac{d}{dt}\int_{B_R(0)} |\nabla \rho^m(x,t)|^2 dx = 2\int_{B_R(0)} \nabla \rho^m \cdot \nabla (\rho^m)_t dx \\
&= -2\int_{B_R(0)} \Delta \rho^m (\rho^m)_t dx = -2m \int_{B_R(0)} \rho^{m-1}\rho_t^2 dx \\
&= -\frac{8}{(m+1)^2}\int_{B_R(0)} \left(\rho^{(m+1)/2}\right)_t^2 dx,
\end{aligned}$$

and use such estimate as follows:

$$\begin{aligned}
&\int_0^T t \int_{B_R(0)} \left(\rho^{(m+1)/2}\right)_t^2 dx dt = -\frac{(m+1)^2}{8}\int_0^T t\frac{d}{dt}\int_{B_R(0)} |\nabla \rho^m|^2 dx dt \\
&= \frac{(m+1)^2}{8}\int_0^T \int_{B_R(0)} |\nabla \rho^m|^2 dx dt - \frac{(m+1)^2}{8} T \int_{B_R(0)} |\nabla \rho^m(x,T)|^2 dx \\
&\leq \frac{(m+1)^2}{8}\int_0^T \int_{B_R(0)} |\nabla \rho^m|^2 dx dt,
\end{aligned}$$

and the last term above is finite in view of the above estimates. All this implies a uniform control of $\rho_R^{(m+1)/2}$ on the space

$$H^1_{x,t}([\delta,T]\times B_R(0)),$$

for all $\delta > 0$. By standard Sobolev compact embedding, we get that $\rho_R^{(m+1)/2}$ is strongly compact in $L^2([\delta,T]\times B_R(0))$, and therefore, up to a subsequence, $\rho_R^{(m+1)/2} \to \rho^{(m+1)/2}$ almost everywhere and in $L^2_{loc}([\delta,T]\times\mathbb{R}^d)$, and consequently $\rho_R \to \rho$ almost everywhere on $[\delta,T]\times\mathbb{R}^d$ for some measurable function ρ. By taking a diagonal procedure, i.e. by choosing $\delta = 1/R$, we easily get almost everywhere convergence of ρ_R up to the initial line $t = 0$. Actually, this convergence holds in all $L^p_{loc}([0,T]\times\mathbb{R}^d)$ spaces by L^p interpolation taking into account that ρ_R is uniformly bounded in L^∞. The details are left as an exercise.

Since ρ_R solves the approximated version of the PDE, consider a test function $\psi \in C^2_c([0,T)\times\mathbb{R}^d)$ and compute

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^d} \rho\psi_t dx dt + \int_0^T \int_{\mathbb{R}^d} \rho^m \Delta\psi dx dt + \int_{\mathbb{R}^d} \rho_0(x)\psi(x,0) dx \\
&= \lim_{R\to+\infty} \int_0^T \int_{\mathbb{R}^d} \rho_R\psi_t dx dt + \int_0^T \int_{\mathbb{R}^d} \rho_R^m \Delta\psi dx dt + \int_{\mathbb{R}^d} \rho_0(x)\psi(x,0) dx = 0.
\end{aligned}$$

Hence, ρ is a weak distributional solution to the Cauchy problem (7.1.1).

7.2 Intermediate asymptotic behaviour

We now pose the problem of the asymptotic behavior for large times for the Cauchy problem (7.1.1). Similarly to what happens for the linear diffusion equation, due to the fact that the total mass is conserved and that the effect of diffusion is that of stabilizing the solution around a constant state, we expect u to converge to zero as $t \to +\infty$. Now, we know a special solution to the nonlinear diffusion equation given by the Barenblatt profile. This is actually a self-similar profile. Although we cannot solve the Cauchy problem (7.1.1), we will prove that *all solutions to* (7.1.1) *actually behave as the Barenblatt profile.* The actual statement is the following. Given u the solution to (7.1.1), given $B(\cdot,t)$ the Barenblatt profile, we aim at proving that

$$\|u(\cdot,)-B(\cdot,t)\|_{L^1(\mathbb{R}^d)} \to 0 \qquad \text{as } t \to +\infty\,. \tag{7.2.1}$$

This result is very important, because it says that no matter what the initial condition is, from a qualitative point of view, all solutions behave as the Barenblatt solution for large times, with an error that goes to zero in the L^1 norm.

We now outline the proof of (7.2.1) under the condition on the initial datum $u(x,0) = u_0(x)$ with $u_0 \in L^\infty(\mathbb{R}^d)$ and u_0 compactly supported. We perform the scaling

$$u(x,t) = \alpha(t)^d v(\alpha(t)x, \beta(t)) \tag{7.2.2}$$

with α and β to be determined. Note that the above ansatz provides a mass preserving function u with respect to time. We have to choose α and β such that the equation for $v = v(y,\tau)$ is totally self-consistent, with $y = \alpha(t)x$ and $\tau = \beta(t)$. We compute

$$\begin{aligned} u_t &= d\alpha^{d-1}\alpha' v + \alpha^d\alpha'\nabla_y v\cdot x + \alpha^d v_\tau\beta' = \alpha^{d-1}\left[\alpha'\mathrm{div}_y(yv) + \alpha\beta' v_\tau\right] \\ \Delta_x u^m &= \alpha^{dm+2}\Delta v^m \end{aligned}$$

and impose

$$\alpha'(t) = -\alpha^{d(m-1)+3}(t)\,, \qquad \beta'(t) = \alpha^{d(m-1)+2}(t)\,,$$

with initial conditions $\alpha(0) = 1$ and $\beta(0) = 0$. The differential equation for α yields

$$\alpha(t) = (1+\lambda t)^{-1/\lambda}\,, \qquad \lambda = d(m-1)+2\,.$$

Hence, we obtain

$$\beta'(t) = (1+\lambda t)^{-1}\,, \qquad \beta(0) = 0\,,$$

which implies

$$\beta(t) = \frac{1}{\lambda}\log(1+\lambda t)\,.$$

In particular, we obtain

$$\tau = \frac{1}{\lambda}e^{\lambda t} - 1\,.$$

The equation in (7.1.1) then becomes

$$v_\tau = \mathrm{div}(yv + \nabla v^m), \tag{7.2.3}$$

which can also be rewritten as

$$v_\tau = \mathrm{div}\left(v\nabla\left(\frac{|y|^2}{2} + \frac{m}{m-1}v^{m-1}\right)\right)\,.$$

The equation (7.2.3) is called *nonlinear Fokker-Planck* equation. Why are we considering this equation? The reason is all in the scaling (7.2.2), which actually reads

$$u(x,t) = (1+\lambda t)^{-d/\lambda} v(y,\tau). \tag{7.2.4}$$

One can easily check that this is the same scaling factor we put in front of the solution in order to recover the Barenblatt self-similar profile. Moreover, by looking for a stationary state for (7.2.3) we recover, as done for the self-similar profile,

$$\frac{|y|^2}{2} + \frac{m}{m-1} v^{m-1} = C$$

which gives

$$v(y) = v_\infty(y) = \left(C - \frac{2m}{m-1}|y|^2 \right)_+^{1/(m-1)} .$$

A simple computation shows that

$$u_\infty(x,t) = (1+\lambda t)^{-d/\lambda} v_\infty \left((1+\lambda t)^{-1/\lambda} x \right)$$

coincides with a time-translated version of the Barenblatt solution to the nonlinear diffusion equation. More precisely, instead of featuring a Dirac's delta as initial condition, u_∞ is equal to v_∞ as $t=0$, which coincides with the original Barenblatt profile at some fixed positive time. We observe that the nonlinear diffusion equation is invariant after time translations, therefore u_∞ still is a solution. As for the initial conditions, by observing that $\beta(0)=0$, we get

$$v(\cdot,0) = u_0(\cdot) ,$$

that is, v and u have the same initial condition.

Now, our goal is to use some kind of functional estimate to prove that $v(\cdot,\tau)$ converges to v_∞ as $\tau \to +\infty$. We define the *entropy functional*

$$E[v] := \frac{1}{m-1} \int v^m(y)dy + \frac{1}{2} \int |y|^2 v(y)dy ,$$

and the *relative entropy*

$$RE[v] := E[v] - E[v_\infty] .$$

The estimate below is somehow formal, because in it we differentiate the solution v to (7.2.3) assuming it is a classical solution. The way to make this estimate rigorous relies on similar approximations we used to construct the solution to the Cauchy problem (7.1.1), we omit the details. Let us then compute

$$\begin{aligned}
&\frac{d}{d\tau} E[v(\cdot,\tau)] = \frac{d}{d\tau} \left(\frac{1}{m-1} \int v^m(y)dy + \frac{1}{2} \int |y|^2 v(y)dy \right) \\
&= \frac{m}{m-1} \int v^{m-1} v_\tau dy + \frac{1}{2} \int |y|^2 v_\tau dy \\
&= \int \left(\frac{m}{m-1} v^{m-1} + \frac{1}{2}|y|^2 \right) \operatorname{div} \left(v \nabla \left(\frac{|y|^2}{2} + \frac{m}{m-1} v^{m-1} \right) \right) dy .
\end{aligned}$$

Integrating by parts and neglecting the boundary conditions, which could be made rigorous as mentioned above by prescribing zero Neumann conditions, we obtain

$$\frac{d}{d\tau}E[v(\cdot,\tau)] = -\int v\left|\nabla\left(\frac{|y|^2}{2}+\frac{m}{m-1}v^{m-1}\right)\right|^2 dy\,.$$

We now set

$$I[v] := \int v\left|\nabla\left(\frac{|y|^2}{2}+\frac{m}{m-1}v^{m-1}\right)\right|^2 dy$$

and write

$$\frac{d}{d\tau}RE[v(\cdot,\tau)] = -I[v(\cdot,\tau)]\,.$$

We integrate in τ and obtain

$$RE[v(\cdot,\tau)] + \int_0^\tau I[v(\cdot,s)]ds = RE[u_0]\,, \tag{7.2.5}$$

which implies

$$\int_0^\tau I[v(\cdot,s)]ds \le E[u_0] - E[v(\cdot,\tau)]\,.$$

It is easily seen that $E[v] \ge 0$ for all $v \ge 0$, but $v \ge 0$ is a consequence of $u_0 \ge 0$ and of the fact that v is a rescaled version of u and the latter is nonnegative. Hence, by imposing

$$E[u_0] < +\infty\,,$$

which is equivalent to

$$u_0 \in L^m(\mathbb{R}^d) \qquad \text{and}\,, \qquad \int_{\mathbb{R}^d} |y|^2 u_0(y)dy < +\infty\,,$$

we get

$$\int_0^\tau I[v(\cdot,s)]ds \le E[u_0] < +\infty\,.$$

As the above right hand side is independent of τ we obtain

$$\int_0^{+\infty} I[v(\cdot,s)]ds < +\infty\,. \tag{7.2.6}$$

We now claim that

$$I[v] = 0 \quad \text{and} \quad \int v(y)dy = M > 0 \qquad \text{if and only if} \quad v = v_\infty \quad \text{and} \quad \int v(y)dy = M > 0\,.$$

Indeed, if $I[v] = 0$ then the assumption on the mass implies v is not identically zero, therefore

$$\nabla\left(\frac{|y|^2}{2}+\frac{m}{m-1}v^{m-1}\right) = 0$$

and this implies v equals v_∞ with the constant $C > 0$ chosen such that the mass is equal to M. Vice-versa, it is immediately seen that $I[v_\infty] = 0$. Now, (7.2.6) implies that there exists a sequence $\tau_n \to +\infty$ such that $I[v(\cdot,\tau_n)] \to 0$. To lighten the notation we denote $v_n = v(\cdot,\tau_n)$. We compute

$$\begin{aligned}
I[v_n] &= \int v_n \left|\nabla\left(\frac{m}{m-1}v_n^{m-1}\right) + y\right|^2 dy \\
&= \int v_n \left|\nabla\left(\frac{m}{m-1}v_n^{m-1}\right)\right|^2 dy + \int v_n|y|^2 dy + 2\int v_n y\cdot\nabla\left(\frac{m}{m-1}v_n^{m-1}\right) dy \\
&= \int v_n \left|m v_n^{m-2}\nabla v_n\right|^2 dy + \int v_n|y|^2 dy + 2m\int v_n^{m-1} y\cdot\nabla v_n dy \\
&= m^2\int \left|v_n^{m-3/2}\nabla v_n\right|^2 dy + \int v_n|y|^2 dy + 2\int y\cdot\nabla v_n^m dy \\
&= \frac{m^2}{(m-1/2)^2}\int \left|\nabla v_n^{m-1/2}\right|^2 dy + \int v_n|y|^2 dy - 2d\int v_n^m dy\,,
\end{aligned}$$

where we have integrated by parts in the last identity. Now, the identity (7.2.5) implies

$$\frac{1}{m-1}\int v_n(y,\tau)^m dy + \frac{1}{2}\int |y|^2 v(y,\tau) dy$$

is uniformly bounded with respect to τ, and since both terms in the above estimate are nonnegative, both terms are uniformly bounded. Hence, since $I[v_n(\cdot)]$ is also bounded as $n \to +\infty$, we get that the quantity

$$\int \left|\nabla v_n^{m-1/2}\right|^2 dy$$

is uniformly bounded. We observe

$$\begin{aligned}
\nabla v_n^m &= \nabla\left(v_n^{m-1/2}\right)^{\frac{m}{m-1/2}} = \frac{2m}{2m-1}\left(v_n^{m-1/2}\right)^{\frac{2m}{2m-1}-1}\nabla v_n^{m-1/2} \\
&= \frac{2m}{2m-1}\left(v_n^{(2m-1)/2}\right)^{\frac{1}{2m-1}}\nabla v_n^{m-1/2} = \frac{2m}{2m-1}\sqrt{v_n}\nabla v_n^{m-1/2}\,.
\end{aligned}$$

Hence, a simple Hoelder inequality implies

$$\int |\nabla v_n^m| dy = \frac{2m}{2m-1}\int \sqrt{v_n}\left|\nabla v_n^{m-1/2}\right| dy \le \frac{2m}{2m-1}\left(\int v_n dy\right)^{1/2}\left(\int \left|\nabla v_n^{m-1/2}\right|^2 dy\right)^{1/2},$$

and since the total mass of v_n is constantly equal to M we get that

$$\int |\nabla v_n^m| dy$$

is uniformly bounded. Since the $\|\cdot\|_{L^m}$ norm of v_n is also uniformly bounded, we get that v_n^m is uniformly bounded in the Sobolev space $W^{1,1}(\mathbb{R}^d)$. Since the embedding of L^1 in $W^{1,1}$ is compact on compact subsets of $\mathbb{R}^d$, we get the existence of a subsequence of v_n^m - which by abuse of notation we shall still refer to as v_n^m - converging to some $h \in W^{1,1}$ almost everywhere and locally strongly in L^1. The almost everywhere convergence implies

$$v_n \to h^{1/m} := v.$$

Moreover, we easily get that $v_n \to v$ strongly in L^1 on compact sets.

The above computation has shown the strong convergence of v_n to some limit v strongly locally on compact sets. This is not enough to conclude our convergence argument because of the locality of the estimate, in particular, we are not yet guaranteed that the *tails* of v_n behave nicely and, as a consequence, the total mass of v_n converges to the total mass of v. We address this task in the following. We observe that v_n is uniformly bounded in $W^{m,1}(\mathbb{R}^d)$. Now, according to Morrey's inequality and to Gagliardo-Nirenberg's inequality, we may easily find $q > m$ such that v_n is uniformly bounded in L^q (exercise). We now use the uniform control on $\int |y| v_n(y)dy$ as follows. For some $\delta > 0$ we get

$$\begin{aligned} \int |y|^\delta v_n^m dy &= \int |y|^\delta v_n^{\delta/2} v_n^{m-\delta/2} dy \\ &= \int \left(|y|^2 v_n\right)^{\delta/2} v_n^{m-\delta/2} dy \end{aligned}$$

and by Hoelder's inequality

$$\int |y|^\delta v_n^m dy \leq \left(\int |y|^2 v_n dy\right)^{\delta/2} \left(\int \left(v_n^{m-\delta/2}\right)^{\frac{2}{2-\delta}} dy\right)^{\frac{2-\delta}{2}} .$$

We compute the exponent

$$(m - \delta/2)\frac{2}{2-\delta} = \frac{2m-\delta}{2-\delta} =: p$$

and observe that $p > m$ for every $\delta > 0$. We choose δ such that $p = q$ and we obtain that

$$\int |y|^\delta v_n^m dy$$

is uniformly bounded. Now, we apply Dunford-Pettis theorem, according to which if a sequence U_n is uniformly bounded in L^1 and is equi-integrable, that is, if for every $\varepsilon > 0$ there exist $\delta > 0$ and $R > 0$ such that for all $x_0 \in \mathbb{R}^d$,

$$\int_{B_\delta(x_0)} |U_n(x)|dx < \varepsilon \qquad \text{and} \qquad \int_{|x|\geq R} |U_n(x)|dx < \varepsilon \qquad \text{for all } n \in \mathbb{N},$$

then the sequence U_n is weakly compact in L^1. We apply the theorem to $U_n := v_n^m$. Given $q > m$ such that v_n is uniformly bounded in L^q, frmo Hoelder's inequality we get

$$\int_{B_\delta(x_0)} v_n^m dy \leq \left(\int_{B_\delta(x_0)} v_n^q dy\right)^{\frac{m}{q}} m(B_\delta(x_0))^{1-\frac{m}{q}} \leq C m(B_\delta(x_0))^{1-\frac{m}{q}}$$

for some C, and the first equi-integrability estimate follows by choosing δ such that $Cm(B_\delta(x_0))^{1-\frac{m}{q}} < \varepsilon$. As for the second estimate, we compute

$$\int_{|y|\geq R} v_n^m dy = \int_{|y|\geq R} \frac{1}{|y|^\delta} |y|^\delta v_n^m dy \leq \frac{1}{R^\delta} \int_{\mathbb{R}^d} |y|^\delta v_n^m dy \leq \frac{C}{R^\delta}$$

for some suitable constant C as from the above estimates. Hence, we choose R such that $\frac{C}{R^\delta} < \varepsilon$ to obtain equi-integrability. we have therefore proven that v_n^m is weakly compact in

$L^1(\mathbb{R}^d)$. In a similar way and using that $\int |y|^2 v_n dy$ and $\int v_n^m dy$ are uniformly bounded, one can prove that v_n is also weakly compact in $L^1(\mathbb{R}^d)$. As a consequence, we obtain

$$\int_{\mathbb{R}^d} v_n^m dy \to \int_{\mathbb{R}^d} v^m dy$$
$$\int_{\mathbb{R}^d} v_n dy \to \int_{\mathbb{R}^d} v dy\,.$$

The uniform bound of $\nabla v_n^{m-1/2}$ in L^2 gives (possibly by extracting another subsequence) that $\nabla v_n^{m-1/2}$ converges weakly in $L^2(\mathbb{R}^d)$ to some function W. We claim that $W = \nabla v^{m-1/2}$. The weak convergence gives, for all test functions φ,

$$-\int_{\mathbb{R}^d} v_n^{m-1/2} \mathrm{div}\varphi dy \to \int_{\mathbb{R}^d} W\varphi dy\,.$$

Moreover, $v_n^{m-1/2}$ has a uniformly bounded L^p norm with $p = \frac{m}{m-1/2} > 1$, and is therefore weakly compact in L^p. Therefore,

$$\int_{\mathbb{R}^d} v_n^{m-1/2} \mathrm{div}\varphi dy \to \int v^{m-1/2} \mathrm{div}\varphi dy$$

and hence $\nabla v^{m-1/2} = W$ in the weak sense. Gathering all the above information, we get by Fatou's Lemma and by weak-lower semicontinuity of the L^2 norm,

$$\begin{aligned}
0 \le I[v] &= \frac{m^2}{(m-1/2)^2} \int \left|\nabla v^{m-1/2}\right|^2 dy + \int v|y|^2 dy - 2d \int v^m dy \\
&\le \liminf_{n\to+\infty} \left(\frac{m^2}{(m-1/2)^2} \int \left|\nabla v_n^{m-1/2}\right|^2 dy + \int v_n |y|^2 dy \right) - 2d \lim_{\to+\infty} \int v_n^m dy \\
&= \liminf_{n\to+\infty} I[v_n] = 0\,,
\end{aligned}$$

and the additional information that $\int v_n dy \to M$, which implies that v coincides with the Barenblatt stationary state v_∞ with mass equal to M. Hence, we have proven v_n converges up to a subsequence to v_∞ strongly in $L^1_{loc}(\mathbb{R}^d)$. To prove convergence in the whole of $L^1(\mathbb{R}^d)$ we fix $\varepsilon > 0$ arbitrary and compute

$$\begin{aligned}
\int_{\mathbb{R}^d} |v_n - v_\infty| dy &= \int_{|y|\le R} |v_n - v_\infty| dy + \int_{|y|>R} \frac{1}{|y|^2} |y|^2 |v_n - v_\infty| dy \\
&\le \int_{|y|\le R} |v_n - v_\infty| dy + \frac{1}{R^2} \left(\int_{\mathbb{R}^d} |y|^2 v_n(y) dy + \int_{\mathbb{R}^d} |y|^2 v_\infty(y) dy \right) \\
&\le CR^{-2} + \int_{|y|\le R} |v_n - v_\infty| dy,
\end{aligned}$$

for some suitable constant C, and letting $n \to +\infty$ the strong $L^1 loc$ convergence implies

$$\limsup_{n\to+\infty} \int_{\mathbb{R}^d} |v_n - v_\infty| dy \le CR^{-2}$$

and the latter can be made smaller than ε for suitably large R. Hence,

$$\lim_{n\to+\infty} \int_{\mathbb{R}^d} |v_n - v_\infty| dy = 0\,.$$

Now, the assertion is almost proven. We have found a sequence of times $\tau_n \to +\infty$ such that $v(\tau_n) \to v_\infty$ strongly in $L^1(\mathbb{R}^d)$. In order to prove that this is in fact true for all $t \geq 0$, namely that $\lim_{\tau\to+\infty} v(\tau) = v_\infty$ in $L^1(\mathbb{R}^d)$, we need to use the monotonicity of the relative entropy functional $RE[v(\cdot,\tau)]$. Assuming that there exists another subsequence $\tau_n' \to +\infty$ such that $v_n' := v(\cdot,\tau_n') \to V'$ with $V' \neq v_\infty$, the monotonicity of the relative entropy would imply a uniform bound for the L^m norm of v_n' and for the second moment $\int |y|^2 v_n' dy$, and hence the continuity of the L^m-norm of v_n' with an argument similar to above. Hence,

$$\int (v_n')^m dy \to \int (V')^m dy\,.$$

We claim that

$$\int |y|^2 v_n' dy \to \int |y|^2 (V')^m dy\,.$$

To see this, we leave as an exercise the estimate

$$\frac{d}{dt}\int_{\mathbb{R}^d} |y|^{2+\delta} v(y,\tau)dy = (2+\delta)(d+\delta)\int_{\mathbb{R}^d} |y|^\delta v^m(y,\tau)dy - (2+\delta)\int_{\mathbb{R}^d} |y|^{2+\delta} v(y,\tau)dy$$

which can be proven by using directly the Fokker-Planck equations and integrating by parts. Since $\int_{\mathbb{R}^d} |y|^\delta v^m(y,\tau)dy$ is uniformly bounded (see the above estimates), a simple variations of constants formula and comparison principle implies that

$$\int_{\mathbb{R}^d} |y|^{2+\delta} v(y,\tau)dy$$

is uniformly bounded in time. Hence, similarly to above we can show that $|y|^2 v_n'(y)$ is weakly compact in L^1 by the Dunford-Pettis theorem, and then

$$\int_{\mathbb{R}^d} |y|^2 v_n'(y)dy \to \int_{\mathbb{R}^d} |y|^2 V'(y)dy\,.$$

We have therefore proven that

$$RE[v_n'] \to RE[V'] \neq 0\,.$$

With exactly the same argument, though, one has that

$$RE[v_n] \to RE[v_\infty] = 0$$

for the previous sequence, but this is not possible, because $RE[v(\cdot,\tau)]$ is *monotone* in τ, and hence it cannot have two distinct limit points. Hence, the assertion

$$\lim_{\tau\to+\infty} \|v(\cdot,\tau) - v_\infty\|_{L^1(\mathbb{R}^d)} = 0$$

is proven. To prove (7.2.1), we use the change of variable $y = (1+\lambda t)^{-1/\lambda} x$

$$\begin{aligned}&\int_{\mathbb{R}^d} |u(x,t) - u_\infty(x,t)|dx \\ &= (1+\lambda t)^{-d/\lambda}\int_{\mathbb{R}^d} |v((1+\lambda t)^{-1/\lambda}x,\tau) - v_\infty((1+\lambda t)^{-1/\lambda}x)|dx = \int_{\mathbb{R}^d} |v(y,\tau) - v_\infty(y)|dy \\ &\to 0\end{aligned}$$

as $t \to +\infty$, which is equivalent to $\tau \to +\infty$.

Chapter 8

Reaction–diffusion systems

When more than one species interact via reaction mechanisms such as predator–prey, mutualism, competition, reaction kinetics etc. *and* they all diffuse with different (in general) diffusivity constants, we are in a situation to be modeled via a *reaction–diffusion system.* As a general rule, when two or more species interact and diffuse, the latter mechanism may be affected, for each species, by the gradient of any species. More precisely, a given species may experience a *self-diffusion* mechanism, through which individuals are "repelled" by individual of the same species, and a *cross-diffusion* mechanism, by which their movement is biased (negatively or positively) by the change of concentration of the other species.

In this course we will just deal with the case of two species, because this case is representative of all the mathematical challenges that arise with many species systems with diffusion. In order to understand at a preliminary stage the impact of self-diffusion and cross-diffusion on the mathematical properties of the system, we consider the following simple one-space dimensional system

$$\begin{cases} \dfrac{\partial \rho_1(x,t)}{\partial t} = \dfrac{\partial}{\partial x}\left(d_{11}\dfrac{\partial \rho_1(x,t)}{\partial x} + d_{12}\dfrac{\partial \rho_2(x,t)}{\partial x}\right) + R_1(\rho_1,\rho_2) \\ \dfrac{\partial \rho_2(x,t)}{\partial t} = \dfrac{\partial}{\partial x}\left(d_{22}\dfrac{\partial \rho_2(x,t)}{\partial x} + d_{21}\dfrac{\partial \rho_1(x,t)}{\partial x}\right) + R_2(\rho_1,\rho_2)\,, \end{cases} \tag{8.0.1}$$

where $d_{11} > 0$ and $d_{22} > 0$ are self-diffusion constants, whereas d_{12} and d_{21} are cross-diffusion constants. Here R_1 and R_2 are reaction terms, which we will assume to be zero for the time being because we just want to understand the structure of the diffusion part. Let us assume $x \in \mathbb{R}$ and compute, via integration by parts and neglecting the boundary terms as usual,

$$\begin{aligned} &\frac{d}{dt}\left[\int \rho_1^2(x,t)dx + \int \rho_2^2(x,t)dx\right] = 2\int \rho_1 \partial_t \rho_1 dx + 2\int \rho_2 \partial_t \rho_2 dx \\ &= 2\int \rho_1 \partial_x(d_{11}\partial_x\rho_1 + d_{12}\partial_x\rho_2)dx + 2\int \rho_2\partial_x(d_{22}\partial_x\rho_2 + d_{21}\partial_x\rho_1)dx \\ &= -2d_{11}\int (\rho_1)_x^2 dx - 2d_{22}\int(\rho_2)_x^2 dx - 2d_{12}d_{21}\int (\rho_1)_x(\rho_2)_x dx\,. \end{aligned}$$

In order to get global stability in the L^2 sense, we need the right hand side to be non-positive. The latter can be written as minus a quadratic form with matrix

$$A = \begin{pmatrix} d_{11} & \frac{d_{12}+d_{21}}{2} \\ \frac{d_{12}+d_{21}}{2} & d_{22} \end{pmatrix}\,.$$

Since the trace of the above matrix is positive, the associated quadratic form is positive definite if and only if $\det A > 0$, which reads

$$d_{11}d_{22} - \frac{1}{4}(d_{12} + d_{21})^2 > 0\,.$$

Such a condition is quite enlightening, as it unveils the fact that self-diffusion is a stabilising term, where as cross-diffusion is potentially producing instability even in case of positive cross-diffusion constants. In some sense, a "dominating" self-diffusion part produces stability for the system.

8.1 Turing instability and spatial patterns

Turing (1952) suggested that, under certain conditions, chemicals can react and diffuse in such a way as to produce steady state heterogeneous spatial patterns of chemical or morphogen concentration. This section is mainly concerned with models for two chemical species, $\rho_1(x,t)$ and $\rho_2(x,t)$ say. The equation system is then of the form

$$\frac{\partial \rho_1}{\partial t} = d_1 \Delta \rho_1 + R_1(\rho_1, \rho_2) \tag{8.1.1}$$

$$\frac{\partial \rho_2}{\partial t} = d_2 \Delta \rho_2 + R_2(\rho_1, \rho_2) \tag{8.1.2}$$

where R_1 and R_2 are the reaction kinetics, and $d_1, d_2 > 0$ are the diffusion constants. Notice that this model only features self-diffusion terms.

Turing's (1952) idea is a simple but profound one. He said that, if in the absence of diffusion (effectively $d_1 = d_2 = 0$), ρ_1 and ρ_2 tend to a linearly stable uniform steady state then, under certain conditions, which we shall derive, spatially inhomogeneous patterns can evolve by *diffusion driven instability* if $d_1 \neq d_2$. Diffusion (in particular self-diffusion, as from the previous section) is usually considered a stabilizing process which is why this was such a novel concept. To see intuitively how diffusion can be destabilizing consider the following, albeit unrealistic, but informative analogy.

Consider a field of dry grass in which there is a large number of grasshoppers which can generate a lot of moisture by sweating if they get warm. Now suppose the grass is set alight at some point and a flame front starts to propagate. We can think of the grasshopper as an inhibitor and the fire as an activator. If there were no moisture to quench the flames the fire would simply spread over the whole field which would result in a uniform charred area. Suppose, however, that when the grasshoppers get warm enough they can generate enough moisture to dampen the grass so that when the flames reach such a pre-moistened area the grass will not burn. The scenario for spatial pattern is then as follows. The fire starts to spread - it is one of the "reactants", the activator, with a "diffusion" coefficient D_F say. When the grasshoppers, the inhibitor "reactant", ahead of the flame front feel it coming they move quickly well ahead of it; that is, they have a "diffusion" coefficient, D_G say, which is much larger than D_F. The grasshoppers then sweat profusely and generate enough moisture to prevent the fire spreading into the moistened area. In this way the charred area is restricted to a finite domain which depends on the diffusion coefficients of the reactants - fire and grasshoppers - and various reaction parameters. If, instead of a single initial fire, there were a random scattering of them we can see how this process would result in a final spatially heterogeneous steady state distribution of charred and uncharred regions in the field and a spatial distribution of grasshoppers, since around each fire the above scenario would take place. If the grasshoppers and flame front diffused at the same speed no such spatial pattern could evolve. It is clear how to construct other analogies.

In the following section we describe the process in terms of reacting and diffusing morphogens and derive the necessary conditions on the reaction kinetics and diffusion coefficients.

A reaction diffusion system exhibits diffusion-driven instability, sometimes called *Turing instability*, if the homogeneous steady state is stable to small perturbations in the absence of diffusion but unstable to small spatial perturbations when diffusion is present. The concept of instability in biology is often in the context of ecology, where a uniform steady state becomes unstable to small perturbations and the populations typically exhibit some temporal oscillatory behavior. The instability we are concerned with here is of a quite different kind. The main process driving the spatially inhomogeneous instability is diffusion: the mechanism determines the spatial pattern that evolves. How the pattern or mode is selected is an important aspect of the analysis, a topic we do not have time to develop here.

We derive here the necessary and sufficient conditions for diffusion-driven instability of the steady state and the initiation of spatial pattern for the system (8.1.2). For simplicity, we shall deal with a *linear* system, which will translate into something to be applied to linearised versions of nonlinear reactions. We consider then

$$\begin{cases} \rho_{1,t} = d_1 \rho_{1,xx} + \alpha\rho_1 + \beta\rho_2 \\ \rho_{2,t} = d_2 \rho_{2,xx} + \gamma\rho_1 + \delta\rho_2 \quad . \end{cases} \tag{8.1.3}$$

We shall work on a one-dimensional bounded interval $x \in [0, L]$ with zero-flux boundary conditions

$$\rho_{1,x}(0,t) = \rho_{1,x}(L,t) = \rho_{2,x}(0,t) = \rho_{2,x}(L,t) = 0\,.$$

The system can be written in vector form as follows. Set

$$U = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} \qquad A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \qquad D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}\,.$$

We get

$$U_t = DU_{xx} + AU\,.$$

To have a Turing instability we first need to require that A is a stable reaction matrix. The eigenvalues of A solve

$$\lambda^2 - (\mathrm{tr}A)\lambda + \det A = 0$$

which implies that stability holds provided that both

$$\mathrm{tr}A = \alpha + \delta < 0$$

and

$$\det A = \alpha\delta - \beta\gamma > 0\,.$$

Now, we are going to look for solutions in a way that is very similar to the separation of variables method used for the single species case. We start with the ansatz

$$\rho_i(x,t) = \sum_{k=0}^{+\infty} c_k^i e^{\lambda_k t} \cos(\omega_k x)\,, \qquad \text{for } i = 1, 2\,.$$

By superposition principle, we substitute each of the above terms in the system in order to detect the values of ω_k and λ_k. As usual, the constants c_k^i will depend on the initial data. Let us substitute first

$$U_k(x,t) = e^{\lambda_k t}(\cos(\omega_k x), \cos(\omega_k x))$$

into the boundary conditions and obtain, with $\partial_x \cos(\omega_k x) = -\omega_k \sin(\omega_k x)$,

$$-\omega_k \sin(\omega_k 0) = 0$$

which is always satisfied, and

$$-\omega_k \sin(\omega_k L) = 0$$

which implies

$$\omega_k = \frac{k\pi}{L}, \qquad k = 0, 1, 2, \ldots$$

In order to obtain the values of λ_k, we now substitute U_k in the system and, with simple calculations and using

$$\partial_t e^{\lambda_k t} \cos(\omega_k x) = e^{\lambda_k t} \cos(\omega_k x)$$

and

$$\partial_{xx}^2 e^{\lambda_k t} \cos(\omega_k x) = -\omega_k^2 e^{\lambda_k t} \cos(\omega_k x)\,,$$

we obtain

$$\lambda_k U_k = -\omega_k^2 \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} U_k + \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} U_k\,.$$

The above cam be re-written as

$$\left[\lambda_k \mathbb{I} + \omega_k^2 \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} - \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\right] U_k = 0$$

which means that λ_k is an eigenvalue for the matrix

$$B = -\omega_k^2 \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} + \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \alpha - \omega_k^2 d_1 & \beta \\ \gamma & \delta - \omega_k^2 d_2 \end{pmatrix}.$$

A simple computation for the eigenvalues of B yields they are roots to the equation

$$\lambda^2 + b_k \lambda + h_k = 0$$

with

$$\begin{aligned} b_k &= \omega_k^2(d_1 + d_2) - \mathrm{tr} A \\ h_k &= \omega_k^4 d_1 d_2 - \omega_k^2(\delta d_1 + \alpha d_2) + \det A\,. \end{aligned}$$

The roots are

$$\lambda_\pm = \frac{-b_k \pm \sqrt{b_k^2 - 4h_k}}{2}\,.$$

Now, the values $\lambda = \lambda_k$ impact on the solution U_k in the exponential time dependent factor. In particular, $\Re\lambda > 0$ for at least one of the two roots corresponds to *instability*, which is what we want to achieve here. Clearly, this is possible if one of two conditions below are satisfied:

$$b_k < 0 \qquad h_k < 0\,.$$

Since the trace of A is negative (due to the stability of the reaction matrix), we clearly have $b_k > 0$. Therefore, instability can only be obtained if $h_k < 0$. Now, with the notation $\xi ? = \omega_k^2$, $\xi \geq 0$, we can write $h_k = h_k(\xi)$ as the parabola

$$h_k(\xi) = d_1 d_2 \xi^2 - (\delta d_1 + \alpha d_2)\xi + \det A \, .$$

Since $\det A > 0$, the above parabola may attain some negative values for some $\xi > 0$ only if

$$\delta d_1 + \alpha d_2 > 0 \, .$$

The above can be seen as a necessary condition for instability. It is immediately clear that such a condition cannot be achieved if $d_1 = d_2$, because we know $\alpha + \delta < 0$ due to the stability of the reaction part. Hence, already at this stage we see how Turing instability can appear only if the two diffusion coefficients are different, consistently with the grasshoppers example.

We now aim at deriving a sufficient condition. A simple computation shows that the minimum of the above parabola is

$$h_{\min} = \det A - \frac{1}{4 d_1 d_2} (\delta d_1 + \alpha d_2)^2 \, .$$

Hence, said parabola will achieve negative values if and only if

$$4 d_1 d_2 (\alpha\delta - \beta\gamma) < \delta^2 d_1^2 + \alpha^2 d_2^2 + 2\alpha\delta d_1 d_2 \, .$$

Now, remember that $\xi = \omega_k^2$. Hence, only some discrete values of the above parabola are actually achieved. Hence, further to the above condition we must ensure that there are "enough discrete values" of ω_k, in particular to make sure that some of them falls into the interval in which h_k is negative. Recalling the expression of ω_k, we see that this is possible if L is *large enough.*

Some final remarks about this model. This is a linear model. Is it possible for this model to produce actual patterns which do not blow up at infinity? The answer is: yes as long as $\mathrm{Re}\lambda_k = 0$. This is something that happens very unlikely, as one of the discrete values of ω_k must coincide with a zero of the parabola h_k. There is another, more efficient interpretation behind the formation of patterns as a consequence of a Turing instability, which is the fact that with no cross-diffusion in the system and a stable nonlinear reaction, a steady state would be globally stable, so there would be no blow up. On the other hand, the local instability of a steady state would give that no homogeneous constant states are formed, so the formation of patterns is in fact the only possibility in the large time limit, we omit the details.

Chapter 9

Chemotaxis

A large number of insects and animals (including humans) rely on an acute sense of smell for conveying information between members of the species. Chemicals which are involved in this process are called *pheromones.* For example, the female silk moth Bombyx mori exudes a pheromone, called bombykol, as a sex attractant for the male, which has a remarkably efficient antenna filter to measure the bombykol concentration, and it moves in the direction of increasing concentration.

The acute sense of smell of many deep sea fish is particularly important for communication and predation. Other than for territorial demarcation one of the simplest and important exploitations of pheromone release is the directed movement it can generate in a population. Here we model this chemically directed movement, which is called *chemotaxis*, which, unlike diffusion, directs the motion *up* a concentration gradient.

It is not only in animal and insect ecology that chemotaxis is important. It can be equally crucial in biological processes where there are numerous examples. For example, when a bacterial infection invades the body it may be attacked by movement of cells towards the source as a result of chemotaxis. Convincing evidence suggests that leukocyte cells in the blood move towards a region of bacterial inflammation, to counter it, by moving up a chemical gradient caused by the infection.

A widely studied chemotactic phenomenon is that exhibited by the slime mould *Dictyostelium discoideum* where single–cell amoebae move towards regions of relatively high concentrations of a chemical called *cyclic-AMP* which is produced by the amoebae themselves. The kinetics involved have been modeled by several authors. As more was found out about the biological system the models necessarily changed.

Most mathematical models for spatial patterning in Dictyostelium discoideum are based on continuum models for the chemoattractants and the cells. Let us suppose that the presence of a gradient in an attractant, $a(x,t)$, gives rise to a movement, of the cells say, up the gradient. The flux of cells will increase with the number of cells, $n(x,t)$, present. Thus we may reasonably take as the chemotactic flux

$$J = n\chi(a)\nabla a \tag{9.0.1}$$

where $\chi(a)$ is a function of the attractant concentration. In the general conservation equation for $n(x,t)$, namely,

$$\frac{\partial n}{\partial t} + \mathrm{div} J = f(n)$$

where $f(n)$ represents the growth term for the cells, the flux

$$J = J_{diffusion} + J_{chemotaxis},$$

where the diffusion contribution is given by Fick's law $J_{diffusion} = -D\nabla n$ with the chemotaxis flux from (9.0.1).

Thus a basic *reaction–diffusion–chemotaxis* equation is

$$\frac{\partial n}{\partial t} = f(n) + D\Delta n - \mathrm{div}(n\chi(a)\nabla a) \tag{9.0.2}$$

where D is the diffusion coefficient of the cells.

Since the attractant $a(x,t)$ is a chemical it also diffuses and is produced, by the amoebae, for example, so we need a further equation for $a(x,t)$. Typically

$$\frac{\partial a}{\partial t} = g(a,n) + \mathrm{div} D_a \nabla a, \tag{9.0.3}$$

where D_a is the diffusion coefficient of a and $g(a,n)$ is the kinetics/source term, which may depend on n and a. Normally we would expect $D_a > D$. If several species or cell types all respond to the attractant the governing equation for the species vector is an obvious generalization of (9.0.2) to a vector form with $\chi(a)$ probably different for each species.

In the seminal model of Keller and Segel (1971) for slime mould, $g(a,n) = hn - ka$, where h, k are positive constants. Here hn represents the spontaneous production of the attractant and is proportional to the number of amoebae n, while $-ka$ represents decay of attractant activity; that is, there is an exponential decay if the attractant is not produced by the cells. One simple version of the model has $f(n) = 0$; that is, the amoebae production rate is negligible. This is the case during the pattern formation phase in the mould's life cycle. The chemotactic term $\chi(a)$ is taken to be a positive constant χ_0. The form of this term has to be determined from experiment. With constant diffusion coefficients, together with the above linear form for $g(a,n)$, the model becomes the nonlinear system

$$\begin{aligned} \frac{\partial n}{\partial t} &= D\Delta n - \chi \mathrm{div}\left(n\nabla a\right), \\ \frac{\partial a}{\partial t} &= D_a \Delta a + hn - ka. \end{aligned} \tag{9.0.4}$$

There we consider n to be a bacterial population and a the food which it consumes.

Before leaving this introduction, note the difference in sign in (9.0.2) and (9.0.4) in the diffusion and chemotaxis terms. More precisely, expanding the equation for n gives

$$\frac{\partial n}{\partial t} = D\Delta n - \chi n \Delta a - \chi \nabla n \cdot \nabla c.$$

Here, both n and c have a Laplacian contribution but with different sign. This suggests that whereas diffusion is generally a stabilizing force, chemotaxis is generally destabilizing, like a kind of negative diffusion. At this stage, therefore, it is reasonable to suppose that the balance between stabilizing and destabilizing forces in the model system (9.0.4) could result in some steady state spatial patterns in n and a, or in some unsteady wavelike spatially heterogeneous structure. That is, nonuniform spatial patterns in the cell density appear. On the other hand if the chemotactic effect is sufficiently strong there could be a possibility of solution blow-up. This in fact can happen in certain cases, as we shall see below.

9.1 Diffusion vs. Chemotaxis: stability vs. instability

Let us consider system (9.0.4) in one space dimension on the whole $\mathbb{R}$, namely

$$\begin{cases} n_t = Dn_{xx} - \chi\,(na_x)_x \\ a_t = D_a a_{xx} + hn - ka. \end{cases} \tag{9.1.1}$$

We want to analyze the linear stability of constant steady states. The procedure will be similar to that used for Turing instability. Since we are on the whole space, we can make such procedure simpler by employing the Fourier's transform

$$\widehat{f}(\xi,t) = \int_{\mathbb{R}} e^{-2\pi i x\xi} f(x,t)dx.$$

Let (n^∞, a^∞) be a constant steady state. The second equation in (9.1.1) yields

$$hn^\infty - ka^\infty = 0,$$

and therefore the steady state is of the form

$$(n^\infty, a^\infty) = \left(n^\infty, \frac{hn^\infty}{k}\right),$$

and the value n^∞ is arbitrary. Let us linearize system (9.1.1) around (n^∞, a^∞), namely let $n = n^\infty + u$, $a = a^\infty + v$. We easily obtain the following system as first order approximation

$$\begin{cases} u_t = Du_{xx} - \chi n^\infty v_{xx} \\ v_t = D_a v_{xx} + hu - kv. \end{cases} \tag{9.1.2}$$

Let us first point out that, in case $D = D_a = \chi = 0$, i. e. no diffusion and no chemotaxis (chemotaxis can be seen as a cross-diffusion), the system becomes

$$\begin{cases} u_t = 0 \\ v_t = hu - kv, \end{cases}$$

which can be solved explicitly to give the solution

$$u(t) \equiv u_0, \quad v(t) = v_0 e^{-kt} + \frac{hu_0}{k}(1 - e^{-kt})$$

which clearly proves that the steady state is stable (neutrally, not asymptotically) without diffusion (Exercise!).

Let us now apply the Fourier transform to (9.1.2) in order to analyze the linear stability of system (9.1.1). We have the following system for $\widehat{u}(\xi,t)$ and $\widehat{v}(\xi,t)$:

$$\begin{cases} \widehat{u}_t = -D\xi^2\widehat{u} + \chi n^\infty \xi^2 \widehat{v} \\ \widehat{v}_t = -D_a\xi^2\widehat{v} + h\widehat{u} - k\widehat{v}, \end{cases} \tag{9.1.3}$$

which can be rewritten in matrix form as follows

$$\widehat{U} = (\widehat{u}, \widehat{v}), \qquad A(\xi) := \begin{pmatrix} -D\xi^2 & \chi n^\infty \xi^2 \\ h & -D_a\xi^2 - k \end{pmatrix} \qquad \partial_t \widehat{U}(\xi,t) = A(\xi)\widehat{U}(\xi,t). \tag{9.1.4}$$

The linear stability of the steady state (n^∞, a^∞) for system (9.1.1) in L^2 is clearly equivalent to the L^2 stability of the zero state $\widehat{U} = (0,0)$ for system (9.1.4). This fact is a trivial consequence of the Plancherel Theorem $\|f\|_{L^2} = \|\widehat{f}\|_{L^2}$, which can be found in any PDE textbook (cf. for instance the book by L.C. Evans). Therefore, since the right-hand-side in the equation in (9.1.4) does not contain derivatives with respect to ξ, all we have to do is compute the eigenvalues of the matrix $A(\xi)$ and check whether they imply stability of instability. We immediately see that

$$\mathrm{Tr}A(\xi) < 0 \qquad \text{for all } \xi \in \mathbb{R}.$$

Let us compute

$$\det A(\xi) = \xi^2(DD_a\xi^2 + kD - \chi h n^\infty).$$

If

$$n^\infty < \frac{kD}{\chi h},$$

then $\det A(\xi) > 0$ for all $\xi \in \mathbb{R}$ and therefore the eigenvalues of $A(\xi)$ are both strictly negative for all ξ and therefore we have stability. On the other hand, if

$$n^\infty > \frac{kD}{\chi h},$$

then there exists an interval $\xi \in [0, \bar{\xi}]$ on which $A(\xi)$ has one positive eigenvalue, which implies linear instability. We have therefore obtained a threshold condition for stability, which involves n^∞, D, χ and k. One way to see such a condition is that if the ratio D/χ is large enough, then diffusion dominates and we have stability, whereas if D/χ is small enough then chemotaxis dominates and we have instability.

Let us remark than, in the case in which we have instability, this is actually a Touring instability since the same steady state is always stable without diffusion. The dichotomy between diffusion and chemotaxis seen as a dichotomy between stability and instability will be more clear in the next paragraph, in which we shall consider a simplified version of Keller–Segel system.

9.2 Diffusion vs. Chemotaxis: stability and blow–up

In this section we shall consider the following Keller–Segel simplified version of the Keller–Segel system,

$$\begin{cases} n_t = \Delta n - \mathrm{div}(n\nabla c) \\ 0 = \Delta c + n \\ n(t,0) = n_0(x). \end{cases} \tag{9.2.1}$$

With respect to (9.0.4), we have dropped the time derivative of the chemoattractant c. This can be justified in a similar way as we saw in the chapter devoted to chemical reaction kinetics (Michaelis–Menten approximation). Actually, in many applications the rate of change in the chemical has a fast scale in which it stabilizes to zero and it can be reasonably considered at equilibrium. Moreover, we have dropped the decay term for c and set all parameters equal to 1 for simplicity.

We shall study the model in dimension 2 (this is consistent with many applications) and on the whole $\mathbb{R}^2$ without boundary conditions. In this case, c can be recovered via the convolution

$$c(t,x) = -\frac{1}{2\pi}\int \log|x-y|n(y)dy,$$

which implies

$$\nabla_x c(t,x) = -\frac{1}{2\pi}\int \frac{x-y}{|x-y|^2}n(y)dy.$$

A key parameter in this context is total mass of n, which is conserved along the flow, namely,

$$M := \int_{\mathbb{R}^2} n(x,t)dx = \int_{\mathbb{R}^2} n(x,0)dx, \qquad \text{for all } t.$$

Another important parameter here is the *center of mass* of n

$$\int xn(t,x)dx$$

which is also conserved in time, as easily seen in the following computation.

$$\frac{d}{dt}\int xn(t,x)dx = \int xn_t(x,t)dx = \int x\mathrm{div}(\nabla n)dx + \frac{1}{2\pi}\int\int x\mathrm{div}\left(\frac{x-y}{|x-y|^2}n(y)n(x)\right)dydx.$$

Here we integrate by parts in both the integral terms above and implicitly assume that n vanish as $|x| \to +\infty$. This is quite reasonable because of the assumption of finite total mass of n. We also use that $\mathrm{div}x = 2$ in dimension 2 to get

$$\frac{d}{dt}\int xn(t,x)dx = -2\int \nabla n dx - \frac{1}{\pi}\int\int\left(\frac{x-y}{|x-y|^2}n(y)n(x)\right)dydx.$$

Now, since the first integral is clearly zero because of the fundamental theorem of calculus and because n vanish as $|x| \to +\infty$. The second term can be easily seen to be zero by expanding the ratio as the sum of two terms. This proves the assertion

$$\frac{d}{dt}\int xn(t,x)dx = 0.$$

For simplicity, we assume that $\int xn(t,x)dx = 0$ initially, and therefore for all times.

Now, we want to show here a very interesting phenomena, which does not appear in any of the PDE models considered before. Namely, the asymptotic behavior here depends crucially on the *total mass of n*. More precisely, if the initial total number of cells is lower than a certain value, diffusion dominates to create stability. When the total mass is higher than such value, then chemotaxis dominates to create aggregation of the cells to a single point. This facts reflects the behavior of the linearized system as seen in the previous section.

We account first of the former case, namely we prove that when the mass is small enough, the zero stationary state is stable in some *energy* sense, which is actually what happens for a diffusion equation without chemotaxis). To see this, let us compute the evolution of the L^2-norm of n:

$$\begin{aligned}\frac{d}{dt}\int n^2(x,t)dx &= 2\int nn_t dx = 2\int n(\Delta n - \mathrm{div}(n\nabla c))dx \\ &= -2\int|\nabla n|^2dx + 2\int n\nabla c\cdot\nabla n dx. \end{aligned} \tag{9.2.2}$$

We have used the integration by parts and the fact that everything vanishes as $|x| \to +\infty$ as usual. Now, we expand the last term in the above computation as follows (integrating by parts once again in the second step):

$$2\int n\nabla c \cdot \nabla n dx = \int \nabla n^2 \cdot \nabla c dx = -\int n^2 \Delta c dx = \int n^3(x,t)dx.$$

In the last step we have used the Poisson's equation in (9.2.1). Putting all together in (9.2.2) we get

$$\frac{d}{dt}\int n^2(x,t)dx = -2\int |\nabla n|^2 dx + \int n^3(x,t)dx. \tag{9.2.3}$$

Now we use the following Gagliardo–Nirenberg interpolation inequality (see, for instance, the PDE book of Evans):

$$\int_{\mathbb{R}^2} f^3 \leq C_{gn}\int_{\mathbb{R}^2} |\nabla f|^2 dx \int_{\mathbb{R}^2} f dx, \qquad f \geq 0. \tag{9.2.4}$$

Using (9.2.4) in the last step of (9.2.3) implies

$$\frac{d}{dt}\int n^2(x,t)dx \leq -2\int |\nabla n|^2 dx + C_{gn}\int |\nabla n|^2 dx \int n dx = (MC_{gn} - 2)\int |\nabla n|^2 dx.$$

Therefore, if the total mass M is smaller than $2/C_{gn}$, then the right hand side above is negative and the L^2 norm of n is decreasing in time, which proves stability in the L^2 sense. Since the stability in L^2 is a typical behavior of the diffusion equation, this shows that when the initial mass is small enough, cells are not enough to produce a significant aggregation phenomenon, and therefore diffusion prevails.

We now want to prove that the solution n exhibits some *concentration* phenomenon in case the initial mass is too high. To see this, we first introduce the second moment (or variance) of n, namely

$$I(t) := \int |x|^2 n(t,x)dx.$$

This quantity accounts for how much the distribution of the particles of the species n is 'far' from its (zero) center of mass. To be more precise with such statement, let us consider the following example. Consider the following family of Gaussian distributions indexed by $\lambda > 0$

$$G_\lambda(x) := \frac{1}{4\pi\lambda}e^{-\frac{|x|^2}{4\pi\lambda}}, \qquad x \in \mathbb{R}^2.$$

It is an easy exercise to prove that $\int_{\mathbb{R}^2} G_\lambda(x)dx = 1$ for all $\lambda > 0$. Since G_λ is radial, we clearly have

$$\int xG_\lambda(x)dx = 0$$

for all $\lambda > 0$. Now, let us compute the second moment of G_λ:

$$M(\lambda) := \int_{\mathbb{R}^2} |x|^2 G_\lambda(x)dx = \frac{1}{4\pi\lambda}\int_{\mathbb{R}^2} |x|^2 e^{-\frac{|x|^2}{4\pi\lambda}}dx = 4\pi\lambda\int_{\mathbb{R}^2} |y|^2 e^{-|y|^2}dy =: C\lambda$$

where $C := 4\pi\int_{\mathbb{R}^2} |y|^2 e^{-|y|^2}dy > 0$. Now, this computation shows that the second moment of G_λ is proportional to λ, which implies for instance that $M(\lambda) \to 0$ as $\lambda \to 0$. Drawing the graph of a Gaussian shows that G_λ is more and more *concentrated* around its center of mass (zero, in this case) for small λ's. This examples shows that the second moment can be taken in general as a tool to measure concentrations around the center of mass. When concentration arises, we have *aggregation* prevailing with respect to diffusion (remember that diffusion implies *spreading*, which is opposite to aggregation).

Another important point is that no function $f(x) \geq 0$ with positive mass $\int f(x) = M > 0$ can have zero second moment. Indeed, if this would be the case we would have

$$\int_{\mathbb{R}^2} |x|^2 f(x) dx = 0$$

and the integrand above is nonnegative and not identically zero. Therefore, we have once again to use the theory of distribution in order to represent a *function* with unit mass and zero second moment. In fact, the limiting case $\lambda = 0$ in the above example gives $G_0 = \delta_0$ where δ_0 is the Dirac's delta distribution.

Let us now go back to our Keller–Segel model and let us compute the second moment of the solution n, namely $M(t) = \int_{\mathbb{R}^2} |x|^2 n(x,t) dx$.

$$\begin{aligned}
\frac{d}{dt} M(t) &= \int |x|^2 n_t dx = \int |x|^2 (\Delta n - \operatorname{div}(n \nabla c)) dx = -2 \int x \cdot \nabla n dx + 2 \int n x \cdot \nabla c dx \\
&= 4 \int n dx - \frac{1}{\pi} \int \int x \cdot \frac{x-y}{|x-y|^2} n(y) n(x) dy dx \\
&= 4M - \frac{1}{2\pi} \int \int (x-y) \cdot \frac{x-y}{|x-y|^2} n(y) n(x) dy dx \\
&= 4M - \frac{M^2}{2\pi} = M \left(4 - \frac{M}{2\pi} \right).
\end{aligned}$$

Therefore, if $M > 8\pi$, we have $\frac{d}{dt} M(t) < -c < 0$ and $c > 0$ is a fixed constant depending on M. Hence, $M(t)$, initially positive, will reach the value 0 in a finite time $T > 0$ (which can be computed explicitly: Exercise!).

Therefore, we have proven that, if $M > 8\pi$ the second moment (variance) of n becomes zero in a finite time. This phenomenon intuitively implies that the solution is concentrating to a Dirac's delta as $t \to T$ for some finite time $T > 0$. Such phenomenon is also known as *blow-up*, because one can also see that the L^∞ norm of n diverges as $t \to T$. To see this, suppose by contradiction that

$$\sup_{t \in [0,T]} \|n(t)\|_{L^\infty} \leq C.$$

Then,

$$0 = \int |x|^2 n(x,T) dx$$

implies that $n(x,T) = 0$ everywhere except in $x = 0$. But then this implies that the mass of n at time T is zero, since n is zero everywhere except at one point. This is a contradiction with the conservation of the mass (the mass is initially nonzero). The contradiction is due to the fact that we assumed n to be a bounded function near $t = T$. This is actually not the case, and therefore n blows up at $x = 0$ when t approaches T.

The blow-up in a finite time is a typical nonlinear phenomenon which accounts for instability. The great advantage in the previous computation is that it is purely nonlinear (without linearizing!). Blow-up and Touring instability are different phenomena in principle, but in this context they are somewhat related as it is shown by the result in this section and in the previous one. In both cases one gets instability vs. stability by means of a threshold phenomena which involves the total amount of individuals (the total mass in this section, the steady state in the previous section).

9.3 Chemotaxis with nonlinear diffusion

In section 5.5 we have seen that a more realistic way to model diffusion is provided by the Darcy's law

$$J = -D(n)\nabla n$$

for the diffusive flux of individuals in some animal populations. Here $D(n)$ increases with respect to n. In this framework, the diffusivity depends on the density in such a way that individuals tend to spread faster for higher densities. Such a phenomenon has been experimentally observed in some specific situations in which the *repulsive* effect due to diffusion becomes stronger and stronger when the distance between two individuals gets smaller and smaller. Moreover, such an assumption can be motivated by a volume filling effect, namely, when the density is very high individuals *sense* each other more intensely due to their positive volume.

Several authors have recently addressed the use of nonlinear diffusion also in the context of chemotaxis, with the aim of achieving a more refined balance between diffusion and chemotaxis, in which (possibly) the competition between the two effects result in possible existence of steady states (or non trivial patterns) when chemotaxis prevails. In a few words, one would like to see a pattern formation instead of a blow up effect as a consequence of a chemotaxis–dominated regime. We shall see that the use of the Darcy's law is one of the possible ways to produce (at least) a solution to the Keller–Segel model which never blows up (no matter how large the initial mass is) and which is always bounded uniformly in time. This suggests the formation of nontrivial patterns for large times.

Let us then consider once again the Keller–Segel system described in the above section with a nonlinear diffusion instead of a linear one. We shall choose

$$D(n) = Cn^{m-1}, \qquad m > 1, \qquad C > 0,$$

which represents a general case of a diffusivity with polynomial growth for large densities. Since $\mathrm{div}(n^{m-1}\nabla m) = \frac{1}{m}\Delta n^m$, we can normalize constants to obtain

$$\begin{cases} n_t = \Delta n^m - \mathrm{div}(n\nabla c) \\ 0 = \Delta c + n \\ n(t,0) = n_0(x). \end{cases} \tag{9.3.1}$$

We somewhat expect that being $D(n)$ larger and larger for high densities helps the diffusion effects to prevail, thus pushing down the density in case of strong aggregation.

We shall prove here that this is actually the case, namely the density n does not blow-up as it did in the case with linear diffusion. For simplicity we shall work once again in the whole space $x \in \mathbb{R}^2$. We shall assume further that the solution n is initially *compactly supported.* Due to the fact that the diffusion is nonlinear (like a porous medium type diffusion), we expect that the support remains compact at any time t. This fact should be actually proven rigorously, but the proof goes beyond our purposes. We shall therefore assume that $\mathrm{supp}(n(t))$ is finite at any $t \geq 0$. Let us now consider the *energy functional*

$$E(n,c) := \frac{1}{m-1}\int n^m dx - \frac{\chi}{2}\int ncdx.$$

Let us compute the evolution of $E(n(t),c(t))$:

$$\frac{d}{dt}E(n(t),c(t)) = \frac{m}{m-1}\int_{\mathbb{R}^2} n^{m-1}n_t dx - \frac{\chi}{2}\int n_t c dx - \frac{\chi}{2}\int n c_t dx.$$

Now, let us recall that we can actually solve the Poisson equation $\Delta c = -n$ on $\mathbb{R}^2$ as follows:

$$c(x,t) = -\frac{1}{2\pi}\int \log|x-y| n(y,t)dy.$$

Therefore, we have

$$\int n(x,t)c_t(x,t)dx = -\frac{1}{2\pi}\int\int \partial_t\left(\log|x-y|n(y,t)\right)n(x,t)dydx$$
$$= -\frac{1}{2\pi}\int\int \log|x-y|n_t(y,t)n(x,t)dydx$$

and since $\log|x-y|$ is invariant after the change of variable $(x,y)\mapsto(y,x)$, we have

$$\int n(x,t)c_t(x,t)dx = -\frac{1}{2\pi}\int\int \log|x-y|n(y,t)n_t(x,t)dydx = \int n_t(x,t)c(x,t)dx.$$

Therefore, going back to the evolution of the energy,

$$\frac{d}{dt}E(n(t),c(t)) = \frac{m}{m-1}\int_{\mathbb{R}^2} n^{m-1}n_t dx - \chi\int n_t c dx = \int_{\mathbb{R}^2} n_t\left(\frac{m}{m-1}n^{m-1} - \chi c\right)dx. \quad (9.3.2)$$

Now let us write the continuity equation for n as follows:

$$n_t = \mathrm{div}\left(\nabla n^m - \chi n\nabla c\right) = \mathrm{div}\left(n\nabla\left(\frac{m}{m-1}n^{m-1} - \chi c\right)\right),$$

which on substituting into (9.3.2) yields

$$\frac{d}{dt}E(n(t),c(t)) = \int_{\mathbb{R}^2}\left(\frac{m}{m-1}n^{m-1} - \chi c\right)\mathrm{div}\left(n\nabla\left(\frac{m}{m-1}n^{m-1} - \chi c\right)\right)dx.$$

Now, assuming (as usual) that the quantity $\left(\frac{m}{m-1}n^{m-1} - \chi c\right)$ goes to zero as $|x|\to+\infty$, we can integrate by parts in the above expression to get

$$\frac{d}{dt}E(n(t),c(t)) = -\int_{\mathbb{R}^2} n\left|\nabla\left(\frac{m}{m-1}n^{m-1} - \chi c\right)\right|^2 dx \leq 0.$$

This means that the energy is non increasing along the solutions to (9.3.1). Assuming further that the initial energy is finite, we have

$$\frac{1}{m-1}\int n^m(x,t)dx - \frac{\chi}{2}\int n(x,t)c(x,t)dx \leq E(n_0,c_0) < +\infty$$

which shows that the energy $E(n(t),c(t))$ is bounded from above by a finite quantity uniformly in time. We want to prove that actually $E(n(t),c(t))$ is *bounded*. We shall actually prove a stronger result, namely that the quantity

$$\frac{1}{m-1}\int n^m(x,t)dx$$

is bounded from above. In order to prove that we shall make use of the following *logarithmic Hardy–Littlewood–Sobolev* inequality

$$-\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f(x)\log|x-y|f(y)dydx \le \frac{M}{2}\int_{\mathbb{R}^2} f(x)\log f(x) + C,$$

which holds for all $f \ge 0$ such that $\int f\log f < +\infty$. Here, M is the total mass of f and C is a positive constant depending on M. With such inequality at hand, we obtain

$$\begin{aligned} E(n(t),c(t)) &\ge \frac{1}{m-1}\int n^m(x,t)dx - \frac{M\chi}{8\pi}\int n(x,t)\log n(x,t)dx - \frac{\chi C}{4\pi} \\ &= \int_{\mathbb{R}^2}\left(\frac{1}{m-1}n^m(x,t) - \frac{M\chi}{8\pi}n(x,t)\log n(x,t)\right)dx - \frac{\chi C}{4\pi}. \end{aligned}$$

Now, we observe that the function

$$[0,+\infty) \ni n \mapsto g(n) := \frac{1}{m-1}n^{m-1} - \frac{M\chi}{8\pi}\log n$$

satisfies

$$\lim_{n\to+\infty} g(n) = +\infty.$$

Then, there exists a constant $K \in \mathbb{R}$ such that

$$g(n) \ge K \qquad \text{for all} \quad n \ge 0.$$

Hence, we can write

$$E[n,c] \ge \int ng(n)dx - \frac{\chi C}{4\pi} \ge KM - \frac{\chi C}{4\pi} > -\infty$$

which proves E is bounded from below too. Now, for a similar reason, there exist two constants $L,R > 0$ such that

$$g(n) \ge Ln^{m-1} \qquad \text{for all } n > R\,.$$

Hence,

$$\begin{aligned} L\int n(x,t)^m dx &= L\int_{x\,:\,n(x,t)\le R} nn^{m-1}dx + L\int_{x\,:\,n(x,t)>R} nn^{m-1}dx \\ &\le LR^{m-1}M + \int ng(n)dx - \int_{x\,:\,n(x,t)\le R} ng(n)dx \\ &\le LR^{m-1}M - KM + \int ng(n)dx \le LR^{m-1}M - KM + \frac{C\xi}{4\pi} + E[n_0,c_0]\,. \end{aligned}$$

We have therefore obtained a global-in-time uniform bound for the L^m norm of $n(\cdot, t)$. This last assertion is a valid argument to deduce that n cannot become a Dirac's delta in a finite time. Indeed, otherwise n could not remain bounded in L^m for $m > 1$.

With the above argument have therefore proven (although in a quite sketchy way) that no finite-time blow up can occur if we replace the classical linear diffusion with a porous medium type diffusion in the Keller–Segel system. Unfortunately, such an argument does not provide any information about possible formation of patterns in the large time. Moreover, we don't know whether stationary solutions exists or not. We still would like to see some *complexity* in the model, namely, to produce multiple behavior by changing the initial data or the parameters of the model.

9.4 Models with maximal density

Another recent approach (due to Hillen and Painter) aiming at preventing concentration of the density of individuals in chemotaxis systems consists in prescribing a *maximal density*. More precisely, one assumes there exists a value n_{max} such that

$$0 \le n(x,t) \le n_{max}, \qquad \text{for all} \quad (x,t) \in \mathbb{R}^d \times [0, +\infty).$$

As a matter of fact, the model (and the initial conditions) should be modified in such a way that such a property is satisfied by the solution for all times. One possible way to perform this task is to consider the following modified equation for the density of individuals

$$n_t = D\Delta n - \chi \operatorname{div}\left(n(n_{max} - n)\nabla c\right)$$

in which we have taken the chemotactic sensitivity to be depending on the density n in a way that the chemotaxis effect shuts off when n reach the density n_{max}. In fact one can easily prove (we shall not do it here) that if the initial datum satisfies $0 \le n_0(x) \le n_{max}$, then the solution satisfies $0 \le n(x,t) \le n_{max}$ for all x and t.

Although possibly more realistic, such a model turns out to be quite uninteresting from the point of view of complexity. Indeed, one can easily prove, for instance, that the one dimensional model

$$\begin{cases} n_t = n_{xx} - \chi(n(n_{max} - n)c_x)_x \\ 0 = c_{xx} - \alpha c + \beta n, \qquad \chi, \alpha, \beta > 0, \end{cases}$$

does not admit nontrivial steady states on the whole $\mathbb{R}$. To prove that, suppose by contradiction that (n^∞, c^∞) is a steady state. Then

$$0 = n^\infty_{xx} - \chi(n^\infty(n_{max} - n^\infty)c^\infty_x)_x = \left(n^\infty(n_{max} - n^\infty)\left(\log \frac{n^\infty}{n_{max} - n^\infty} - \chi c\right)_x\right)_x .$$

Since n^∞ should be integrable at $|x| \to +\infty$, the term

$$n^\infty(n_{max} - n^\infty)\left(\log \frac{n^\infty}{n_{max} - n^\infty} - \chi c\right)_x$$

should be zero (it is constant, but the constant can only be zero). Therefore, either $n^\infty = 0$, or $n^\infty = n_{max}$, or

$$\log \frac{n^\infty}{n_{max} - n^\infty} = \chi c^\infty + C$$

for some constant $C \in \mathbb{R}$. Now, clearly the steady state n^∞ should be continuous (this is a consequence of the smoothing effect of the diffusion operator, we omit the details). Therefore, in case $n^\infty(x) = n_{max}$ at some point x, there must be a sequence of points $x_k \to x$ such that $n^\infty(x_k) \to n_{max}$ and $n^\infty(x_k) < n_{max}$. By substituting above we get

$$\log \frac{n^\infty(x_k)}{n_{max} - n^\infty(x_k)} = \chi c^\infty(x_k) + C$$

and the term in the left hand side diverges to $+\infty$. Therefore $c(x_k) \to +\infty$. This is impossible, because we can prove that c is uniformly bounded. To see this, multiply the equation for c by c and integrate on $\mathbb{R}$:

$$0 = \int c c_{xx} dx - \alpha \int c^2 dx + \beta \int c n dx.$$

Integration by parts implies

$$\int c_x^2 dx + \alpha \int c^2 dx = \beta \int c n dx \leq \frac{\alpha}{2} \int c^2 dx + C(\alpha, \beta) \int n^2 dx,$$

where we have used the weighted Young's inequality on the right hand side. Recalling the definition of the Sobolev H^1 norm we obtain

$$\|c(t)\|_{H^1}^2 \leq C_0(\alpha, \beta) \int n^2 dx$$

and the last term is uniformly bounded. Recalling the Sobolev inequality 6.3.1, we obtain

$$\|c(t)\|_{L^\infty} \qquad \text{uniformly bounded for all } t \geq 0.$$

We have therefore proven that n^∞ can never touch the value n_{max}. On the other hand, with a similar argument we can prove that n^∞ cannot touch the value 0, because otherwise we would have the logarithmic term above unbounded once again. This cannot happen also when $|x| \to +\infty$. Therefore, the steady state should be such that $n^\infty \in [\epsilon, 1 - \epsilon]$ for some $\epsilon > 0$, but this fact contradicts the integrability of n^∞.

A valid alternative of the above model is the following one

$$\begin{cases} n_t = \operatorname{div}\left(n(n_{max} - n)\nabla\left(Dn - \chi c\right)\right) \\ 0 = \Delta c - \alpha c + \beta n. \end{cases} \tag{9.4.1}$$

Let us point out that the chemical c can be recovered by solving the elliptic equation above as follows. For simplicity, we shall solve such a problem in one space dimension. One can easily prove (we shall not do it here) that if the initial datum n_0 satisfies $0 \leq n_0 \leq n_{max}$, then $0 \leq n(x,t) \leq n_{max}$.

Let us apply the Fourier's transform to the equation

$$0 = c_{xx} - \alpha c + \beta n,$$

to obtain

$$0 = -(4\pi^2\xi^2 + \alpha)\widehat{c} + \beta \widehat{n}.$$

This implies that

$$\widehat{c}(\xi,t)=\frac{\beta}{4\pi^2\xi^2+\alpha}\widehat{n}=\frac{\beta}{\alpha}\cdot\frac{1}{4\pi^2(\xi/\sqrt{\alpha})^2+1}\widehat{n}.$$

Now, we leave as en exercise for the reader to compute the inverse Fourier transform

$$\mathfrak{F}^{-1}\left(\frac{1}{4\pi^2(\xi/\sqrt{\alpha})^2+1}\right)(x)=\int_{-\infty}^{+\infty}e^{ix\xi}\frac{1}{4\pi^2(\xi/\sqrt{\alpha})^2+1}d\xi=\frac{\sqrt{\alpha}}{2}e^{-|\sqrt{\alpha}x|}=:K_\alpha(x).$$

Then, the convolution property of the Fourier transform implies

$$\begin{aligned}c(x,t)&=\frac{\beta}{\alpha}\mathfrak{F}^{-1}\left(\widehat{K_\alpha}(\xi)\widehat{n}(\xi,t)\right)\\&=\frac{\beta}{\alpha}K_\alpha*n(x,t)=\frac{\beta}{2\sqrt{\alpha}}\int_{-\infty}^{+\infty}e^{-\sqrt{\alpha}|x-y|}n(y,t)dy.\end{aligned}$$

Let us define the energy

$$E(n,c)=\int n(Dn-\chi c)dx,$$

and compute its evolution

$$\begin{aligned}&\frac{d}{dt}E(n(t),c(t))=2D\int nn_tdx-2\chi\int cn_tdx=2\int n_t(Dn-\chi c)dx\\&=2\int(Dn-\chi c)\mathrm{div}\left(n(n_{max}-n)\nabla\left(Dn-\chi c\right)\right)=\\&=-2\int n(n_{max}-n)\left|\nabla\left(Dn-\chi c\right)\right|^2dx\le 0,\end{aligned}$$

where we have used integration by parts in the last step and the fact that $\int n_tcdx=\int nc_t$, which analogous to what we have seen for the Keller–Segel system with nonlinear diffusion. The term on the right-hand-side is called *energy production*, and we shall denote it by $-I(n(t),c(t))$. More precisely, we shall write

$$\frac{d}{dt}E(n(t),c(t))=-2I(n(t),c(t)).$$

Let us now consider two separate cases, namely $D<\frac{\beta\chi}{\alpha}$ and $D>\frac{\beta\chi}{\alpha}$.

Case $D<\frac{\beta\chi}{\alpha}$. In this case we can prove (or, at least, give a sketch of the proof) that there exists a non-trivial, continuous stationary solution to the model (9.4.1) in one space dimension. To see that, decouple system (9.4.1) by using the above convolution method and consider a stationary solution n^∞ (c^∞ can be recovered by convolution):

$$0=(n^\infty(n_{\max}-n^\infty)(Dn^\infty-\frac{\beta\chi}{\alpha}K_\alpha*n^\infty)_x)_x.$$

By the usual argument, we impose the argument of the above space derivative equals zero. This is possible if one of the three following conditions hold: either $n^\infty=0$, or $n^\infty=n_{\max}$, or

$$\varepsilon n^\infty-K_\alpha n^\infty=C\,,$$

with

$$\varepsilon=\frac{D\alpha}{\beta\chi}<1\,.$$

Assuming n^∞ is not identically zero (which is obvious given that we are looking for non trivial steady states), even if we assume $n = n^\infty$ on some interval, due to the continuity of the steady state there must be a non trivial interval $I = [a, b]$ on which $\varepsilon n^\infty - K_\alpha n^\infty = C$ on I. Assuming n^∞ is an even symmetric function which is also "radially decreasing", that is $n_x^\infty \leq 0$ on $[0, +\infty)$, we may assume the above equation holds only on two intervals I and $-I$. Given $u = n_x^\infty$, we get on the whole $\mathbb{R}$,

$$\varepsilon u - K_\alpha u = 0\,.$$

Under the above mentioned conditions, the last equation is in fact equivalent to $u = n_x^\infty$ and n^∞ being a steady state. We may apply the Fourier transform and get

$$(\varepsilon - \hat{K_\alpha})\hat{u} = 0\,.$$

Multiplying by $\hat{u}$ and integrating on $\mathbb{R}$ we get

$$\int (\varepsilon - \hat{K_\alpha}(\xi))\hat{u}^2(\xi)d\xi = 0\,.$$

Since $\hat{K_\alpha}(0) = 1 = \int K_\alpha dx$, the weight function $\varepsilon - \hat{K_\alpha}(\xi)$ is negative for small ξ, and since $K_\alpha \to 0$ at infinity, the same weight is positive for large ξ. It is therefore possible to find a nontrivial $\hat{u}$ that satisfies the above equation by choosing two "bumps", suitably supported on the negative and on the positive parts of $\varepsilon - \hat{K_\alpha}(\xi)$ respectively, we omit the details.

Case $D > \frac{\beta\chi}{\alpha}$. In this case we can (once again, by a sketch) prove that there exist no compactly supported stationary solutions other than $n^\infty = 0$. To see this, with the computations in the previous case we obtain

$$\int (\varepsilon - \hat{K_\alpha}(\xi))\hat{u}^2(\xi)d\xi = 0\,,$$

this time with the weight $\varepsilon - \hat{K_\alpha}(\xi) > 0$ for all $\xi \in \mathbb{R}$. Therefore, the only possibility to satisfy that equation is $\hat{u} = 0$, which corresponds to n^∞ constant, and the only possible constant is zero due to the finite mass.

Chapter 10

Nonlocal transport models for swarms

In this chapter we describe nonlocal models aiming at describing the dynamics of animal "swarms" (herds, flocks, schools, etc.) in population dynamics. This is a rapidly emerging research fields, with many applications in robotics too, and most recently also in the dynamics of epidemics. We refer to the paper by Mogilner and Edelstein-Keshet on J. Math. Biology in 1999 for an introduction to the subject. These equations can describe interactions at a distance, e.g. due to vision, hearing, and other senses.

We consider an integro-differential equation model that is simple enough to be treated analytically. The model captures the idea of attraction-repulsion interactions between organisms.

Let us assume we have a group of individuals, whose density is described by $\rho = \rho(x,t)$. The individuals are subject to diffusion, seen as a repulsive effect, which for simplicity we shall assume to be linear with constant $D > 0$. The "attraction" part is modelled through an interaction kernel $K : \mathbb{R}^2 \to [0,+\infty)$ which is assumed to be smooth, at least C^2 in this chapter, radial, that is $K(x) = k(|x|)$, and radially decreasing, that is $k'(r) < 0$ if $r > 0$. Putting all this together we get

$$\rho_t = D\Delta\rho - \mathrm{div}(\rho\nabla K * \rho)\,. \tag{10.0.1}$$

The equation is as usual coupled with an initial condition $\rho_0 \in L^1_+(\mathbb{R}^d)$. We shall work on $x \in \mathbb{R}^d$. The nonlocal part is understood as follows. The velocity in the continuity equation has a nonlocal part

$$v(x,t) = -\nabla K * \rho(x,t) = \int_{\mathbb{R}^d} \nabla K(x-y)\rho(y,t)dy\,.$$

In short, every individual set on position x interacts with all other individuals at all positions y depending on the distance $|x-y|$, which is in fact the only quantity affecting the kernel.

10.1 The purely nonlocal case $D = 0$

. In this case, the models looks like

$$\rho_t = -\mathrm{div}(\rho\nabla K * \rho)\,. \tag{10.1.1}$$

There are many ways to solve this model. We shall adopt the one which in fact one of the main reasons why this model is so popular. We shall prove that the solution ρ can be expressed as the approximation of the "empirical measure" of a finite number of individuals. Let us explain this in detail.

Consider N individuals with positions $x_1(t), \dots, x_N(t) \in \mathbb{R}^d$ and set $v_i(t) = \dot{x}_i(t)$ as their velocities. We need some way to describe the evolution of the whole set of individuals "jointly". To this purpose, we introduce the *empirical measure*

$$\mu_N(t) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i(t)} \,.$$

To be more precise, μ_N is the distribution on the space $\mathcal{D}'((0,+\infty) \times \mathbb{R}^d)$ acting as follows on test functions $\varphi \in C_c^\infty((0,+\infty) \times \mathbb{R}^d)$:

$$\langle \mu_N, \varphi \rangle = \frac{1}{N} \int_0^{+\infty} \sum_{i=1}^{N} \varphi(x_i(t), t) dt \,.$$

Our purpose is to set up an evolution for the set of individuals such that the empirical measure μ_N satisfies (10.1.1). The correct way to do that, inspired by the particular form of the velocity field in the continuity equation (10.1.1), is the following one:

$$\dot{x}_i(t) = \frac{1}{N} \sum_{k=1}^{N} \nabla W(x_i(t) - x_j(t)) \,, \qquad i = 1, \dots, N \,. \tag{10.1.2}$$

We claim that, assuming x_i satisfy (10.1.2), the empirical measure μ_N is a solution to (10.1.1) in the sense of distributions. We first compute the convolution

$$\begin{aligned} \nabla W * \mu_N(x,t) &= \int_{\mathbb{R}^d} \nabla W(x-y) d\mu(y) = \langle \mu_N, \nabla W(x - \cdot) \rangle \\ &= \frac{1}{N} \sum_{j=1}^{N} \langle \delta_{x_i(t)}, \nabla W(x - \cdot) \rangle = \frac{1}{N} \sum_{j=1}^{N} \nabla W(x - x_j(t)) \,. \end{aligned}$$

Now, in order to prove that $\mu_N(t)$ satisfies (10.1.1), by a formal integration by parts we need to show that, for all $\varphi \in C_c^\infty((0,+\infty) \times \mathbb{R}^d)$,

$$-\int_0^{+\infty} \langle \mu, \varphi_t \rangle dt - \int_0^{+\infty} \langle \mu, \nabla W * \mu \cdot \nabla \varphi \rangle dt = 0 \,.$$

The left hand side reads

$$-\int_0^{+\infty} \frac{1}{N} \sum_{i=1}^{N} \varphi_t(x_i(t), t) dt - \int_0^{+\infty} \frac{1}{N} \sum_{i=1}^{N} \nabla W * \mu_N(x_i(t), t) \cdot \nabla \varphi(x_i(t), t) dt,$$

and the above computation on the convolution implies the right hand side equals

$$-\int_0^{+\infty} \frac{1}{N} \sum_{i=1}^{N} \varphi_t(x_i(t), t) dt - \int_0^{+\infty} \frac{1}{N} \sum_{i=1}^{N} \dot{x}_i(t) \cdot \nabla \varphi(x_i(t), t) dt$$

and the above equals

$$-\frac{1}{N}\sum_{i=1}^{N}\int_{0}^{+\infty}\frac{D}{Dt}\varphi(x_i(t),t)dt = -\frac{1}{N}\sum_{i=1}^{N}(\varphi(x_i(+\infty),+\infty) - \varpi(x_i(0),0))$$

which equals zero since φ has compact support.

Hence, if the initial condition of the PDE (10.1.1) has the form of an empirical measure $\mu = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_{i,0}}$, we know how to solve the model by considering the empirical measure μ_N with x_i solving (10.1.2) and having initial conditions $x_i(0) = x_{i,0}$. By the way, the ODE system (10.1.2) has a unique local solution provided W is C^2 according to Cauchy-Lipschitz theorem.

Now, in order to solve the PDE for a general initial condition $\rho_0 \in L^\infty$ with compact support we adopt a formal "particle method". We approximate ρ_0 by an empirical measure $\mu_{N,0}$. We omit the details, but this is possible for example in one space dimension by dividing the subgraph of ρ_0 by N regions with equal mass, and place a particle at each left (or each right) edge of the resulting intervals. The details are in the hand-written notes.

Then, we prove the family of corresponding empirical measure is compact in the space of measures thanks to a uniform estimate of the support, see the hand-written notes. Finally, we prove consistency in the limit thanks to Prokhorov's Theorem, see the hand-written notes.

10.2 The diffusive case $D > 0$

.

See the hand-written notes.

10.3 Collapse in finite time for the case $D = 0$

.

See the hand-written notes.

Chapter 11

Structured population dynamics

Many problems arising in biology may be described, at a first stage, using differential equations. This means that the model has been elaborated after averaging some population in order to keep only the time variable. This average usually hides some character that can vary from an individual to the other. Taking in to account this character leads to the so–called *structured population* dynamic equations.

We shall give some example from ecology, of population structured by a parameter describing a biological character of the individuals. When this character is inherent to the individual, i. e. it is fixed at the very beginning of its life, we refer to it as a *trait* (*phenotype* alternatively). The theory which focuses on phenotypic evolution driven by small mutations in replication, while ignoring both sex and genes, is known by the name *Adaptive Dynamics* and is part of *Evolution theory.*

The two main ingredients in this theory are (i) the selection principle which prescribes the population with best adapted trait, and (ii) mutations which allow off-springs to have slightly different traits than their mother. The combination of the two effects is studied by adaptive dynamics. This turns out to be an extremely complicated theory on which several possible mathematical approaches are possible. One of the reasons is that it is merely impossible to consider this problem without introducing small parameters (mutations can be small or rare for instance, population should be large in any case but relative death rates can vary). Therefore adaptive evolution can be studied with various mathematical tools.

In this chapter, we give a first and very elementary point of view based on structuring an ODE in this context. Our main goal is to perform the corresponding asymptotic theory and show how the concept of monomorphic population arises naturally in the limit of small mutations over a long time compared to one generation length. Moreover, we shall prove a similar result in a continuous model.

11.1 An example in ecology: competition for resources

We consider the dynamics of several micro-organisms under the action of a *chemostat*, i. e. a single substrate acting as a nutrient. An inflow with rate R of pure chemostat with concentration S_0 is compensated by an outflow with the same rate containing both the micro-organisms and the nutrient. In this situation, the system is given by

$$\begin{cases} \dot{S}(t) = R(S_0 - S) - \sum_{j=1}^{I} \eta_j(S) N_j \\ \dot{N}_i(t) = N_i(\eta_i(S) - R) \\ S(0) = S^0 > 0 \\ N_i(0) = N_i^0 > 0, \quad i = 1, \dots, I. \end{cases} \tag{11.1.1}$$

The variable $S(t)$ denotes the single substrate of nutrient (in terms of the mass of a given representative constituent), $N_i(t)$ is the biomass of the i-th micro-organism in the chemostat (written in terms of the same constituent), R is the dilution rate of the input flow of nutrient concentration S^0. Finally, the ability for the i-th organism to use the nutrient S depends only on S in the simplest model, and is denoted by $\eta_i(S)$ (uptake rates).

One usually assumes that η_i is increasing with respect to S, more precisely we shall require

$$\eta_i'(S) \geq \alpha > 0, \qquad \text{for all} \quad i = 1, \dots, I, \quad \text{and} \quad S \geq 0. \tag{11.1.2}$$

Moreover, we assume that $\eta_i(S_0) > R$ for all $i = 1, \dots, I$. The latter indicates that the initial amount S_0 of chemostat is enough to produce a growth for all the species. Also, assume that the numbers $\eta_i^{-1}(R)$ are all different. Then there are $I + 1$ steady states. The first one is the trivial one $N_i \equiv 0$, $S = S_0$. To find the non trivial ones, suppose

$$\eta_k(S) = R$$

for some $k \in \{1, \dots, I\}$, which corresponds to have N_k non zero. Then, since $\eta_j(S) \neq R$ for all $j \neq k$, we have $N_j = 0$ for all $j \neq k$. Therefore, the steady equation for S implies

$$0 = R(S_0 - S) - RN_k$$

which gives

$$N_k = S_0 - \eta_k^{-1}(R).$$

Hence, besides the trivial steady state we have the I states composed of a single micro-organism

$$(0, \dots, 0, \overline{N}_i, 0, \dots, 0) \qquad \overline{S} = \eta_i^{-1}(R) < S_0,$$

with $\overline{N}_i = S_0 - \eta_i^{-1}(R)$.

We shall prove now that, among these steady states, only one is asymptotically stable. More precisely, among the I species of micro-organisms, only one will survive in the large time asymptotics. Such species is somehow determined as the one with the 'largest' growth rate for a given amount of nutrient S. More precisely, let us denote by i_0 the index such that

$$S^* := \min_{1 \leq i \leq I} \eta_i^{-1}(R) = \eta_{i_0}^{-1}(R).$$

Then, we shall prove that

$$N_{i_0}(t) \to \overline{N}_{i_0} = S_0 - S^*, \qquad S(t) \to S^*$$

as $t \to +\infty$.

As a first step let us prove the following balance law. Let us add all equations in (11.1.1). We obtain

$$\frac{d}{dt}\left[S(t)+\sum_{i=1}^{I}N_i(t)\right]=R(S_0-S)-\sum_{j=1}^{I}\eta_j(S)N_j+\sum_{i=1}^{I}N_i(\eta_i(S)-R)$$
$$=-R\left[S(t)+\sum_{i=1}^{I}N_i(t)\right]+RS_0.$$

Therefore, by solving the above linear ODE for the quantity in the squared bracket, we obtain

$$S(t)+\sum_{i=1}^{I}N_i(t)=e^{-Rt}\left(S^0+\sum_{i=1}^{I}N_i^0+RS_0\int_0^t e^{Rs}ds\right)$$
$$=e^{-Rt}\left(S^0+\sum_{i=1}^{I}N_i^0+S_0(e^{Rs}-1)\right)=S_0+Q_0e^{-Rt}, \tag{11.1.3}$$

with

$$Q_0=-S_0+S^0+\sum_{i=1}^{I}N_i^0.$$

The above property also proves that all the quantities $N_1,\ldots,N_I,S$ are uniformly bounded in t.

As a second step, we prove that the sum of all the populations of micro-organisms is bounded from below for large times. To see this, let us take the sum of all equations for N_i:

$$\frac{d}{dt}\sum_{i=1}^{I}N_i(t)=\sum_{i=1}^{I}N_i(t)(\eta_i(S)-R).$$

Let us define

$$\underline{\eta}(S):=\min_{i=1,\ldots,I}\eta_i(S),$$

with $\underline{\eta}(S)>0$ for all S (check!). Then, we have

$$\frac{d}{dt}\sum_{i=1}^{I}N_i(t)\geq(\underline{\eta}(S)-R)\sum_{i=1}^{I}N_i(t).$$

Now, if $\sum_{i=1}^{I}N_i(t)$ tends to zero as $t\to+\infty$, then (11.1.3) implies that $S(t)\to S_0$. Since, due to the assumption $\eta_i(S_0)>R$ for all $i=1,\ldots,I$ we have $(\underline{\eta}(S_0)-R>0$, then we can conclude that $(\underline{\eta}(S)-R)$ is uniformly positive for large times and therefore $\sum_{i=1}^{I}N_i(t)$ grows exponentially for large times, which contradicts the fact that all N_i are uniformly bounded. Therefore, we have proven that

$$\liminf_{t\to+\infty}\sum_{i=1}^{I}N_i(t)=M>0. \tag{11.1.4}$$

We now prove that $S(t)$ has a limit as $t\to+\infty$. Such a procedure is usually referred to as *convergence*, which must be complemented with *consistency* in order to prove that a certain time dependent quantity has a certain limit as $t \to +\infty$. The former only solve the problem of the existence of a limit, i. e. of a unique limit point. From the first step we know that

$$\dot{S} + \sum_{i=1}^{I} \dot{N}_i = -RQ_0 e^{-Rt},$$

which yields, after taking the time derivative of $\dot{S}$,

$$\begin{aligned}
\frac{d}{dt}\dot{S}(t) &= \frac{d}{dt}\left(R(S_0 - S) - \sum_{j=1}^{I} \eta_j(S) N_j \right) \\
&= -R\dot{S} - \sum_{i=1}^{I} \eta_i'(S)\dot{S} N_i - \sum_{i=1}^{I} \eta_i(S)\dot{N}_i \\
&= -\sum_{i=1}^{I} (\eta_i(S) - R)\dot{N}_i - \dot{S}\sum_{i=1}^{I} \eta_i'(S) N_i + R^2 Q_0 e^{-Rt} \\
&= -\sum_{i=1}^{I} (\eta_i(S) - R)^2 N_i - \dot{S}\sum_{i=1}^{I} \eta_i'(S) N_i + R^2 Q_0 e^{-Rt}.
\end{aligned}$$

We now multiply the above identity by

$$\operatorname{sign}_+(\dot{S}) = \begin{cases} 1 & \text{if } \dot{S} \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

to obtain

$$\begin{aligned}
\frac{d}{dt}(\dot{S})_+ &= \operatorname{sign}_+(\dot{S}) \frac{d}{dt}\dot{S} \\
&= \operatorname{sign}_+(\dot{S}) \left(-\sum_{i=1}^{I} (\eta_i(S) - R)^2 N_i - \dot{S}\sum_{i=1}^{I} \eta_i'(S) N_i + R^2 Q_0 e^{-Rt} \right) \\
&\leq -(\dot{S})_+ \sum_{i=1}^{I} \eta_i'(S) N_i + R^2 Q_0 e^{-Rt}.
\end{aligned}$$

We recall that, for $z \in \mathbb{R}$,

$$(z)_+ = \max\{z, 0\}.$$

Due to the assumption (11.1.2) and to (11.1.4), therefore obtain

$$\frac{d}{dt}(\dot{S})_+ \leq -\alpha M (\dot{S})_+ + R^2 Q_0 e^{-Rt}.$$

Let us integrate the last inequality on the time interval $[0, T]$:

$$\alpha M \int_0^T (\dot{S}(T))_+ dt \leq R^2 Q_0 \int_0^T e^{-Rt} dt + (\dot{S}(0))_+ - (\dot{S}(T))_+.$$

We then send $T \to +\infty$ and recover that

$$\int_0^{+\infty} (\dot{S}(T))_+ dt < +\infty.$$

In a similar way (we omit the details) one can prove that

$$\int_0^{+\infty} (\dot{S}(T))_- dt < +\infty,$$

where

$$(z)_- = \max\{-z, 0\}.$$

We also recall that, for all $z \in \mathbb{R}$,

$$|z| = (z)_+ + (z)_-.$$

Hence, we have proven that

$$\int_0^{+\infty} |\dot{S}(T))| dt < +\infty,$$

which is equivalent to require that $S(t)$ is a function with bounded variation on the time half line $[0, +\infty)$. Standard results in functional analysis (cf. the book by Perthame and the PDE book by Evans) imply that there exists the limit $\lim_{t\to+\infty} S(t)$.

The next step is the consistency, i. e. the identification of the limit for $S(t)$ as $t \to +\infty$. We aim to prove that

$$\lim_{t\to+\infty} S(t) = S^*.$$

Assume first that

$$\lim_{t\to+\infty} S(t) > S^*.$$

Then, by definition of S^*,

$$\lim_{t\to+\infty} \eta_{i_0}(S) > R$$

and therefore, using the equation for N_{i_0} we deduce that N_{i_0} grows exponentially fast to $+\infty$, which is a contradiction with all species being uniformly bounded. Let us then assume that

$$\lim_{t\to+\infty} S(t) < S^*.$$

Now, the definition of η_{i_0} and the fact that all η_i are strictly increasing easily imply that

$$\eta_i(S^*) \leq \eta_{i_0}(S^*) = R, \qquad \text{for all} \quad i = 1, \ldots, I,$$

and therefore

$$\lim_{t\to+\infty} \eta_i(S) \leq \eta_i(S^*) < R$$

and using the equation for N_i we deduce that all species N_i decay to zero as $t \to +\infty$. This is clear because

$$\dot{N_i} = N_i(\eta_i(S) - R)$$

and $\lim_{\to+\infty}(\eta_i(S) - R) = -l < 0$ for some positive l, which implies that N_i behaves like the solution to

$$\dot{N_i} = -lN_i$$

for large times. This fact is a contradiction because of (11.1.4): the sum of the species $\sum N_i$ cannot converge to zero for large times. Therefore, the unique possible behavior for S is

$$\lim_{t\to+\infty} S(t) = S^*.$$

As a consequence of that, $\eta_{i_0}(S) - R$ tends to zero as $t \to +\infty$, whereas all the values $\eta_i(S) - R$ for $i \neq i_0$ achieve negative values in the limit. This implies that N_{i_0} is the only species not converging to zero in the limit. Since the limit of the whole dynamical system (11.1.1) is a steady state, the only possible limiting value for N_{i_0} is N^*.

11.2 Continuous traits

A natural generalization of the model (11.1.1) is that of a family of species indexed by a *continuous* trait (uptake ability, in this case) $x > 0$. The system (11.1.1) is replaced by the following system of integro–partial differential equations, in which the discrete variable i is replaced by a continuous variable x,

$$\begin{cases} \dot{S} = R(S_0 - S) - \displaystyle\int_0^{+\infty} \eta(x,S)n(x,t)dx, & t \geq 0,\ x \geq 0 \\ \dfrac{\partial}{\partial t}n(x,t) = n(x,t)(\eta(x,S) - R), & \\ S(0) = S^0 > 0, & \\ n(x,0) = n^0(x) > 0, & n^0 \in L^1 \cap L^\infty([0,+\infty)) \end{cases} \tag{11.2.1}$$

Under similar assumptions to those prescribed for the discrete model (11.1.1), one can prove that the solution $n(x,t)$ to (11.2.1) *concentrates* around one single trait x_0. More precisely, assuming that $\eta(x,S)$ is increasing with respect to S, that $\eta(x,S_0) > R$ for all $x \geq 0$, and assuming the existence of a unique point $x_0 > 0$ such that $S^* := \eta^{-1}(x_0,R) < \eta^{-1}(x,S)$ for all $x \neq x_0$ (here η^{-1} denotes the inverse with respect to S), one can prove that, as $t \to +\infty$,

$$n(x,t) \to (S_0 - S^*)\delta_{x_0}(x), \qquad S(t) \to S^*.$$

Such a limiting population (with a single trait) is called *monomorphic.*

We shall not prove the above result. On the other hand, in the next subsection we shall tackle a similar problem from evolution theory with a mathematical structure very similar to (11.2.1).

11.3 Evolutionary stable strategy in a continuous model

The next question we address here is to give a mathematical description of the process in which some specific trait is selected in a given environment. It is the best adapted trait in terms of using resources and that trait is called an *Evolutionary Stable Strategy* (ESS in short). The origin of this denomination comes from evolution theory; no other mutant with a different trait can invade a population with the trait corresponding to an ESS. An example of the selection principle has already been mentioned, in the case of the chemostat. Here we give an easy example that can be treated by explicit computations.

For simplicity we shall consider a variant of the logistic equation which we structure with a trait $x \in \mathbb{R}$ and we illustrate the selection principle on this very simple example. We consider that the reproduction rate depends on the trait, i.e., $b = b(x) > 0$ (b a continuous function), and that the death rate is proportional to the total population number. We arrive at

$$\begin{cases} \dfrac{\partial}{\partial t} n(x,t) = b(x)n(x,t) - \rho(t)n(x,t), \\ \rho(t) = \displaystyle\int_{-\infty}^{+\infty} n(x,t)dx, \\ n(x,0) = n_0(x) \geq 0, \end{cases} \tag{11.3.1}$$

and we assume that

$$n_0(x) > 0 \quad \text{if and only if} \quad x \in (x_m, x_M),$$

for given $x_m, x_M \in \mathbb{R}$. Such an assumption implies that only the range of traits (x_m, x_M) is initially present. We state that such a set of traits is actually *invariant*. More precisely, we prove that $n(x,t) = 0$ for all x which do not belong to the set (x_m, x_M). Let $x \notin [x_m, x_M]$, then we have

$$\frac{\partial}{\partial t} n(x,t) \leq Cn(x,t)$$

because b is bounded on $[x_m, x_M]$. Therefore, by comparison principle (exercise!) we have

$$n(x,t) \leq n_0(x)e^{Ct} = 0$$

because $n_0(x) = 0$ in this case. This proves the assertion.

Before proving our asymptotic result on (11.3.1), let us (partially) justify the use of such an equation in the context of models with a chemostat of the form (11.2.1). As already pointed out about chemotaxis models, it is customary to assume that the chemostat diffuse much faster than the other species. Therefore, the evolution of the chemostat can be supposed to be approximately at equilibrium, i. e. the system (11.2.1) can be replaced by

$$\begin{cases} 0 = R(S_0 - S) - \displaystyle\int_0^{+\infty} \eta(x,S)n(x,t)dx, \quad t \geq 0,\ x \geq 0 \\ \dfrac{\partial}{\partial t} n(x,t) = n(x,t)(\eta(x,S) - R), \\ S(0) = S^0 > 0, \\ n(x,0) = n^0(x) > 0, \qquad\qquad n^0 \in L^1 \cap L^\infty([0,+\infty)). \end{cases} \tag{11.3.2}$$

Let us assume the simple case of a function η linearly depending on S as follows

$$\eta(S,x) = \xi(x)S.$$

We can therefore decouple the above system in the following way

$$\begin{aligned} & S(t) = S_0 - \frac{S(t)}{R} \int_{-\infty}^{+\infty} \xi(x)n(x,t)dx \\ \Rightarrow \quad & S(t) = \frac{S_0}{1 + \dfrac{1}{R}\displaystyle\int_{-\infty}^{+\infty} \xi(x)n(x,t)dx} =: Q\left(\int_{-\infty}^{+\infty} \xi(x)n(x,t)dx\right). \end{aligned}$$

Hence, the system (11.3.2) reduces to

$$\frac{\partial}{\partial t} n(x,t) = n(x,t)\left(\xi(x)Q\left(\int_{-\infty}^{+\infty} \xi(x)n(x,t)dx\right) - R\right).$$

In the case of a constant function ξ (for instance $\xi(x) \equiv 1$) one would recover a nonlocal equation in which the evolution of n depend on the total population, in a similar fashion as (11.3.1), namely

$$\frac{\partial}{\partial t} n(x,t) = n(x,t) \left(\xi(x) Q \left(\rho(t) \right) - R \right), \qquad \rho(t) = \int_{-\infty}^{+\infty} n(x,t) dx.$$

We shall now devote on detecting the asymptotic behavior of (11.3.1) as $t \to +\infty$. In order to prove our result, we shall require the following conditions: b continuous, $b(x) \geq \bar{b} > 0$, $n^0 > 0$, and

$$b(x_0) = \max_{x \in [x_m, x_M]} b(x) \qquad \text{is attained for a single } \bar{x} \in [x_m, x_M]. \tag{11.3.3}$$

Please notice that the condition (11.3.3) depends not only on b but also on the support of the initial condition n^0. We shall prove that only the trait $\bar{x}$ will survive at $t \to +\infty$. More precisely, we shall prove that

$$\rho(t) \to \bar{\rho} := b(\bar{x}), \quad n(x,t) \to b(\bar{x}) \delta_{\bar{x}}(x), \qquad \text{as } \ t \to +\infty. \tag{11.3.4}$$

Notice that the equation ((11.3.1)) admits many steady states, namely $n(x) = b(y)\delta_y$ for any y, therefore the result (11.3.4) really selects the best trait, the ESS. One should understand it as the trait that realizes

$$\max_x [b(x) - \bar{\rho}] = b(\bar{x}) - \bar{\rho} = 0.$$

This result also indicates that the natural setting for structured population models should differ from that for differential equations because functional spaces (measures here) appear to play a role.

We give a proof of (11.3.4) that relies on a simple computation, another proof is possible based on more general arguments. We define

$$N(x,t) = n(x,t) e^{\int_0^t \rho(s) ds}. \tag{11.3.5}$$

This satisfies

$$\frac{dN(x,t)}{dt} = b(x) N(x,t),$$

and thus $N(x,t) = n^0(x) e^{b(x)t}$. We deduce from (11.3.5) that

$$\frac{d}{dt} e^{\int_0^t \rho(s) ds} = \rho(t) e^{\int_0^t \rho(s) ds} = \int_{-\infty}^{+\infty} N(x,t) dx = \int_{-\infty}^{+\infty} n^0(x) e^{b(x)t} dx.$$

Therefore, integrating the above identity along the time interval $[0,t]$, using Fubini's theorem we obtain

$$\begin{aligned}
& e^{\int_0^t \rho(s) ds} = \int_0^t \int_{-\infty}^{+\infty} n^0(x) e^{b(x)s} dx ds = \int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)} e^{b(x)t} dx + K, \\
& K = 1 - \int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)} dx. \\
& \int_0^t \rho(s) ds = \log \left(\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)} e^{b(x)t} dx + K \right), \\
& \rho(t) = \frac{\int_{-\infty}^{+\infty} n^0(x) e^{b(x)t} dx}{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)} e^{b(x)t} dx + K},
\end{aligned}$$

and we notice that the constant K may be negative but the denominator above is larger than 1. This is a Laplace type of formula which we can analyze as follows. We have

$$\rho(t) = \frac{\int_{-\infty}^{+\infty} b(x)\frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx + K} \leq b(\bar{x})\frac{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx + K}$$

and since the integral term above diverges to $+\infty$ as $t \to +\infty$, we have

$$\limsup_{t\to+\infty} \rho(t) \leq b(\bar{x}) \lim_{t\to+\infty} \frac{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx + K} = b(\bar{x}).$$

To prove the reverse inequality, we fix an $\varepsilon > 0$ and define

$$I(\varepsilon) := \left\{x \in \mathbb{R} : \; b(x) \geq b(\bar{x}) - \varepsilon\right\}.$$

Then,

$$\begin{aligned}
\rho(t) &\geq \frac{\int_{I(\varepsilon)} n^0(x)e^{b(x)t}dx}{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx + K} \\
&= \frac{\int_{I(\varepsilon)} b(x)\frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx + K} \\
&\geq (b(\bar{x}) - \varepsilon)\frac{\int_{I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{-\infty}^{+\infty} \frac{n^0(x)}{b(x)}e^{b(x)t}dx + K}.
\end{aligned} \tag{11.3.6}$$

Now, it can be easily seen that

$$\frac{\int_{\mathbb{R}\setminus I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx} \to 0 \quad \text{as} \quad t \to +\infty. \tag{11.3.7}$$

Indeed,

$$\frac{\int_{\mathbb{R}\setminus I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx} = \frac{\int_{\mathbb{R}\setminus I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{(-b(\bar{x})+\varepsilon+b(x))t}dx}{\int_{I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{(-b(\bar{x})+\varepsilon+b(x))t}dx}$$

and the definition of $I(\varepsilon)$ gives (11.3.7). Therefore, we can take the $\liminf_{t\to+\infty}$ in (11.3.6) to obtain

$$\liminf_{t\to+\infty} \rho(t) \geq \liminf_{t\to+\infty}(b(\bar{x}) - \varepsilon)\left(1 + \frac{\int_{\mathbb{R}\setminus I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx}{\int_{I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx} + \frac{K}{\int_{I(\varepsilon)} \frac{n^0(x)}{b(x)}e^{b(x)t}dx}\right)^{-1} = b(\bar{x}) - \varepsilon.$$

Since ε is arbitrary, we obtain

$$\limsup_{t\to+\infty} \rho(t) \leq b(\bar{x}) \leq \liminf_{t\to+\infty} \rho(t)$$

which proves the first formula of (11.3.4). Finally, from (11.3.5) and the expression of $N(x,t)$, we deduce

$$n(x,t) = n^0(x)e^{b(x)t - \int_0^t \rho(s)ds}.$$

Now, let $x \neq \bar{x}$. Then, it is clear that $b(x)t - \int_0^t \rho(s)ds$ tends to $-\infty$ as $t \to +\infty$. Therefore $n(x,t) \to 0$ as $t \to +\infty$. This proves (11.3.4).

Bibliography

[1] T. R. Malthus, N. O'Flaherty, D. Valenze, E. A. Wrigley, K. Binmore, and K. O'Brien. *Front Matter*, pages i–vi. Yale University Press, 2018.

[2] P. Verhulst. Notice sur la loi que la population poursuit dans son accroissement. *Corresp. mathématique et physique*, 10, 1838.