CHM1485: Molecular Dynamics and Chemical Dynamics in Liquids

Stochastic Models - Part I

October 18, 2005

1 Stochastic Modeling

Stochastic systems are systems with randomness, i.e. with a non-deterministic element, but in which the statistics of the randomness is somehow known.

A simple example of a stochastic process in the coin toss. When tossing an unbiased coin there is a 50% chance for head or tail and one cannot predict which one from the previous outcomes.

This example already shows the need for a statistical description of such a system.

Note: The words *stochastic* and *random* mean practically the same. Sometimes saying that something is *random* means it has a probability distribution, whereas something *stochastic* can mean anything pertaining to a system with randomness. More often than not, however, the two terms are interchangeable.

If we have a system that is microscopically deterministic (from a classical point of view), why would we describe it with a stochastic model?

There are several possible reasons:

1. *Quantum mechanics: which tells us nothing is really predictable.*
   This is a false argument, as the coin tossing example already shows: if it were quantum mechanics that makes it random, then why does Planck's constant not appear in the probabilities? If fact completely deterministic systems (as those from classical mechanics) can appear random on a coarse grained level.
2. *Ignorance: We cannot know all the variables so we treat only some explicitly and the rest as random.*
   This is more an excuse than an argument.
3. *Experience: tells us that certain processes behave as if they are random.*
   This is true but not a first principle explanation for why stochastic modeling works.
4. *Separation of time scales: The system under investigation can naturally be divided into a relevant or slow part and an irrelevant or fast part that may be treated as random.*
   This is the best case scenario but it is rarely possible to prove that this is the case. Furthermore, even for very simple systems effective randomness can have different sources, such a chaotic microscopic behaviour or a large number of particles.

In practice some combination of the last two is used to motivate the use of a stochastic model, i.e. from experience, experiment or physio-chemical intuition one can identify the relevant slow variables, and this assumption can be tested by investigating the fastness of the irrelevant fast variables.

Often, the following separation of times scales occurs in chemical and physical systems:

1. Fast behaviour takes place at *microscopic* length scales.
2. Stochastic behaviour takes places at larger length and time scales, the *mesoscopic scale*.
3. At even larger length and time scales fluctuations are no longer visible so that random variables behave as if they are deterministic. This is the *macroscopic scale*.

A separation of time scales of this sort is essential for the stochastic approach to work.

2 Some examples

Example 1: Coin Toss

- 50% chance of finding head or tail
- It is not possible to predict the outcome from previous outcomes.
- One can only predict probabilities and averages.

Example 2: Brownian Motion

- Large (micron sized) pollen particle
- Suspended in a solvent composed of much smaller molecules (water)
- Interactions or collisions of solvent molecules constitute a random force acting on the Brownian particle
- The assumption of randomness works here because the Brownian particle is much heavier than the solvent molecules.
- Remarkably, even though the average of the random forces is zero, their effect is still visible under a microscope.

To understand this last point better we will consider a simplified variant of the Brownian particle, namely the random walk.

Example 3: Random Walk

- The *random walker* (traditionally a drunkard) can move in one dimension only and only on the points of a *lattice* which are a distance δx apart (lamp posts).
- Time is also chosen to proceed in discrete intervals δt (but it is not essential to do so).
- At any instant the walker moves to a neighbouring point, choosing randomly between left and right.
- Thus, if $x(t)$ is the walker's position at time t,

$$\begin{aligned} x(t+\delta t) &= x(t) + \Delta x(t) && \text{(1a)} \\ \Delta x(t) &= \pm\delta x \text{ (each with probability 1/2)} && \text{(1b)} \end{aligned}$$

- This equation is not deterministic. It is a *stochastic difference equation*.
- On average the walker does not get anywhere:

$$\langle x(t+\delta t)\rangle = \langle x(t)\rangle + \langle \Delta x(t)\rangle = \langle x(t)\rangle.$$

  Here, the average is over *realizations* of the process, i.e. summing or integrating over all ways the process may work out (with proper statistical weights).

- The different realizations may be said to form an *ensemble*.
- Although the mean displacement of the walker is zero, the *mean square displacement* grows in time:

$$\begin{aligned} \langle [x(t+\delta t)]^2\rangle &= \langle [x(t)+\Delta x(t)]^2\rangle \\ &= \langle [x(t)]^2 + 2x(t)\Delta x(t) + [\Delta x(t)]^2\rangle \\ &= \langle [x(t)]^2\rangle + 2\langle x(t)\rangle\langle \Delta x(t)\rangle + \langle [\Delta x(t)]^2\rangle \\ &= \langle [x(t)]^2\rangle + \delta x^2 \end{aligned}$$

$$\Rightarrow \langle [x(t+\tau)]^2\rangle = \langle [x(t)]^2\rangle + 2D\tau \qquad .$$

  where $D = \dfrac{\delta x^2}{2\delta t}$ is called the *diffusion coefficient*.

- Thus, an ensemble of (independent) walkers starting at the same point at time zero (closing time of the bar?) would spread out over the lattice. The spread at time τ would be $\sqrt{2D\tau}$.
- Behaviour where the mean square displacement grows *linearly in time* such as above, is generally called *diffusive*. In an isotropic system in d dimensions one has

$$\langle |\mathbf{r}(\tau) - \mathbf{r}(0)|^2\rangle = \langle |r_1(\tau) - r_1(0)|^2\rangle + \ldots + \langle |r_d(\tau) - r_d(0)|^2\rangle = 2dD\tau. \qquad (2)$$

- An alternative description to equation (1), which describes how the *position* $x(t)$ evolves, is an equation which describes the time evolution of $p(x,t)$, i.e., the *probability* that the walker is at point x at time step t.

- This $p(x,t)$ satisfies:

$$p(x, t+\delta t) = \frac{p(x-\delta x, t) + p(x+\delta x, t)}{2}. \tag{3}$$

- This equation in words: The total probability to be at point x equals [(the probability to have been at point $x - \delta x$ one time step δt earlier) times (the probability to have moved from there to its right neighbour)] plus [(the probability to have been at point $x + \delta x$) times (the probability to have moved from there to its left neighbour)].

- Equation (3) is an example of a *master equation* (sometimes also called the Frobenius-Perron equation).

- Equations (1) and (3) above are equivalent. This is analogous to the Heisenberg vs. Schrödinger pictures in quantum mechanics. The former equation has the time evolution of the $x(t)$ ($\sim$ 'observable'), while the latter has the evolution of the probability distribution $p(x,t)$ ($\sim$ 'wave function').

- *Continuum limit:* Take the limit of small δx and δt while keeping the diffusion constant $D = \delta x^2/\delta t$ fixed. Expanding equation (3) gives

$$\begin{aligned}
&p(x,t) + \delta t \frac{\partial p(x,t)}{\partial t} + \mathcal{O}(\delta t^2) \\
&= \frac{1}{2}\left[p(x,t) - \frac{\partial p(x,t)}{\partial x}\delta x + \frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2}\delta x^2 - \frac{1}{6}\frac{\partial^3 p(x,t)}{\partial x^3}\delta x^3 \right.\\
&\qquad \left. + p(x,t) + \frac{\partial p(x,t)}{\partial x}\delta x + \frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2}\delta x^2 + \frac{1}{6}\frac{\partial^3 p(x,t)}{\partial x^3}\delta x^3\right] + \mathcal{O}(\delta x^4) \\
&= p(x,t) + \frac{\partial^2 p(x,t)}{\partial x^2}\frac{\delta x^2}{2} + \mathcal{O}(\delta x^4) \\
&= p(x,t) + \frac{\partial^2 p(x,t)}{\partial x^2} D\delta t + \mathcal{O}(\delta t^2),
\end{aligned}$$

whence in the limit $\delta t \to 0$

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^2 p(x,t)}{\partial x^2}. \tag{4}$$

- This is called the *diffusion equation*.

3 Markov Processes and Master Equations

- In a *Markov process* the probability at time t depends at most on one earlier time, while even earlier times have no influence (a more rigorous definition can be found in Van Kampen's book).
- The random walk and diffusion are examples of Markov processes. I.e. for the random walk equations (1) and (3) only involve $t + \delta t$ and t. Likewise, the diffusion equation (4) only involves $p(x,t)$ and $\partial p(x,t)/\partial t = \lim_{\delta t \to 0}[p(x,t+\delta t) - p(x,t)]/\delta t$.
- Instead of a position x, in general, $\mathbf{x}$ represents a *state* with multiple components.
- The states may have a continuous or discontinuous range (or even a combination).
- Continuous states can represent positions of molecules, non-bounded energy levels, etc.
- Discontinuous states can be used to represent the number of molecules of various kinds, discrete bounded energy levels, etc.
- Time may be continuous or discrete (although the latter is rarely physical).
- A Markov process is completely characterized by the function $G(\mathbf{x}, t|\mathbf{x}_0, t_0)$, which is the probability (or probability density) to be in state $\mathbf{x}$ at time t given that at time t_0 the system was in state $\mathbf{x}_0$.
- $G(\mathbf{x}, t|\mathbf{x}_0, t_0)$ can be viewed as the probability $P(\mathbf{x}, t)$ to be in $\mathbf{x}$ at time t with the initial condition that the system was in $\mathbf{x}_0$ at time t_0: $G(\mathbf{x}, t|\mathbf{x}_0, t_0) = \delta(\mathbf{x} - \mathbf{x}_0)$.

  Example:
  The solution of the diffusion equation (4) in one dimension with the initial condition that the particle is at position x_0 at time t_0, is given by

  $$G(x, t|x,t_0) = \frac{1}{\sqrt{4\pi D(t-t_0)}} \exp\left[-\frac{(x-x_0)^2}{4D(t-t_0)}\right].$$

  So in this case G is a *Gaussian* or *normal distribution*. We could have expected this since the diffusion equation is the limit of a random walk, which consists of adding independent increments $\pm\delta x$ to the position of the walker and the *central limit theorem* tells us that the probability distribution for the sum of N independent random variables approaches (in properly scaled form) a Gaussian as $N \to \infty$.
- For continuous time and continuous states $G(\mathbf{x}, t|\mathbf{x}_0, t_0)$ and $P(\mathbf{x}, t)$ satisfy the general *master equation:*

  $$\frac{\partial P(\mathbf{x}, t)}{\partial t} = \int d\mathbf{x}' \Big[W(\mathbf{x}' \to \mathbf{x})P(\mathbf{x}', t) - W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x}, t)\Big], \tag{5}$$

  where $W(\mathbf{x} \to \mathbf{x}')d\mathbf{x}'$ is the transition probability per unit time to go from state $\mathbf{x}$ to another state between $\mathbf{x}'$ and $\mathbf{x}' + d\mathbf{x}'$.

- Discrete time and/or space master equations can be found in the following table:

time		discrete state space	continuous state space
discrete	$P(\mathbf{x}, t+\delta t) - P(\mathbf{x}, t) =$	$\sum_{\mathbf{x}'}[W(\mathbf{x}' \to \mathbf{x})P(\mathbf{x}', t) - W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x}, t)]$	$\int d\mathbf{x}'\ [W(\mathbf{x}' \to \mathbf{x})P(\mathbf{x}', t) - W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x}, t)]$
continuous	$\frac{\partial P(\mathbf{x}, t)}{\partial t} =$	$\sum_{\mathbf{x}'}[W(\mathbf{x}' \to \mathbf{x})P(\mathbf{x}', t) - W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x}, t)]$	$\int d\mathbf{x}'\ [W(\mathbf{x}' \to \mathbf{x})P(\mathbf{x}', t) - W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x}, t)]$

where for discrete time $W(\mathbf{x} \to \mathbf{x}')$ is the probability to go from $\mathbf{x}$ to $\mathbf{x}'$ in the time δt.

- The random walk equation (3) is an example of the discrete time, discrete space master equation with $W(x \to x') = \frac{1}{2}\delta_{|x-x'|,\delta x}$.
- The diffusion equation (4) is an example of the continuous time, continuous space master equation (5) with $W(x \to x') = \lim_{\delta x \to 0} \frac{D}{2\delta x^3}\delta_{|x-x'|,\delta x} = D\frac{\partial^2 \delta(x-x')}{\partial x^2}$.
- The master equation requires the assumption of Markovianness and the knowledge of the transition probabilities (per unit time) $W(\mathbf{x} \to \mathbf{x}')$. Given these properties it is an exact equation and can therefore be the starting point of a rigorous theory.
- We can cast the right hand side of equation (5) in a more familiar form by defining the *stochastic Liouville operator:*

$$-\mathcal{L}P(\mathbf{x}) \equiv \int d\mathbf{x}' \Big[W(\mathbf{x}' \to \mathbf{x})P(\mathbf{x}') - W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x})\Big]$$

so that equation (5) can be written as

$$\frac{\partial P}{\partial t} = -\mathcal{L}P.$$

The formal solution of this equation is

$$P(t) = e^{-\mathcal{L}(t-t_0)}P(t_0).$$

4 Equilibrium, Detailed Balance and Reversibility

- A solution of the master equation which is independent of time is called a *stationary solution:*

$$P(\mathbf{x}, t) = P^{\mathrm{s}}(\mathbf{x}).$$

- There is always at least one stationary solution of the master equation.
- In general there can be multiple stationary solutions of the master equation.
- If the state space is discrete and finite, one can prove that any initial distribution goes to one of the stationary solutions for $t \to \infty$.
  For a continuous or infinite state space this is typically (but not always) true as well, e.g. when we know that the system will approach equilibrium.

- If there is a unique stationary solution, it can often be identified with the *equilibrium solution:*

  $$P^{\mathrm{s}}(\mathbf{x}) = P^{\mathrm{e}}(\mathbf{x}).$$

  (Caution: whether this is really so depends on other considerations, such as whether the system is closed and there are no external fields or forcings.)

- From the master equation (5) above the stationarity of P^{e} implies that $\mathcal{L}P^{\mathrm{e}} = 0$ or

  $$\int d\mathbf{x}' \; [W(\mathbf{x}' \to \mathbf{x})P^{\mathrm{e}}(\mathbf{x}') - W(\mathbf{x} \to \mathbf{x}')P^{\mathrm{e}}(\mathbf{x})] = 0. \tag{6}$$

- The simplest (but by far not the only!) way to get stationarity is if

  $$W(\mathbf{x}' \to \mathbf{x})P^{\mathrm{e}}(\mathbf{x}') = W(\mathbf{x} \to \mathbf{x}')P^{\mathrm{e}}(\mathbf{x}) \tag{7}$$

  This property is known as *detailed balance*.

- Because we know the equilibrium distribution from statistical mechanical considerations ($e^{-\beta H}$), *detailed balance is a property of the transition probabilities W*.

- Detailed balance is often encountered in stochastic models in chemistry and physics, for one of two very distinct reasons:

  1. *Microscopic reversibility:*
     For closed, isolated, physical systems, detailed balance in the mesoscopic, stochastic description of that system is a consequence of the reversibility of the deterministic dynamics of that same system on the microscopic level, provided that the Hamiltonian and the states $\mathbf{x}$ are even functions of the momenta and the $\mathbf{x}$ behave Markovian. This therefore constitutes a *fundamental* reason to impose detailed balance.
  2. *Theoretical convenience:*
     It can be shown that a system in which all states can be reached and which obeys detailed balance, will ultimately relax to the equilibrium P^{e}. Thus, if one is constructing a stochastic model for the purpose of sampling an equilibrium ensemble (as in Monte Carlo simulations), adjusting the transition probabilities so that they obey detailed balance helps to guarantee that the right distribution is sampled.

     Furthermore, from detailed balance other properties follow, such as that the stochastic Liouville operator $\mathcal{L}$ is diagonalizable and has real eigenvalues.

- Obviously, detailed balance is a nice property to have.

- Note that in *non-equilibrium systems* (which may be open and not isolated and be subject to time-dependent external fields), there is in general *no fundamental justification for detailed balance* to be obeyed (i.e. microscopic reversibility alone is not enough). Somehow this does not seem to deter people from using detailed balance in such systems without motivating it, even in far-from-equilibrium situations.

5 Macroscopic Behaviour

- On a macroscopic level one expects fluctuations to be relatively unimportant (away from critical points) and one should be able to describe the behaviour of the system in terms of averages.
- Indeed, if $\mathbf{x}$ is a set of macroscopic variables (e.g. the numbers of molecules of different kinds), we know from statistical mechanics that *in equilibrium* the fluctuations in $\mathbf{x}$ are small.
- In other words, in equilibrium the microscopic probability distribution P^{e} of such quantities is highly peaked around the average.
- Consider now a non-equilibrium initial state where $\mathbf{x}$ has a given initial value $\mathbf{x}_0$ so that its initial distribution function is equal to $P(\mathbf{x},0) = \delta(\mathbf{x}-\mathbf{x}_0)$. Alternatively, one may take an initial distribution which is very peaked around $\mathbf{x}_0$.
- As time progresses, $P(\mathbf{x},0) \to P(\mathbf{x},t) \to P(\mathbf{x},\infty) = P^{\mathrm{e}}(\mathbf{x})$. Because both $P(\mathbf{x},0)$ and $P(\mathbf{x},\infty) = P^{\mathrm{e}}(\mathbf{x})$ are highly peaked, it is reasonable to assume that $P(\mathbf{x},t)$ is highly peaked around the average throughout the entire process.
- We therefore *identify* the average of $\mathbf{x}$ at any instant with the macroscopic value $\mathbf{x}^{\mathrm{macro}}$ at time t:
$$\mathbf{x}^{\mathrm{macro}}(t) = \langle \mathbf{x}(t) \rangle = \int d\mathbf{x}\, \mathbf{x}\, P(\mathbf{x},t).$$
- To find the behaviour of the macroscopic variable in time, we take its time derivative
$$\frac{d\mathbf{x}^{\mathrm{macro}}(t)}{dt} = \int d\mathbf{x}\, \mathbf{x}\, \frac{\partial P(\mathbf{x},t)}{\partial t}.$$

  and substitute the derivative of the distribution function from the master equation (5), giving
$$\begin{aligned}
\frac{d\mathbf{x}^{\mathrm{macro}}(t)}{dt} &= \iint d\mathbf{x}\, d\mathbf{x}'\, \mathbf{x}\, [W(\mathbf{x}' \to \mathbf{x})P(\mathbf{x}',t) - W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x},t)] \\
&= \iint d\mathbf{x}\, d\mathbf{x}'\, (\mathbf{x}'-\mathbf{x})W(\mathbf{x} \to \mathbf{x}')P(\mathbf{x},t) \\
&= \langle \mathbf{a}_1(\mathbf{x}(t)) \rangle,
\end{aligned}$$

  where we have defined the *first jump moment* $\mathbf{a}_1(\mathbf{x}(t))$ as
$$\mathbf{a}_1(\mathbf{x}(t)) = \int d\mathbf{x}'\, [\mathbf{x}' - \mathbf{x}(t)] W(\mathbf{x}(t) \to \mathbf{x}'). \tag{8}$$

  This is the average change of $\mathbf{x}$ per unit time given that its value is $\mathbf{x}(t)$.

- $P(\mathbf{x},t)$ is highly peaked around $\mathbf{x}^{\mathrm{macro}}(t)$ so we can expand $\mathbf{a}_1(\mathbf{x}(t))$ around $\mathbf{x}(t) = \mathbf{x}^{\mathrm{macro}}(t)$. Neglecting fluctuations this then gives the *macroscopic equation*:

$$\frac{d\mathbf{x}^{\mathrm{macro}}(t)}{dt} = \mathbf{a}_1(\mathbf{x}^{\mathrm{macro}}(t)). \tag{9}$$

6 Chemical Reactions

- Stochastic methods are very useful for modeling *chemical reactions in dilute solutions*, since in dilute solutions the time and place when molecules meet and react is practically random.
- We will consider an idealized situation in which the system is *uniform* so that any reaction products diffuse rapidly through the whole volume V. In less ideal situations this may be accomplished by stirring.
- Let the system now contain (at most) J different kinds of chemical compound, denoted by X_j $(j = 1, 2, \ldots, J)$.
  ◇ *As a running example, we consider a system with* $J = 4$ *compounds* A*,* B*,* C *and* D*.*
- Denote the number of molecules of each compound by N_j.
  ◇ *Thus in our running example, there are* N_{A} *molecules* A*,* N_{B} *molecules* B*,* N_{C} *molecules* C *and* N_{D} *molecules* D*.*
- The collection $\mathbf{x} = (N_1, N_2, \ldots, N_J)$ of the numbers of molecules for all compound specifies the state of the system.
  ◇ *In our example,* $\mathbf{x} = (N_{\mathrm{A}}, N_{\mathrm{B}}, N_{\mathrm{C}}, N_{\mathrm{D}})$.
- Note that this is makes for a discrete state space, which can be interpreted as a J dimensional lattice on which the state vectors $\mathbf{x}$ must lie. Because the N_j are nonnegative, only an 'octant' of the lattice is available.
  ◇ *In our example, we have a four-dimensional lattice and* $N_{\mathrm{A}} = 0, 1, 2, \ldots$, $N_{\mathrm{B}} = 0, 1, 2, \ldots$, *etc.*
- Let the possible chemical reactions be specified through the forms

$$s_1^{(\rho)} X_1 + s_2^{(\rho)} X_2 + \ldots \longrightarrow r_1^{(\rho)} X_1 + r_2^{(\rho)} X_2 + \ldots$$

  where $s_j^{(\rho)}$ and $r_j^{(\rho)}$ are the *stoichiometric coefficients*, which are positive integers representing the number of molecules of the respective compound involved in reaction ρ.
  ◇ *For example, consider that in our system we have only one reaction*

$$2\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C} + \mathrm{D}$$

  then $s_{\mathrm{A}} = 2$*,* $s_{\mathrm{B}} = 1$*,* $s_{\mathrm{C}} = 0$*,* $s_{\mathrm{D}} = 0$ *and* $r_{\mathrm{A}} = 0$*,* $r_{\mathrm{B}} = 0$*,* $r_{\mathrm{C}} = 1$*,* $r_{\mathrm{D}} = 1$.

Note that with each reaction is associated the reverse reaction

$$r_1^{(\rho)} X_1 + r_2^{(\rho)} X_2 + \ldots \ \rightarrow \ s_1^{(\rho)} X_1 + s_2^{(\rho)} X_2 + \ldots$$

Allowing furthermore some of the stoichiometric coefficients to take on a value of zero, we may write

$$\sum_{j=1}^{J} s_j^{(\rho)} X_j \ \rightleftharpoons \ \sum_{j=1}^{J} r_j^{(\rho)} X_j .$$

◇ *For our example, we have simply:*

$$2\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C} + \mathrm{D}.$$

- These equations specify the possible transitions between states. When a reaction ρ happens, it changes the number of molecules from N_j to $N_j + r_j^{(\rho)} - s_j^{(\rho)}$, changing the state vector from $\mathbf{x}$ to $\mathbf{x} + \Delta\mathbf{x}^{(\rho)}$, where

  $$\Delta\mathbf{x}^{(\rho)} = (r_1^{(\rho)} - s_1^{(\rho)}, r_2^{(\rho)} - s_2^{(\rho)}, \ldots, r_J^{(\rho)} - s_J^{(\rho)})$$

  ◇ *For the example,*

  $$\begin{aligned} \Delta\mathbf{x} &= (\Delta N_{\mathrm{A}}, \Delta N_{\mathrm{B}}, \Delta N_{\mathrm{C}}, \Delta N_{\mathrm{D}}) \\ &= (-2, -1, 1, 1). \end{aligned}$$

- Note: Because only an integer number of reactions can take place, the values of the state vector $\mathbf{x}$ are at any time equal to

  $$\mathbf{x} = \mathbf{x}_0 + \sum_{\rho} \xi_\rho \Delta\mathbf{x}^{(\rho)}.$$

  where $\mathbf{x}_0$ is the initial state vector and the ξ_ρ are integers (which may also be negative because of the reverse reactions).
  It is quite typical that these values of the state vector do not cover the full lattice, but instead form a sub-lattice, determined by the initial state $\mathbf{x}_0$ of the system. The dynamics will then take place on this sub-lattice only.
  ◇ *In our running example,* $(N_{\mathrm{A}}, N_{\mathrm{B}}, N_{\mathrm{C}}, N_{\mathrm{D}}) = (N_{\mathrm{A}}, N_{\mathrm{B}}, N_{\mathrm{C}}, N_{\mathrm{D}})_0 + \xi(-2, -1, 1, 1)$ *with* ξ *integer. These points form a one-dimensional lattice on the four dimensional lattice.*

- So far we have only specified the possible transitions, but not their transition rates W. The rate for the forward reaction ρ is given by

  $$W(\mathbf{x} \rightarrow \mathbf{x} + \Delta\mathbf{x}^{(\rho)}) = k_+^{(\rho)} V \prod_{j=1}^{J} \frac{N_j!}{(N_j - s_j^{(\rho)})!} \frac{1}{V^{s_j}}, \tag{10}$$

and for the reverse reaction by

$$W(\mathbf{x} \to \mathbf{x} - \Delta\mathbf{x}^{(\rho)}) = k_-^{(\rho)} V \prod_{j=1}^{J} \frac{N_j!}{(N_j - r_j^{(\rho)})!} \frac{1}{V^{r_j}} \tag{11}$$

(derivation: problem set). Here $k_+^{(\rho)}$ and $k_-^{(\rho)}$ are the forward and reverse reaction rates, respectively, which do not depend on the state $\mathbf{x}$, but rather on the cross section for a collision of the required molecules and the probability that such a collision results in a reaction.

◇ *For the example, this means we have*

$$W\big((N_\mathrm{A}, N_\mathrm{B}, N_\mathrm{C}, N_\mathrm{D}) \to (N_\mathrm{A}-2, N_\mathrm{B}-1, N_\mathrm{C}+1, N_\mathrm{D}+1)\big) = \frac{k_+ N_A(N_A-1)N_B}{V^2}$$

and

$$W\big((N_\mathrm{A}, N_\mathrm{B}, N_\mathrm{C}, N_\mathrm{D}) \to (N_\mathrm{A}+2, N_\mathrm{B}+1, N_\mathrm{C}-1, N_\mathrm{D}-1)\big) = \frac{k_- N_C N_D}{V}.$$

- The master equation can now be written as

$$\begin{aligned} \frac{\partial P(\mathbf{x},t)}{\partial t} \; &= \; \sum_\rho \Big\{ W(\mathbf{x} + \Delta\mathbf{x}^{(\rho)} \to \mathbf{x}) P(\mathbf{x} + \Delta\mathbf{x}^{(\rho)}, t) \\ &\quad + W(\mathbf{x} - \Delta\mathbf{x}^{(\rho)} \to \mathbf{x}) P(\mathbf{x} - \Delta\mathbf{x}^{(\rho)}, t) \\ &\quad - \big[W(\mathbf{x} \to \mathbf{x} + \Delta\mathbf{x}^{(\rho)}) + W(\mathbf{x} \to \mathbf{x} - \Delta\mathbf{x}^{(\rho)})\big] P(\mathbf{x},t) \Big\}. \end{aligned}$$

◇ *Writing this out for our example, we get*

$$\begin{aligned} &\frac{\partial P(N_\mathrm{A}, N_\mathrm{B}, N_\mathrm{C}, N_\mathrm{D}, t)}{\partial t} \\ &= \frac{k_-(N_\mathrm{C}+1)(N_\mathrm{D}+1)}{V} P(N_\mathrm{A}-2, N_\mathrm{B}-1, N_\mathrm{C}+1, N_\mathrm{D}+1, t) \\ &\quad + \frac{k_+(N_\mathrm{A}+2)(N_\mathrm{A}+1)(N_\mathrm{B}+1)}{V^2} P(N_\mathrm{A}+1, N_\mathrm{B}+1, N_\mathrm{C}-1, N_\mathrm{D}-1, t) \\ &\quad - \left[\frac{k_+ N_\mathrm{A}(N_\mathrm{A}+1)N_\mathrm{B}}{V^2} + \frac{k_- N_\mathrm{C} N_\mathrm{D}}{V}\right] P(N_\mathrm{A}, N_\mathrm{B}, N_\mathrm{C}, N_\mathrm{D}, t). \end{aligned}$$

- The corresponding macroscopic equations for the <u>*concentrations*</u> $c_i(t) \equiv \langle N_j(t)\rangle / V$ follow from equations (8) and (9);

$$\begin{aligned} \frac{dc_i(t)}{dt} \; &= \; a_{1i}\big(\{c_j(t)\}\big) \; = \; \sum_\rho a_{1i\rho}\big(\{c_j(t)\}\big) \\ &= \; \sum_\rho (r_i^{(\rho)} - s_i^{(\rho)}) \left\{ k_+^{(\rho)} \prod_{j=1}^{J} [c_j(t)]^{s_j} - k_-^{(\rho)} \prod_{j=1}^{J} [c_j(t)]^{r_j} \right\}. \end{aligned} \tag{12}$$

where the thermodynamic limit $\langle N_j(t)\rangle \to \infty$, $V \to \infty$ keeping $c_i(t) = \langle N_i(t)\rangle/V$ finite was taken.

⋄ *Writing this out for our example, we get*

$$\begin{aligned}
\frac{dc_{\mathrm{A}}(t)}{dt} &= -2\left\{k_+ c_{\mathrm{A}}^2(t)\, c_{\mathrm{B}}(t) - k_-\, c_{\mathrm{C}}(t)\, c_{\mathrm{D}}(t)\right\}\\
\frac{dc_{\mathrm{B}}(t)}{dt} &= -\left\{k_+ c_{\mathrm{A}}^2(t)\, c_{\mathrm{B}}(t) - k_-\, c_{\mathrm{C}}(t)\, c_{\mathrm{D}}(t)\right\}\\
\frac{dc_{\mathrm{C}}(t)}{dt} &= k_+ c_{\mathrm{A}}^2(t)\, c_{\mathrm{B}}(t) - k_-\, c_{\mathrm{C}}(t)\, c_{\mathrm{D}}(t)\\
\frac{dc_{\mathrm{D}}(t)}{dt} &= k_+ c_{\mathrm{A}}^2(t)\, c_{\mathrm{B}}(t) - k_-\, c_{\mathrm{C}}(t)\, c_{\mathrm{D}}(t).
\end{aligned}$$

- The macroscopic equations for the concentrations in equations (12) are called the *rate equations* (but confusingly they are sometimes also called master equations).
- As $t \to \infty$, the system approaches equilibrium and $dc_i/dt \to 0$. One can argue that each reaction ρ on the right hand side of equation (12) balanced out to zero separately (because one can hypothetically switch off all other reactions). Setting each term in the sum over ρ to zero leads to the *law of mass action*:

$$\frac{k_+^{(\rho)}}{k_-^{(\rho)}} = \prod_{j=1}^{J} c_j^{r_j^{(\rho)} - s_j^{(\rho)}}.$$

  where the c_j are the equilibrium values of the concentrations. These could be found e.g. from equilibrium statistical mechanics.

  ⋄ *For our example, the law of mass action states that*

$$\frac{k_+}{k_-} = \frac{c_C c_D}{c_A^2 c_B}.$$

Notes:

- The law of mass action may also be derived from the master equation using the (grand canonical) equilibrium distribution of concentrations in dilute systems.
- In the above, only reactions in a closed volume were discussed, but the stochastic approach can be adapted to deal with open systems without much difficulty.
- Also systems which are non-uniform may be treated, leading to reaction-diffusion equations and the like, but this requires substantial additional effort.
- To obtain the reaction rates entering in the above description of chemical reactions, we need to look on a deeper level at the dynamics *during* a single reaction. This dynamics can sometimes also be studied using stochastic models, as we will do later using the concept of a *reaction coordinate.*
- However for that purpose, we first need to study the previously mentioned Brownian motion in more detail: Next time.