1 Introduction to modeling by ordinary differential equations, by PDE and by stochastic processes.

The course has a goal to give a relatively elementary and "problem oriented" introduction to three main (from the point of view of the lecturer) approaches to mathematical modeling. The course supplies examples and exercises connecting theory with mathematical models mainly in biology but also in chemistry and in physics.

In the first lecture we are going to use chemical kinetics as the main example for demonstrating various techniques for mathematical modeling. The word kinetics comes from Greek 'kinetikos' that means 'moving' and means both microscopic interactions and the spacial transport of particles.

1.1 Modeling of chemical reactions by ODE. Law of mass action.

Law of mass action. The rate of reactions is proportional to the active concentrations of the reactants.

To illustrate and make clear the meaning of the law consider the irreversible reaction described by

$$A + B \to C.$$

If we denote concentrations of reagents by A, B, C the law of mass action says that

$$\frac{dA}{dt} = -k\,A\,B,$$

where t denotes time and the constant k is called the rate constant of this reaction.

Other two equations corresponding this reaction are

$$\frac{dB}{dt} = -k\,A\,B, \qquad \frac{dC}{dt} = k\,A\,B.$$

There are as many equations as reactants. If the reaction is reversible with forward and backward rate constants k_1 and k_{-1} respectively,

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C \Longrightarrow \frac{dA}{dt} = -k_{\,1} A\,B + -k_{\,-1} C.$$

The fact that one molecule of A combines with a one molecule of B to form C means that reaction is a *bimolecular* one: these are by far the commonest.

Most of *chemical reactions in biology* include reactants of different types having different functions. It is typical that a part of reactants serve as catalysts or inhibitors.They are usually proteins. Such proteins and are called enzymes. They act remarkably efficient in that they operate at very low concentrations. Enzymes react very selectively in combination with definite compounds called substrates (or ligands) and therefore are highly specific.

Enzyme-substrate-product *Michaelis Menten reaction*= (1913) is the following model. Free enzyme and substrate first combine via a reversible reaction a complex which in turn breaks down irreversibly to form the free enzyme again and a product. It can be represented by the following diagram:

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} SE$$

$$SE \overset{k_2}{\to} P + E. \tag{1}$$

We can now use the law of mass action for writing down a system of equations for Michaelis Menten reaction

$$\begin{aligned} \frac{ds}{dt} &= -k_1 s\, e + k_{-1} c, \qquad (2)\\ \frac{de}{dt} &= -k_1 s\, e + (k_{-1} + k_2) c,\\ \frac{dc}{dt} &= k_1 s\, e - (k_{-1} + k_2) c,\\ \frac{dp}{dt} &= k_2 c, \end{aligned}$$

where c - concentration of the complex SE, s - concentration of S, e - concentration of E, p -concentration of the product.

Relevant initial conditions at time $t = 0$ are $c(0) = 0$, $p(0) = 0$, concentrations $e(0)$ and $s(0)$ are given and non-zero.

Adding equations for $\frac{de}{dt}$ and $\frac{dc}{dt}$ we get a *conservation law* for e and c namely

$$e + c = const = e(0) = e_0 \tag{3}$$

So if c is known the conservation law gives e and the equation for $\frac{dp}{dt}$ gives p by integration:

$$p(t) = k_2 \int_0^t c(t')dt' \tag{4}$$

These observations imply a system of two equations for s and c only, namely

$$\begin{aligned} \frac{ds}{dt} &= -k_1 e_0\, s + (k_1 s + k_{-1})c, & (5)\\ \frac{dc}{dt} &= -k_1\, e_0\, s - (k_1 s + k_{-1} + k_2)c & (6) \end{aligned}$$

with initial conditions $s(0) = s_0$, $c(0) = 0$. This system cannot be solved analytically, but can be analyzed and in many interesting cases can be approximately reduced to one also non-linear equation that can be solved analytically. This reduction is done not for mathematical reasons themselves (the system is rather simple), but for clarifying typical qualitative properties of the reaction.

Before analyzing 5 and 6 the equations must be written in a non-dimensional form. This is crucial for the analysis of any mathematical model. Qualitatively it means that relative magnitude of various terms in the equations must be made clear.This only can be done with confidence if all quantities are made dimensionless otherwise the words large and small have no relevance. Therefore *dimensions analysis and scaling of a problem* is a general requirement for a proper mathematical modeling.

For example in 5 each term must have dimension $concentration \times (time)^{-1}$. Thus k_1 must have dimension $concentration^{-1} \times (time)^{-1}$ whereas k_{-1} must have dimension $(time)^{-1}$. A non-dimensional model is independent of the specific system of units used. Introduce the non-dimensional quantities

$$\begin{aligned} \tau &= k_1 e_0 t, \quad \lambda = k_2/k_1 s_0, \quad \kappa = (k_{-1} + k_2)/k_1 s_0, \\ x(\tau) &= \frac{s(t)}{s_0}, \quad y(\tau) = \frac{c(t)}{e_0}, \quad \varepsilon = \frac{e_0}{s_0}. \end{aligned}$$

Substituting new dimensionless variables into the system of equations we get the following non-dimensional system for $x(\tau)$ and $y(\tau)$:

$$\begin{aligned} \frac{dx}{d\tau} &= -x + (x + \kappa - \lambda)y, \\ \varepsilon\frac{dy}{d\tau} &= x - (x + \kappa)y, \\ x(0) &= 1, \quad y(0) = 0. \end{aligned} \tag{7}$$

In most biological situations the ratio of initial enzyme to initial substrate is very small that is $\varepsilon = \frac{e_0}{s_0} \ll 1$. So the system (7) is a *singular perturbation problem* which is recognized here by the fact that setting $\varepsilon = 0$ the order of the system is reduced but the initial conditions cannot in general both be satisfied. If we try to do so we get

$$\begin{aligned} y &= \frac{x}{x + \kappa}, \implies \\ \frac{dx}{d\tau} &= \frac{-\lambda x}{x + \kappa}, \implies \\ x + \kappa \ln x &= 1 - \lambda\tau \end{aligned} \tag{8}$$

where x satisfies the first initial condition in 7 but y does not satisfy the initial condition. The equations (8) are called the Michaelis-Menten kinetic law.

1.2 Asymptotic of solutions to ODE

The very unsatisfactory argument which is frequently used to get the dimensional form of the Michaelis-Menten kinetic law is the following. Since the enzyme is present only in small quantities $de/dt \approx 0$ and so from the conservation law $e + c = const$ it follows that $dc/dt \approx 0$ and therefore from 6 it follows $c = \frac{k_1 e_0 s}{k_1 s + k_{-1} + k_2}$ and the following dimensional form of the Michaelis-Menten kinetic law:

$$\frac{ds}{dt} = \frac{-k_2 e_0 s}{K + s}, \quad c = \frac{k_1 e_0 s}{k_1 s + k_{-1} + k_2}, \quad K = \frac{k_{-1} + k_2}{k_1}. \tag{9}$$

The disadvantage of this naive approach is that the quantitative criteria $\varepsilon = \frac{e_0}{s_0} \ll 1$ for validity of this approximation is lost and also no way of resolving the problem with initial data can be suggested.

Therefore the rigorous treating of the *asymptotics of solutions to differential equations* is an important aspect in mathematical modeling. One chapter of the course will be devoted to asymptotic methods.

1.3 Stationary states and periodic solutions

It is easy to see by numerics that solutions to the system 7 go to a "stable" stationary state independently of initial data. There might be in general stationary states that are not stable: the system runs away from such a state if being turned a bit away from it. A typical example for this from mechanics is the upper position of a pendelum. It is a stationary state, but evidently unstable. There is a well developed *theory of stability for stationary states* of dynamic systems that we are going to study in the course. Relative positions and stability properties of stationary points of a dynamic system determine in many respects also the global properties of the system.

It is now believed in the light of experimental evidence that many cellular processes tend not to stationary states but to oscillatory regimes, that are *periodic solutions* and it is intrinsic rhythm behavior which provides a robust dynamic basis for self-organization in cellular development. Material on oscillatory and in particular on *periodical solutions* to ordinary differential equations constitute another chapter of the course.

We discuss for pedagogical reasons the following model system of reactions called Brusselator (Prigogine 1977):

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} X$$

$$B + X \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} Y + D$$

$$2X + Y \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} 3X$$

$$X \underset{k_{-4}}{\overset{k_4}{\rightleftarrows}} E$$

Here reactants A, B, D, E are kept constant, so the system is opened and the only intermediates are X and Y.

$$\begin{aligned} \frac{dX}{dt} &= k_1 A - k_2 BX + k_3 YX^2 - k_4 X \\ \frac{dY}{dt} &= k_2 BX - k_3 YX^2 \end{aligned}$$

Prigogines Brusselator

The non-dimensional form of this system is

$$\begin{aligned}\frac{du}{d\tau} &= 1-(b+1)u+a\,u^2v\\ \frac{dv}{d\tau} &= bu-a\,u^2v\end{aligned}$$

We will consider a slightly simplified model introduced by Quan et. al.(2002):

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} X$$

$$B + X \overset{k_2}{\rightarrow} Y + D$$

$$2X + Y \overset{k_3}{\rightarrow} 3X$$

Quans model

$$\begin{aligned}\frac{dX}{dt} &= k_1A - k_{-1}X + k_3\,X^2Y \qquad (10)\\ \frac{dY}{dt} &= k_2B - k_3\,X^2Y\end{aligned}$$

Quans model

The non-dimensional form of this system is

$$\begin{aligned}\frac{du}{d\tau} &= a-u+u^2v\\ \frac{dv}{d\tau} &= b-u^2v \qquad (11)\end{aligned}$$

where $u=\sigma X,\ v=\sigma Y$, $\sigma=\sqrt{k_3/k_{-1}}, \tau=k_{-1}t$. The amount of parameters in the system is reduced to two:$a=(k_1/k_{-1})\sqrt{k_3/k_{-1}}A$, $b=\sqrt{k_2/k_{-1}}B$.

Properties of the system depends essentially on what kind of stationary points there are in the phase plane (u,v).By stationary points we mean points where $\frac{du}{d\tau}=0$, $\frac{dv}{d\tau}=0$ because of the systems

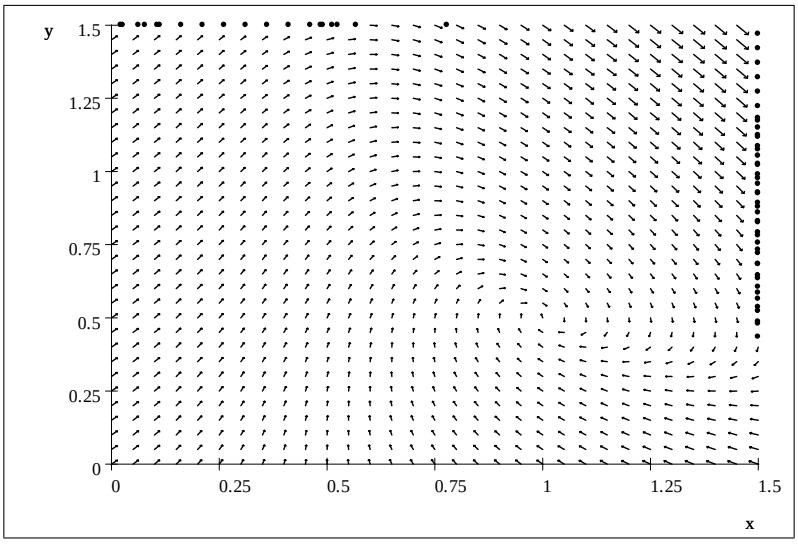
of equations, namely in this particular case right hand sides are zero: $a - u + u^2 v = 0$ and $b - u^2 v = 0$.

The system has just one stationary point (u^0, v^0) :

$$u^0 = a + b, \quad v^0 = \frac{b}{(a+b)^2}.$$

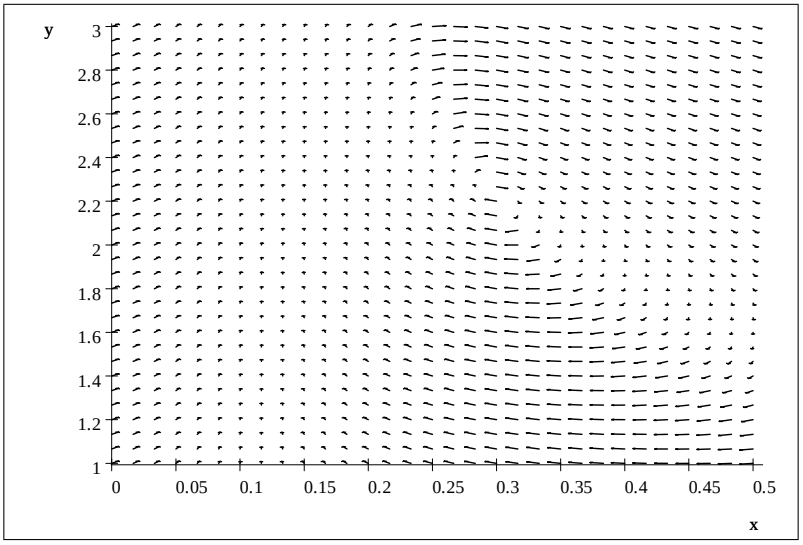
The graph below shows the directions of velocities in the *phase plane* for the case $a = b = 0.5$

with the stable stationary point $u^0 = 1$, $v^0 = 0.5$.



Another example is for $a = 0.1; b = 0.2$,

with the unstable stationary point $u^0 = 0.3$; $v^0 = 0.2/(0.1 + 0.2)^2 = 2.\,222\,2$



We consider the linear approximation of the right hand sides $a - u + u^2 v$ and $b - u^2 v$ by first terms of the Taylor expansion around the stationary point (u^0,v^0) and the corresponding linear system of ODE for deviations $x = u - u^0$, $y = v - v^0$ of u, v from (u^0, v^0),. In our particular case it looks as

$$\frac{d}{d\tau}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{b-a}{b+a} & (a+b)^2 \\ -\frac{2b}{b+a} & -(a+b)^2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}. \tag{12}$$

This approximate system helps to investigate the *stability properties of the stationary point*.

The stationary point is called *stable* if for any initial data close to it the solution will stay close to it. There is an algebraic theory on stability of linear systems of ODE and also a theory about connection between the stability properties of stationary points to non-linear systems of ODE and stability of their linear approximations. In many times (but not always!) the phase portraits are locally similar.

Properties of linear system in the form $\frac{d\overrightarrow{r}}{dt} = A\overrightarrow{r}$ the stability of the stationary point at the origin is totally determined by eigenvalues and eigenvectors of the matrix A. Eigenvalues are solutions of the characteristic equation $\det(A - \lambda I) = 0$. Eigenvalues with positive real parts are the sign of instability. In our example with 2-dimensional system the characteristic equation has the form $\lambda^2 - Trace(A) + \det(A) = 0$. The determinant of the matrix A 12

is $(a+b)^2 > 0$ that means that real parts of the eigenvalues are the same. The stability is therefore determined by the trace of the matrix, the sum of eigenvalues:

$Trace(A) = \frac{(b-a)}{(b+a)} - (a+b)^2 = (a+b)^2 \left(\frac{b-a}{(b+a)^3} - 1 \right) = (a+b)^2(\alpha - 1)$.

$\alpha = \frac{b-a}{(b+a)^3} = \frac{1-a/b}{b^2(1+a/b)^3}$.

When the trace is negative, the stationary point is stable. Therefore for $\alpha < 1$ the stationary point is stable and for the positive - unstable. Looking on examples above we see that for $a = 0.5$, $b = 0.5$, $\alpha = 0 < 1$ (stationary point is stable) ; for $a = 0.1$, $b = 0.2$, $\alpha = 3.703\,7 > 1$ (stationary point is unstable).

More refined analysis shows that solutions of this nonlinear system cannot go infinitely far from some neighborhood of (u^0, v^0). A *theorem by Poincare and Bendixson* implies that in the case of two-dimensional systems of ODE the presence of a unique unstable stationary point in a domain that cannot be left by solutions implies the existence of a periodical (oscillating) solution. Therefore the system of equations 10 has a periodic solution for $\alpha > 1$.

1.4 Stochastic models and processes.

The analysis of chemical kinetics above is based on the law of mass action. It requires that $\sim 10^4$ molecules take part in a reaction. It is not always the case in biology. The reason is that a chemical reaction is a stochastic process that happens with some probability. If we have a large number of well mixed reactants the amount of reacting molecules will be proportional to their densities (the law of mass action). If the number of of molecules participating in the reaction is small one cannot longer describe the system in a deterministic macroscopic fashion. Rather one describes the system as stochastic and microscopic by the *probability* of having n_x of X and n_y of Y at time t in terms of a probability distribution function $P(n_x, n_y, t)$.

The stochastic kinetics can be depicted by the scheme below:

$$
\begin{array}{ccccccc}
 & \searrow & & & & & \\
\rightleftarrows & (n-1,m+1) & \overset{k_1 n_A}{\rightleftarrows} & (n,m+1) & \overset{k_1 n_A}{\rightleftarrows} & (n+1,m+1) & \\
 & \searrow & \overset{k_{-1}(n)}{\underset{k_3(n-2)(n-1)(m+1)}{\searrow}} & \uparrow{\scriptstyle k_2 n_B} & \overset{k_{-1}(n+1)}{\underset{k_3(n-1)n(m+1)}{\searrow}} & & \searrow \\
\rightleftarrows & (n-1,m) & \overset{k_1 n_A}{\rightleftarrows} & (n,m) & \overset{k_1 n_A}{\rightleftarrows} & (n+1,m) & \rightleftarrows \\
 & & \overset{k_{-1}(n)}{\underset{k_3(n-2)(n-1)(m)}{\searrow}} & \uparrow{\scriptstyle k_2 n_B} & \overset{k_{-1}(n+1)}{\underset{k_3(n-1)n(m+1)}{\searrow}} & & \searrow \\
 & & \rightleftarrows & (n,m-1) & \overset{k_1 n_A}{\underset{k_{-1}(n+1)}{\rightleftarrows}} & (n+1,m-1) & \rightleftarrows \\
 & & & & & & \searrow
\end{array}
$$

At any time t the system has the numbers (n_x, n_y) of X and Y molecules. The random variables n_x and n_y can change only by 1 or -1. Therefore, the stochastic kinetics resembles a random walk on the two-dimensional lattice.

From each state (a point in the lattice) the system can undergo transition governed by specific rate constants to any of four possible neighboring states. Therefore writing the kinetic equation for this stochastic system is straightforward:

$$
\begin{aligned}
\frac{dP(n_x,n_y)}{dt} = & \\
& -[k_1 n_A + k_{-1} n_x + k_2 n_B + k_3 n_x (n_x-1) n_y] P(n_x,n_y) \\
& + k_1 n_A P(n_x-1,n_y) + k_2 n_B P(n_x,n_y-1) \\
& + k_{-1}(n_x+1) P(n_x+1,n_y) \\
& + k_3 (n_x-1)(n_x-2)(n_y+1) P(n_x-1,n_y+1) \quad (13)
\end{aligned}
$$

Stochastic processes of such type are called the birth-death processes in the theory of probability.They are also known as stochastic compartmental systems in biology.

The system 13 is a large (or indefinite!) system of ordinary differential equations that can in principle be solved straightforwardly by a generic numerical method. On the other hand it might become very time and memory consuming even for a modern computer.

1.5 Fokker-Planck equation and other PDE models

We can consider a smooth probability distribution function $P(n_x, n_y)$

defined not only in the nodes with integer coordinates but in the whole plane.

If the function $P(n_x, n_y)$ is smooth we can notice that

$$\begin{aligned}
P(n_x+1,n_y)-P(n_x-1,n_y) &\approx 2\frac{\partial P(n_x,n_y)}{\partial n_x}\\
P(n_x,n_y+1)-P(n_x,n_y-1) &\approx 2\frac{\partial P(n_x,n_y)}{\partial n_y}\\
P(n_x+1,n_y)-2P(n_x,n_y)+P(n_x-1,n_y) &\approx \frac{\partial^2 P(n_x,n_y)}{\partial n_x^2}\\
P(n_x,n_y+1)-2P(n_x,n_y)+P(n_x,n_y-1) &\approx \frac{\partial^2 P(n_x,n_y)}{\partial n_y^2}
\end{aligned}$$

The equation 13 for the stochastic process can be in this case approximated by an *equation with partial derivatives (PDE)* that have certain advantages both for analytic treating and for numerical solution.

To clarify how to use this idea we consider first a simpler system with reactions described by the scheme

$$\begin{array}{ccccc}
 & & (n,m+1) & & \\
 & & {\scriptstyle\kappa}\uparrow\downarrow{\scriptstyle\kappa} & & \\
(n-1,m) & \underset{\kappa}{\overset{\kappa}{\rightleftarrows}} & (n,m) & \underset{\kappa}{\overset{\kappa}{\rightleftarrows}} & (n+1,m)\ .\\
 & & {\scriptstyle\kappa}\uparrow\downarrow{\scriptstyle\kappa} & & \\
 & & (n,m-1) & &
\end{array}$$

In this case the equation for probability $P(n_x, n_y)$ is particularly simple:

$$\frac{dP(n_x,n_y)}{dt} = \kappa\left[-4P(n_x,n_y)+P(n_x+1,n_y)+P(n_x-1,n_y)+P(n_x,n_y+1)+P(n_x,n_y-1)\right] \tag{14}$$

Using the above relation for finite differences we immediately come to the approximate differential equation with partial derivatives:

$$\frac{dP(n_x,n_y)}{dt} = \kappa\left[\frac{\partial^2 P(n_x,n_y)}{\partial n_x^2}+\frac{\partial^2 P(n_x,n_y)}{\partial n_y^2}\right] = \kappa div(\nabla P), \tag{15}$$

that is the classical diffusion equation.

Doing similar procedure for our main example we also get a second order PDE. Introducing non-dimensional variables u, v, τ, a as for ODE before and $\sigma = u/n_x = v/n_y$ we get an approximate PDE for continuous $P(u, v)$ in the form

$$\frac{\partial P(u, v, \tau)}{\partial \tau} = \nabla(\boldsymbol{D}\,\nabla P - \overrightarrow{\boldsymbol{F}}\,P) \tag{16}$$

where $u \geq 0$, $v \geq 0$, the matrix $\boldsymbol{D}$ and the vector $\overrightarrow{\boldsymbol{F}}$ are of the following form:

$$\boldsymbol{D} = \frac{\sigma}{2}\begin{bmatrix} a + u + u^2 v & -u^2 v \\ -u^2 v & b + u^2 v \end{bmatrix},$$

$$\overrightarrow{\boldsymbol{F}} = \begin{bmatrix} a - u + u^2 v - \sigma(1/2 + 2uv - u^2/2) \\ b - u^2 v + \sigma(2uv - u^2/2) \end{bmatrix}.$$

This equation is called sometimes *Fokker-Planck equation*. Similar type of equations appear in the description of transport of particles or energy in stochastic environments with the simplest example the diffusion equation. In such situations coordinates u, and v are the space coordinates and the function $P(u, v, \tau)$ has the sense of the density of particles or temperature at the point (u, v) at time τ.

For $\sigma = 0$ the motion by 16 reduces to the deterministic motion as for the law of mass action:

$$\frac{\partial P(u, v, \tau)}{\partial \tau} = \nabla \overrightarrow{\boldsymbol{F}}\,P, \tag{17}$$

$$\overrightarrow{\boldsymbol{F}} = \begin{bmatrix} a - u + u^2 v \\ b - u^2 v \end{bmatrix}. \tag{18}$$

$\overrightarrow{\boldsymbol{F}}$ is a right hand side of the corresponding deterministic system 11. The meaning of the equation 17 is simple: the distribution function $P(u, v, \tau)$ is constant along the trajectories of the deterministic system 11 with right hand side $\overrightarrow{\boldsymbol{F}}$.

1.6 Monte-Carlo method

One more approach to modeling stochastic processes and chemical kinetics in particular (two other were the equation for distribution

function and the Fokker-Planck equation) is the so called *Monte Carlo method* or the stochastic modeling method.

It consists shortly of the following procedure. One models not the whole distribution function $P(n_x, n_y, \tau)$ but many individual "stories" or stochastic trajectories of molecules in the phase plane undergoing stochastic transformations corresponding to the dynamics described by the equation 13. After getting these stochastic trajectories one looks for each time t on how many of these individual trajectories come to a particular node (n_x, n_y) at this time. The mathematical theory of *Monte Carlo method* describes how the modeling of the stochastic trajectories should be organized so that the averaged over the trajectories population of the node (n_x, n_y) in the lattice approximates the value of the distribution function $P(n_x, n_y)$.

We give here a description of the method by Gillespie for stochastic modeling of chemical reactions and apply it to the same type of problems as we considered above.

The fundamental hypothesis is that for a reaction R_μ of a set of certain reactants, there is a constant c_μ such that

$c_\mu \delta t$ is equal up to the first order in δt to averaged probability that such particular combination of reactant molecules will react in the next time interval δt. For example if R_μ is of type $2X+Y$ as one of reactions before then the probability that such set of molecules will react in th next time interval δt is equal to $c_\mu \delta t + o(\delta t)$.

The connection between c_μ and the rate constants k_μ in ODE equations for the number of molecules is the following. We remind that the reaction rate k_μ is the average reaction rate in unit volume divided by the densities of all reactants. The averaged reaction rate in the particular case is c_μ times the average number of distinct combinations of $2X$ an Y that is in this particular case equal to $< X(X-1)/2Y >$ where $< ... >$ denotes the statistical average.

Therefore $k_\mu =< X(X-1)/2Y > c_\mu/V$. Rewriting this expression for densities $x = X/V$ and $y = Y/V$ and taking into account that for a large system $X \approx X-1$ we come to the expression $k_\mu =< x^2 y/2 > V^2 c_\mu/(< x >^2 < y >)$.Using that for a large number of molecules $< x^2 y > \approx < x >^2 < y >$ we come to the expression $k_\mu = V^2 c_\mu/2$.For other combinations of molecules in a reaction one can get connection formulated in a similar way.

The main quantity in a stochastic modeling method by Gillespie is *the reaction probability function* $P(\tau, \mu)$ depending on a continuous parameter τ and a discrete number μ, that is the number of reaction. $P(\tau, \mu)$ is the probability that starting at time t the next reaction in the volume V will occur in the time interval $(t+\tau, t+\tau+\delta\tau)$ and will be the reaction R_μ with number μ.

Similar approaches were independently formulated also for other

types of kinetics: transport of neutrons in nuclear reactors, light scattering in atmospheres of stars, transport in porous media etc.

Let h_μ be the average number of distinct combinations of reactant molecules for reaction R_μ at time t in the volume V.

Therefore $h_\mu c_\mu \delta t$ is to the first order of δt the probability that the reaction R_μ will occur in V in the next interval δt.

$h_\mu = XY$ for a bimolecular reaction $X + Y$ with two different types of molecules; X and Y;

$h_\mu = X(X-1)/2$ for a bimolecular reaction $X + X$ with two similar molecules;

$h_\mu = X(X-1)Y/2$ for a three-molecular reaction $2X + Y$. Other combinations can also be calculated by elementary combinatorics.

Therefore the probability $P(\tau, \mu)d\tau$ above is the product of probability $P_0(\tau)$, the probability at time t, that no reactions will occur in the time period $(t, t+\tau)$, and

the probability $h_\mu c_\mu \delta\tau$, the subsequent probability that an R_μ reaction will occur in the next differential time interval $(t+\tau, t+\tau+d\tau)$.

$$P(\tau, \mu)d\tau = P_0(\tau) h_\mu c_\mu \delta\tau$$

To calculate $P_0(\tau)$ we divide interval $(t, t+\tau)$ into $K >> 1$ small subintervals of length ε. The provability that no reactions of any type R_ν will occur in one subinterval is

$1 - \sum h_\nu c_\nu \varepsilon + o(\varepsilon) = 1 - \sum h_\nu c_\nu \tau/K + o(\tau/K)$. By multiplying K such expressions we get the following expression for $P_0(\tau)$:

$$P_0 = \left[1 - \sum h_\nu c_\nu \tau/K + o(\tau/K)\right]^K \to \exp\left[-\sum h_\nu c_\nu \tau\right]$$

$$P(\tau, \mu)d\tau = h_\mu c_\mu \exp\left[-\sum h_\nu c_\nu \tau\right]$$

We illustrate the idea of of using the modeling algorithm on an example:

Example:

1. Betrakta följande reaktioner: $X + Z \underset{c_2}{\overset{c_1}{\rightleftarrows}} W, \qquad W \overset{c_3}{\to} P$

   där $c_i\, dt$ är sannolikheten att under tiden dt reaktionen med index i skall äga rum. $i = 1, 2, 3$.

a) Ange differentialekvationer för antalet partiklar för dessa reaktioner. **(2p)**

b) Ange formler för algoritmen som skulle stokastiskt modellera dessa reaktioner enligt Gillespies metod. **(2p)**

Lösning: a)

$X' = -c_1 XZ + c_2 W$

$Z' = -c_1 XZ + c_2 W$

$W' = c_1 XZ - (c_2 + c_3)W$

$P' = c_3 W$

b) **Gillespies metod.**

$P(\tau, \mu)d\tau$ är sannolikheten att ha en reaktion av typ μ under ett tidsintervall $d\tau$ efter tiden τ då inga andra reaktioner ägde rum.

$P(\tau, \mu) = P_0(\tau) h_\mu c_\mu d\tau.$

Här $P_0(\tau)$ är sannolikheten att inga reaktioner skall äga rum under tiden τ.

$h_\mu c_\mu d\tau$ är sannolikheten att just reaktionen μ skall gå under tidsintervall $d\tau$.

h_μ är antalet olika kombinationer av partiklar för aktuella X, Z, W, P som kan delta i rekationen μ. För reaktionen 1 i exemplet är det $h_1 = X \cdot Z$, för reaktion 2 är det $h_2 = W$, för reaktion 3 är det $h_3 = W$.

$P_0(\tau) = exp(-a\tau)$ med $a = \sum_{\mu=1}^{3} h_\mu c_\mu$.

Algoritmen för att stokastiskt modellera givna reaktioner.

0) inicialisera variabler X, Z, W, P och tiden $t = 0$.

1) Beräkna h_i, a.

2) Generera två slumptal r och p jämt fördelade över intervallet $(0, 1)$.

Tag tiden τ före nästa reaktion som $\tau = 1/a \ln(1/r)$.

Välj nästa reaktion μ så att $\sum_{i=1}^{\mu} h_i c_i \leq p\, a \leq \sum_{i=\mu+1}^{3} h_i c_i$.

3) Tillägg τ till tidsvariabeln t. Ändra antalet partiklar enligt väljad reaktion:

$\mu = 1 \quad \rightarrow X = X - 1,\ Z = Z - 1,\ W = W + 1.$

$\mu = 2 \quad \rightarrow X = X + 1,\ Z = Z + 1,\ W = W - 1.$

$\mu = 3 \quad \rightarrow P = P + 1,\ W = W - 1.$

3) Om tiden är större änd den maximala tiden vi är intresserade av, sluta beräkningar, annars gå till steg 1.