# Final Theme: Reaction Kinetics

Rules:

- Each student should hand in an individually completed report at latest **January 14, 2010**.

- You may discuss the problem among fellow students. If you receive considerable help from someone, say so in your solutions.

- Do not copy solutions or code from others. Do not lend your solution or your code to other students.

Introduction to Reaction Kinetics

Quantitative interpretation of chemical reaction systems is frequently based on the construction and analysis of kinetic mechanisms or models. These models are based on the idea of breaking up the overall reaction into a number of components or *elementary steps*. An example is the following between a hydrogen molecule and an oxygen molecule:

$$H_2 + O_2 \rightarrow 2OH.$$
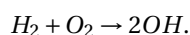
By the law of mass action, the *rate* (the number of reactions per time unit) of an elementary step is given in terms of the product of the concentrations of the participating species. For the reaction above the rate is given by $k[H_2][O_2]$, where $[H_2]$ is the concentration of hydrogen molecules, $[O_2]$ is the concentration of dioxygen, and the constant k is known as the *reaction rate coefficient* or the *reaction rate constant*.[1] Thus for our current system $2k[H_2][O_2]$ mol of hydroxide is formed each second; here, the factor 2 is due to that two molecules of hydroxide are formed in each reaction. The number $k[H_2][O_2]$ is called the *propensity* of the reaction. Throughout this theme, the propensities are proportional to the reactant concentrations following the law of mass action. That is, the reaction rate is given by a constant multiplying the product of the reacting concentrations (the concentrations of the species on the left hand side). This is true for reactions involving only a single step (elementary reacations).[2] For the elementary step above, we can write down ordinary differential equations describing the change in concentration of the participating species as

$$\frac{d[H_2]}{dt} = -k[H_2][O_2], \quad \frac{d[O_2]}{dt} = -k[H_2][O_2], \text{ and } \frac{d[OH]}{dt} = 2k[H_2][O_2].$$

Typically, elementary steps are *bimolecular*, that is, they involve two species on the left hand side. The reaction is considered to occur in a single collision between the reactants. However, the formation of a hydrogen molecule from two hydrogen atoms is written as

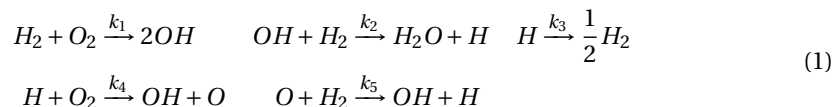$$H \xrightarrow{k} \frac{1}{2}H_2,$$

---
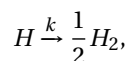
[1] Typically, rate constants are independent of concentration but depend on the temperature of the reacting mixture.

[2] For more general reactions, including more steps, the propensities depend nonlinearly on the reactant concentrations.

where the k over the arrow denotes the reaction rate constant. Here the reaction rate is simply given by $k[H]$.[3] The *stoichiometric factor* $\frac{1}{2}$ tells us that we need 1 mol of H to form $\frac{1}{2}$ mol of H_2 and is used to specify that the reaction rate depends linearly on the concentration of hydrogen atoms.

By combining all individual elementary steps, we get a description of the full system. For the hydrogen–oxygen system, an acceptable mechanism[4] is obtained using the following steps:

$$H_2 + O_2 \xrightarrow{k_1} 2OH \qquad OH + H_2 \xrightarrow{k_2} H_2O + H \quad H \xrightarrow{k_3} \frac{1}{2}H_2$$

$$H + O_2 \xrightarrow{k_4} OH + O \qquad O + H_2 \xrightarrow{k_5} OH + H \tag{1}$$

From the reactions above, we can construct an ordinary differential equation describing the evolution of the concentration of the included species. Let us now consider the change of concentration of hydrogen atoms per time unit, summing up the contributions from the five elementary steps listed above we arrive at the following equation

$$\frac{d[H]}{dt} = k_2[OH][H_2] - k_3[H] - k_4[H][O_2] + k_5[O][H_2].$$

Note that all terms that remove hydrogen atoms depend on the concentration $[H]$, thus if the concentration is zero then the derivative $d[H]/dt$ is also zero ensuring that the concentration stays non-negative.

Tasks

1. In this first problem, we study the the three step reaction

$$A \xrightarrow{k_1} B, \quad 2B \xrightarrow{k_2} B + C, \quad B + C \xrightarrow{k_3} A + C.$$
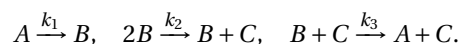
Write the above reactions as a three dimensional ODE system on the form $y' = f(y)$. Solve the resulting ODE in the time interval $0 \le t \le 60$ using the following reaction rate constants $k_1 = 0.04$, $k_2 = 3 \cdot 10^7$, $k_3 = 10^4$ and the initial condition is $y(0) = (1,0,0)^\mathsf{T}$. Try various ODE solvers, try at least MATLAB's build in solvers `ode45` and `ode15s`. Plot the concentrations for each component separately and all method simultaneously (there will be only small amounts of B throughout the reaction) using a logaritmic scale for the time variable (`semilogx`). For each solver report the number of function evaluations/timesteps required to simulate the system during the first minute.

2. Perhaps the most widely used model for oscillation amongst chemists is the Brusselator scheme[5] of Prigogine[6] and Lefever. The model has two reactant species A and B that are converted to products D and E via a sequence of irreversible reaction processes involving two intermediates X and Y. The model has the form

$$A \xrightarrow{k_1} X \qquad B + X \xrightarrow{k_2} D + Y \qquad Y + 2X \xrightarrow{k_3} 3X \qquad X \xrightarrow{k_4} E.$$
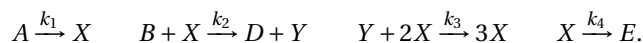
Assuming that the concentrations of A and B are constant, and that $k_1 = k_2 = k_3 = k_4 = 1$, we want to study the concentrations of the intermediates X and Y.

---

[3] If the reaction would have been $2H \xrightarrow{k} H_2$, the reaction rate would have been $k[H]^2$.

[4] According to Scott, Chemical Chaos, Oxford University Press 1991.

[5] The model was published in the article, Symmetry Breaking Instabilities in Dissipative Systems II, in The Journal of Chemical Physics, 1968.

[6] Ilya Prigogine was awarded the Nobel Prize in chemistry in 1977 for his contributions to nonequilibrium thermodynamics, particularly the theory of dissipative structures.

(a) Set up ODE system for the concentrations of X and Y.

(b) Find the stationary solutions to the system you formed in (a).

(c) Simulate the system varying the concentrations of A and B, try for example the pairs $[A,B]^\mathsf{T} = [1,2]^\mathsf{T}$, $[A,B]^\mathsf{T} = [1,4]^\mathsf{T}$, and $[A,B]^\mathsf{T} = [2,4]^\mathsf{T}$.

Small concentrations

In cases where the species are present in low copy numbers, it does not make sense to study the concentrations. Instead we study the probabilitiy that the system is at a certain state at any given time. First let us consider the birth–death problem,

$$\emptyset \to A, \quad A \to \emptyset.$$
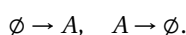
Here, we assume that at each moment in time, A can be formed at a constant probability p_1, and that an A vanishes with a probability $p_2|A|$, where the copy number $|A|$ denotes the number of copies of A present in the system. Let $P(t,x)$ denote the probability that we have x copies of A at time t, then

$$\frac{\mathrm{d}P}{\mathrm{d}t}(t,x) = p_1 P(t,x-1) - p_1 P(t,x) + p_2(x+1)P(t,x+1) - p_2 x P(t,x). \tag{2}$$

The first two terms correspond to the creation of a new A; the probability that we will have x copies of A increases proportionally to the probability that we have $x-1$ copies (the creation of a new copy will give us x copies of A) and decreases proportionally to the probability that we have x (here the creation of a new A results in a copy number of $x+1$). Similarly the two last terms pertain the death of an A. System (2) can be viewed as an infinite dimensional ODE system. An attractive alternative to the ODE model above is to use a stochastic model to simulate the system. The mathematical model we will use is a continuous time discrete space Markov model—a fundamental mathematical formalism with applications in numerous fields of applied mathematics, physics, chemistry, biology, and computer science.

A Markov process has the important property of not having memory. That is, if we know the state at time t, the state at a time $\tilde{t}$ later than t only depends on the state at time t and not on any state before t. In the stochastic model, we study the copy number of each species in the system. That is, we are dealing with integers, and not reals as was the case for concentrations. Each reaction is considered to be a discrete jump from one state to another, changing the state by integer amounts. In the stochastic model, the propensity function is interpreted as the probability per time unit that the reaction occurs.

Monte Carlo Method

There are many definitions of Monte Carlo methods and algorithms. However, they all have in common that they use random sampling of a process to compute the result. The algorithms typically rely on *pseudo random numbers*[7] to compute or generate a *realization*, one possible outcome, of the process. The main idea is to generate and observe the results from multiple realizations of the process. The observed data is then used when performing statistical analysis to answer various questions about the process. Monte Carlo relies heavily on the following two important theorems from probability.

Theorem 1 ((weak) Law of Large Numbers). *Let $X_1, X_2, \ldots, X_n$ be independent random variables with mean μ and define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then, for any $\delta > 0$, $P(\bar{X}_n - \delta < \mu < \bar{X}_n + \delta) \to 1$ as $n \to \infty$.*

---

[7]Computer generated numbers mimicking true random numbers.

Theorem 2 (Central Limit Theorem). *Let $X_1, X_2, \ldots, X_n$ be a sequence of independent identically distributed random variables with finite mean μ and variance $\sigma^2 > 0$ and let $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Then $\widetilde{X}_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma$ converges in distribution to the standard normal distribution $N(0, 1)$ as $n \to \infty$.*

In essence, the two above theorems state that if we repeatedly sample a stochastic variable, the sample mean converges to the expected value. Moreover, the central limit theorem implies that the error in the approximation of the mean decays slowly ($\sim 1/\sqrt{N}$) as a function of the sampe size N. Thus Monte Carlo methods converge very slowly, and are often computationally intensive. However, the convergence rate does not depend on the dimensionality of the problem.

Monte Carlo methods can be used to solve both stochastic and deterministic problems. Monte Carlo methods have been extensively used to evaluate definite integrals. Here, the integral is computed by repeatedly sampling the integrand and by identifying the integral with a certain expected value. More precisely, we can approximate the integral by computing the mean of the samples and multiplying this mean with the area of the integration domain. For smooth problems in low dimensions traditional quadrature method outperform Monte Carlo integration. However, Monte Carlo is more computationally efficient than traditional quadrature rules in very high dimensions, especially for non-smooth problems.

Stochastic Simulation Algorithm

A popular use of Monte Carlo methods is to simulate chemical and biochemical reactions. These simulations are particularly useful when the number of molecules of each kind is small. Without going into any details the algorithm first samples the time until the next reaction and then finds the next reaction that occurs. It can be shown that the inter event time is exponentially distributed with mean $1/$(sum of propensities) and that the probability for a given reaction to occur is proportional to its propensity. The algorithm below simulates a system with R reactions from time T_0 to final time T_f. In the listing ω_r denotes the propensity of the of the rth reaction and n_r denotes the *stoichiometry vector* (describing the change in state) for reaction r

Set the initial state $x = x_0$ and time $t = T_0$
while $t < T_f$ **do**
    Compute the sum of propensities $a = \sum_{r=1}^{R} \omega_r(x)$
    Draw two random numbers u and v between 0 and 1 from a uniform distribution.
    Compute the time $\tau = -\frac{\log(u)}{a}$ to the next reaction.
    Determine the next reaction k, by finding the smallest k such that $av < \sum_{r=1}^{k} \omega_r(x)$.
    Update the time $t = t + \tau$ and state $x = x + n_k$
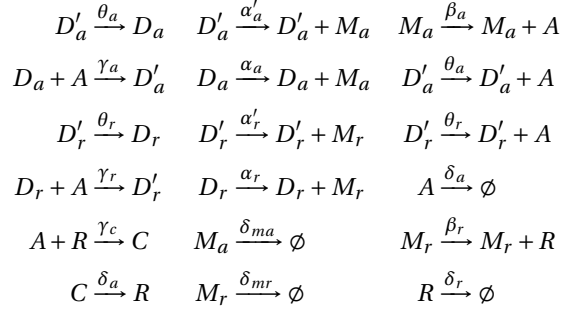**end**

Tasks

3. Numerically compute the integral

$$\int_0^1 x^3 e^{x^2} \, \mathrm{d}x = \frac{1}{2}$$

using Monte Carlo integration as well as the composite trapeziodal rule. Compare errors and run times.[8]

---

[8]The Matlab commands `tic` and `toc` constitute a stopwatch timer.

4. In this final task, we study a circadian rhythm. This system is taken from a paper by Vilar et al.[9] and consists of 9 species and 18 reactions

$$D'_a \xrightarrow{\theta_a} D_a \quad D'_a \xrightarrow{\alpha'_a} D'_a + M_a \quad M_a \xrightarrow{\beta_a} M_a + A$$

$$D_a + A \xrightarrow{\gamma_a} D'_a \quad D_a \xrightarrow{\alpha_a} D_a + M_a \quad D'_a \xrightarrow{\theta_a} D'_a + A$$

$$D'_r \xrightarrow{\theta_r} D_r \quad D'_r \xrightarrow{\alpha'_r} D'_r + M_r \quad D'_r \xrightarrow{\theta_r} D'_r + A$$

$$D_r + A \xrightarrow{\gamma_r} D'_r \quad D_r \xrightarrow{\alpha_r} D_r + M_r \quad A \xrightarrow{\delta_a} \emptyset$$

$$A + R \xrightarrow{\gamma_c} C \quad M_a \xrightarrow{\delta_{ma}} \emptyset \quad M_r \xrightarrow{\beta_r} M_r + R$$

$$C \xrightarrow{\delta_a} R \quad M_r \xrightarrow{\delta_{mr}} \emptyset \quad R \xrightarrow{\delta_r} \emptyset$$

The parameters of the above model are:

α_a	α'_a	α_r	α'_r	β_a	β_r	δ_{ma}	δ_{mr}	δ_a	δ_r	γ_a	γ_r	γ_c	θ_a	θ_r
50	500	0.01	50	50	5	10	0.5	1	δ_r	1	1	2	50	100

Simulate the circadian rhythm using the ODE model as well as an implementation of the stochastic simulation algorithm. Use the same initial concentration/population as proposed by Vilar et al., that is, $D_a = D_r = 1$ and $D'_a = M_a = D'_r = M_r = A = C = R = 0$. Plot trajectories of A, C, and R corresponding to $\delta_r = 0.2$, $\delta_r = 0.08$, and $\delta_r = 0.01$.

---

[9]Vilar, Kueh, Barkai, and Leibler, Mechanisms of noise-resistance in genetic oscillators, *Proc. Nat. Acad. Sci.*, 99(9):5988-5992, 2002.