8
Introduction to master equations

In this chapter we will briefly present the main results about master equations. They are differential equations that describe the evolution of the probabilities for Markov processes for systems that jump from one to other state in a continuous time. In this sense they are the continuous time version of the recurrence relations for Markov chains mentioned at the end of chapter 1. We will emphasize their use in the case that the number of available states is discrete, as it corresponds to applications to chemical reactions, radioactive substances, epidemic spreading and many other cases of interest. We give also a brief account of the generating function technique to solve master equations and some approximated methods of solution.

8.1
A two-state system with constant rates

Let us consider now a situation in which something can switch between two states that we name "1" and "2". There are many examples of this situation:
-A radioactive atom which is in state 1 has not yet disintegrated, whereas the state 2 is the atom after the disintegration.
-A person can be healthy (state 1) or sick (state 2).
-In a chemical reaction, say the combination of sodium and chlorine to produce salt, Na+Cl→NaCl, an atom of sodium can be not bound (state 1) or bound (state 2) to a chlorine atom.
-A bilingual person can be speaking one language (state 1) or another (state 2).

In these and many other examples, one can represent the transitions between the two states as random events that happen at some rates. We denote by $\omega(1 \to 2)$ the rate at which one particle[1)] jumps from state 1 to state 2 and we assume that those transitions occur uniformly and randomly at the constant rate $\omega(1 \to 2)$. This means that there is a probability $\omega(1 \to 2)dt$ that the particle jumps from 1 to 2 in the time interval $(t, t + dt)$.

The inverse process, that of switching from 2 to 1, might or might not occur. For

1) For the sake of brevity, we will call "particles" the agents, systems, persons, atoms or whatever that is making the jumps between the states.

example, a person can get a flu (so it goes from $1 \to 2$) but it can usually recover from the flu (so it goes from $2 \to 1$). A radioactive substance, however, can not switch back from the decayed atom to the original atom. In the reversible chemical reactions, the product molecule (NaCl, for instance) can be broken up in the constituent atoms. The bilingual person can be switching many times a day between the using of both languages. When the inverse switching does occur, its rate $\omega(2 \to 1)$ might or might not be related to the rate $\omega(1 \to 2)$. If a particle in state 1 adopts the form X and a particle in state 2 the form Y, we can indicate this process schematically by:

$$X \rightleftharpoons Y. \tag{8.1}$$

We will use sometimes, for brevity, the notation $\omega_{i\to j} \equiv \omega(i \to j)$ to indicate the rate at which one particle goes from state i to state j.

8.1.1 The particle point of view

Starting at some initial time t_0, we ask the probabilities $P_1(t)$ and $P_2(t)$ that one given particle is in state 1 or 2, respectively, at time t. Obviously, they must satisfy $P_1(t) + P_2(t) = 1$. We will derive now a differential equation for $P_1(t)$. We do that by relating the probabilities at two closed times, t and $t + dt$. In doing so, we are implicitly using the Markov assumption as the transition rate from one state to the other during the time interval $(t, t + dt)$ does not depend at all on the previous history, but only on the state at time t. The probability $P_1(t + dt)$ that the particle is in state 1 at time $t + dt$ has two contributions: that of being in 1 at time t and not having jumped to state 2 during the interval $(t, t+dt)$, and that of being at 2 at time t and having made a jump from 2 to 1 in the interval $(t, t + dt)$. Summing up all these cases and using the rules of conditional probability, we have

$$P_1(t + dt) = P_1(t)\text{Prob(staying in 1)} + P_2(t)\text{Prob(jumping from 2 to 1)}. \tag{8.2}$$

By the definition of the rate, the probability of jumping from 2 to 1 in the time interval $(t, t + dt)$ is $\omega(2 \to 1)dt$ whereas the probability of staying in state 1 is one minus the probability of leaving state 1 in the same time interval, or $1-\omega(1 \to 2)dt$. This leads to:

$$P_1(t + dt) = P_1(t)[1 - \omega(1 \to 2)dt] + P_2(t)\omega(2 \to 1)dt + O(dt^2). \tag{8.3}$$

The terms of order $O(dt^2)$ could arise if the particle is in state 1 at time $t+dt$ because it was in state 1 at time t and made two jumps, one from 1 to 2 and another from 2 to 1, during the time interval $(t, t + dt)$. In this way, it could end up again in state 1, so contributing to Prob(staying in 1). This would happen with probability $\omega(1 \to 2)dt \times \omega(2 \to 1)dt = O(dt^2)$. Similarly, there could be higher order contributions from particles jumping from 2 to 1 and then back from 1 to 2. All these multiple events happen with vanishing probability in the limit $dt \to 0$. Rearranging, and

taking the limit $dt \to 0$ we get the differential equation:

$$\frac{dP_1(t)}{dt} = -\omega(1 \to 2)P_1(t) + \omega(2 \to 1)P_2(t). \tag{8.4}$$

A similar reasoning leads to the equivalent equation for $P_2(t)$:

$$\frac{dP_2(t)}{dt} = -\omega(2 \to 1)P_2(t) + \omega(1 \to 2)P_1(t). \tag{8.5}$$

Eqs. (8.4)-(8.5) are a very simple example of **master equations**: equations for the probability that a stochastic particle than can jump between different states is in one of these states at a time t. These first-order differential equations have to be solved under the assumption of some initial conditions at some initial time t_0: $P_1(t_0)$ and $P_2(t_0)$. Note that

$$\frac{d}{dt}[P_1(t) + P_2(t)] = 0, \tag{8.6}$$

implying that $P_1(t) + P_2(t) = 1$ at all times t provided that the initial condition satisfies, as it should, $P_1(t_0) + P_2(t_0) = 1$. One defines the probability current, $J(1 \to 2)$ from state 1 to 2 as:

$$J(1 \to 2) = -\omega(1 \to 2)P_1(t) + \omega(2 \to 1)P_2(t). \tag{8.7}$$

It has a negative contribution coming from those jumps leaving state 1 to go to state 2, and a positive contribution from those jumps from state 2 to state 1. A similar definition leads to $J(2 \to 1) = -J(1 \to 2)$. In terms of these probability currents, the master equations are

$$\frac{dP_1(t)}{dt} = J_1(t), \quad \frac{dP_2(t)}{dt} = J_2(t). \tag{8.8}$$

In the case of constant rates $\omega(1 \to 2)$ and $\omega(2 \to 1)$ that we are considering throughout this section it is possible to find the explicit solution of Eqs. (8.4)-(8.5):

$$\begin{aligned} P_1(t) &= P_1(t_0)\frac{\omega_{2\to1} + \omega_{1\to2}\mathrm{e}^{-(\omega_{2\to1}+\omega_{1\to2})(t-t_0)}}{\omega_{2\to1} + \omega_{1\to2}} \\ &\quad + P_2(t_0)\frac{\omega_{2\to1}(1 - \mathrm{e}^{-(\omega_{2\to1}+\omega_{1\to2})(t-t_0)})}{\omega_{2\to1} + \omega_{1\to2}}, \end{aligned} \tag{8.9}$$

$$\begin{aligned} P_2(t) &= P_1(t_0)\frac{\omega_{1\to2}(1 - \mathrm{e}^{-(\omega_{2\to1}+\omega_{1\to2})(t-t_0)})}{\omega_{2\to1} + \omega_{1\to2}} \\ &\quad + P_2(t_0)\frac{\omega_{1\to2} + \omega_{2\to1}\mathrm{e}^{-(\omega_{2\to1}+\omega_{1\to2})(t-t_0)})}{\omega_{2\to1} + \omega_{1\to2}}. \end{aligned} \tag{8.10}$$

From here, and using $P_1(t_0) + P_2(t_0) = 1$, we can obtain the stationary distribution as the limit $t \to \infty$:

$$P_1^{\mathrm{st}} = \frac{\omega_{2\to1}}{\omega_{2\to1} + \omega_{1\to2}}, \tag{8.11}$$

$$P_2^{\mathrm{st}} = \frac{\omega_{1\to2}}{\omega_{2\to1} + \omega_{1\to2}}, \tag{8.12}$$

a particularly simple solution. Note that in this case the stationary distribution satisfies

$$\omega(1 \to 2)P_1^{\rm st} = \omega(2 \to 1)P_2^{\rm st}, \tag{8.13}$$

showing that in the case of two states the stationary distributions satisfy the *detailed balance condition*, equivalent to (1.146) for Markov chains.

An interesting quantity is the probability $P(i,t|j,t_0)$ that the particle is in state i at time t given that it was in state j at time t_0 for $i,j=1,2$. In general, it is difficult to compute $P(i,t|j,t_0)$ directly from the rules of the process. The reason has already been mentioned before: in a finite interval (t_0,t) there might have been many jumps to intermediate states and all these have to be included in the calculation of the conditional probability. For instance, to compute $P(1,t|1,t_0)$ one has to include the jumps 1→2→1, in which the particle has left state 1 to go to state 2 and returned from it, *any number of times*, from zero to infinity.

Luckily, there is a simple alternative way to proceed in the case of two states and constant rates. We can obtain $P(1,t|1,t_0)$ from $P_1(t)$ setting as a initial condition $P_1(t_0)=1,\ P_2(t_0)=0$, as we know (probability 1) that the particle is in state 1 at time t_0. Taking the explicit solution (8.9) with $P_1(t_0)=1,\ P_2(t_0)=0$, we obtain

$$P(1,t|1,t_0) = \frac{\omega_{2\to 1} + \omega_{1\to 2}{\rm e}^{-(\omega_{2\to 1}+\omega_{1\to 2})(t-t_0)}}{\omega_{2\to 1}+\omega_{1\to 2}}, \tag{8.14}$$

and, of course,

$$P(2,t|1,t_0) = 1 - P(1,t|1,t_0) = \frac{\omega_{1\to 2}\left(1-{\rm e}^{-(\omega_{2\to 1}+\omega_{1\to 2})(t-t_0)}\right)}{\omega_{2\to 1}+\omega_{1\to 2}}, \tag{8.15}$$

with equivalent expressions for $P(1,t|2,t_0)$ and $P(2,t|2,t_0)$:

$$P(1,t|2,t_0) = \frac{\omega_{2\to 1}\left(1-{\rm e}^{-(\omega_{1\to 2}+\omega_{1\to 2})(t-t_0)}\right)}{\omega_{2\to 1}+\omega_{1\to 2}}, \tag{8.16}$$

$$P(2,t|2,t_0) = \frac{\omega_{1\to 2} + \omega_{2\to 1}{\rm e}^{-(\omega_{2\to 1}+\omega_{1\to 2})(t-t_0)}}{\omega_{2\to 1}+\omega_{1\to 2}}. \tag{8.17}$$

In terms of these conditional probabilities, we can reason that the probability that the particle is in state 1 at time t is the probability that it was in state 1 at time t_0 times the probability $P(1,t|1,t_0)$ plus the probability that it was in state 2 at t_0 times the probability $P(1,t|2,t_0)$:

$$P_1(t) = P_1(t_0)P(1,t|1,t_0) + P_2(t_0)P(1,t|2,t_0), \tag{8.18}$$

and similarly:

$$P_2(t) = P_1(t_0)P(2,t|1,t_0) + P_2(t_0)P(2,t|2,t_0). \tag{8.19}$$

In fact, the conditional probabilities $P(i,t|j,t_0)$ of a Markov process can not be arbitrary functions. If we consider an intermediate time $t_0 < t_1 < t$, the probability that the particle is in state 1 at time t provided it was in state 2 at time t_0 can be

computed using the probabilities that the particle was at the intermediate time t_1 in one of the two states. Namely,

$$P(1,t|2,t_0) = P(1,t|1,t_1)P(1,t_1|2,t_0) + P(1,t|2,t_1)P(2,t_1|2,t_0), \tag{8.20}$$

and similar equations for $P(1,t|1,t_0)$, $P(2,t|2,t_0)$, $P(2,t|1,t_0)$. The first term in the right-hand-side is the probability that the particle went from 2 at time t_0 to 1 at time t and passed trough 1 at time t_1 and a similar interpretation for the second term. These relations are called the **Chapman-Kolmogorov equations** for this process. If we knew the conditional probabilities $P(i,t|j,t_0)$ for arbitrary times t, t_0, the solution of these two functional equations (8.18)-(8.19) would allow one to compute $P_1(t)$ and $P_2(t)$. However, (1) it is difficult to obtain directly from the rules of the process the conditional probabilities $P(i,t|j,t_0)$, and (2) it is difficult to solve a functional equation. We have circumvented these problems in our previous treatment by considering a small time increment $t - t_0 = dt$. In this limit of small dt we find two advantages: (1) the conditional probabilities can be obtained from the rates, i.e. $P(1,t+dt|2,t) = \omega(2\to 1)dt + O(dt^2)$, etc. and (2) by expanding $P_1(t+dt) = P_1(t) + \dfrac{dP_1(t)}{dt}dt + O(dt^2)$ we obtain a differential instead of a functional equation.

When we will discuss in the next chapter the use of numerical algorithms to simulate numerically master equations, we will need to use the probability density $f^{\text{1st}}(j,t|i,t_0)$ of the first jump from $i \to j$. This is defined such that $f^{\text{1st}}(j,t|i,t_0)dt$ is the probability that the particle is at state i at time t_0 and stays there until it jumps to state j in the time interval $(t, t+dt)$, **with no intermediate jumps in the whole interval** (t_0, t). Let us now derive directly this function for the particular case of the two-state system. We divide the time interval (t_0, t) in subintervals of length dt: $(t_0, t_0+dt), (t_0+dt, t_0+2dt), \ldots, (t-dt, t)$. The particle is in, say, state $i = 1$ at time t_0 and does not make any jump to state 2 until the time interval $(t, t+dt)$. This means that it does not jump during any of the intervals $(t_0+(k-1)dt, t_0+kdt)$ for $k = 1, \ldots, K = (t-t_0)/dt$. The probability that it does not jump during the interval $(t_0+(k-1)dt, t_0+kdt)$ is $1-\omega(1\to 2)dt$. Therefore, the probability that it does not jump during any of these intervals is the product of all these probabilities, $\prod_{k=1}^{K}(1-\omega(1\to 2)dt) = (1-\omega(1\to 2)dt)^K = (1-\omega(1\to 2)dt)^{(t-t_0)/dt}$. In the limit of $dt \to 0$ this becomes[2] equal to $e^{-\omega(1\to 2)(t-t_0)}$. Finally, we have to multiply by the probability that there is a jump from 1 to 2 during the interval $(t, t+dt)$ which is $\omega(1\to 2)dt$. This yields, finally:

$$f^{\text{1st}}(2,t|1,t_0) = \omega(1\to 2)\mathrm{e}^{-\omega(1\to 2)(t-t_0)}. \tag{8.21}$$

Similarly, we obtain

$$f^{\text{1st}}(1,t|2,t_0) = \omega(2\to 1)\mathrm{e}^{-\omega(2\to 1)(t-t_0)}. \tag{8.22}$$

2) We use $\lim_{x\to 0}(1+x)^{a/x} = e^a$ with $x = -\omega(1\to 2)dt$ and $a = -\omega(1\to 2)(t-t_0)$.

8.1.2 The occupation numbers point of view

We have considered so far a particle that can jump between two states and asked for the probability for that particle to be in state 1 or in state 2. We now consider a system which is composed by N of such particles, each one of them making the random jumps from state 1 to 2 or vice-versa. We introduce the occupation number n of state 1 and ask now for the probability $p(n;t)$ that at a given time t exactly n of the N particles are in that state. If the N particles are independent of each other and they all start in the same state, such that $P_1(t)$ is the same for all particles, then $p(n;t)$ is given by a binomial distribution:

$$p(n;t) = \binom{N}{n} P_1(t)^n (1-P_1(t))^{N-n}. \tag{8.23}$$

We will now find directly a differential equation satisfied by $p(n;t)$. This is not really necessary in this simple case, but the techniques we will develop will be useful in the cases in which the jumps between states do not occur independently for each of the particles, or particles start in different initial conditions.

As before, we will relate the probabilities at times t and $t+dt$. Recall first that the total number of particles is N and if there are n particles in state 1, then the remainder $N-n$ are in state 2. How can we have n particles in state 1 at time $t+dt$? There are three possibilities:
(1) There were n particles in state 1 and $N-n$ in state 2 at time t and no one left the state it was in.
(2) There were $n-1$ particles in state 1 and one of the $N-(n-1)$ particles that were in state 2 jumped from 2 to 1 during that interval.
(3) There were $n+1$ particles in state 1 and one particle jumped from 1 to 2 during that interval.

As we are considering the limit of a small time interval dt, other possible processes (for instance, the jump of two or more particles from one state to another) need not be included as their contribution will be of a higher order in dt. So, we write:

$$\begin{aligned} p(n;t+dt) = {} & p(n;t) \times \text{Prob(no particle jumped)} \\ & + p(n-1;t) \times \text{Prob(any of the } N-n+1 \text{ particles jumps from } 2 \to 1) \\ & + p(n+1;t) \times \text{Prob(any of the } n+1 \text{ particles jumps from } 1 \to 2) \\ & + O(dt^2). \end{aligned} \tag{8.24}$$

Let us analyze each term one by one:
(1) The probability that no particle jumped is the product of probabilities that none of the n particles in state 1 jumped to 2 and none of the $N-n$ particles in state 2 jumped to 1. The probability that one particle does not jump from 1 to 2 is $1-\omega_{1\to 2}dt$, hence, the probability that none of the n particles jumps from 1 to 2 is the product for all particles of this probability, or $[1-\omega_{1\to 2}dt]^n$. Expanding to first order in dt, this is equal to $1-n\omega_{1\to 2}dt+O(dt^2)$. Similarly, the probability that none of

the $N-n$ particles in 2 jumps to 1 is $1-(N-n)\omega_{2\to 1}dt+O(dt^2)$. Finally, the probability that no jump occurs whatsoever is the product of these two quantities, or $1-(n\omega_{1\to 2}+(N-n)\omega_{2\to 1})dt+O(dt^2)$.
(2) The probability that one particle jumps from 2 to 1 is $\omega_{2\to 1}dt$, hence the probability that any of the $N-n+1$ particles jumps from $2\to 1$ is the sum of all these probabilities, or $(N-n+1)\omega_{2\to 1}dt$.
(3) The probability that one particle jumps from 1 to 2 is $\omega_{1\to 2}dt$, hence the probability that any of the $n+1$ particles jumps from $1\to 2$ is the sum of all these probabilities, or $(n+1)\omega_{1\to 2}dt$.

Again, in all these expressions could be higher-order terms corresponding to multiple, intermediate, jumps. Putting all these terms together, we obtain:

$$\begin{aligned} p(n;t+dt) &= p(n;t)(1-(n\omega_{1\to 2}+(N-n)\omega_{2\to 1})dt) \\ &\quad +p(n-1;t)(N-n+1)\omega_{2\to 1}dt \\ &\quad +p(n+1;t)(n+1)\omega_{1\to 2}dt+O(dt^2). \end{aligned} \tag{8.25}$$

Rearranging and taking the limit $dt\to 0$ we obtain the differential equations:

$$\begin{aligned} \frac{\partial p(n;t)}{\partial t} &= -(n\omega_{1\to 2}+(N-n)\omega_{2\to 1})p(n;t) \\ &\quad +(N-n+1)\omega_{2\to 1}p(n-1;t)+(n+1)\omega_{1\to 2}p(n+1;t), \end{aligned} \tag{8.26}$$

This set of $N+1$ equations for the functions $p(n;t)$, valid for $n=0,1,\ldots,N$, constitute the master equation from the occupation number point of view, complementary to the one-particle point of view considered in the previous subsection.

As this set of $N+1$ coupled differential equations is certainly much more complicated than (8.4)-(8.5) and given that we can reconstruct $p(n;t)$ from $P_1(t)$ using (8.23), one can ask which is the usefulness of such approach. The answer is that (8.23) is only valid in the case of independent particles, an assumption which is not always correct. For example, it is true in the case of the radioactive atoms in which each atom decays independently of others. But it is not correct in the vast majority of cases of interest. Consider the example in which state 1 is "healthy" and state 2 is "sick" for a contagious disease. The rate $\omega_{1\to 2}$ at which a healthy person gets infected (the rate at which he goes from 1 to 2) depends naturally on the number $N-n$ of sick persons, as the more infected people there is around one person, the higher the probability that this person gets the disease, whereas we might consider that the recovery rate $\omega_{2\to 1}$ is independent on how many people are infected. In a chemical reaction, the rate at which Cl and Na react to form NaCl depends on how many free (not combined) atoms exist, and so on. In all these cases, it is not possible to write down closed equations for the probability that one particle is in state 1 independently on how many particles share this state. When this happens, master equations of the form of (8.26) are the necessary starting point.

The general structure of the master equation obtained here is not so different of the one derived in the previous subsection taking the particle point of view. We can introduce global rates to jump between global states with n particles in state 1. In

this case we define the rates:

$$\Omega(n \to n+1) = (N-n)\omega_{2\to 1} \tag{8.27}$$

$$\Omega(n \to n-1) = n\omega_{1\to 2} \tag{8.28}$$

and write the master equation as:

$$\begin{aligned}\frac{\partial p(n;t)}{\partial t} = & -(\Omega(n\to n-1)+\Omega(n\to n+1))p(n;t) \\ & +\Omega(n-1\to n)p(n-1;t)+\Omega(n+1\to n)p(n+1;t),\end{aligned} \tag{8.29}$$

the same kind of balance relation that we can find in (8.4-8.5). The first term in the right-hand-side is the loss of probability from n to either $n-1$ or $n+1$, whereas the second and third represent the gain of probability from states with $n-1$ or $n+1$ particles, respectively.

Before we generalize these results to particles that can be in more than two states, let us introduce a simplifying notation. We introduce the so-called "step" operator E. This is a linear operator defined for any function $f(n)$ which depends on an integer variable n and it is defined as

$$E[f(n)] = f(n+1), \tag{8.30}$$

which allows us to define powers of this operator as

$$E^{\ell}[f(n)] = f(n+\ell), \tag{8.31}$$

for $\ell \in \mathbb{Z}$ an integer number, in particular $E^{-1}[f(n)] = f(n-1)$. Using this notation, (8.26) can be written in the compact form

$$\frac{\partial p(n;t)}{dt} = (E-1)[\Omega(n\to n-1)p(n;t)] + (E^{-1}-1)[\Omega(n\to n+1)p(n;t)]. \tag{8.32}$$

Note that the first term, the one with the operator $(E-1)$, represents the loss of probability of processes in which particles leave state 1 to go to state 2 and hence, decrease the population of state 1. On the other hand, the term including the operator $(E^{-1}-1)$ represents those processes that increase the population of 1 by including the contribution of the jumps from 2 to 1. This will be a common feature of more general master equations.

8.2 The general case

We now briefly introduce the notation and equations in the case that there are more than two states, as the generalization is straightforward. If one particle can be in many, but numerable, discrete states $i = \ldots, -2, -1, 0, 1, 2, \ldots$ we denote by $P_i(t)$ the probability that it is in state i at time t. From this state it can jump to state j with

rate $\omega(i \to j)$. As it can jump, in principle, to any other state, the probability that it leaves state i in the time interval $(t, t+dt)$ is $\sum_{j\neq i}\omega(j\to i)dt$ and the probability that it remains in state i during the same time interval is $1-\sum_{j\neq i}\omega(j\to i)dt$. At the same time, the probability that there is a jump from any other state j into state i at this time interval is $\sum_{j\neq i}\omega(j\to i)dt$. So we can write for the probability of being in state i at time $t+dt$:

$$P_i(t+dt) = P_i(t)[1-\sum_{j\neq i}\omega(i\to j)dt] + \sum_{j\neq i}P_j(t)\omega(j\to i)dt + O(dt^2). \quad (8.33)$$

Again, after rearranging and taking the limit $dt\to 0$ we obtain the **master equation** for a discrete process:

$$\frac{dP_i(t)}{dt} = \sum_{j\neq i}\left[-\omega(i\to j)P_i(t) + \omega(j\to i)P_j(t)\right], \quad (8.34)$$

or, in terms of the currents $J(i\to j) = -\omega(i\to j)P_i(t) + \omega(j\to i)P_j(t)$:

$$\frac{dP_i(t)}{dt} = \sum_{j\neq i}J(i\to j). \quad (8.35)$$

Although it is not very common in practice, nothing prevents us from considering the more general case where the transition rates depend on time. Hence, a more general master equation is:

$$\frac{dP_i(t)}{dt} = \sum_{j\neq i}\left[-\omega(i\to j;t)P_i(t) + \omega(j\to i;t)P_j(t)\right]. \quad (8.36)$$

To find the solution of this set of equations we need to specify an initial condition $P_i(t=t_0), \forall i$.

A natural extension is to consider that the states i in which a particle can be form a continuum, instead of a discrete set. We can think, for instance, in the position of a Brownian particle than can jump randomly from point x to point y. In this case, we need to define $f(x;t)$ as the probability density of being in location x at time t. In other words, $f(x;t)dx$ is the probability that one particle is found in the interval $(x, x+dx)$. Similarly, we need to define the transition rate $w(y\to x)$ from y to x such that $w(y\to x)dxdt$ is the probability of jumping to the interval $(x, x+dx)$ during the time interval $(t, t+dt)$ given that it was at the point y at time t. With these definitions, it is not difficult to generalize the master equation as:

$$\frac{\partial f(x;t)}{\partial t} = \int dy\ \left[w(y\to x)f(y;t) - w(x\to y)f(x;t)\right], \quad (8.37)$$

which was already mentioned at the end of chapter 1.

Let us go back to the discrete case. We stress again that the transition rates $\omega_{i\to j}\equiv\omega(i\to j)$ do not need to satisfy any relation amongst them[3)]. Remember also that

3) The elements $\omega(i\to i)$ are not defined and one usually takes $\omega(i\to i)=0$ although their precise value is irrelevant in the majority of formulas. Note also that the sum in (8.36) can be replaced by $\sum_{\forall j}$ since the term $j=i$ does not contribute to this sum whatever the value of $\omega_{i\to i}$.

$\omega(i \to j)$ are rates, not probabilities, and besides having units of [time]$^{-1}$ do not need to be bounded to the interval $[0, 1]$ (although they are non-negative quantities). It is easy now to verify that whatever the coefficients $\omega(i \to j)$ it follows from (8.36) the conservation of the total probability,

$$\frac{d}{dt}\sum_i P_i(t) = 0, \tag{8.38}$$

and, again, we have the normalization condition $\sum_i P_i(t) = 1$ for all times t provided that $\sum_i P_i(t_0) = 1$.

When the total number of states N is finite, it is possible, and useful sometimes, to define the matrix $\mathbf{W}$ as

$$\begin{array}{lcll} W_{ij} & = & \omega(j \to i) & \text{if } i \neq j \\ W_{ii} & = & -\sum_{j\neq i}\omega(i \to j) & \end{array} \tag{8.39}$$

$W_i \equiv |W_{ii}|$ is nothing but the total scape rate from the state i as the sum of all rates to all possible states. In terms of the coefficients W_{ij} the rate equations admit the matrix form:

$$\frac{dP_i(t)}{dt} = \sum_j W_{ij}P_j(t) \tag{8.40}$$

The matrix W is such that the rows add to zero. This property ensures that the solutions $P_i(t)$ respect the positivity condition $P_i(t) \geq 0$ provided that $P_i(t_0) \geq 0$.

We now determine the pdf $f^{\text{1st}}(j, t|i, t_0)$ of the first jump from $i \to j$. The definition is the same as before: $f^{\text{1st}}(j, t|i, t_0)dt$ is the probability that the particle is at state i at time t_0 and stays there until it jumps to state j in the time interval $(t, t + dt)$, **with no intermediate jumps to any other state in the interval** (t_0, t). We need not to repeat the proof we gave in the case of two states. Since the probability that the system jumps from i to any other state in the time interval $(t, t + dt)$ is $W_i dt$, the probability that there have been no jumps in the interval (t_0, t) is $\mathrm{e}^{-W_i(t-t_0)}$. As the probability that there is a jump from i to j in the time interval $(t, t + dt)$ is $\omega(i \to j)dt$, the required probability density function is:

$$f^{\text{1st}}(j, t|i, t_0) = \omega(i \to j)\mathrm{e}^{-W_i(t-t_0)} \tag{8.41}$$

This will be useful when discussing the numerical methods for simulating a master equation.

We can adopt also the occupation numbers point of view and ask for the probability $p(n_1, n_2, \ldots; t)$ that there are n_1 particles in state 1, n_2 in state 2, and so on. Instead of writing general formulas, we will now consider some specific examples.

8.3 Examples

Radioactive decay

Let us take as an example a β-radioactive substance. The events are the emission of electrons by an atom at an individual rate ω. Schematically:

$$\mathrm{X} \longrightarrow \mathrm{Y} \tag{8.42}$$

where X (state 1) denotes a radioactive atom and Y (state 2) the product of the disintegration. In fact, this example is nothing but a simplification of the two states case considered in section 8.1 as the reverse transition does not occur. All we need to do then is to take $\omega(1 \to 2) = \omega$ and $\omega(2 \to 1) = 0$. The initial condition at $t_0 = 0$ is that the atom is in the 1 state, or $P_1(t_0) = 1$. Then, the probability that this atom has not yet disintegrated at time t is

$$P_1(t) = e^{-\omega t}. \tag{8.43}$$

If at time $t = 0$ there are N atoms which have not yet disintegrated and, as atoms disintegrate independently of each other, the probability that at time t there are n atoms not yet disintegrated is given by the binomial distribution (8.23), or

$$p(n;t) = \binom{N}{n} \mathrm{e}^{-n\omega t}(1 - \mathrm{e}^{-\omega t})^{N-n}. \tag{8.44}$$

The average value of the binomial distribution is:

$$\langle n(t) \rangle = \sum_n np(n;t) = N\mathrm{e}^{-\omega t}, \tag{8.45}$$

the law of radioactive decay.

The probability $p(n;t)$ satisfies the master equation (8.32) with $\Omega(n \to n+1) = 0$, as it follows from $\omega_{2\to 1} = 0$ consistent with the intuitive condition that there is no mechanism by which the number of non-disintegrated atoms can be increased:

$$\frac{\partial p(n;t)}{\partial t} = (E-1)\left[\Omega(n \to n-1)p(n;t)\right], \tag{8.46}$$

and it is a simple exercise to check that (8.44) indeed satisfies this equation.

Birth (from a reservoir) and death process

An important class of master equations respond to the *birth and death* scheme. In one of its simplest forms, we assume that particles are created at a constant rate ω_A out of N_A sources. Once created these particles can disappear at another constant rate ω. This is schematized as:

$$A \longrightarrow X \longrightarrow \emptyset \tag{8.47}$$

The sources A are assumed to be endless, so that the rate of production of the X particles is always constant, independently on how many particles have been already created. The set of N_A particles are the "reservoir" out of which the X-particles are created.

We take the occupation numbers point of view and focus on the probability $p(n;t)$ that there are n X-particles at time t. We have now three elementary contributions to $P(n;t+dt)$ according to what happened in the time interval $(t,t+dt)$:
(1) There were n X-particles at time t and none was lost and none was created from the reservoir. The probability that one source A does not create a particle in the interval $(t,t+dt)$ is $1-\omega_A dt$ and the probability that none of the N_A sources creates a particle is $(1-\omega_A dt)^{N_A} = 1-\Omega_A dt + O(dt^2)$, with $\Omega_A \equiv N_A\omega_A$. The probability that one of the X-particles does not disappear is $1-\omega dt$ and the probability that none of the n X-particles disappears is $(1-\omega dt)^n = 1-n\omega dt + O(dt^2)$. Hence, the probability that nothing occurs in the interval $(t,t+dt)$ is the product of these two probabilities.
(2) There were $n+1$ X-particles at time t and one particle disappeared. The probability of one particle disappearing is ωdt, and the probability that any of the $n+1$ particles disappears is the sum of these probabilities, or $(n+1)\omega dt$.
(3) There were $n-1$ X-particles and one was created from the reservoir. As each one of the N_A sources has a probability $\omega_A dt$ of creating a particle, the total probability for this event is $N_A\omega_A dt = \Omega_A dt$.

Combining the probabilities of these events we get :

$$\begin{aligned} p(n;t+dt) &= p(n;t)[1-\omega n dt][1-\Omega_A dt] \\ &+ p(n+1;t)(n+1)\omega dt \\ &+ p(n-1;t)\Omega_A dt + O(dt^2). \end{aligned} \tag{8.48}$$

Rearranging, and taking the limit $dt \to 0$ we get the master equation:

$$\frac{\partial p(n;t)}{\partial t} = -(n\omega+\Omega_A)p(n;t) + (n+1)\omega p(n+1;t) + \Omega_A p(n-1;t), \tag{8.49}$$

or in terms of the step operator:

$$\frac{\partial p(n;t)}{\partial t} = (E-1)[\Omega(n\to n-1)p(n;t)] + (E^{-1}-1)[\Omega(n\to n+1)p(n;t)]. \tag{8.50}$$

with

$$\Omega(n\to n-1) = n\,\omega, \tag{8.51}$$

$$\Omega(n\to n+1) = \Omega_A = N_A\omega_A. \tag{8.52}$$

Again, the term $E-1$ corresponds to the destruction of X-particles, whereas the term $E^{-1}-1$ corresponds to their creation. This time, however, there is no *a priori* upper limit for the value of the number of X-particles and therefore $n = 0,1,\ldots,\infty$. The master equation consists in infinite coupled equations which have to be solved using the initial conditions $P(n;0) = \delta_{n,N_0}$, $n = 0,1,\ldots$, being N_0 the number of X-particles present at the initial time $t=0$. We will find its solution in a later section.

A chemical reaction

We consider the simple chemical reaction in which an atom A and an atom B combine to give the molecule AB. The reaction is reversible, so the molecule AB can break up in the constituent atoms:

$$A + B \rightleftarrows AB. \tag{8.53}$$

For simplicity, let us consider a situation in which initially there are the same number N of A-atoms and B-atoms and no AB molecules. An A-atom can be not bound (state 1) or bound to B to form AB (state 2). We denote by $n(t)$ the number of A-atoms in state 1 at time t. As one atom A combines with an atom B, the number of B-atoms at time t is also $n(t)$, and the number of AB-molecules is $N - n(t)$. The combination of an A and a B-atom to form a molecule is a complicated process for which we adopt a probabilistic point of view. We assume that the rate at which an individual A-atom combines with a B-atom to form the product AB is $\omega_{1\to 2}$, while the reverse reaction $AB \to A + B$ happens at a rate $\omega_{2\to 1}$.

We could focus on the particle point of view and write down equations for the probability than one A-atom is in state 1 (unbound) or in state 2 (bound). These equations for $P_1(n;t)$ and $P_2(n;t)$ would look similar to (8.4)-(8.5). A new ingredient appears in this case. It is not reasonable to assume that the rate $\omega_{1\to 2}$ is a constant independent on how many B-atoms there are. For A to react with B, they first have to meet. We can imagine that the reaction takes place in a container of volume V. If the atoms move freely throughout the whole volume we can assume that the reaction rate $\omega_{1\to 2}$ is proportional to the particle density n/V of B-atoms. We insist that this is a assumption we take to describe the chemical reaction. This assumption will in general not be fully correct as it assumes that the density of B-atoms at the neighborhood of an A-atom is homogeneous. If we take this assumption, though, we are led to taking the dependence $\omega_{1\to 2}[n] = k_{12} n V^{-1}$, being k_{12} this time a constant, independent of n and V. On the other hand, it seems reasonable to assume that the breaking up of an AB-molecule into its constituent atoms, being an event involving only one molecule, does not depend on the density of atoms or molecules, so the rate $\omega_{2\to 1}$ is independent of n.

To cope with this difficulty, we adopt the occupation numbers point of view and consider the individual events that increase of decrease the number n of A-atoms in order to analyze the change in probability during the time interval $(t, t + dt)$. The terms that contribute to $p(n; t + dt)$, the probability that at time $t + dt$ there are n A-atoms are as follows:
(1) There are n A-atoms at time t and no one of them reacts with a B-atom and no one of the $N-n$ molecules breaks up. Probability: $(1-\omega_{1\to 2}[n]dt)^n (1-\omega_{2\to 1}dt)^{N-n} = 1 - (n\omega_{1\to 2}[n] + (N-n)\omega_{2\to 1})dt + O(dt^2)$.
(2) There are $n + 1$ A-atoms at time t and one of them reacts with any of the $n + 1$ B-atoms. The probability that one A-atom reacts is $\omega_{1\to 2}[n + 1]dt$. The probability that any of the $(n + 1)$ A-atoms reacts is $(n + 1)\omega_{1\to 2}[n + 1]dt$.
(3) There are $n-1$ A-atoms at time t and one of the $N-(n-1)$ AB-molecules breaks

up in its constituent atoms. This event happens with probability $(N-n+1)\omega_{2\to 1}dt$. Combining all these event, we get:

$$\begin{aligned} p(n;t+dt) &= p(n;t)\times(1-(n\omega_{1\to 2}[n]+(N-n)\omega_{2\to 1})dt) \\ &\quad +p(n+1;t)\times(n+1)\omega_{1\to 2}[n+1]dt \\ &\quad +p(n-1;t)\times(N-n+1)\omega_{2\to 1}dt. \end{aligned} \tag{8.54}$$

Rearranging, taking the limit $dt\to 0$ and introducing the step operators, we arrive at:

$$\frac{\partial p(n;t)}{dt}=(E-1)[\Omega(n\to n-1)p(n;t)]+(E^{-1}-1)[\Omega(n\to n+1)p(n;t)]. \tag{8.55}$$

with

$$\Omega(n\to n-1) = n\omega_{1\to 2}[n]=k_{12}V^{-1}n^2, \tag{8.56}$$

$$\Omega(n\to n+1) = (N-n)\omega_{2\to 1}. \tag{8.57}$$

The fact that the rate $\Omega(n\to n-1)\propto n^2$ is an example of the law of mass action. The process $n\to n-1$ requires that an A-atom meets a B-atom, an event that is postulated to happen with a probability proportional to the product n_An_B of the number n_A of A-atoms and n_B of B-atoms, which (in our simplified treatment) are exactly the same $n_A=n_B=n$. More generally, if we consider the chemical reaction

$$aA+bB \rightleftharpoons cC+dD, \tag{8.58}$$

where a A-molecules and b B-molecules react to form c C-molecules and d D-molecules, the global rates are assumed to be:

$$\Omega(n_A\to n_A-a, n_B\to n_B-b, n_C\to n_C+c, n_D\to n_D+d) = kn_A^a n_B^b, \tag{8.59}$$

$$\Omega(n_A\to n_A+a, n_B\to n_B+b, n_C\to n_C-c, n_D\to n_D-d) = k'n_C^c n_D^d, \tag{8.60}$$

with k and k' constants.

Self-annihilation

In this case, we consider that the X-particles are created out of an endless reservoir at a rate ω_A, but disappear in pairs. Schematically,

$$\begin{aligned} A &\longrightarrow X, \\ X+X &\longrightarrow \emptyset. \end{aligned} \tag{8.61}$$

We denote by ω the individual rate at which one particle encounters another and leads to the annihilation of both. This example combines ingredients from the two previous examples. While the creation part from the reservoir is identical to the one analyzed in 8.3, the same arguments used in the previous chemical reaction case

lead us to assume that the individual reaction rate depends on the density of available particles with which a given particle can interact with, or $\omega[n] = kV^{-1}(n-1)$, for a system containing n particles.

To find the master equation for $p(n;t)$, the probability of having n X-particles at time t, we focus on the elementary processes that can occur between t and $t+dt$:
(1) There were n X-particles at time t and nothing happened during the time interval $(t, t+dt)$. This includes that no particle was created, probability $1-\Omega_A dt + O(dt^2)$ and no X-particles were annihilated, with probability $1 - n\omega[n]dt + O(dt^2)$, for a final probability $1-(n\omega[n]+\Omega_A)dt + O(dt^2)$.
(2) There were $n+2$ particles at time t. Two particles were annihilated with probability $(n+2)\omega[n+2]dt$, corresponding to the product of the number of particles $n+2$ with the rate at which any of then can be annihilated, $\omega[n+2]$.
(3) There were $n-1$ X-particles and one was created from the reservoir with probability $\Omega_A dt$.
Combining all these terms, we can write:

$$\begin{array}{rcll} p(n;t+dt) & = & p(n;t)[1-(n\omega[n]+\Omega_A)dt] & \text{case (i)} \\ & + & p(n+2;t)(n+2)\omega[n+2]dt & \text{case (ii)} \\ & + & p(n-1;t)\Omega_A dt + O(dt^2) & \text{case (iii)} \end{array} \quad (8.62)$$

Rearranging, taking the limit $dt \to 0$ and introducing the steps operators, it can be written as:

$$\frac{\partial p(n;t)}{\partial t} = (E^2-1)[\Omega(n\to n-2)p(n;t)] + (E^{-1}-1)[\Omega(n\to n+1)p(n;t)]. \quad (8.63)$$

with

$$\Omega(n\to n-2) = n\omega[n] = kV^{-1}n(n-1), \quad (8.64)$$

$$\Omega(n\to n+1) = \Omega_A. \quad (8.65)$$

Where the term E^2-1 represents the annihilation of the two particles, and the $E^{-1}-1$ the creation of one particle.

If we now added to the annihilation of particles the death of a single particle at a rate ω', i.e. the scheme:

$$\begin{array}{c} A \longrightarrow X, \\ X + X \longrightarrow \emptyset, \\ X \longrightarrow \emptyset, \end{array} \quad (8.66)$$

we could do all the detailed algebra again, but the result

$$\begin{aligned} \frac{\partial p(n;t)}{\partial t} =& (E^2-1)[\Omega(n\to n-2)p(n;t)] + (E^{-1}-1)[\Omega(n\to n+1)p(n;t)] \\ & + (E-1)[\Omega(n\to n-1)p(n;t)]. \end{aligned} \quad (8.67)$$

with $\Omega(n\to n-1) = n\omega'$ could have been guessed given the interpretation of the terms $E^\ell - 1$. The reader can fill in the gaps of the proof if he finds it necessary.

The prey-predator Lotka-Volterra model

Let us see now an example of the use of a master equations in the context of population dynamics. We consider the so-called predator-prey Lotka-Volterra model. In this model a predator (e.g. a fox) survives and can reproduce thanks to the eating of a prey (e.g. a rabbit) which, in turn, survives and reproduces by eating a natural resource (e.g. grass). There are many simplifications assumed in this model and it can only considered to be a sketch of the real dynamics. But this simple modeling is very popular as it can grasp the essential (but not all) details of the process.

So, we consider an animal species X (the prey) which reproduces by eating an unlimited natural resource, A. The schematic reaction is as follows[4)]:

$$A + X \longrightarrow 2X, \tag{8.68}$$

with some rate ω_0. This means that there is a probability $\omega_0 dt$ that a rabbit gives rise to another rabbit at the time interval $(t, t+dt)$. As we have assumed that the resources (grass) are unlimited, the rate ω_0 is considered to be a constant. The population of rabbits at time t is $n_1(t)$. At the same time, the species Y (the predator, the foxes) reproduces by eating species X. Again schematically:

$$X + Y \longrightarrow 2Y, \tag{8.69}$$

with a individual rate ω_1. This means that there is a probability $\omega_1 dt$ that a fox eats a rabbit and reproduces during the time interval $(t, t + dt)$. As this bears similarities with previous examples, it seems reasonable to assume that this individual rate depends on the density of rabbits present at time t, so we take the dependence[5)] $\omega_1[n_1] = k_1 V^{-1} n_1$. Finally, the species Y can die of natural causes at a rate ω_2:

$$Y \longrightarrow \emptyset. \tag{8.70}$$

It is possible to add other processes such as the spontaneous death of the prey X, but let us not complicate the model and study the consequences of these simple steps.

We denote by $p(n_1, n_2; t)$ the probability that there are n_1 animals of species X and n_2 animals of species Y at time t. The master equation can be obtained by enumerating the elementary processes occurring in the time interval $(t, t + dt)$ that might contribute to $p(n_1, n_2; t + dt)$ namely:
(i) The population was (n_1, n_2) at time t and no rabbit reproduced and no rabbit was eaten and no fox died.
(ii) The population was $(n_1 - 1, n_2)$ at time t and a rabbit reproduced.
(iii) The population was $(n_1, n_2 + 1)$ at time t and a fox died.
(iv) The population was $(n_1 + 1, n_2 - 1)$ at time t and a fox ate a rabbit and reproduced.

4) Note that this assumes some sort of "asexual" reproduction as it is not necessary that two rabbits meet in order to have offspring.

5) As foxes and rabbits live in a two-dimensional space, V has to be considered as a measure of the area, rather than the volume, where they live.

The contributions to the probability are, respectively:

$$\begin{aligned} p(n_1,n_2;t+dt) =& p(n_1,n_2;t)[1-n_1\omega_0 dt][1-n_2\omega_1[n_1]dt][1-n_2\omega_2 dt] \\ &+ p(n_1-1,n_2;t)(n_1-1)\omega_0 dt \\ &+ p(n_1,n_2+1;t)(n_2+1)\omega_2 dt \\ &+ p(n_1+1,n_2-1;t)(n_2-1)\omega_1[n_1+1]dt + O(dt^2) \end{aligned} \tag{8.71}$$

Rearranging and taking the limit $dt \to 0$ we obtain the desired master equation:

$$\begin{aligned} \frac{\partial p(n_1,n_2;t)}{\partial t} =& -(n_1\omega_0 + n_2\omega_1[n_1] + n_2\omega_2)p(n_1,n_2;t) \\ &+(n_1-1)\omega_0 p(n_1-1,n_2;t) \\ &+(n_2+1)\omega_2 p(n_1,n_2+1;t) \\ &+(n_2-1)\omega_1[n_1+1]p(n_1+1,n_2-1;t). \end{aligned} \tag{8.72}$$

It can also be written using the step operators E_1 and E_2 acting on variables n_1 and n_2 respectively:

$$\begin{aligned} \frac{\partial p(n_1,n_2;t)}{\partial t} =& (E_1^{-1}-1)[\Omega\left((n_1,n_2)\to(n_1+1,n_2)\right)p(n_1,n_2;t) \\ &+ (E_2-1)[\Omega\left((n_1,n_2)\to(n_1,n_2-1)\right)p(n_1,n_2;t), \\ &+ (E_1E_2^{-1}-1)[\Omega\left((n_1,n_2)\to(n_1-1,n_2+1)\right)p(n_1,n_2;t)], \end{aligned} \tag{8.73}$$

with:

$$\Omega\left((n_1,n_2)\to(n_1+1,n_2)\right) = n_1\omega_0, \tag{8.74}$$

$$\Omega\left((n_1,n_2)\to(n_1,n_2-1)\right) = n_2\omega_2, \tag{8.75}$$

$$\Omega\left((n_1,n_2)\to(n_1-1,n_2+1)\right) = n_2\omega_1[n_1] = k_1V^{-1}n_1n_2. \tag{8.76}$$

The term $(E_1^{-1}-1)$ represents the creation of an X-particle, (E_2-1) the annihilation of an Y-particle, and $(E_1E_2^{-1}-1)$ the simultaneous annihilation of an X-particle and the creation of an Y-particle.

It should be clear by now which is the general structure of a master equation and, after some practice, the reader should be able to write down the final expression in terms of the step operators without the need to go through all the detailed steps of the proof.

Now that we have learnt how to derive master equations, we have to solve them. This is not an easy task and that is the main reason why numerical methods to simulate a master equation are so widespread. We will review them in the next chapter. Before, however, we will see a powerful technique and some approximated methods of solution.

8.4 The generating function method for solving master equations

We now introduce an analytical method to solve master equations. We will start by the example of the birth and death process and consider the set of equations (8.32).

We define the *generating function* $G(s,t)$ by means of:

$$G(s,t) = \sum_{n=-\infty}^{\infty} s^n p(n;t) \tag{8.77}$$

Note that the sum runs for all values of n. This is a technical point which is not always necessary, but simplifies the derivation of the solution. We note that, although equations (8.32) have been derived for a situation in which the variable n can only take values between 0 and N, we can consider them valid for all integer values of n. All we need to do is to set the initial condition such that $p(n;0) = 0$ for $n \notin [0,N]$. One can check that $\dfrac{\partial p(n;t)}{\partial t} = 0$ for $n \notin [0,N]$, and hence it is $p(n;t) = 0$ for $n \notin [0,N]$.

If we know the generating function, we can expand it in series and, using (8.77) identify the coefficients of the series expansion with the probabilities $p(n;t)$. Note the property:

$$G(1,t) = \sum_{n=-\infty}^{\infty} p(n;t) = 1, \tag{8.78}$$

coming from the normalization condition and valid at all times t. Moreover, the knowledge of the generating function allows the easy determination of the moments of the random variable n. For the first moment, we use the trick $n = \left.\frac{\partial s^n}{\partial s}\right|_{s=1}$:

$$\begin{aligned} \langle n(t)\rangle &= \sum_n np(n;t) = \sum_n \left.\frac{\partial s^n}{\partial s}\right|_{s=1} p(n;t) = \left.\frac{\partial \sum_n s^n p(n;t)}{\partial s}\right|_{s=1} \\ &= \left.\frac{\partial G(s,t)}{\partial s}\right|_{s=1}. \end{aligned} \tag{8.79}$$

For the second moment, we use $n^2 = \left.\frac{\partial}{\partial s}\left(s\frac{\partial s^n}{\partial s}\right)\right|_{s=1}$ to obtain:

$$\langle n^2(t)\rangle = \sum_n n^2 p(n;t) = \left.\frac{\partial}{\partial s}\left(s\frac{\partial G(s,t)}{\partial s}\right)\right|_{s=1}. \tag{8.80}$$

We next find a differential equation for $G(s,t)$. We begin by taking the time derivative of (8.78) and replacing (8.32) and the rates (8.27)-(8.28):

$$\begin{aligned} \frac{\partial G(s,t)}{\partial t} &= \sum_{n=-\infty}^{\infty} s^n \frac{\partial p(n;t)}{\partial t} \\ &= \sum_{n=-\infty}^{\infty} s^n \left[(E-1)[n\omega_{1\to 2}p(n;t)] + (E^{-1}-1)[(N-n)\omega_{2\to 1}p(n;t)]\right]. \end{aligned} \tag{8.81}$$

Now we use the following result, valid for $k \in \mathbb{Z}$:

$$\begin{aligned}\sum_n s^n (E^k - 1)[f(n)] &= \sum_n s^n (f(n+k) - f(n)) \\ &= \sum_n s^n f(n+k) - \sum_n s^n f(n) \\ &= s^{-k} \sum_n s^{n+k} f(n+k) - \sum_n s^n f(n) \\ &= (s^{-k} - 1) \sum_n s^n f(n), \end{aligned} \tag{8.82}$$

where in the third line we have use the change of variables $n + k \to n$.

We also use the result:

$$\sum_n s^n n p(n;t) = \sum_n s \frac{\partial s^n}{\partial s} p(n;t) = s \frac{\partial \left(\sum_n s^n p(n;t)\right)}{\partial s} = s \frac{\partial G(s,t)}{\partial s}. \tag{8.83}$$

Substitution of (8.82)-(8.83) in (8.81), leads after some simple algebra to

$$\frac{\partial G(s,t)}{\partial t} = (1-s) \left[(\omega_{1\to 2} + \omega_{2\to 1} s) \frac{\partial G(s,t)}{\partial s} - \omega_{2\to 1} N G(s,t) \right]. \tag{8.84}$$

This is a partial differential equation for the generating function $G(s,t)$. We have hence reduced the problem of solving a set of $N+1$ ordinary differential equations (8.81) to that of solving one partial differential equation subject to the initial condition $G(s, t=0) = G_0(s) \equiv \sum_n p(n;0)$. The solution of (8.84) can be found by the method of the characteristics and the reader is addressed in this point to the wide bibliography in this vast area (see exercise 4).

We take a simpler and limited approach here. Imagine we want to study just the stationary distribution in which the probabilities $p_{\rm st}(n)$ are no longer a function of time and the generating function is $G_{\rm st}(s) = \sum_n s^n p_{\rm st}(n) = \lim_{t\to\infty} G(s,t)$. This function, being independent of time, satisfies (8.84) with the time derivative equal to zero or, after simplifying the $(1-s)$ factor:

$$0 = (\omega_{1\to 2} + \omega_{2\to 1} s) \frac{dG_{\rm st}(s)}{ds} - \omega_{2\to 1} N G_{\rm st}(s). \tag{8.85}$$

This is an ordinary differential equation whose solution under the initial condition following from (8.78), $G_{\rm st}(1) = 1$ is

$$G_{\rm st}(s) = \left[\frac{\omega_{1\to 2} + \omega_{2\to 1} s}{\omega_{1\to 2} + \omega_{2\to 1}} \right]^N. \tag{8.86}$$

All that remains is to expand this in powers of s using Newton's binomial theorem:

$$G_{\rm st}(s) = \sum_{n=0}^{N} \binom{N}{n} \left(\frac{\omega_{2\to 1} s}{\omega_{1\to 2} + \omega_{2\to 1}} \right)^n \left(\frac{\omega_{1\to 2}}{\omega_{1\to 2} + \omega_{2\to 1}} \right)^{N-n}, \tag{8.87}$$

which gives

$$p_{\rm st}(n) = \binom{N}{n} \left(\frac{\omega_{2\to 1}}{\omega_{1\to 2} + \omega_{2\to 1}} \right)^n \left(\frac{\omega_{1\to 2}}{\omega_{1\to 2} + \omega_{2\to 1}} \right)^{N-n} \tag{8.88}$$

a binomial distribution of parameter $P_1^{\rm st} = \dfrac{\omega_{2\to 1}}{\omega_{1\to 2}+\omega_{2\to 1}}$, in full agreement with (8.23) in the steady state using (8.11).

Let us apply, as a last example, the technique of the generating function to the birth and death process of section 8.3. The resulting partial differential equation is:

$$\frac{\partial G(s,t)}{\partial t} = (s-1)\left[\Omega_A G(s,t) - \omega \frac{\partial G(s,t)}{\partial s}\right]. \tag{8.89}$$

It is possible to find the solution of this equation using the method of characteristics. For simplicity, let us focus again in the stationary state in which the time-derivative is equal to zero:

$$0 = \Omega_A G_{\rm st}(s) - \omega \frac{dG_{\rm st}(s)}{ds}. \tag{8.90}$$

The solution with the initial condition $G_{\rm st}(1) = 1$ is

$$G_{\rm st}(s) = {\rm e}^{\frac{\Omega_A}{\omega}(s-1)}. \tag{8.91}$$

If we expand it in power series of its argument:

$$G_{\rm st}(s) = {\rm e}^{-\frac{\Omega_A}{\omega}} \sum_{n=0}^{\infty} \left(\frac{\Omega_A}{\omega}\right)^n \frac{s^n}{n!} \tag{8.92}$$

hence

$$p_{\rm st}(n) = \frac{{\rm e}^{-\frac{\Omega_A}{\omega}}}{n!}\left(\frac{\Omega_A}{\omega}\right)^n, \tag{8.93}$$

a Poisson distribution of parameter $\lambda = \dfrac{\Omega_A}{\omega}$.

8.5 The mean-field theory

Sometimes (more often than desired) it is not possible to solve the master equation using the generating function technique or any other. However, it is possible to obtain very easily approximated equations for the first moments of the probability $p(n;t)$. In many occasions the knowledge of the first moment $\langle n(t)\rangle$, gives important information about the underlying stochastic process.

Let us then the consider the general master equation:

$$\frac{\partial p(n;t)}{\partial t} = \sum_{\ell} (E^{\ell} - 1)\left[\Omega_{n\to n-\ell}\, p(n;t)\right], \tag{8.94}$$

Multiplying by n and summing over n, one gets after some algebra the (exact) equations for the first moment, as:

$$\frac{d\langle n(t)\rangle}{dt} = -\sum_{\ell} \langle \ell\, \Omega_{n\to n-\ell}\rangle. \tag{8.95}$$

Let us apply this result to the radioactive decay discussed in 8.3. The only contribution to the master equation comes from the term $\ell = 1$ with $\Omega(n \to n-1) = \omega n$, therefore:

$$\frac{d\langle n(t)\rangle}{dt} = -\langle \omega n\rangle = -\omega\langle n\rangle \tag{8.96}$$

A closed equation whose solution $\langle n(t)\rangle = N\mathrm{e}^{-\omega t}$ agrees, of course, with (8.45), obtained by other methods.

Other times we are not so lucky. If we take the master equation of the self-annihilation process of 8.3, as given in (8.63), the contributions come from the terms $\ell = 2$ and $\ell = -1$, with the rates (8.64)-(8.65). This gives:

$$\frac{d\langle n(t)\rangle}{dt} = -\langle 2kV^{-1}n(n-1)\rangle + \langle \Omega_A\rangle \tag{8.97}$$

$$= -2kV^{-1}\langle n^2\rangle + 2kV^{-1}\langle n\rangle + \Omega_A. \tag{8.98}$$

But, alas!, this is a not closed equation, as the evolution of the first moment $\langle n\rangle$ depends on the second moment $\langle n^2\rangle$. It is possible now to derive an equation for the evolution of the second moment using the general result

$$\frac{d\langle n^2(t)\rangle}{dt} = -\sum_\ell \langle \ell(\ell-2n)\Omega_{n\to n-\ell}\rangle\,. \tag{8.99}$$

However, when we replace the rates (8.64)-(8.65), the evolution of the second moment depends on the third moment, and so on in an infinite hierarchy of equations.

The simplest simplification scheme to break the hierarchy is to assume that the stochastic process is such that the fluctuations of the n variable can be neglected, or $\sigma^2[n] = \langle n^2\rangle - \langle n\rangle^2 \approx 0$, which implies $\langle n^2\rangle \approx \langle n\rangle^2$. This is the mean-field approximation[6]. Replacing in (8.98) we obtain the closed equation:

$$\frac{d\langle n(t)\rangle}{dt} = -2kV^{-1}\langle n\rangle^2 + 2kV^{-1}\langle n\rangle + \Omega_A. \tag{8.100}$$

It is convenient to consider the average density of the X-particles, defined as $x(t) = V^{-1}\langle n(t)\rangle$. A simple manipulation leads to

$$\frac{dx(t)}{dt} = -2kx^2 + 2kV^{-1}x + V^{-1}\Omega_A. \tag{8.101}$$

The last term $V^{-1}\Omega_A = V^{-1}\omega_A N_A = \omega_A x_A$, being $x_A = V^{-1}N_A$ the density of the reservoir. In the thermodynamic limit, $V \to \infty$, we can neglect the second term in the right-hand-side and arrive at the macroscopic equation[7]:

$$\frac{dx(t)}{dt} = -2kx^2 + \omega_A x_A. \tag{8.102}$$

6) The words "mean-field" have different meanings in different contexts. Here, it refers specifically to the assumption that the fluctuations can be neglected.

7) The explicit solution is $x(t) = x_{\rm st}\tanh\left(t/\tau + \mathrm{arctanh}(x(0)/x_{\rm st})\right)$ with $x_{\rm st} = \sqrt{\omega_A x_A/2k}$ and $\tau = 1/\sqrt{2k\omega_A x_A}$.

Let us turn now to the prey-predator Lotka-Volterra model. We begin by computing the average values of the number of prey and predators, $\langle n_1(t)\rangle = \sum_{n_1,n_2} n_1 p(n_1,n_2;t)$ and $\langle n_2(t)\rangle = \sum_{n_1,n_2} n_2 p(n_1,n_2;t)$. Taking the time derivative and replacing the master equation (8.73) and the rates (8.74)-(8.76), one obtains after some algebra:

$$\frac{d\langle n_1(t)\rangle}{dt} = \omega_0\langle n_1\rangle - k_1 V^{-1}\langle n_1 n_2\rangle. \tag{8.103}$$

$$\frac{d\langle n_2(t)\rangle}{dt} = k_1 V^{-1}\langle n_1 n_2\rangle - \omega_2\langle n_2\rangle. \tag{8.104}$$

These equations are again not closed. We could now compute the evolution the time evolution of $\langle n_1 n_2\rangle = \sum_{n_1,n_2} n_1 n_2 p(n_1,n_2;t)$ but then it would be coupled to higher and higher order moments, a complete mess! We use again the mean-field approach and assume that the correlations between the populations of prey and predator can be neglected, or $\langle n_1 n_2\rangle = \langle n_1\rangle\langle n_2\rangle$. Under this approximation, we obtain the closed equations:

$$\frac{d\langle n_1(t)\rangle}{dt} = \omega_0\langle n_1\rangle - k_1 V^{-1}\langle n_1\rangle\langle n_2\rangle, \tag{8.105}$$

$$\frac{d\langle n_2(t)\rangle}{dt} = k_1 V^{-1}\langle n_1\rangle\langle n_2\rangle - \omega_2\langle n_2\rangle. \tag{8.106}$$

Which can be written in terms of the density of the different species $x_1(t) = V^{-1}\langle n_1(t)\rangle$, $x_2(t) = V^{-1}\langle n_2(t)\rangle$,

$$\frac{dx_1(t)}{dt} = \omega_0 x_1 - k_1 x_1 x_2, \tag{8.107}$$

$$\frac{dx_2(t)}{dt} = k_1 x_1 x_2 - \omega_2 x_2. \tag{8.108}$$

These are the celebrated Lotka-Volterra equations.

8.6 The Fokker-Planck equation

The master equation is a complicated set of many coupled differential equations. We have already seen that it can be analyzed in terms of a partial differential equation for the generating function $G(s,t)$. We will now find an approximated partial differential equation valid directly for the probability $p(n;t)$. It is possible to develop a rigorous derivation of this equation estimating the order of the approximation that occurs in the truncation. Here, we offer a very simplified derivation in which it is not possible to determine precisely the error of the approximation. We limit ourselves to master equations for one variable of the form (8.94), but similar expansions can be carried out in the case of having more than one variable. The idea is to consider that n is a large, macroscopic, variable and can be treated as a continuous variable[8]. With

8) It is possible to formalize this idea by introducing $x = n/V$, being V a large parameter (not necessarily related to the volume V). The expansion becomes then a power series in V.

this is mind, we use a Taylor series expansion in the definition of the step operator applied to any function $f(n)$:

$$E^{\ell}[f(n)] = f(n+\ell) = f(n) + \ell \frac{df(n)}{dn} + \frac{\ell^2}{2!}\frac{d^2 f(n)}{dn^2} + \ldots, \tag{8.109}$$

and

$$(E^{\ell} - 1)[f(n)] = E^{\ell}[f(n)] - f(n) = \ell \frac{df(n)}{dn} + \frac{\ell^2}{2!}\frac{d^2 f(n)}{dn^2} + \ldots, \tag{8.110}$$

where (without much justification) we restrict the expansion to the second order in ℓ. The lack of justification of this truncation is one of the weak points of this simple derivation[9)]. Replacing in (8.94) we obtain:

$$\frac{\partial p(n;t)}{\partial t} = \sum_{\ell} \left(\ell \frac{\partial (\Omega_{n\to n-\ell} p(n;t))}{\partial n} + \frac{\ell^2}{2!}\frac{\partial^2 (\Omega_{n\to n-\ell} p(n;t))}{\partial n^2} \right), \tag{8.111}$$

or, rearranging terms,

$$\frac{\partial p(n;t)}{\partial t} = \frac{\partial}{\partial n}\left(F(n)p(n;t) + \frac{1}{2}\frac{\partial}{\partial n}\left(G(n)p(n;t)\right)\right), \tag{8.112}$$

with

$$F(n) = \sum_{\ell} \ell\, \Omega_{n\to n-\ell}, \tag{8.113}$$

$$G(n) = \sum_{\ell} \ell^2 \Omega_{n\to n-\ell}. \tag{8.114}$$

Equation (8.112) has the form of a Fokker-Planck equation for the probability $p(n;t)$ with $F(n)$ and $G(n)$ the drift and diffusion coefficients, that we have already encounter in previous chapters. Just to finish this chapter, we mention that one can write down the associated Langevin equation (in the Itô interpretation):

$$\frac{dn(t)}{dt} = F(n) + \sqrt{G(n)}\xi(t), \tag{8.115}$$

being $\xi(t)$ the usual, zero-mean, delta-correlated white noise.

Further reading and references

The theory of master equations can be found, explained to a much deeper level than the one used here, in the book by van Kampen[21]. Pawula's theorem appeared in [37]. The classic Lotka-Volterra equations and their solutions are explained in many books; see, for example [38].

9) An interesting result is Pawula's theorem that states, basically, that only by truncating at second order does the resulting equation (8.112) be ensured to have the property that the probability $p(n;t)$ remains positive at all times.

Exercises

1) Check that the conditional probabilities (8.14-8.17) satisfy the Chapman-Kolmogorov equations (8.20) and that the probabilities (8.9-8.10) satisfy the functional relations (8.18-8.19).
2) Derive the pdf of the first jump in the case of time dependent rates $\omega_{ij}(t)$ as

$$f^{1\text{st}}(j,t|i,t_0) = \omega_{ij}(t)\mathrm{e}^{-\int_{t_0}^{t} ds\, W_{ii}(s)}$$

with $W_{ii}(t) = \sum_{j\neq i}\omega_{ij}$.
3) Consider the general master equation:

$$\frac{\partial P(n,t)}{\partial t} = \sum_{\ell}(E^{\ell}-1)\left[C_{\ell}(n)\,P(n,t)\right],$$

where we use the notation $C_\ell(n) = \Omega_{n,n-\ell}$. Assume that the transition rates can be expanded in a Taylor series $C_\ell(n) = \sum_k C_\ell^k n^k$ and use that $s^{n+k} = \left(s\frac{\partial}{\partial s}\right)^k s^n$, to derive the equation for the generating function:

$$\frac{\partial G}{\partial t} = \sum_{\ell}(s^{-\ell}-1)C_{\ell}\left(s\frac{\partial}{\partial s}\right)G(s,t).$$

4) Check by direct substitution that

$$G(s,t) = \left[\frac{\omega_1+\omega_2 s+\omega_2 e^{-\omega t}(1-s)}{\omega}\right]^N G_0\left(\frac{\omega_1+\omega_2 s-\omega_1 e^{-\omega t}(1-s)}{\omega_1+\omega_2 s+\omega_2 e^{-\omega t}(1-s)}\right)$$

with $\omega \equiv \omega_1+\omega_2$, is the general solution of (8.84). Set as initial conditions (i) $p(n;t=0)=\delta_{n,0}$ and (ii) $p(n;t=0)=\delta_{n,N}$, determine the function G_0 and check that in both cases $p(n;t)$ is given by a binomial distribution.
5) Check by direct substitution that

$$G(s,t) = \mathrm{e}^{\frac{\Omega_A}{\omega}(s-1)(1-\mathrm{e}^{-\omega t})}G_0\left(1+(s-1)\mathrm{e}^{-\omega t}\right)$$

is the general solution of (8.32). Use the derivatives of this function to find the first two moments $\langle n(t)\rangle$ and the variance $\sigma^2[n(t)]$ that follows from the master equation.
6) Derive the mean field equation

$$\dot{x}(t) = \omega_{21}(v^{-1}-x) - k_{12}x^2$$

for the chemical reaction indicated in section 8.3. Here $v = V/N$ is the volume per atom and $x = n/V$ is the density of NaCl molecules.
7) Derive the Lotka-Volterra mean-field equations in the case that we consider separately and with different rates the processes $X+Y \longrightarrow 2Y$ and $X+Y \longrightarrow Y$. Show that these general equations:

$$\begin{aligned}\frac{dx_1(t)}{dt} &= \omega_0 x_1 - k_1 x_1 x_2,\\ \frac{dx_2(t)}{dt} &= k_2 x_1 x_2 - \omega_2 x_2,\end{aligned}$$

admit the constant of movement $x_1^{\omega_2} x_2^{\omega_0} \mathrm{e}^{-k_1 x_2 - k_2 x_1}$. Conclude that the trajectories $x_1(t)$, $x_2(t)$ are periodic and happen on the orbit:

$$x_2 = -\frac{\omega_0}{k_1} W\left[-\frac{k_1}{\omega_0} x_2^0 \mathrm{e}^{-\frac{k_1}{\omega_0} x_2^0} \left(\frac{x_1^0}{x_1}\right)^{\omega_2/\omega_0} \mathrm{e}^{\frac{k_2}{\omega_2}(x_1 - x_1^0)} \right],$$

depending on the initial condition (x_1^0, x_2^0), and the convenient branch of the Lambert function $W[x]$ has to be used as necessary to ensure continuity.

8) **The enzymatic reaction.** L Michaelis and M.L. Mentis analyzed the enzymatic reaction in which a "substance" S becomes a "product" with the help of an enzyme E. The enzyme works by binding to the product and facilitating the transformation of S into P, but E it is left unchanged afterwards. Schematically,

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} P + E,$$

being ES an intermediate molecule and we have indicated by k_i the corresponding rates. By assuming a high energy barrier for the combination of a product with an enzyme the backwards rate k_{-2} can be neglected. Write down the mean-field equations for the evolution of the concentration of substance $s(t)$, enzyme $e(t)$, intermediate molecule $c(t)$ and product $p(t)$. Study the steady state limit in which the intermediate molecule concentration is a constant, $\dot{c}(t) = 0$.

9) A simple model for the spread of an epidemics is the so-called SIR model: S (for susceptible), I (for infectious) and R (for recovered). In its simplest form a population of N individuals is split into these three groups: n_S susceptible people can get the disease; n_I infectious people have the disease and can hence pass the infection to susceptible people; n_R recovered people from the infection who then become immune to new contagion. In this simple version, there are no death or birth of individuals and the total number $N = n_R + n_I + n_R$ remains constant. The process is schematized as:

$$S + I \overset{\beta}{\longrightarrow} 2I,$$
$$I \overset{\nu}{\longrightarrow} R,$$

with individual rates β and ν. Write down the mean-field equations for this model and solve them numerically. Assume that the total contagion rate is $\beta n_S n_I / V$ being V a measure of the volume available to the population.

10) Derive, using the method explained in the text, the Langevin equation corresponding to the birth and death process of section 8.3. Solve it numerically and check the precision with which the Poisson distribution is obtained in the steady state.

11) Write down the master equation corresponding to the SIR model of the previous problem using as independent variables n_S and n_I, the number of susceptible and infected people. Expand the master equation up to second-order derivatives and obtain the 2-d Fokker-Planck equation and, from there, the Langevin equation.

12) Find the general expression for the stationary solution of the Fokker-Planck equation (8.112). Apply the result to the birth and death process with rates (8.51-8.52) and compare the result with the exact stationary distribution given by (8.93).