On the mathematical structure and algorithmic implementation of biochemical network models

Paola Lecca

Abstract—Modeling and simulation of biochemical reactions is of great interest in the context of system biology. The central dogma of this re-emerging area states that it is system dynamics and organizing principles of complex biological phenomena that give rise to functioning and function of cells. Cell functions, such as growth, division, differentiation and apoptosis are temporal processes, that can be understood if they are treated as dynamic systems. System biology focuses on an understanding of functional activity from a system-wide perspective and, consequently, it is defined by two hey questions: (i) how do the components within a cell interact, so as to bring about its structure and functioning? (ii) How do cells interact, so as to develop and maintain higher levels of organization and functions? In recent years, wet-lab biologists embraced mathematical modeling and simulation as two essential means toward answering the above questions. The credo of dynamics system theory is that the behavior of a biological system is given by the temporal evolution of its state. Our understanding of the time behavior of a biological system can be measured by the extent to which a simulation mimics the real behavior of that system. Deviations of a simulation indicate either limitations or errors in our knowledge.

The aim of this paper is to summarize and review the main conceptual frameworks in which models of biochemical networks can be developed. In particular, we review the stochastic molecular modelling approaches, by reporting the principal conceptualizations suggested by A. A. Markov, P. Langevin, A. Fokker, M. Planck, D. T. Gillespie, N. G. van Kampfen, and recently by D. Wilkinson, O. Wolkenhauer, P. S. Jöberg and by the author.

I. INTRODUCTION

We can distinguish four fields of application of mathematical models to biology: 1. population dynamics; 2. cell and molecular biology; 3. physiological systems; 4. spatial modeling.

Different formalisms are usually applied to describe the dynamics of these different fields. In general the mathematical structure of a model of a physical phenomenon depends on the nature of the determination, of the time, and of the space state. The determination of a model can be deterministic or stochastic, or also hybrid deterministic and stochastic. The time course can be continuous of discrete, and the state space can also be continuous of discrete. The combination of the these characteristics give rise to different mathematical approaches to the modeling the dynamics of the phenomenon. Here following we list some of the most common mathematical formalism and approaches to specify the dynamics of a system with respect to the four categories listed above.

1) **Deterministic processes** (Newtonian dynamical systems). A fixed mapping between an initial state and a final state. Starting from an initial condition and moving forward in time, a deterministic process will always

Authors affiliation: The Microsoft Research - University of Trento, CoSBi, Trento (Italy). E-mail:lecca@cosbi.eu

TABLE I
CLASSES OF BIOLOGICAL PHENOMENA AND MOST USED FORMALISMS TO DESCRIBE THEM.

Population dynamics	Deterministic processes Ordinary differential equations.
Cell and molecular biology	Stochastic processes: Jump Markov processes and continuous Markov processes.
Physiological systems Spatial modeling (epidemiology)	Deterministic processes Deterministic processes. Partial differential equations.

generate the same trajectory and no two trajectories cross in state space.

- *Ordinary differential equations* (Continuous time. Continuous state space. No spatial derivatives.)
- *Partial differential equations* (Continuous time. Continuous state space. Spatial derivatives.)
- *Maps* (Discrete time. Continuous state space)

2) **Stochastic processes** (random dynamical systems) A random mapping between an initial state and a final state, making the state of the system a random variable with a corresponding probability distribution.

- *Jump Markov process* – Master equation (Continuous time with no memory of past events. Discrete state space. Waiting times between events discretely occur and are exponentially distributed.)
- *Continuous Markov process* – stochastic differential equations or a Fokker-Planck equation (Continuous time. Continuous state space. Events occur continuously according to a random Wiener process.)
- *Non-Markovian processes* – Generalized master equation (Continuous time with memory of past events. Discrete state space. Waiting times of events (or transitions between states) discretely occur and have a generalized probability distribution.)
- *Stochastic simulation algorithms*: Gillespie exact simulation and StochSim

3) Hybrid stochastic/deterministic systems (metabolic and signaling pathways)

- *Gillespie τ-leap algorithm* – Differential equations for the simulation of fats reactions and Gillespie algorithm for the exact simulation of slow reactions.

With regard to modeling the chemistry of intracellular dynamics, the two most popular frameworks are the deterministic modeling and the stochastic modeling. The deterministic modeling is based on the construction of a set of rate equations to describe the reactions in the biochemical pathways of interest. These rate equations are ordinary differential equa-

tions with concentrations of chemical species as variables. In general, given the complexity of biological pathways we have to deal with non-linear differential equations. Deterministic simulations produce the time course of the concentrations by solving the differential equations.

In its most known aspect stochastic modeling involves the formation of a set of chemical master equations with probabilities as variables [32]. Stochastic simulation produces counts of molecules of some chemical species as realizations of random variables drawn from the probability distribution described by the master equations.

Which framework is appropriate for a given biological system is not only a question of what biological phenomena are investigated but also influenced by assumptions one makes to simplify the analysis. For instance, the scale, and thus the level of granularity at which a phenomenon is investigated may be parameters to choose a deterministic of a stochastic approach [26].

In this paper we firstly review the deterministic approach to chemical kinetics, then we examine thoroughly and critically discuss the main concepts of stochastic chemical kinetics and highlight the necessary re-formulations to adapt them to the biological simulation context.

II. CHEMICAL REACTIONS

Here we report some introductory considerations of Wilkinson [33]. There are many ways one could represent a model of a biological system. Biologist have traditionally favored diagrammatic schemes conveying qualitative information of the depicted mechanisms. At the other extreme, applied mathematicians traditionally prefer to work with systems of ordinary or partial differential equations. These are more precise and fully quantitative, but also have a number of disadvantages. The differential equations models are too low level description, as they not only encode the essential features of the model, but also a wealth of accompanying baggage associated with a particular interpretations of chemical kinetics, that is not always well suited to application in the molecular biology context. Between these two extremes, the biochemist view systems as networks of coupled chemical reactions. These networks are sufficiently general that they can be simulated in different ways using different algorithms depending on assumptions made about the underlying kinetics. Furthermore, they are sufficiently detailed so that, once the kinetics have been specified, they can be used directly to construct full dynamic simulations of the system behavior on a computer.

A general chemical reaction takes the form

$$s_1 X_1 + s_2 X_2 + \cdots + s_n X_n \longrightarrow r_1 Y_1 + r_2 Y_2 + \cdots + r_m Y_n \quad (1)$$
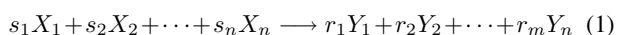
where n is the number of reactants and m is the number of products. X_i represents the ith reactant molecule and Y_i is the jth product molecule. s_i is the number of molecules of X_i consumed in a single reaction step, and r_j is the number of molecules of Y_j produced in a single reaction step. The coefficients s_i and r_j are known as *stoichiometries*, and they

are usually (tough not always) integer numbers. there is no assumption that X_i and Y_j are distinct, i. e. a given molecule can be both consumed and produced in by a single reaction. In particular, if a chemical species occurs on both the left and the right hand side, is referred to as a *modifier*. In this case the reaction will have no effect on the amount of this species, that is usually included in the system because the velocity at which the reaction proceeds depends on the level of this species.

Let n_j be the number of molecules of species X_j The reaction equation describe precisely which chemical species react together and in what proportions, along with what is produced. for instance consider the dimerization of a molecule M, that is written as follows.

$$2M \longrightarrow M_2$$

Two molecules of P react together to produce a single molecule of P_2. Here P has a stoichiometry of 2 and P_2 has a stoichiometry of 1, that usually is not written. Similarly, the reaction for the dissociation of the dimer is written as

$$M_2 \longrightarrow 2M$$

A reaction that can happen in both directions is known as *reversible*. Reversible reactions are quit common in biology. They are written explicitly adding a reverse arrow for the backward reaction. This notation is simply a shorthand for the two separate reaction processes taking place. In the context of the stochastic models to be studied in this thesis, it will be not acceptable to replace the two separate reactions by a single reaction proceeding with a velocity given by some kind of combination of the velocities of the two separated reactions.

$$2M \rightleftharpoons M_2$$
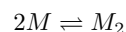
III. KINETICS OF CHEMICAL REACTIONS

Chemical kinetics is concerned with the time-evolution of a reaction system specified by a set of coupled chemical reactions. In particular, it is concerned with the system behavior away from equilibrium. Although the reaction equations capture the key interactions between the competing species, on their own they are not enough to determine the full system *dynamics*. Solving the dynamics of a chemical system means solving the following general problem: if a fixed volume V contains a spatially uniform mixture of N chemical species which can interact through M chemical reaction channels, then given the numbers of molecules of each species present at some initial time, what will these molecular population levels be at any later time?

For answering this question we need to know the *rates* at which each of the reactions occurs together with the initial concentration of the reacting species. The rate of a reaction is a measure of how concentration of the involved substances changes with time. Consider a closed volume V containing a mixture of chemical compounds X_j $(j = 1, 2, \cdots, J)$ and a typical reaction as in the following

$$s_1 X_1 + s_2 X_2 + \cdots \longrightarrow r_1 X_1 + r_2 X_2 + \cdots \quad (2)$$
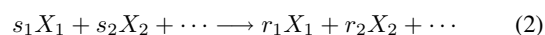
Let x_j be the number of molecules X_j. It is convenient to represent the set $\{n_j\}$ geometrically by a vector $\vec{n}$ in a J-dimensional *state-space*. The integral values of x_j constitute a lattice. Every lattice point in the octant of non-negative values corresponds to a state of the mixture and vice versa (Fig. III).

Fig. 1. The state space of a binary mixture.

The state of the mixture changes when a chemical reaction occurs. Both sides can be written as a sum over all j when zero values of s_j and r_j are admitted.

$$\sum s_j X_j \longrightarrow \sum r_j X_j$$

If for any k one has $s_k = r_k \neq 0$ the corresponding X_k is a catalyst. If $r_k > s_k > 0$ then X_k is an autocatalyst. s_j is the actual number of molecules needed for a reactive collision. A reaction that proceeds through intermediate steps (chain reaction) has to be written as a sequence of single collision reactions, the intermediate products being included as separate items among the X_j. As three-body collisions are rare to meet in practice only reactions with $\sum s_j$ equal to 1 or 2; or possibly 3 if a catalyst is involved.

Each reactive collision of type (2) changes the state $\{x_j\}$ of the mixture into $x_j + s_j - r_j$. In the geometrical representation it means that it changes the state vector $\vec{x}$ by adding to it a vector $\vec{v}$ with components $v_j = r_j - s_j$. As the reaction proceeds the state vector runs over a sequence of lattice points lying on a straight line. This line cannot extend to the infinity and must therefore end on one of the boundaries of the physical octant.

The reverse reaction

$$\sum r_j X_j \longrightarrow \sum s_j X_j$$

will have instead the effect of subtracting $\vec{v}$ from the state vector. Thus, starting from an initial state $\vec{x}_0$, the direct and inverse reactions together cause the state vector to move over a discrete chain of lattice points lying on a straight line between two boundaries of the physical octant. The accessible points are

$$\vec{x} = \vec{x}^0 + \xi \vec{v} \tag{3}$$

where ξ takes all integer values between an upper and a lower bound.

Suppose now that another reactions s_j', r_j' and its reverse are possible. Starting from $\vec{x}_0$ a second chain of lattice points becomes accessible. Together with the previous reaction, a network of points can now be reached,

$$\vec{n}_0 + \xi \vec{v} + \xi' \vec{v}' \qquad (\xi, \ \xi' = \ldots, -1, 0, 1, \ldots) \tag{4}$$

When in this way all possible reactions are taken into account a sublattice is generated of points accessible from $\vec{x}^0$. Since $\sum_j x_j$ is bounded, it cannot cover all the octant. As the

reactions take place in a closed volume, there is no other way by which n_j can vary. Thence, this bounded sublattice is the set of all accessible states of the systems. The physical octant decomposes in such sublattices and the system is confined to that sublattice on which its initial state $\vec{x}^0$ happens to lie. Using the expression (4), it is possible to parametrize the accessible sublattice in the following way. Each possible reaction ρ has a vector $\vec{v}^{(\rho)}$, and by construction all lattice points accessible from $\vec{x}^0$ are

$$\vec{x} = \vec{x}^0 + \sum_\rho \xi_\rho \vec{v}^{(\rho)} \tag{5}$$

Each parameter ξ_ρ takes the integer values $\ldots, -2, -1, 0, 1, 2, \ldots$ and it is called *degree of advancement*, because it indicates how far the reaction ρ has advanced[1].

Suppose that representation (5) is unique, i. e. for each $\vec{x}^0$ each accessible point $\vec{x}$ is represented by a single set of values $\{\xi_\rho\}$. If that is so, then (5) maps the accessible sublattice onto the integral value lattice in the space with coordinates ξ_ρ. Each lattice point in the accessible part of this space corresponds to a one and only one state of the mixture. Each reactive collision corresponds to a unit step parallel to one of the coordinates axes ξ. However, in general, there is no reason why (5) should be unique. There may be two different sets of ξ_ρ that lead from $\vec{x}^0$ to the same $\vec{x}$. That implies that there is a set of integers ζ_ρ, not all zero, such that

$$\sum_\rho \zeta_\rho \vec{v}^{(\rho)} \tag{6}$$

In this case it is still possible to find a smaller set of lattice vectors $\vec{w}^{(\rho)}$, such that each point of the accessible sublattice is uniquely represented by

$$\vec{x} = \vec{x}^0 + \sum_\rho \eta_\rho \vec{w}^{(\rho)}, \tag{7}$$

with integer η_ρ. Each lattice point in the space with coordinates η_ρ corresponds to one and only one state of the mixture, but while in ξ-state space reactions correspond to unit steps, in η-state space they do not. Hence not much has been gained with respect to the original representation in the space of state vectors $\vec{x}$.

The reactions that are possible in a closed volume are restricted by conservation laws for the atoms involved. Let α label the various kinds of atoms and suppose X_j contains m_j^α atoms of kind α, where $m_j^\alpha = 0, 1, 2, \ldots$. Then the stoichiometric coefficients of (2) obey for each α

$$\sum_j s_j m_j^\alpha = \sum_j r_j m_j^\alpha$$

Since this holds for all reactions, the accessible sublattice lies entirely on the intersection of hyperplanes given by

$$\vec{x} \cdot \vec{m}^\alpha = C^\alpha \tag{8}$$

[1]Other names for ξ that are not rare to find in literature are "progress variable"[32], "extent of the reaction", and "reaction parameter".

Fig. 2.   Accessible states for the reactions $2A \rightleftharpoons 2B$ with C = 7.

where C^α is the total number of available atoms α.

The conservation laws (8) are not all necessary independent. If a group of different atoms are bounded in molecule through all reactions, it gives rise to a single conservation law. For example the reaction

$$2NO + Cl_2 \rightleftharpoons 2NOCl$$
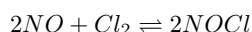
involves three kinds of molecules, but the conservation laws for N and O coincide, because N and O atoms stay together in both direct and inverse reactions. In addition to the laws expressing the conservation of atoms, there may be also other conservation laws. For instance if X_k only occurs as a catalyst the corresponding stoichiometric coefficient is conserved by itself.

All conservation laws together define a linear subspace of lattice points. The *accessible* subspace lies in this subspace and usually it is identical with it, but not necessarily so. A counterexample would be

$$2A \rightleftharpoons 2B$$
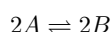
in which two molecules X by colliding may change into a different modification Y. The conservation law for this reaction is

$$x_A + x_B = C$$

and it defines a straight line in the 2-dimensional state space (Fig. 2), but only every other lattice point is accessible from a given $\vec{x}^0$.

A. Mass-action kinetics

The rate of a reaction is a measure of how the concentrations of the involved substances changes with time. For the rate at which a reaction as (2) occurs, one takes the Van't Hoff expression

$$k \cdot \Pi_{j=1}^{J} c_j^{s_j} \qquad (9)$$

Here k is a constant, which involves the cross-section for a collision of the required molecules, time the probability for the collision to result in a reaction. This probability is calculated as the product of the reactants concentrations, $c_j = x_j/V$, raised to the power of their stoichiometries. The Van't Hoff expression in (9) gives the number of collision per unit time per unit volume in which $\{x_j\} \longrightarrow \{x_j + s_j - r_j\}$. The rate equations are therefore

$$\frac{dx_i}{dt} = Vk(r_i - s_i)\Pi_{j=1}^{J}\left(\frac{x_j}{V}\right)^{s_j} \qquad (10)$$

This equation is not an universal truth, but holds when the following physical requirements are satisfied.

1) The mixture must be homogeneous, so that the its density at each point of V equals x_j/V.

2) The elastic, non-reactive collisions must be sufficiently frequent to ensure that the Maxwell velocity distribution is maintained. Otherwise the collision frequency could not be proportional to the product of densities, but more details of the velocities distribution would enter. This requirement will be satisfied in the presence of a solvent or an inert gas.

3) The internal degrees of freedom of the molecules are also supposed to be in thermal equilibrium, with the same temperature T as the velocities. Otherwise the fraction of collisions that result in a reaction would depend on the details of the distribution over internal states, and not just on the concentrations. Long-lived excited states, however, may be taken into account by listing them among the X_j as a separate species, but a clear-cut difference in time scales is indispensable.

4) The temperature must be constant in space and time in order that one may treat the reaction rate coefficients as constants.

Although these assumptions may be not very realistic in many actual chemical reactions, they do not violate any physical law and their validity can therefore be approximated to any desired accuracy in suitable experiments. They assure that the state of the mixture is fully described by the state vector $\vec{x}$.

In the following two sections, we present two simulations of the chemical kinetics specified by Eq. (10): the first is the Lotka-Volterra system and the second is the enzymatic Michaelis-Menten catalysis. These two examples allow us to introduce the main concepts related to the analysis of differential equation and to the time-scale analysis methods used to simply the molecular models.

B. Example 1: the Lotka-Volterra system

Chemical kinetics is concerned with the time-behavior of a system of coupled chemical reactions away from the equilibrium. As example let consider the Lotka-Volterra (LV) predator-prey system for two interacting species [33].

$$\begin{aligned} Y_1 &\longrightarrow 2Y_1 \\ Y_1 + Y_2 &\longrightarrow 2Y_2 \\ Y_2 &\longrightarrow \varnothing \end{aligned}$$
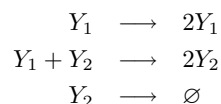
This model is the simplest model exhibiting a non-linear auto-regulatory feedback behavior. Y_1 represents a prey species (such as rabbits) and Y_2 represents a predator species (such a foxes) [2]. The first reaction is the representation of prey reproduction. the second reaction is an attempt to capture

---

[2]The use of the reactions to model the interaction of species in a population dynamics context explains the use of the term "species" to refer to a particular type of chemical molecule in a set of coupled chemical reactions.

predator-prey interaction (consumption of prey by predator, in turn influencing predator reproduction rate). The third reaction represents death of predators due to natural causes.

The LV model encourages to think about the number of prey and predators as integers, which can change only by discrete integer amounts when a reaction event occurs. However in the classical continuous deterministic chemical kinetics, the amounts of reactants and products are expressed as concentration, measured in moles per liter (M), which can vary continuously when as the reaction progresses. Conventionally, the concentration of a chemical species X is denoted $[X]$. The equation (10) states that the instantaneous rate of a reaction is directly proportional to the concentration (in turn directly proportional to mass) of each reactant to the power of its stoichiometry. This kinetic law is know as *mass-action kinetics*. So for the LV system, the second reaction will proceed at a rate proportional to $[Y_1][Y_2]$. Consequently, due to the effect of this reaction, $[Y_1]$ will decrease at instantaneous rate $k_2[Y_1][Y_2]$, where k_2 is the constant of proportionality for this reaction. $[Y_2]$ will increase at the same rate, because the overall effect of the reaction is to decrease $[Y_1]$ at the same rate $[Y_2]$ increases. The expression $k_2[Y_1][Y_2]$ is the rate law of the reaction, and k_2 is the *rate constant*. Considering all three reactions, we can write down a set of ordinary differential equations (ODEs) for the system.

$$\frac{[Y_1]}{dt} = k_1[Y_1] - k_2[Y_1][Y_2] \qquad (11)$$

$$\frac{[Y_2]}{dt} = k_2[Y_1][Y_2] - k_3[Y_2] \qquad (12)$$

The three rate constants k_1, k_2, and k_3 must be specified, as well as the initial concentrations of each species. Once this has been dome, the entire dynamics of the system are completely determined and can be revealed by solving the set of the ODEs, either analytically (in the rare case where it is possible), or numerically using a computer. Fig. 3 shows the time behaviors of the solutions when the initial values of $[Y_1]$ and $[Y_2]$ are 4 and 10, respectively and the rate constants are $k_1 = 1$, $k_2 = k_3 = 0.1$. The solutions have been obtained using the ODEs solution and analysis package XPPAUT [6].

Fig. 3. Lotka-Volterra dynamics for $[Y_1]_{t=0}$, $[Y_2]_{t=0}$, $k_1 = 1$, and $k_2 = k_3 = 0.1$.

An alternative way to display the dynamics of the system is an orbit in a *phase space*, where the values of one variable is plotted against the values of the other variables. Fig. 4 shows the dynamics in this way.

Fig. 4. Lotka-Volterra dynamics for $[Y_1]_{t=0}$, $[Y_2]_{t=0}$, $k_1 = 1$, and $k_2 = k_3 = 0.1$. The equilibrium solution for this combination of parameters is $[Y_1] = 1$ and $[Y_2] = 10$. These values correspond to the coordinates of the nullclines intersection points.

Phase plane analysis is a powerful way to determine how the behavior of a system will change with changes in the various parameters. Several types of plots are utilized as part of what is generically called phase plane analysis:

- a *phase portrait* consists of the variables describing a system plotted against each other rather then as a function of time to produce a trajectory in phase space. A phase portrait tells us how the variables interact for a given set of parameters.
- A *vector field* shows in the direction in which a system will evolve from any location in phase space.
- *Nullclines* are plotted in phase space, and show the values of a pair of variables at which one of the variables does not change. For a system of coupled equations $X(x, y, t)$ and $Y(x, y, t)$ nullclines are the solutions of the equations

$$\frac{dX}{dt} = 0 \qquad (13)$$

$$\frac{dY}{dt} = 0 \qquad (14)$$

There is a nullcline for each variable. The points of intersection of two nullclines are called *fixed points*, and represent *stable steady states*, also known as *equilibrium points*.

1) Equilibrium and conservation law: Even when the set of O.D.E.s is not analytically intractable, it may be possible to discover an *equilibrium* solution of the system by analytic (or simple numerical) means. An equilibrium solution is a set of concentrations which will not change over time, and hence can be found by solving the set of simultaneous equations formed by setting the right-hand side of the O.D.Es to zero. For the Lotka-Volterra example, this is

$$k_1[Y_1] - k_2[Y_1][Y_2] = 0$$
$$k_2[Y_1][Y_2] - k_3[Y_2] = 0$$

Solving these for $[Y_1]$ and $[Y_2]$ in terms of k_1, k_2 and k_3 gives two solutions. The first is

$$[Y_1] = 0, \quad [Y_2] = 0$$

and the second is

$$[Y_1] = \frac{k_3}{k_2}, \quad [Y_2] = \frac{k_1}{k_2}$$

Further analysis (rather tangential to the scope of this work) reveals that this second solution is not unstable, and hence corresponds to a realistic stable state of the system. Moreover, it is not "attractive" stable state, and so there is no reason to suppose that the system will tend to this state irrespective of the starting conditions. Despite knowing the existence of an equilibrium solution of this system, there is no reason to suppose that any particular set of initial conditions will lead to this equilibrium, and even if we suppose it, it would say nothing about how the system reaches this equilibrium. To answer this question we have to reduce to a particular set of initial conditions and integrate the O.D.Es to uncover the full dynamics.

In the context of chemical kinetics it is worth considering the conservation laws, that are useful to reduce the dimension

of the system under investigation. Consider for example a reversible dimerization reaction

$$2P \rightleftharpoons P_2$$
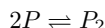
If we make the very strong assumption that neither of these species are involved in any other reactions, then we get the O.D.Es

$$\frac{d[P]}{dt} = 2k_2[P_2] - 2k_1[P]^2 \tag{15}$$

$$\frac{d[P_2]}{dt} = k_1[P]^2 - k_2[P] \tag{16}$$

where k_1 and k_2 are the forward and backward rate constants, respectively. the system is at equilibrium whenever

$$k_2[P_2] = k_1[P]^2.$$

that can be re-written as

$$\frac{[P_2]}{[P]^2} = \frac{k_1}{k_2} \equiv K_{eq} \tag{17}$$

where K_{eq} is the *equilibrium constant* of the system. This equilibrium is stable and attractive.

Note now that $[P]$ and $[P_2]$ are deterministically related in this system. One way to see this is to add twice the second O.D.E. of the system in (15) to the first to get

$$\frac{d[P]}{dt} + 2\frac{d[P_2]}{dt} = 0 \Rightarrow \frac{d}{dt}\left([P] + C\right) = 0 \Rightarrow [P] + [P_2] = c \tag{18}$$

where c is the concentration of $[P]$ if the dimers were fully dissociated. Equation (18) is known as *conservation equation*, as the value of left-hand side is conserved by the reaction system. Solving the conservation equation for $[P_2]$ and substituting back into the equilibrium relation (17) we find the equilibrium concentration of $[P]$ as a solution of the quadratic equation

$$2K_{eq}[P]^2 + [P] - c = 0$$

that has a single real positive root give by

$$[P]_{eq} = \frac{\sqrt{8cK_{[eq} + 1} - 1}}{4K_{eq}}$$

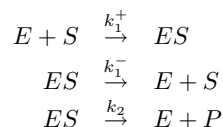the conservation equation can be alternatively used to reduce the pair of O.D.Es to a single first-order ODE

$$\frac{d[P]}{dt} = k_2(c - [P]) - 2k_1[P]^2. \tag{19}$$

It turn out that (19) has an analitycal solution derivable by solving in $[P]$ the following equation

$$0.49 \times \ln\left|\frac{-4[P] - 1.56}{-4[P] + 2.56}\right| = t + const.$$

C. Example 2: the Michaelis-Menten kinetics

Let consider the following chemical reactions

$$E + S \xrightarrow{k_1^+} ES$$
$$ES \xrightarrow{k_1^-} E + S$$
$$ES \xrightarrow{k_2} E + P$$

This model involves four variables E, S, ES and P. However, the total concentration of enzyme

$$[E]_{tot} = [E] + [ES]$$

and of the substrate

$$[S]_{tot} = [S] + [ES] + [P]$$

are conserved, so that only two of the concentrations change independently. In this analysis we choose the concentration of substrate $[S]$ and enzyme-substrate complex $[ES]$ as variables and eliminate the concentration of enzyme using the conservation law

$$[E] = [E]_{tot} - [ES].$$

Because the catalytic process is irreversible, the concentration of the product $[P]$ does not appear in the equations for $[E]$ and $[S]$, that are obtained applying the mass-action laws applied to the chemical equations of the system

$$\frac{d[S]}{dt} = -k_1^+[E]_{tot}[S] + (k_1^- + k_1^+[S])[ES] \tag{20}$$

$$\frac{d[S]}{dt} = k_1^+[E]_{tot}[S] - (k_1^- + k_2^+ + k_1^+[S])[ES] \tag{21}$$

The two important time scales in the Michaelis-Menten model are the time that the substrate needs to be converted into product, and the time scale on which enzyme-substrate complex forms. Therefore the important rates are $k_1^+[E]_{tot}$ and $k_1^+[S]$. Michaelis and Menten assumed that the quantity of the enzyme if very little compared to the quantity of substrate. Under this condition we expect there to be very little complex compared to substrate. This means that, at least at the beginning before a lot of product has been made, the rate $k_1^+[E]_{tot}$ is much smaller than the rate $k_1^+[S]$. As consequence, it is the small ratio of the concentration of catalyst to total concentration of substrate, i. e.

$$\epsilon = \frac{[E]_{tot}}{[S]_{tot}}$$

that makes the two time scales widely different.

IV. THE STRUCTURE OF KINETIC MODELS

The most part of the mathematical models of chemical reactions is based on the assumption of spatial homogeneity. This means that in these models diffusion and other transport processes can be neglected. Thence, from the formal point of view chemical reaction is handled as a temporal process and

a network of chemical interactions is considered a dynamical system.

A dynamics system is an ordered pair: $(\mathbb{A}, \phi)$, where $\mathbb{A}$ is the state space, and $\phi : \mathbb{T} \times \mathbb{A} \to \mathbb{A}$ is a function which assigns to an arbitrary point $x_0 \in \mathbb{A}$ the point $x \in \mathbb{A}$, that characterizes the state at the time t, assuming that the system was in x_0 at $t = 0$. A fundamental property of ϕ is the validity of the identity

$$\phi((t + s), x_0) = \phi(s, \phi(t, x_0)). \qquad (22)$$

The motion of a dynamics system is the one variable function

$$\phi_{x_0} \quad : \quad T \to A \qquad (23)$$
$$\phi_{x_0} \quad \equiv \quad \phi(\cdot, x_0) \qquad (24)$$

where $\mathbb{T} \subset \mathbb{R}$ and $\mathbb{A} \subset \mathbb{R}^M$, $(M \in \mathbb{N})$, or $\mathbb{A}$ consists of random variables taking their values from $\mathbb{R}^M$. For every $t \in T$ $\phi(t, \cdot) : \mathbb{A} \to \mathbb{A}$ is an automorphism.

The process, or equivalently the chemical reaction, to be described can be classified either by the properties of the *process-time*, or by the structure of the *state space*, or by the nature of *determination*.

A. Properties of process-time

The time can be chosen as continuous $(T \subset \mathbb{R})$ or a discrete $(T \subset \mathbb{Z})$ variable. Both the continuous and the discrete time models presents advantages, disadvantages, arguments in favor and arguments in disfavor. The arguments generally adopted for choosing a continuous time variable are:

1) Calculation with continuous time models have greater tradition. continuous models have the advantage over discrete time models in that they are more amenable to algebraic manipulation, although they are slightly harder to implement on a computer.

2) Most physical processes are inherently continuous in time. In particular, some physico-chemical quantities can be transduced continuously. Thus, the parameters in the models are strongly correlated with the physical properties of the systems; something that is very appealing to an engineer. Moreover, as cost of computation becomes cheaper, today's data acquisition equipment can provide nearly continuous-time measurements. Fast sampled data can be more naturally dealt with using continuous-time models than discrete-time models.

Arguments for selecting a discrete time variable are the following.

1) Time is really discrete. The idea that time has no objective existence but depends on events led some scientists to abandon the assumption that it is a continuous variable. Moreover, we perceive temporal intervals of finite duration rather than durationless instants; and the researcher prefer to assume that the nature has properties that can be verified.

2) The notion of 'immediate next time" can be easily interpreted, ant this is non so easy in the case of continuous time

3) the experimentalists measure at discrete points only.

B. Properties of state-space

The state space can be chosen either continuous or discrete. To emphasize the existence of elementary particles of a population as in reaction kinetics a discrete state space formalism is preferred.

The notion of *state* was derived from the theory of mechanics and of thermodynamics and generalized by mathematical system theory. The quantities of a model can be classified into two categories: state variables and constitutive quantities. State variables are functions such that their values specify the state of the system. the constitutive quantities are functions of the state, in the sense that their value is univocally determined once the state of the system has been assigned. Thus, a constitutive quantity Ω can be expressed as follows

$$\Omega(t) = \omega(g(t), t) \qquad (25)$$

where g denotes the state of the system and $\omega : \mathbb{A} \times T \to \mathbb{R}'$ is the constitutive functional[3] mapping the state into a constitutive quantity $r \in \mathbb{N}$. The case $r = 1$ means that the value of the constitutive quantity is a scalar.

As we already introduced in the previous section, the state of an M-component chemical system is described by a vector:

$$\vec{x} : \mathbb{T} \to \mathbb{R}^M, \quad t \mapsto \vec{x}(t) \in \mathbb{R}^M \qquad (26)$$

In this section we also said that a state is described by function. The two statements are not in contradiction, namely a finite-dimensional vector can also be interpreted as a function: $\mathbb{R}^M$ can be considered as an abbreviation

$$\mathbb{R}^M := \mathbb{R}^{\{1,2,\cdots,M\}} = \{f; f : \{1, 2, \cdots, M\} \to \mathbb{R}\}.$$

The state of the system with continuously changing components is described at a fixed point of time by a (not necessarily scalar valued) function $f : \mathbb{R}^M \to \mathbb{R} * m$. The state of the system is $\tilde{n}$, where $\tilde{n} : \mathbb{T} \to (\mathbb{R}^m)^{\mathbb{R}^M}$, or it is an element of the set

$$[(\mathbb{R}^m)^{\mathbb{R}^M}]^{\mathbb{T}} = \{f; f : \mathbb{T} \to (\mathbb{R}^m)^{\mathbb{R}^M}$$

According to the convectional treatment of pure homogeneous reaction kinetics the state is a finite-dimensional vector and the only constitutive quantities are the reaction rates.

The theory of thermodynamics adopts the concept of 'particles with memory'. According to this concept, the constitutive quantities depend on the history of the independent variables, and not only on their present value. This means that it is not definite that the instantaneous value of state variables (i. e. state) completely determines the state.

---

[3]The functional assigns a number to a function. Here the term refers to every mapping having the function as argument.

Let introduce the *site* function $h : \mathbb{T} \to \mathbb{R}^M$. Since the state is determined by earlier values of the site function, therefore the state g is interpreted as

$$g : \mathbb{T} \to G, \quad t \mapsto h'$$

where h' is known as *history* function defined as

$$h'(s) = h(t - s), \quad s > 0.$$

Knowing the history, the state can be set up

$$\mathscr{H}(h, \cdot) = g,$$

i. e.

$$\mathscr{H}(h, t) = g(t) = \mathscr{H}(h') = h'$$

$\mathscr{H}$ is a mapping assigning a function to a function and to a number. If we assume that the history of the site does not influence the state, then the constitutive functional reduces to a function. Furthermore, if we also assume the invertibility of this function then the differences between the state variables and constitutive quantities are not significant. These two assumptions are tacitly adopted in the classical theories of the thermodynamics.

The stochastic version of a memory-free deterministic process is a Markov process (more precisely, a first-order Markov process).

C. Nature of determination

An $(\mathbb{A}, \phi)$ dynamic system is deterministic if knowing the state of the system at one time means that the system is uniquely specified for all $t \in \mathbb{T}$.

When the state of the system can be assigned to a set of values with a certain probability distribution, the future behavior of the system can be determined stochastically. Discrete time, discrete state space (first order) Markov processes (i. e. Markov *chain*) are defined by the formula

$$\mathscr{P}(\xi_{t+1} = a | \xi_0 = a, \xi_1 = a_1, \dots, \xi_t = a_t) = \mathscr{P}(\xi_{t+1} = a | \xi_t = a_t)$$

(27)

where the set $\{\xi_t | t = 0, 1, 2, \dots\}$ is a discrete time stochastic process.

Knowing the total history of the process we can extrapolate its future behavior with the same probability as if we knew only the actual current state. Put another way, a Markov process is a stochastic process which possesses the property that the future behavior depends only on the current state of the system. Thus, given information about the current state of the system, information about the past behavior of the system is no help in predicting the time-evolution of the process. The behavior of the chain is therefore determined by $\mathscr{P}(\xi_{t+1} = a | \xi_t = a_t)$, and thus it depends on a and t. However, if there is no t dependence, so that

$$\mathscr{P}(\xi_s = x | \xi_t = y) = \mathscr{P}_{xy}(s - t),$$

i. e. the transition probabilities are stationary, the Markov chain is said to be *time homogeneous*. In this case the law of evolution of the system does not depend explicitly on time and consequently, the time origin can be defined arbitrarily. Deterministic dynamics systems generated by ordinary differential equations

$$\frac{dx(t)}{dt} = f(x(t))$$

can be associated with the time homogeneous Markov processes.

Markov processes are particularly amenable for to both theoretical and computational analysis and the dynamic behavior of biochemical networks can be effectively modeled by a Markov chain. Moreover, a Markovian description can be introduced by generalizing deterministic systems modeled by ordinary differential equations, since the stochastic version of a deterministic process without 'after-effect' is a Markov process.

However, the Markov character of the chemical process represented by the state vector has not been derived from microscopic models of the chemical dynamics. Therefore Markovicity is not more and not less than a plausible assumption.

D. XYZ models

At least eight different kinetic models can be defined, depending on the specification of time (X), state space (Y) and nature of determination (Z). As was explained earlier, time can be discrete (D) or continuous (C), the state space can be also discrete (D) or continuous (C), and the nature of determination can be deterministic (D) or stochastic (S).

Mass-action type kinetic differential equations can be identified with the CCD model, while the more often used stochastic model is the CDS model. DCD models have achieved a significance in the last decade in connection with chaotic phenomena. There are at least two distinct methods of relating DCD models to CCD models. The first is the discretization of time. An autonomous differential equation

$$\frac{dx}{dt} = f(x(t), t), \quad x_{t=0} = x_0$$

can be transformed as

$$x(t + h) = x(t) + f(x(t), t)h + o(h)$$

The second method can be applied if the differential equation has a periodic solution. take a hyperplane of dimension $n - 1$ traverse to the curve $t \longrightarrow x(t)$ through x_0. A map

$$F : U \longrightarrow \mathbb{R}^{n-1}$$

is induced by associating with t_0 the nearest intersection of the trajectory (with initial condition $x_{t=0} = x_0$) with the given hyperplane. If the first such intersection occurs at x_1, we define $F(x_0) \equiv x_1$. Since the form of F is independent of the index of the series and also of the coordinates, we can specify

$$x_{n+1} = F(x_n).$$

Thus a difference equation has been obtained from a differential system.

V. Markov processes

The story of the master equation must begin with Markov processes. A Markov process is a special case of a stochastic process. Stochastic processes are often used in physics, biology and economy to model randomness. In particular, Markov processes are often used to model randomness, since it is much more tractable than a general stochastic process. A general stochastic process is a random function $f(X; t)$, where X is a stochastic variable and t is time. The definition of a stochastic variable consists in specifying

- a set of possible values (called "set of states" or "sample space")
- a probability distribution over this set.

The set of states may be *discrete*, e. g. : the number of molecules of a certain component in a reacting mixture. Or the set may be *continuous* in a given interval, e. g: one velocity component of a Brownian particle and the kinetic energy of that particle. Finally the set may be partly discrete and partly continuous, e. g. the energy of an electron in the presence of binding centers. Moreover the set of states may be *multidimensional*: in this case X is written as a vector $\vec{X}$. Examples: $\vec{X}$ may stand for the three velocity components of a Brownian particle or for the collection of all numbers of molecules of the various components in a reacting mixture.

The probability distribution, in the case of a continuous one-dimensional range, is given by a function $P(x)$ that is non-negative

$$P(x) \geq 0$$

and normalized in the sense

$$\int P(x)dx = 1$$

where the integral extends over the whole range. The probability that X has a value between x and $x + dx$ is

$$P(x)dx$$

Often in physical and biological sciences a probability distribution is visualized by an "ensemble". From this point of view, a fictitious set of an arbitrary large number $\mathcal{N}$ of quantities, all having different values in the given range, is introduced.In such a way the number of these quantities having a value between x and $x + dx$ is $\mathcal{N}P(x)dx$. Thus the

probability distribution is replaced with a density distribution of a large number of "samples". This does not affect any simulation result, since it is merely a convenience in talking about probabilities, and in this work we will use this language. It may be added that it can happen that a biochemical system does consists of a large number of identical replica, which to a certain extent constitute a physical realization of an ensemble. For instance, the molecules of an ideal gas may serve as an ensemble representing the Maxwell probability distribution for the velocity. The use of an ensemble is not limited to such cases, nor based on them, but serves as a more concrete visualization of a probability distribution.

Finally, we remark that in a continuous range it is possible for $P(x)$ to involve delta functions,

$$P(x) = \sum_n p_n \delta(x - x_n) + \tilde{P}(x),$$

where $\tilde{P}$ is finite or at least integrable and non-negative, $p_n > 0$, and

$$\sum_n p_n + \int \tilde{P}(x)dx = 1$$

Physically this may visualized as a set of discrete states x_n with probability p_n embedded in a continuous range. If $P(x)$ consists of δ functions alone (i. e. $\tilde{P}(x) = 0$, then it can also be considered as a probability distribution p_n on the discrete set of states x_n.

A general way to specify a stochastic process is to define the joint probability densities for values $x_1, x_2, x_3, \ldots$ at times $t_1, t_2, t_3, \ldots$ respectively

$$p(x_1, t_1; x_2, t_2; x_3, t_3; \ldots) \qquad (28)$$

If all such probabilities are known, the stochastic process is fully specified, (but, in general, it is not an easy task to find all such distributions). Using (28) the conditional probabilities can be defined as usual

$$p(x_1, t_1; x_2, t_2; \ldots | y_1, \tau_1; y_2, \tau_2; \ldots) = \frac{p(x_1, t_1; x_2, t_2; \ldots | y_1, \tau_1; y_2, \tau_2; \ldots)}{p(y_1, \tau_1; y_2, \tau_2; \ldots)}$$

where $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ are values at times $t_1 \geq t_2 \geq \cdots \geq \tau_1 \geq \tau_2 \geq \ldots$. This is where a Markov process has a very attractive property. It has no memory. For a Markov process

$$p(x_1, t_1; x_2, t_2; \ldots | y_1, \tau 1; y_2, \tau_2; \ldots) = p(x_1, t_1; x_2, t_2; \ldots | y_1, \tau_1)$$

the probability to reach a state x_1 at time t_1 and state x_2 at time t_2, if the state is y_1 at time τ_1, is independent of any previous state, with times ordered as before. This property makes it possible to construct any of the probabilities (28) by a *transition probability* $p_\rightarrow(x, t|y, \tau)$, $(t \geq \tau)$, and an initial probability distribution $p(x_n, t_n)$:

219

$$p(x_1, t_1; x_2, t_2; \ldots x_n, t_n) =$$
$$p_\rightarrow(x_1, t_1 | x_2, t_2) p_\rightarrow(x_2, t_2 | x_3, t_3) \ldots$$
$$\ldots p_\rightarrow(x_{n-1} t_{n-1} | x_n, t_n) p(x_n, t_n)$$

A consequence of the Markov property is the Chapman-Kolmogorov equation

$$p_\rightarrow(x_1, t_1 | x_3, t_3) = \int p_\rightarrow(x_1, t_1 | x_2, t_2) p_\rightarrow(x_2, t_2 | x_3, t_3) dx_2 \tag{29}$$

VI. THE MASTER EQUATION

The master equation is a differential form of the Chapman-Kolmogorov equation (29). The terminology differs between different authors. Sometimes the term master equation is used only for jump processes. Jump processes are characterized by discontinuous motion, that is there is a bounded and non-vanishing transition probability per unit time

$$w(x|y, t) = \lim_{\Delta t \to 0} \frac{p_\rightarrow(x, t + \Delta t | y, t)}{\Delta t}$$

for some y such that $|x - y| > \epsilon$. Here $w(x|y; t) = w(x|y)$.

The master equation for jump processes can be written

$$\frac{\partial p(x, t)}{\partial t} = \int \big(w(x|x') p(x', t) - w(x'|x) p(x, t) \big) dx' \tag{30}$$

The master equation has a very intuitive interpretation. The first part of the integral is the *gain of probability* from the state x' and the second part is the *loss of probability* to x'. The solution is a probability distribution for the state space. Analytical solutions of the master equation are possible to calculate only for simple special cases.

A. The chemical master equation

A reaction R is defined as a jump to the state $\vec{X}$ to a stare $\vec{X}_R$, where $\vec{X}, \vec{X}_R \in \mathbb{Z}_+^N$. The propensity $w(\vec{X}_R) = \tilde{v}(\vec{X})$ is the probability for transition from $\vec{X}_R$ to $\vec{X}$ per unit time. A reaction can be written as

$$\vec{X}_R \xrightarrow{w(\vec{X}_R)} \vec{X}$$
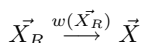
The difference in molecules numbers $\vec{n}_R = \vec{X}_R - \vec{X}$ is used to write the master equation (30) for a system with M reactions

$$\frac{dp(\vec{X}, t)}{dt} = \sum_{i=1}^{M} w(\vec{X} + n) p(\vec{X} + \vec{n}_R, t) - \sum_{i=1}^{M} w(\vec{X}) p(\vec{X}, t) \tag{31}$$

This special case of master equations is called the *chemical master equation* (CME) [32], [25]. It is fairly easy to write: however, solving it is quite another matter. The number of problems for which the CME can be solved analytically is even fewer than the number of problems for which the deterministic reaction-rate equations can be solved analytically. Attempts to use master equation to construct tractable time-evolution equations are also usually unsuccessful, unless all the reaction

in the system are simple monomolecular reactions [9]. Let consider for instance a deterministic model of two metabolites coupled by a bimolecular reaction, as shown in Fig. 5. The set of differential equation describing the dynamic of this model is given in Table II, where the $[A]$ and $[B]$ are the concentrations of metabolite A and metabolite B, while k, K, and μ determine the maximal rate of synthesis, the strength of the feedback, and the rate of degradation, respectively.

Fig. 5. Two metabolites A and B coupled by a bimolecular reactions. Adapted from [16].

In the formalism of the Markov process, the reactions in Table II are written as in Table III. The CME equation for the system of two metabolites of Fig. 5 looks fairly complex as in Table IV.

VII. MOLECULAR APPROACH TO CHEMICAL KINETICS

The solution of the set of differential equations of the form (10), written for each species X_j included in the system, describes the time-evolution of the system, i. e. the changes in time of the state vectors $\vec{x}$ of in the system.

The expression (10) is not the precise number of reactive collisions, but the average. The actual number fluctuates around it and in order to find the resulting fluctuation in x_j around the macroscopic values determined by (10) we need to switch to a molecular approach to the chemical kinetics.

Now, to see how chemical kinetics can be modeled in a stochastic way, first we need to address the difference between the deterministic and the stochastic approach in the representation of the amount of molecular species. In the stochastic model, this is an integer representing the number of molecules of the species, but in the deterministic model, it is a concentration, measured in M (moles per liter). Then for a concentration of X of $[X]$ M in a volume of V liters, there are $[X]V$ moles of X and hence $n_A[X]V$ molecules, where $n_A \simeq 6.023 \times 10^{23}$ is the Avogadro's constant (the number of molecules in a mole). The second issue that needs to be addressed is the rate constant conversion. Much of the literature on biochemical reaction is dominated by a continuous deterministic view of kinetics. Consequently, where rate constants are documented, they are usually deterministic constants k. In the following we review the expression of the reaction propensity and the formulae that convert the deterministic rate constants into stochastic rate constants.

A. Reactions are collisions

For a reaction to take place, molecules must collide with sufficient energy to create a transition state. Ludwig Boltzmann developed a very general idea about how energy was distributed among systems consisting of many particles. He said that the number of particles with energy E would be proportional to the value $\exp[-E/k_B T]$. The Boltzmann distribution predicts the distribution function for the fractional number of particles N_i/N occupying a set of states i which each have energy E_i:

TABLE II
REACTIONS OF THE CHEMICAL MODEL DISPLAYED IN FIG. 5. NO. CORRESPONDS TO THE NUMBER IN THE FIGURE. ADAPTED FROM [16].

No.	Reaction	Rate equation	Type
1	$\emptyset \xrightarrow{v_1([A])} A$	$v_1([A]) = \frac{k_1}{1+[A]K_1}$	synthesis
2	$A \xrightarrow{v_2([A])} \emptyset$	$v_2([A]) = \mu[A]$	degradation
3	$\emptyset \xleftarrow{v_3([B])} B$	$v_3([B]) = \frac{k_2}{1+[B]/K_2}$	synthesis
4	$B \xrightarrow{v_4([B])} \emptyset$	$v_4([B]) = \mu[B]$	degradation
5	$A + B \xrightarrow{v_5([A],[B])} \emptyset$	$v_5([A],[B]) = k_3[A][B]$	bimolecular reaction

TABLE III
REACTIONS OF THE CHEMICAL MODEL DEPICTED IN FIG. 5, THEIR PROPENSITY AND CORRESPONDING "JUMP" OF STATE VECTOR $\vec{n}_R^T$. V IS THE VOLUMES IN WHICH THE REACTIONS OCCUR. ADAPTED FROM [16].

No.	Reaction	$w(\vec{x})$	$\vec{n}_R^T$
1	$\emptyset \xrightarrow{w_1(a)} A$	$w_1(a) = Vk_1/(1 + a/VK_1))$	$(-1,0)$
2	$A \xrightarrow{w_2(a)} \emptyset$	$w_2(a) = \mu a$	$(1,0)$
3	$\emptyset \xrightarrow{w_3(b)} B$	$w_3(b) = VK_2/(1 + b/(VK_2))$	$(0,-1)$
4	$B \xrightarrow{w_4(b)} \emptyset$	$w_4(b) = \mu b$	$(0,1)$
5	$A + B \xrightarrow{w_5(a,b)} \emptyset$	$w_5(a,b) = k_2ab/V$	$(1,1)$

$$\frac{N_i}{N} = \frac{g_i e^{-E_i/k_B T}}{Z(T)}$$

where k_B is the Boltzmann constant, T is temperature (assumed to be a sharply well-defined quantity), g_i is the degeneracy, or number of states having energy E_i, N is the total number of particles:

$$N = \sum_i N_i,$$

and $Z(T)$ is called the *partition function*

$$Z(T) = \sum_i g_i e^{-E_i/k_B T}$$

Alternatively, for a single system at a well-defined temperature, it gives the probability that the system is in the specified state. The Boltzmann distribution applies only to particles at a high enough temperature and low enough density that quantum effects can be ignored.

James Clerk Maxwell used Boltzmann's ideas and applied them to the particles of an ideal gas to produce the distribution bearing both men's names (the Maxwell-Boltzmann distribution). Maxwell also used for the energy E the formula for kinetic energy $E = (1/2)mv^2$, where v is the velocity of the particle. The distribution is best shown as a graph which shows how many particles have a particular speed in the gas. It may also be shown with energy rather than speed along the x axis. Two graphs are shown in Fig. 6 and 7. Consider a

Fig. 6. Since the curve shape is not symmetric, the average kinetic energy will always be greater than the most probable. For the reaction to occur, the particles involved need a minimum amount of energy - the activation energy.

bi-molecular reaction of the form

$$S_1 + S_2 \longrightarrow \ldots \qquad (32)$$

Fig. 7. Maxwell-Boltzmann speed distributions at different temperatures. As temperature increases, the curve will spread to the right and the value of the most probable kinetic energy will decrease. At temperature increases the probability of finding molecules at higher energy increases. Note also that the area under the curve is constant since total probability must be one.

the right-hand side is not important. This reaction means that a molecule of S_1 is able to react with a molecule of S_2 if the pair happen to collide with one another with sufficient energy, while moving around randomly, driven by Brownian motion. Consider a single pair of such molecules in a closed volume V. It is possible to use statistical mechanics arguments to understand the physical meaning of the *propensity* (i. e. hazard) of molecules colliding. Under the assumptions that the volume is not too large or well stirred, an in thermal equilibrium, it can be rigorously demonstrated that the *collision propensity* (also called collision *hazard*, *hazard function* or *reaction hazard*) is constant, provided that the volume is fixed and the temperature is constant. Since the molecules are uniformly distributed throughout the volume and this distribution does not depend on time, then the probability that the molecules are within reaction distance is also independent of time. A comprehensive treatment of this issue is given in Gillespie [9], [11]. Here we briefly review it by highlighting the physical basis of the stochastic formulation of chemical kinetics. Consider now that the system composed of a mixture of the two molecular species, S_1 and S_2 in gas-phase and in thermal, but not necessarily chemical equilibrium inside the volume V. Let assume that the S_1 and S_2 molecules are hard spheres of radii r_1 and r_2, respectively. A collision will occur whenever the center-to-center distance between an S_1 molecule and an S_2 molecule is less than $r_{12} = r_1 + r_2$. To calculate the molecular collision rate, let pick an arbitrary 1-2 molecular pair, and denote by v_{12} the speed of the molecule 1 relative to molecule 2. Then, in the next small time interval δt, molecule 1 will sweep out relative to molecule 2 a collision

TABLE IV
SET OF CHEMICAL MASTER EQUATIONS DESCRIBING THE METABOLITES INTERACTION SHOWED IN FIG. 5 AND DISCUSSED IN [16].

$$\frac{\partial(0,0,t)}{\partial t} = \mu p(1,0,t) + \mu p(0,1,t) + \frac{k_3}{V}p(1,1,t) - V(k_1+k_2)p(0,0,t)$$

$$\frac{\partial(0,b,t)}{\partial t} = V\frac{k_2}{1+\frac{b-1}{VK_2}}p(0,b-1,t) +$$
$$+ \mu p(1,b,t) + \mu(b+1)p(0,b+1,t) + \frac{k_3}{V}(b+1)p(1,b+1,t) -$$
$$- \left(V\left(k_1 + \frac{k_2}{1+\frac{b}{VK_2}}\right) + \mu b\right)p(0,b,t)$$

$$\frac{\partial p(a,0,t)}{\partial t} = V\frac{k_1}{1+\frac{a-1}{VK_1}}p(a-1,0,t) +$$
$$+ \mu(a+1)p(a+1,0,t) + \mu p(a,1,t) +$$
$$+ \frac{k_3}{V}(a+1)p(a+1,1,t) -$$
$$- \left(V\left(\frac{k_1}{1+\frac{a}{VK_1}} + k_2\right) + \mu a\right)p(a,0,t)$$

$$\frac{\partial p(a,b,t)}{\partial t} = V\frac{k_1}{1+\frac{a-1}{VK_1}}p(a-1,b,t) + V\frac{k_2}{1+\frac{b-1}{VK_2}}p(a,b-1,t) +$$
$$+ \mu(a+1)p(a+1,b,t) + \mu(b+1)p(a,b+1,t) +$$
$$+ \frac{k_3}{V}(a+1)(b+1)p(a+1,b+1,t) -$$
$$- \left(V\left(\frac{k_1}{1+\frac{a}{VK_1}} + \frac{k_2}{1+\frac{b}{VK_2}}\right) + \mu(a+b) + \frac{k_3}{V}ab\right)p(a,b,t)$$

volume

$$\delta V_{coll} = \pi r_{12}^2 v_{12}\delta t$$

i. e. if the center of molecule 2 happens to lie inside δV_{coll} at time t, then the two molecules will collide in the time interval $(t, t+\delta t)$. Now, the classical procedure would estimate the number of S_2 molecules whose centers lie inside δV_{coll}, divide the number by δt, and then take the limit $\delta \rightarrow 0$ to obtain the rate at which the S_1 molecule is colliding with S_2 molecules. However, this procedure suffers from the following difficulty: as $\delta V_{coll} \rightarrow 0$, the number of S_2 molecules whose centers lie inside δV_{coll} will be either 1 or 0, with the latter possibility become more and more likely as the limiting process proceeds. Then, in the limit of vanishingly small δt, it is physically meaningless to talk about "the number of molecules whose center lie inside δV_{coll}". To override this difficulty we can exploit the assumption of thermal equilibrium. Since the system is in thermal equilibrium, the molecules will at all times be distributed randomly and uniformly throughout the containing volume V. Therefore, the *probability* that the center of an arbitrary S_2 molecule will be found inside δV_{coll} at time t will be given by the ratio $\delta V_{coll}/V$; note that this is true even in the limit of vanishingly small δV_{coll}. If we now average this ration over the velocity distributions of S_1 and S_2 molecules, we may conclude that the average probability that a particular 1-2 molecular pair will collide in the next vanishingly small time interval δt is

$$\overline{\delta V_{coll}/V} = \frac{\pi r_{12}^2 \overline{v_{12}}\delta t}{V} \qquad (33)$$

For Maxwellian velocity distributions the average relative speed $\overline{v_{12}}$ is

$$\overline{v_{12}} = \left(\frac{8kT}{\pi m_{12}}\right)^{\frac{1}{2}}$$

where k is the Boltzmann's constant, T the absolute temperature, and m_{12} the reduced mass $m_1 m_2/(m_1+m_2)$. If we are given that at time t there are X_1 molecules of the species S_1 and X_2 molecules of the species S_2, making a total of $X_1 X_2$ distinct 1-2 molecular pairs, then if follows from (33) that the probability that a 1-2 collision will occur somewhere inside V in the next infinitesimal time interval $(t, t+dt)$ is

$$\frac{X_1 X_2 \pi r_{12}^2 \overline{v_{12}}dt}{V} \qquad (34)$$

Although we cannot rigorously calculate the number of 1-2 collisions occurring in V in any infinitesimal interval, we can rigorously calculate the probability of a 1-2 collision occurring in V in any infinitesimal time interval. Consequently, we really ought to characterize a system of thermally equilibrized molecules by a collision probability per unit time (namely the coefficient of dt in (34) instead of by a collision rate. This is why these collisions constitute a stochastic Markov process instead of a deterministic rate process.

Fig. 8. The collision volume δV_{coll} which molecule 1 will sweep out relative to molecule 2 in the next small time interval δt. Adapted fro [9].

Then we can conclude that for a bimolecular reaction of the form (32, the probability that a randomly chosen A-B pair will react according to R in next dt is

$$P_{react} = \left\{ \left(\frac{\overline{v_{12}}(\pi r_{12}^2)}{V} \exp(-E/(k_B T) \right) \right\} X_1 X_2 dt \quad (35)$$

B. Reaction rates

The reaction rate for a reactant or product in a particular reaction is defined as the amount of the chemical that is formed or removed (in moles or mass units) per unit time per unit volume. The main factors that influence the reaction rate include: the physical state of the reactants, the volume of the container in which the reaction occurs, the temperature at which the reaction occurs, and whether or not any catalysts are present in the reaction.

Physical state

The physical state (solid, liquid or gas) of a reactant is also an important factor of the rate of change. When reactants are in the same phase, as in aqueous solution, thermal motion brings them into contact. However, when they are in different phases, the reaction is limited to the interface between the reactants. Reaction can only occur at their area of contact, in the case of a liquid and a gas, at the surface of the liquid. Vigorous shaking and stirring may be needed to bring the reaction to completion. This means that the more finely divided a solid or liquid reactant, the greater its surface area per unit volume, and the more contact it makes with the other reactant, thus the faster the reaction.

Volume The reaction propensity is inversely proportional to the volume. We can explain this fact in the following way. Consider two molecules Molecule 1 and Molecule 2. Let the molecules positions in space be denoted by p_1 and p_2 respectively. If p_1 and p_2 are uniformly and independently distributed over the volume V, for a sub-region of space D with volume V', the probability that a molecule is inside D is

$$\Pr(p_i \in D) = \frac{V'}{V} \quad i = 1, \ 2$$

If we are interested in the probability that Molecule 1 and Molecule 2 are within a reacting distance r of one another at any given instant of time (assuming that r is much smaller than the dimensions of the container, so that boundary effects can be ignored), this probability can be calculated as

$$\Pr(|p_1 - p_2| < r) = E(\Pr(|p_1 - p_2| < r|p_2))$$

but the conditional probability will be the same for any p_2 away from the boundary, so that the expectation in redundant, and we can state that

$$E(\Pr(|p_1 - p_2| < r|p_2)) = \Pr(|p_1 - p_2| < r) = \Pr(p_i \in D) = \frac{4\pi r^3}{3V}$$

This probability is inversely proportional to V.

Arrhenius equation

Temperature usually has a major effect on the speed of a reaction. Since a molecule has more energy when it is heated, then the more energy it has, the more chances it has to collide with other reactants. Thus, at a higher temperature, more collisions occur. More importantly however, is the fact that heating a molecules affects its kinetic energy, and therefore the "energy" of the collision.

The reaction rate coefficient k has a temperature dependency, which is usually given by the empirical Arrhenius law:

$$k = A \exp\left[-\frac{E_a}{RT} \right] \quad (36)$$

E_a is the activation energy and R is the gas constant. Since at temperature T the molecules have energies given by a Boltzmann distribution, one can expect the number of collisions with energy greater than E_a to be proportional to $\exp[-E_a/RT]$. A is the frequency factor. This factor indicates how many collisions between reactants have the correct orientation to lead to the products. The values for A and E_a are dependent on the reaction.

It can be seen that either increasing the temperature or decreasing the activation energy (for example through the use of catalysts) will result in an increase in rate of reaction.

While remarkably accurate in a wide range of circumstances, the Arrhenius equation is not exact, and various other expressions are sometimes found to be more useful in particular situations. One example comes from the "collision theory" of chemical reactions, developed by Max Trautz and William Lewis in the years 1916-18. In this theory, molecules react if they collide with a relative kinetic energy along their line-of-centers that exceeds Ea This leads to an expression very similar to the Arrhenius equation, with the difference that the pre-exponential factor "A" is not constant but instead is proportional to the square root of temperature. This reflects the fact that the overall rate of all collisions, reactive or not, is proportional to the average molecular speed which in turn is proportional to $\sqrt{T}$. In practice, the square root temperature dependence of the pre-exponential factor is usually very slow compared to the exponential dependence associated with E_a.

Another Arrhenius-like expression appears in the Transition State Theory of chemical reactions, formulated by Wigner, Eyring, Polanyi and Evans in the 1930's. This takes various forms, but one of the most common is:

$$k = \frac{k_B T}{h} \exp\left[-\frac{\Delta G}{RT} \right]$$

where ΔG is the Gibbs free energy of activation, k_B is Boltzmann's constant, and h is Planck's constant. At first

sight this looks like an exponential multiplied by a factor that is linear in temperature. However, one must remember that free energy is itself a temperature dependent quantity. The free energy of activation includes an entropy term as well as an enthalpy term, both of which depend on temperature, and when all of the details are worked out one ends up with an expression that again takes the form of an Arrhenius exponential multiplied by a slowly varying function of T. The precise form of the temperature dependence depends upon the reaction, and can be calculated using formulas from statistical mechanics (it involves the partition functions of the reactants and of the activated complex).

Catalysts

A catalyst is a substance that accelerates the rate of a chemical reaction but remains unchanged afterward. The catalyst increases rate reaction by providing a different reaction mechanism to occur with a lower activation energy. In autocatalysis a reaction product is itself a catalyst for that reaction possibly leading to a chain reaction. Proteins that act as catalysts in biochemical reactions are called enzymes. Michaelis-Menten kinetics mentioned in Chapter 1 describes the rate of enzyme mediated reactions.

The formulation of stochastic chemical kinetics of Gillespie assumes that temperature and volume of container do not change in time. However in the biological context these assumption are too strong and may lead to obtain wrong simulation results. We will see in the next chapter how to give up this assumption.

C. Zeroth-order reactions

These reactions have the following form

$$R_\mu: \quad \emptyset \xrightarrow{c_\mu} X \tag{37}$$

Although in practice things are not created from nothing, it is sometimes useful to mode a constant rate of production of a chemical species (or influx from another compartment) via a zeroth-order reaction. In this case, c_μ is the propensity of a reaction of this type occurring, and so

$$a_\mu(Y, c_\mu) = c_\mu \tag{38}$$

For a reaction of this nature the deterministic rate law is k $\text{M}s^{-1}$, and thus for a volume V, X is produced at a rate $n_A V k_\mu$ molecules per second, where k_μ is the deterministic rate constant for the reaction R_μ. As the stochastic rate law is just c_μ molecules per second, we have

$$c_\mu = n_A V k_\mu \tag{39}$$

D. First-order reactions

Consider the first-order reaction

$$R_\mu: \quad X_i \xrightarrow{c_\mu} \ldots \tag{40}$$

Here c_μ represents the propensity that a particular molecule of X_i will undergo the reaction. However, if there are x_i molecule

of of X_i, each of which having a propensity of c_μ of reacting, the combined propensity for a reaction of this type is

$$a_\mu(Y, c_\mu) = c_\mu x_i \tag{41}$$

First-order reactions of this nature represent the spontaneous change of a molecule into one or more other molecules or the spontaneous dissociation of a complex molecule into simpler molecules. They are not intended to model the conversion of one molecule into another in presence of a catalyst, as this is really a second-order reaction. However, in the presence of a large pool of catalyst that can be considered not to vary in concentration during the time evolution of the reaction network, a first-order reaction provides a good approximation. For a first-order reaction, the deterministic rate law is $k_\mu[X]$ $\text{M } s^{-1}$, and so for a volume V, a concentration $[X]$ correspond to $x = n_A[X]V$ molecules. Since $[X]$ decreases at rate $n_A k_\mu[X]V = k_\mu x$ molecules per second. Since the stochastic rate law is $c_\mu x$ molecules per second, we have

$$c_\mu = k_\mu \tag{42}$$

i. e. for first-order reactions, the stochastic and the deterministic rate constants are equal.
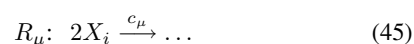
E. Second-order reactions

The form of the second-order reaction is the following

$$R_\mu: \quad X_i + X_k \xrightarrow{c_\mu} \ldots \tag{43}$$
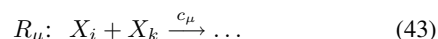
Here, c_μ represents the propensity that a particular pair of molecules of type X_i and X_k will react. But, if there are x_i molecule of X_i and x_k molecules of X_k, there are $x_i x_k$ different pairs of molecules of this type, and so this gives the combined propensity of

$$a_\mu(Y, c_\mu) = c_\mu x_i x_k \tag{44}$$

There is another type of second-order reaction, called *homodimerization* reaction, which needs to be considered:

$$R_\mu: \quad 2X_i \xrightarrow{c_\mu} \ldots \tag{45}$$

Again, c_μ is the propensity of a particular pair of molecules reacting, but here there are only $x_i(x_i - 1)/2$ pairs of molecules of species X_i, and so

$$a_\mu(Y, c_\mu) = c_\mu \frac{x_i(x_i - 1)}{2} \tag{46}$$

For second-order reactions, the deterministic rate law is $k_\mu[X_i][X_k]$ $\text{M } s^{-1}$. Here for a volume V, the reaction proceeds at a rate of $n_A k_\mu[X_i] [X_k]V = k_\mu x_i x_k/(n_A V)$ molecules per second. Since the stochastic rate law is $c_\mu x_i x_k$ molecules per second, we have

$$c_\mu = \frac{k_\mu}{n_A V} \tag{47}$$

For homodimerization reaction, the deterministic law is $k_\mu[X_i]^2$, so the concentration of X_i decreases at rate $n_A 4 k_\mu[X_i]^2 V = 2k_\mu x_i^2/(n_A V)$ molecules per second. The stochastic rate law is $c_\mu x_i(x_i - 1)/2$ so that molecules X_i

are consumed at a rate of $c_\mu x_i(x_i - 1)$ molecules per second. These two laws do not match, but for large x_i, $x_i(x_i - 1)$ can be approximated by x_i^2, and so to the extent that the kinetics match, we have

$$c_\mu = \frac{2k_\mu}{n_A V} \tag{48}$$

Note the additional factor of two in this case.

By equating Eq. (47) with Eq. (35) we obtain the following expression for the deterministic rate of a second-order reaction of type (43)

$$k_\mu = n_A \overline{v_{12}} \pi r_{12}^2 \exp\left[\frac{E_\mu}{k_B T}\right] \tag{49}$$

while for a second-order reaction of type (45), the deterministic rate constant is

$$k_\mu = \frac{1}{2} n_A \overline{v_{12}} \pi r_{12}^2 \exp\left[\frac{E_\mu}{k_B T}\right] \tag{50}$$

F. Higher-order reactions

Most (although not all) reactions that are normally written as a single reaction of order higher than two, in fact represent the combined effect of two or more reactions of order one or two. In these cases it is usually recommended to model the reactions in detail rather than via high-order stochastic kinetics. Consider, for example, the following trimerization reaction

$$c_\mu:\ 3X \xrightarrow{c_\mu} X_3$$
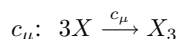
The rate constant c_μ represents the propensity of triples of molecules of X coming together simultaneously and reacting, leading to a combined propensity of the form

$$a_\mu(Y, c_\mu) = c_\mu \binom{x}{3} = c_\mu \frac{x(x-1)(x-2)}{6} \tag{51}$$

However, in most cases it is likely to be more realistic to model the process as the pair of second-order reactions

$$
\begin{aligned}
2X &\longrightarrow X_2 \\
X_2 + X &\longrightarrow X_3
\end{aligned}
$$
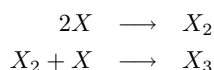
and this system will have a quite different dynamics to the corresponding third-order system.

VIII. Fundamental hypothesis of stochastic chemical kinetics

Let now generalize using a more formal approach the concepts exposed in the previous section. If we apply the foregoing arguments specifically to reactive collisions (i. e. to those collisions which results in an alteration of the state vector), the chemical reactions are more properly characterized by a *reaction probability per unit time* instead of a *reaction rate*. Thus, suppose that S_1 and S_2 molecules can undergo the reactions

$$R_1:\ S_1 + S_2 \to 2S_1 \tag{52}$$

Then in analogy with the Eq. (33), we may assert the existence of a constant c_1, which depends only on the physical properties of the two molecules and the temperature of the system, such that

$$
\begin{aligned}
c_1 dt =\ &\text{average probability that a particular 1-2}\\
&\text{molecular pair will react according to } R_1\\
&\text{in the next infinitesimal time interval } dt
\end{aligned} \tag{53}
$$

More generally if, under the assumption of spatial homogeneity (or thermal equilibrium) the volume V contains a mixture of X_i molecules of chemical species S_i, $(i = 1, 2, \ldots, N)$, and these N species can interact through M specified chemical reaction channels c_μ $(\mu = 1, 2, \ldots, M)$, we may assert the existence of M constants c_μ, depending only on the physical properties of the molecules and the temperature of the system. Formally, we assert that

$$
\begin{aligned}
c_\mu =\ &\text{average probability that a particular combination}\\
&\text{of } c_\mu \text{ reactant molecules will react accordingly to}\\
&c_\mu \text{ in the next infinitesimal time interval } dt.
\end{aligned} \tag{54}
$$

This equation is regarded both as the *definition* of the stochastic reaction constant c_μ, and also as the *fundamental hypothesis* of the stochastic formulation of chemical kinetics. This hypothesis is valid for any molecular system that is kept "well-mixed", either by direct stirring or else by simply requiring that non-reactive collisions occur much more frequently that reactive molecular collisions.

Let finally note that the master equation, that describe the time-evolution of the probability function $P(\{X_1, X_2, \ldots, X_N\}, t)$, may be derived from (54). To derive the master equation from the fundamental hypothesis of stochastic chemical kinetics, $P(\{X_1, X_2, \ldots, X_N\}, t)$ is expressed using the sum and the multiplication laws of probability theory. Thus $P(\{X_1, X_2, \ldots, X_N\}, t)$ is the sum of the probabilities of the $M + 1$ different ways in which the system can reach the state $(\vec{X} = X_1, X_2, \ldots, X_N)$ at time $t + dt$:

$$P(\{X_1, X_2, \ldots, X_N\}, t + dt) =$$
$$P(\{X_1, X_2, \ldots, X_N\}, t)(1 - \sum_{\mu=1}^{M} a_\mu dt) + \sum_{\mu=1}^{M} B_\mu dt \tag{55}$$

where

$$a_\mu dt \equiv$$
$$c_\mu \times \{\text{no. of distinct molecular combinations in the state } \vec{X}\}$$
$$= \text{probability that a } c_\mu \text{ reaction will occur in } V \text{ in } (t = dt),$$
$$\textit{given the system is in the state } (X_1, X_2, \ldots, X_N) \text{ at time } t \tag{56}$$

The first term in Eq. (55) is the probability that the system will be in the state $(X_1, X_2, \ldots, X_N)$ at time t, and then remains in that state (i. e. it undergoes no reactions) in $(t,\ t+dt)$. The quantity $B_\mu dt$ gives the probability that the system is one

c_μ reaction removed from the state $(X_1, X_2, \ldots, X_N)$ at time t, and the n undergoes an c_μ reaction in $(t, \ t + dt)$. Namely, B_μ will be the product of P evaluated at the appropriate once-removed state at t, times c_μ, times the number of c_μ molecular reactant combinations available in that once-removed state. Thus, Eq. (55) leads directly to the master equation

$$\frac{\partial}{\partial t} P(X_1, \ldots, X_N; t) = \sum_{\mu=1}^{M} [B_\mu - a_\mu P(X_1, \ldots, X_N : t)]$$

$$(57)$$

IX. GENERAL DERIVATION OF THE STOCHASTIC RATE CONSTANT

Here, we report the a general derivation for c_μ, developed by Wolkenhauer et. al. in [26]. Then we will compare it with the derivation of Gillespie. Let consider a reaction pathway involving N molecular species S_i. A network, which may include reversible reactions, is decomposed into M unidirectional basic reaction channels R_μ

$$R_\mu\colon l_{\mu1}S_{p(\mu,1)} + l_{\mu2}S_{p(\mu,2)} + \cdots + l_{\mu L_\mu}S_{p(\mu,L_\mu)} \xrightarrow{k_\mu} \ldots$$
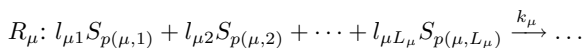
where L_μ is the number of reactant species in channel R_μ, $l_{\mu j}$ is the stoichiometric coefficient of reactant species $S_{p(\mu,j)}$, and the index $p(\mu, j)$ selects those S_i participating in R_μ. k_μ is the rate constant. Assuming a constant temperature and a homogeneous mixture of reactant molecules, the generalized mass action models (GMA) consist of N differential rate equations

$$\frac{d}{dt}[S_i] = \sum_{\mu=1}^{M} \nu_{\mu i} k_\mu \Pi_{j=1}^{L_\mu} [S_{p(\mu,j)}]^{l_{\mu j}} \quad (58)$$

where ν_μ denotes the change in molecules of S_i resulting from a single reactions R_μ. We write for concentrations and count of molecules, respectively
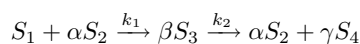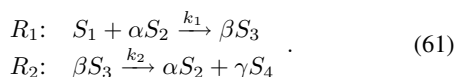
$$[S] = \frac{\langle S \rangle}{V} \quad (59)$$

and

$$\#S = S \times N_A \quad (60)$$

where N_A is the Avogadro's number. The units of $[S]$ are mol per liter, M = mol/liter. In this context, S is the number of moles and $\#S$ is the count of molecules.

Let use the following example for a chemical reaction

$$S_1 + \alpha S_2 \xrightarrow{k_1} \beta S_3 \xrightarrow{k_2} \alpha S_2 + \gamma S_4$$

which for the purpose of a stochastic simulation is split into two reactions channels

$$\begin{matrix} R_1\colon & S_1 + \alpha S_2 \xrightarrow{k_1} \beta S_3 \\ R_2\colon & \beta S_3 \xrightarrow{k_2} \alpha S_2 + \gamma S_4 \end{matrix} \quad (61)$$

The GMA representation of these reactions if given by the following rate equations

$$\begin{cases} \frac{d[S_1]}{dt} = -k_1[S_1][S_2]^\alpha \\ \frac{d[S_2]}{dt} = -\alpha k_1[S_1][S_2]^\alpha + \alpha k_2[S_3]^\beta \\ \frac{d[S_3]}{dt} = \beta k_1[S_1][S_2]^\alpha - \beta k_2[S_3]^\beta \\ \frac{d[S_4]}{dt} = \gamma k_2[S_3]^\beta \end{cases} \quad (62)$$

Substituting (59) and (60) in (58) gives

$$\frac{d}{dt}\langle \#S_i \rangle = \sum_{\mu=1}^{M} \frac{\nu_{\mu i} k_\mu}{(N_A V)^{K_\mu - 1}} \Pi_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} \quad (63)$$

where

$$K_\mu = \sum_{j=1}^{L_\mu} l_{\mu j}$$

denotes the molecularity of the reaction channel R_μ. The differential operator is justified only with the assumption of large numbers of molecules involved, such that near continuous changes are observed. Now, the "particle-O.D.E." for the temporal evolution of $\langle \#S_i \rangle$ is

$$\frac{d}{dt}\langle \#S_i \rangle = \sum_{\mu=1}^{M} \nu_{\mu i} k'_\mu \Pi_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} \quad (64)$$

Comparing (63) with (64) we find

$$k'_\mu = \frac{k_\mu}{(N_A V)^{K_\mu - 1}} \quad (65)$$

This equation than describes the interpretation of the rate constant, dependent on whether we consider concentrations or counts of molecules.

Let us now arrive to a general expression for the propensity a_μ. Note that from (64), the average number of reactions R_μ occurring in $(t, t + dt)$ is

$$\langle R_\mu \rangle = k'_\mu \Pi_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} dt \quad (66)$$

Let $\#R_\mu$ be the number of reaction R_μ. If we consider $\#R_\mu$ a discrete random variable with probability distribution function $p_{r_\mu} = Prob\{\#R_\mu = r_\mu\}$, where r_μ is the value assumed by the random variable $\#R_\mu$, the expectation value $\langle \#R_\mu \rangle$ is given by

$$\langle \#R_\mu \rangle = \sum_{r_\mu} r_\mu \langle p_{r_\mu} \rangle \quad r_\mu = 0, 1, 2, \ldots \quad (67)$$

where

$$p_{r_\mu} = \begin{cases} a_\mu dt + o(dt) & \text{if } r_\mu = 1 \\ 1 - a_\mu dt + o(dt) & \text{if } r_\mu = 0 \\ o(dt) & \text{if } r_\mu > 0 \end{cases} \quad (68)$$

where $o(dt)$ is a negligible probability for more than one R_μ reaction to occur during dt. Since p_{r_μ} is randomly varying and then the average $\langle p_{r_\mu} \rangle$ over the ensemble in (67), the equation (67) becomes

$$\langle \#R_\mu \rangle = 0 \cdot p_0 + 1 \cdot p_1 + \sum_{r_\mu > 1} r_\mu \langle p_{r_\mu} \rangle$$

From (67) and (68) we then have

$$\langle \# R_\mu \rangle = \langle a_\mu dt \rangle + o(dt) \tag{69}$$

where from (66) and (69) the propensity of R_μ reaction to occur in dt is given as

$$\langle a_\mu \rangle = k'_\mu \Pi_{j=1}^{L_\mu} \langle \# S_{p(\mu,j)} \rangle^{l_{\mu j}} \tag{70}$$

As already seen in the previous section, the propensity a_μ for a reaction R_μ is expressed as the product of the stochastic rate constant c_μ and the number h_μ of distinct combination of reactant molecules of R_μ

$$a_\mu = c_\mu \cdot h_\mu \tag{71}$$

In the literature h_μ is knows as *redundancy function*. This function varies over time in the following way

$$h_\mu(n) \begin{cases} \Pi_{j=1}^{L_\mu} \binom{n_{p(\mu,j)}}{l_{\mu j}} & \text{for } n_{p(\mu,j)} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{72}$$

If $n_{p(\mu,j)}$ is large and $l_{\mu j} > 1$, terms like $(n_{p(\mu,j)} - 1), \ldots, (n_{p(\mu,j)} - l_{\mu j} + 1)$ are not much different from $n_{p(\mu,j)}$ and we may write

$$h_\mu \approx \Pi_{j=1}^{L_\mu} \frac{(n_{p(\mu,j)})^{l_{\mu j}}}{l_{\mu j}!} = \frac{\Pi_{j=1}^{L_\mu}(n_{p(\mu,j)})^{l_{\mu j}}}{\Pi_{j=1}^{L_\mu} l_{\mu j}!} \tag{73}$$

We can write an alternative expression for a_μ by substituting (73) into (71) and considering the average

$$\langle a_\mu \rangle = c_\mu \cdot \left\langle \frac{\Pi_{j=1}^{L_\mu}(\# S_{p(\mu,j)})^{l_{\mu j}}}{\Pi_{j=1}^{L_\mu} l_{\mu j}!} \right\rangle \tag{74}$$

where $\# S_{p(\mu,i)}$ is the random variable whose value is $n_{p(\mu,j)}$. Comparing (70) with (74), we obtain

$$k'_\mu \Pi_{j=1}^{L_\mu} \langle \# S_{p(\mu,j)} \rangle^{l_{\mu j}} = \frac{c_\mu \left\langle \Pi_{j=1}^{L_\mu} \langle \# S_{p(\mu,j)} \rangle^{l_{\mu j}} \right\rangle}{\Pi_{j=1}^{L_\mu} l_{\mu j}!}$$

Making the assumption of zero covariance (i. e. $\langle \# S_i \# S_j \rangle = \langle \# S_i \rangle \langle \# S_j \rangle$ means for $i \neq j$ nullifying correlation, and for $i = j$ nullifying random fluctuations) gives

$$k'_\mu = \frac{c_\mu}{\Pi_{j=1}^{L_\mu} l_{\mu j}!} \tag{75}$$

which can be turned into an expression for c_μ

$$c_\mu = k'_\mu \cdot \Pi_{j=1}^{L_\mu} l_{\mu j}! \tag{76}$$

Inserting (65) for k'_μ, we arrive at

$$c_\mu = \left(\frac{k_\mu}{(N_A V)^{K_\mu - 1}} \right) \cdot \Pi_{j=1}^{L_\mu} l_{\mu j}! \tag{77}$$

Equation (77) is the law of conversion of the deterministic rate constant k_μ into the stochastic rate constant c_μ and is used in most implementation of Gillespie-like stochastic simulation algorithms. Note if above we substitute $\langle S \rangle / V$ in (58) for $[S]$ instead of $\langle \# S \rangle / (N_A V)$, the only difference to (65) and (77) is that N_A would not appear in these equations.

This derivation is different from the one given by Gillespie in [10]. The difference is that Wolkenhauer et al. introduced the average number of reactions (Eq. (66)) to move from the general GMA representation (58), which is independent of particular examples, to an expression that allow to derive parameter c_μ of the stochastic simulation (77) without referring to the temporal evolution of moments of CME. This makes the derivation more compact. Moreover, in [10] the temporal evolution of the mean is derived for examples of bi- and tri-molecular reactions only.

Finally, we add some comments to this derivation and its implications in a simulation algorithm. First, using the approximation (73) for $h_m u$ is valid for large numbers of molecules with $l_{\mu j} > 1$. In the simulations presented in this thesis this does not lead to significant differences. More important however is the fact that the derivation (77) relies on the rate constant of the GMA model. Nevertheless, this does not mean that the CME approach relies on the GMA model, since to derive rather than postulate a rate equation, one must first postulate a stochastic mechanism from which the GMA arises as a limit.

The existence of a relationship between deterministic and stochastic models, makes to suppose the existence of a way to compare these two approaches. In principle we can assert that the GMA model (58) has the following advantage with respect to the CME model: its terms and parameters are the direct translation of the biochemical reaction diagrams that capture the biochemical relationships of the molecules involved. On the contrary, rate equations are in virtually all cases simpler than CME. However for any realistic pathway model a formal analysis is not always feasible and a numerical solution (simulation) is the only way to compare two models. In this case the Gillespie algorithm, that will be presented in the following sections, provides an efficient implementation to generate realization of the CME (i. e. it is a realization of a time-continuous Markov process).

X. THE REACTION PROBABILITY DENSITY FUNCTION

In this section we introduce the foundation of the stochastic simulation algorithm of Gillespie. If we are given that the system is in the state $\vec{X} = \{X_1, \ldots, X_N\}$ at time t, computing its stochastic evolution means "moving the system forward in time". In order to do that we need to answer two questions.

1) When will the next reaction occur ?
2) What kind of reaction will it be ?

Because of the essentially random nature of chemical interactions, these two questions are answerable only in a probabilistic way.

Let introduce the function $P(\tau, \mu)$ defined as the probability that, give the state $\vec{X}$ at time t, the next reaction in the volume V will occur in the infinitesimal time interval $(t + \tau, t + \tau + d\tau)$, and will be an R_μ reaction. $P(\tau, \mu)$ is called *reaction probability density function*, because it is a joint probability density function on the space of the *continuous* variable τ ($0 \leq \tau < \infty$) and the *discrete* variable μ ($\mu = 1, 2, \ldots, M$) (i. e., according to the notation introduced in subsection IV-D, we are referring to a CDS model).

The values of the variables τ and μ will gives us answer to the two questions mentioned above. Gillespie showed that from the fundamental hypothesis of stochastic chemical kinetics (see section VIII) it is possible to derive an analytical expression for $P(\tau, \mu)$, and then use it to extract the values for τ and μ. Gillespie showed how to derive from the fundamental hypothesis and from an analytical expression of $P(\tau, \mu)$. First of all, $P(\tau, \mu)$ can be written as the product of $P_0(\tau)$, the probability that given the state $\vec{X}$ at time t, no reaction will occur in the time interval $(t, t + dt)$, times $a_\mu d\tau$, the probability that an R_μ reaction will occur in the time interval $(t + \tau, t + \tau + d\tau)$

$$P(\mu, \tau)d\tau = P_0(\tau)a_\mu dt \tag{78}$$

In turn $P_0(\tau)$ is given by

$$P_0(\tau' + d\tau') = P_0(\tau')\Big[1 - \sum_{i=1}^{M} a_i d\tau'\Big] \tag{79}$$

where $[1 - \sum_{i=1}^{M} a_i d\tau']$ is the probability that no reaction will occur in time $d\tau'$ from the state $\vec{X}$. Therefore

$$P_0(\tau) = \exp\Big[-\sum_{i=1}^{M} a_i \tau \tag{80}$$

Inserting (79) into (78), we find the following expression for the reaction probability density function

$$P(\mu, \tau) = \begin{cases} a_\mu \exp(-a_0\tau) & \text{if } 0 \leq \tau < infty \\ 0 & otherwise \end{cases} \tag{81}$$

where a_μ is given by (71) and

$$a_0 \equiv \sum_{i=1}^{M} a_i \equiv \sum_{i=1}^{M} h_i c_i \tag{82}$$

The expression for $P(\mu, \tau)$ in (81) is, like the master equation in (31), a rigorous mathematical consequence of the fundamental hypothesis (54). Notice finally that $P(\tau, \mu)$ depends on all the reaction constants (not just on c_μ) and on the current numbers of all reactant species (not just on the R_μ reactants).

XI. THE STOCHASTIC SIMULATION ALGORITHMS

In this section we review the three formulations of stochastic simulation variants of Gillespie algorithm: Direct, First Reaction, and Next Reaction Method.

A. Direct Method

On each step the Direct Method generates two random numbers r_1 and r_2 from a set set of uniformly distributed random numbers in the interval $(0, 1)$. The time for the next reaction to occur is given by $t + \tau$, where τ is given by

$$\tau = \frac{1}{a_0} \ln\left(\frac{1}{r_1}\right) \tag{83}$$

The index μ of the occurring reaction is given by the smallest integer satisfying

$$\sum_{j=1}^{\mu} a_j > r_2 a_0 \tag{84}$$

The system states are updated by $X(t + \tau) = X(t) + \nu_\mu$, then the simulation proceeds to the next occurring time.

Algorithm

1) Initialization: set the initial numbers of molecules for each chemical species; input the desired values for the M reaction constants $c_1, c_2, \ldots, c_M$. Set the simulation time variable t to zero and the duration T of the simulation.
2) Calculate and store the propensity functions a_i for all the reaction channels $(i = 1, dots, M)$, and a_0.
3) Generate two random number r_1 and r_2 in $Unif(0, 1)$.
4) Calculate τ according to (83)
5) Search for μ as the smallest integer satisfying (84).
6) Update the states of the species to reflect the execution of μ (e. g. if R_μ:$S_1 + S_2 \to 2S_1$, and there are X_1 molecules of the species S_1 and X_2 molecules of the species S_2, then increase X_1 by 1 and decrease X_2 by 1). Set $t \leftarrow t + \tau$.
7) If $t < T$ then go to step 2, otherwise terminate.

Note that the random pair (τ, μ), where τ is given by (83) and μ by (84), is generated according to the probability density function in (81). A rigorous proof of this fact may be found in [8]. Suffice here to say that (83) generates a random number τ according to the probability density function

$$P_1(\tau) = a_0 \exp(-a_0\tau) \tag{85}$$

while (84) generates an integer μ according to the probability density function

$$P_2(\mu) = \frac{a_\mu}{a_0} \tag{86}$$

and the stated result follows because

$$P(\tau, \mu) = P_1(\tau) \cdot P_2(\mu)$$

Note finally that, to generate random numbers between 0 and 1 we can do as follows. Let $F_X(x)$ be a distribution function of an exponentially distributed variable X and let $U \sim Unif[0, 1)$ denote an uniformly distributed random variable U on the interval 0 to 1.

$$F_X(x) = \begin{cases} 1 - e^{-ax} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{87}$$

$F_X(x)$ is a continuous non-decreasing function and this implies that it has an inverse F_X^{-1}. Now, let $X(U) = F_X^{-1}(U)$ and we get the following

$$\begin{aligned} P(X(U) \leq x) &= P(F_X^{-1}(U) \leq x) \\ &= P(U \leq F_X(x) \tag{88} \\ &= F_X(x) \tag{89} \end{aligned}$$

It follows that

$$F_X^{-1}(U) = -\frac{\ln(1-U)}{a} \sim Exp(a) \qquad (90)$$

In returning to step 1 from step 7, it is necessary to re-calculate only those quantities a_i, corresponding to the reactions R_i whose reactant population levels were altered in step 6; also a_0 must be re-calculated simply by adding to it the difference between each newly changed a_i value and its corresponding old value. This algorithm uses M random numbers per iteration, takes time proportional to M to update the a_i's and takes time proportional to M to identify the smallest putative time.

B. First Reaction Method

The First Reaction Method generates a τ_k for each reaction channel R_μ according to

$$\tau_i = \frac{1}{a_i} \ln\left(\frac{1}{r_i}\right) \qquad (91)$$

where $r_1, r_2, \ldots, r_M$ are M statistically independent samplings of $Unif(0,1)$. Then τ and μ are chosen as

$$\tau = \min\{\tau_1, \tau_2, \ldots, \tau_M\} \qquad (92)$$

and

$$\mu = \text{the index of } \min\{\tau_1, \tau_2, \ldots, \tau_M\} \qquad (93)$$

Algorithm

1) Initialization: set the initial numbers of molecules for each chemical species; input the desired values for the M reaction constants $c_1, c_2, \ldots, c_M$. Set the simulation time variable t to zero and the duration T of the simulation.
2) Calculate and store the propensity functions a_i for all the reaction channels $(i = 1, dots, M)$, and a_0.
3) Generate M independent random numbers from $Unif(0, 10$.
4) Generate the times τ_i, $(i = 1, 2, \ldots, M)$ according to (91).
5) Find τ and μ according to (92) and (93), respectively.
6) Update the states of the species to reflect the execution of reaction μ. Set $t \leftarrow t + \tau$.
7) If $t < T$ then go to step 2, otherwise terminate.

The Direct and the First Reaction methods are fully equivalent to each other [9], [8]. The random pairs (τ, μ) generated by both methods follow the same distribution.

C. Next Reaction Method

Gibson and Bruck [7] transformed the First Reaction Method into an equivalent but more efficient new scheme. The Next Reaction Method is more efficient than the Direct method when the system involves many species and loosely coupled reaction channels. This method can be viewed as an extension of the First Reaction Method in which the unused

$M - 1$ reaction times (92) are suitably modified for reuse. Clever data storage structures are employed to efficiently find τ and μ.

Algorithm

1) Initialize:
   - set the initial numbers of molecules, set the simulation time variable t to zero, generate a dependency graph G;
   - calculate the propensity functions a_i, for all i
   - for each i, $(i = 1, 2, \ldots, M)$, generate a putative time τ_i, according to an exponential distribution with parameter a_i
   - store the τ_i values in an indexed priority queue P.
2) Let μ be the reaction whose putative time τ_μ stored in P, is least. Set $\tau \leftarrow \tau_\mu$.
3) Update the states of the species to reflect the execution of the reaction μ. Set $t \leftarrow \tau_\mu$.
4) For each edge (μ, α) in the dependency graph G
   - update a_0
   - if $\alpha \neq \mu$, set

$$\tau_\alpha \leftarrow \frac{a_{\alpha,old}}{a_{\alpha,new}}(\tau_\alpha - t) + t \qquad (94)$$

   - if $\alpha = \mu$, generate a random number r and compute τ_α according to the following equation

$$\tau_\alpha = \frac{1}{a_\alpha(t)} \ln\left(\frac{1}{r}\right) + t \qquad (95)$$

   - replace the old τ_α value in P with the new value
5) Go to step 2.

Two data structures are used int his method:

- The *dependency graph* G is a data structure that tells precisely which a_i should change when a given reaction is executed. Each reaction channel is denoted as a node in the graph. A direct edge connects R_i to R_j if and only if the execution of R_i affects the reactants in R_j. The dependency graph can be used to recalculate only the minimal number of propensity functions in step 4.
- The indexed priority queue consists of a tree structure of ordered pairs of the form (i, τ_i), where i is a reaction channel index and τ_i is the corresponding time when the next R_i reaction is expected to occur, and an index structure whose ith element points to the position in the tree which contains (i, τ_i). In the tree, each parent has a smaller τ than either of its children. The minimum τ always stays in the top of the node and the order is only vertical. In each step the update changes the value of the node and then bubbles it up or down according to its value to obtain the new priority queue. Theoretically, this procedure takes at most $\ln(M)$ operations. In practice, usually there are a few reactions that occur much more frequently. Thus, the actual update takes less than $\ln(M)$ operations.

The Next Reaction Method takes some CPU time to maintain the two data structures. For a small system, this cost

dominated the simulation. For a large system, the cost of maintaining the data structures may be relatively smaller compared with the savings. The argument for the advantage of the Next Reaction Method over the Direct Method is based on two observations: first, in each step, the Next Reaction Method generates only one uniform random number, while the Direct Method requires two. Second, the search for the index μ of the next reaction channel takes $O(M)$ time for Direct Method, while the corresponding cost for the Next Reaction Method is on the update of the indexed priority queue which is $O(\ln(M))$

XII. TIME-DEPENDENT EXTENSION OF FIRST REACTION METHOD

The Gillespie algorithm has been used on numerous occasion to analyze biochemical kinetics. Its success is due to its proved equivalence with Master Equation and its efficiency and precision: no time is wasted on simulation iterations in which no reactions occur, and the treatment of the time as a continuum allows the generation of exact series of τ values based on rigorously derived probability density function. However, all the formulations of the algorithm consider the reaction rate constant in time and have not taken into account for the effects of temporal changes of volumes, temperature, activation energy and presence of catalysts concentration. In this section, we provide an extension of First Reaction Methods to the case of time-depending rates. Our extension is inspired inspired to [22], and focuses on the time dependence of the kinetic rates on volume and temperature deterministic changes. This re-formulation has been adapted to be incorporated in the framework of stochastic π-calculus and its implementation has been succesfully applied to a sample simulation in biology: the passive glucose cellular transport [19], [18].

Let suppose that the volume $V_s(t)$ contains a mixture of chemical species, X_i ($i = 1, \ldots, N$) which may interact through the reaction channels R_μ, $\mu = 1, \ldots, M$. Let suppose furthermore that a subset of these channels is characterized by the time-dependent propensities

$$a_s(t) = a'_s/V(t), \quad s = 1, \ldots, S \qquad (96)$$

and an other sub-set is characterized by the time-dependent propensities

$$a_q(t) = a'_q/V(t), \quad q = S+1, \ldots, M \qquad (97)$$

where a'_s and (a'_q) are the time-independent propensities, that have to be computed using the Eqs. (38), (41), (44), according to the type of reaction.

Following the Gillespie approach, let introduce these probabilities:

1) $P(\tau, \mu|Y, t)d\tau$: probability that, given the state $Y = (X_1, \ldots, X_N)$ at time t, the next reaction will occur in the infinitesimal time interval $(t + \tau, t + \tau + d\tau)$, at it will be reaction R_μ
2) $a_\mu(t)dt$: probability that, given the state $Y = (X_1, \ldots, X_N)$ at time t, reaction R_μ will occur within the interval $(t, t + dt)$.

$P(\tau, \mu|Y, t)d\tau$ is computed as a product of the probabilities that no reaction will occur within $(t, t+\tau)$ times the probability that R_μ will occur within the subsequent interval $(t + \tau, t + \tau + d\tau)$

$$P(\tau, \mu|Y, t)d\tau = P_0(\tau|Y, t) \cdot a_\mu(\tau + t)d\tau \qquad (98)$$

where, summing over all reaction channels $\mu = 1, \ldots, M$ and splitting the sum in the two terms over s and q

$$P_0(\tau + d\tau|Y, t) =$$
$$P_0(\tau|Y, t)\left[1 - d\tau \sum_{s=1}^{S} a_s(t + \tau) - d\tau \sum_{q=S+1}^{M} a_q(t + \tau)\right]$$

With the initial condition $P_0(\tau = 0|Y, t) = 1$, the solution of this differential equation is

$$P_0(\tau|Y, t) =$$
$$\exp\left[-\sum_s \int_t^{t+\tau} a_s(t + \tau')d\tau' - \sum_q \int_t^{t+\tau} a_q(t + \tau')d\tau'\right] \tag{99}$$

Now, by combining Eq. (98) with the Eq. (99), we obtain

$$P(\tau, \mu|, Y, y) = a_\mu(t + \tau) \times$$
$$\exp\left[-\sum_s \int_t^{t+\tau} a_s(t + \tau') - \sum_q \int_t^{t+\tau} a_q(t + \tau')d\tau'\right] \tag{100}$$

By introducing two functions $f_s(\tau)$ and $f_q(\tau)$ describing the variation of volume in time, the time-dependence of the volumes can be described by these expressions:

$$V_s(t + \tau) = V_s(t)f_s(\tau) \quad \text{and} \quad V_q(t + \tau) = V_q(t)f_q(\tau).$$

Consequently, the propensities are

$$a_s(t + \tau) = a_s(t)/f_s(\tau) \quad \text{and} \quad a_q(t + \tau) = a_q(t)/f_q(\tau).$$

Substituting these expressions in Eq. (100), and introducing, for convenience

$$A_s \equiv \sum_s a_s(t) \quad \text{and} \quad A_q \equiv \sum_q a_q(t)$$

$$F_s(\tau) \equiv \int_t^{t+\tau} \frac{1}{f_s(\tau')}d\tau' \quad \text{and} \quad F_q(\tau) = \int_t^{t+\tau} \frac{1}{f_q(\tau')}d\tau'$$

so that Eq. (100) can be re-written as

$$P(\tau, \mu|Y, t) = \begin{cases} \frac{a_s(t)}{f_s(\tau)} \cdot \exp\left[-A_s F_s(\tau) - A_q F_q(\tau)\right] \\ \frac{a_q(t)}{f_q(\tau)} \cdot \exp\left[-A_s F_s(\tau) - A_q F_q(\tau)\right] \end{cases} \tag{101}$$

Finally, the probability of any reaction occurring between time t and the time $t+T$, is obtained by integrating Eq. (101) over time and summing over all channels:

$$\int_0^T \sum_\mu P(\tau,\mu|Y,t)d\tau = \begin{cases} \int_0^T \sum_{s'=1}^S \frac{a_{s'(t)}}{f_s(\tau)} \cdot \exp\left[-A_s F_s(\tau) - A_q F_q(\tau)\right]d\tau \\ \int_0^T \sum_{q'=S+1}^M \frac{a_{q'(t)}}{f_q(\tau)} \cdot \exp\left[-A_s F_s(\tau) - A_q F_q(\tau)\right]d\tau \end{cases} \tag{102}$$

Generalizing, in systems where the physical reaction space is divided into n sub-spaces whose volumes change in time, the probability density function of reaction is split into n exponential terms multiplied by the ratio between reaction propensity and volume of the subspace. The volume of each sub-spaces can follow a different temporal behavior. Consequently a different reaction probability and a different expression of reaction time are obtained for each sub-regions of the space.

The effect of temperature changes on the probability density function can be simulated by expressing the time dependence of the propensity of a reaction μ as $a_\mu(t+\tau) = a_\mu(t) \cdot T(\tau)$, where $T(\tau) = \exp(1/(a + b\tau))$ models the variation of the propensity function following the Arrhenius formula (for instace, il Lecca [18] $a = 37°C$ and $b = 1\ °C/min$).

XIII. SPATIO-TEMPORAL ALGORITHMS

Previous sections cover the stochastic algorithms for modeling biological pathways with no spatial information. However, the real biological world consists of components which interact in a three dimensional space. Within a cell compartment, the intracellular material is not distributed homogeneously in space and molecular localization plays an important role, e. g. diffusion of ions and molecules across membranes and propagation of an action potential along a nerve fiber's axon. Thus, basic assumption of spatial homogeneity and large concentration diffusion is no longer valid in realistic biological systems [4]. In this context, stochastic spatio-temporal simulation of biological system is required.

The enhancement on the performance of Gillespie Algorithms has made the spatio-temporal simulation tractable. Stundzia and Lumsden [31], and Elf et al. [4], extended the Gillespie Algorithms to model intracellular diffusion. They formalized the reaction-diffusion master equation and the diffusion probability density functions. The entire volume of a model was divided into multiple subvolumes and by treating diffusion processes as chemical reactions, the Gillespie Algorithm was applied without much modification. Stundzia has showcased the application of the algorithm on calcium wave propagation within living cells and has observed regional fluctuations and spatial correlations in the small particles limit. However, this approach requires detailed knowledge about the diffusion processes to be available, in order to estimate the probability density function for diffusion. Furthermore, the algorithms have only been applied to small systems with finite number of molecular species but require large amount of computational power.

Shimizu in [30] also extended the Stochsim algorithm to include spatial effects of the system. In his approach, spatial information was added to the attributes of each molecular species and a simple two dimensional lattice was formed to enable interaction between neighboring nodes. The algorithm was applied to study the action of a complex of signaling proteins associated with the chemotactic receptors of coliform bacteria. He showed that the interactions among receptors could contribute to high sensitivity and wide dynamic range in the bacterial chemotaxis pathway.

Another way of simulating stochastic diffusion is to directly approximate the Brownian movements of the individual molecules (MCell [2]). In this case, the motion and direction of the molecules are determined by using random numbers during the simulation. Similarly, collisions with potential binding sites and surfaces are detected and handled by using only random numbers with a computed binding probability. MCell is capable of treating stochastic and a 3-dimensional biological model that involves a discrete number of molecules. Though MCell incorporates 3D spatial partitioning and parallel computing to increase algorithmic efficiency, the simulation is limited to the microphysiological processes such as synaptic transmission due to high computational requirement.

Apart from the enhancements on various algorithms, the simulation of a spatio-stochastic biological system is still a challenging problem. To address it the author recently proposed a new mathematical treatment of diffusion that can be incorporated in a stochastic algorithm simulating the dynamics of a reaction-diffusion system is presented [20], [21]. The movement of a molecule A from a region i to a region j of the space is represented as a first order reaction $A_i \xrightarrow{k} A_j$, where the rate constant k depends on the diffusion coefficient. The diffusion coefficients are modeled as function of the local concentration of the solutes, their intrinsic viscosities, their frictional coefficients and the temperature of the system. The stochastic time evolution of the system is given by the occurrence of diffusion events and chemical reaction events. At each time step an event (reaction or diffusion) is selected from a probability distribution of waiting times determined by the intrinsic reaction kinetics and diffusion dynamics. To demonstrate the method the simulation results of the reaction-diffusion system of chaperone-assisted protein folding in cytoplasm are shown.

XIV. THE LANGEVIN EQUATION

While internal fluctuations are self-generated int he system, and they can occur in closed and open systems as well, external fluctuations are determined by the environment of the system. We have seen that a characteristic property of internal fluctuations is that they scale with the system size and tend do vanish in the thermodynamics limit. External noise has a crucial role in the formation of ordered biological structures. External noise-induced ordering was introduced to model the ontogenetic development and plastic behavior of certain neural structures [5]. Moreover, it was demonstrated that noise can support the transition of a system from a stable state to another stable state. Since stochastic models might exhibit qualitatively different behavior than their deterministic counterpart, external

noise can support transitions to states which are not available (or even do not exist) in a deterministic framework [15].

In the case of extrinsic stochasticity, the stochasticity is introduced by incorporating multiplicative or additive stochastic terms into the governing reaction equations. These terms, normally viewed as random perturbations to the deterministic system, are also known as stochastic differential equations. The general equation is:

$$\frac{dx}{dt} = f(x) + \xi_x(t) \tag{103}$$

The definition of the additional term ξ_x differs according to the formalism adopted. In Langevin Equations [12], ξ_x is represented by Eq. (104) Other studies [14] adopt a different definition where $\xi_i(t)$ is a rapidly fluctuating term with zero mean ($\{\xi_i(t)\} = 0$). The statistics of i(t) are such that ($\{\xi_i(t)\xi_i(t')\} = 0) = D\delta_{ij}(t - t')$ to maintain independence of random fluctuations between different species (D is proportional to the strength of the fluctuation).

$$\xi_x(t) = \sum_{j=1}^{M} V_{ij} \sqrt{\alpha_j} X(t) N_j(t) \tag{104}$$

where V_{ij} is the change in number of molecules of species i brought by one reaction j and N_j are statistically independent normal random variables with mean 0 and variance 1.

A. Use and abuse of Langevin equation

The way in which Langevin introduced fluctuations into the equation of molecular population level evolution does not carry over nonlinear systems. This section briefly sketch the difficulties to which such a generalization leads. External noise denotes fluctuations created in an otherwise deterministic system by the application of a random force, whose stochastic properties are supposed to be known. Internal noise is due to the fact that the system itself consists of dicrete particles. It is inherent in the mechanism by which the state of the system evolves and cannot be divorced from its evolution equation. A Brownian particle, with its surrounding fluid is a cloded physical system with internal noise. Langevin, however, treated, the particle as a mechanical system subject to the force exerted by the fluid. This force he subdivided in a deterministic damped force and a random force, which he treated as external, i. e. its properties as a function of time were supposed to be known. For the physical pictures, these properties will not be altered if an additional force on the particle is introduced.

In more recent years, however, Eq. (103) has been used also in modeling the evolution of biochemical systems, although the noise source in a chemical reacting network is internal and no physical basis is available for a separation into a mechanical part and a random term with known properties. The strategy used in the application of Langevin equation in modeling the evolution of a system of chemical reacting particles is the following. Suppose to have a system whose evolution is described phenomenologically by a deterministic differential equation

$$\frac{dx}{dt} = f(x) \tag{105}$$

where x stands for a finite set of macroscopic variables, but for simplicity in the present discussion we take the case that x is a single variable. Let suppose to know that for some reason there must also be fluctuations about this macroscopic values. Therefore, we supplement (105) with a Langevin term

$$\frac{dx}{dt} = f(x) + L(t) \tag{106}$$

Note now, that on averaging (106) one does not find that $\langle x \rangle$ obeys to the phenomenological equation (105), rather than

$$\partial_t \langle x \rangle = \langle f(x) \rangle = f(\langle x \rangle) + \frac{1}{2}\langle(x - \langle x \rangle)^2\rangle\partial_t^2(\langle x \rangle) + \ldots$$

It follows that $\langle x \rangle$ does not obey any differntial equation at all. This reveals the basic flaw in the application of the Langevin approach to the internal noise of systems whose phenomenological law is nonlinear. The phenomenological equation (105) holds only in the approximation in which fluctuations are neglected. That implies that $f(x)$ is determined phenomenologically with an inherent margin of uncertainty of the order of fluctuations. If we deduce a certain form of $f(x)$ from a theory or experiment in which fluctuations are ignored there is no justification for postulating that $f(x)$ is to be used in (106). There may be a mismatch between both of the same size as the fluctuations; thta would not show up in macroscopic results, but cannot of course be neglected in the equation of the fluctuations themselves.

XV. Hybrid algorithms

Biological system are stiff by nature in the sense that processes with very different time scales are coupled. Some molecues are quickly synthesized and degenrated (typically metabolites) and take a long time to run over (typically macromolecules). Some biochemical reactions involve a chain of many steps, while other reactions just involve a single association of dissociation event. We have already seen in section III-B1 that this difference in time scales can be exploited by assuming quasi-equilibrium and usign the equilibrium constant to eliminate from the model some components, and thus to reduce its complexity.

Stochastic algorithms suffer from the same "stiffness" problems as that of deterministic algorithms. In order to capture the fast dynamics of the system, entire simulation is slowed down significantly. Hence, the basic idea of hybrid algorithms aims to exploit the advantages of other algorithms to offset the disadvantages of the stochastic algorithms.

Several attempts have been made to illustrate the relevance and feasibility of hybrid algorithms. Bundschuh et al [28], Haseltine and Rawlings [13], and Puchalka and Kierzek [27], have used a similar approach to integrate ODE/Langevin with Gillespie algorithms. In both cases, the modeler has to identify methods and criteria to partition the system into fast dynamics and slow dynamics sub-systems. The fast dynamics subsystem can be handled by either ODE or Langevin Equations while the slow dynamics subsystem can be handled by Gillespie algorithms. In addition, numerical treatment such as the "slow variables" in [28], and the "probability of no reaction" in

Haseltine and Rawlings [13], is required to maintain accuracy of the solutions. The algorithms show promising results and the results are consistent with those from Gillespie algorithms. Haseltine and Rawlings in [13], showed the applicability of hybrid algorithms by simulating the effect of stochasticity to the bi-modality of an intracellular viral infection model using the algorithm. Kiehl et al. [17], also tested the algorithms on the λ phage model.

The relevance of hybrid algorithms has been pointed out in several papers (Alur et al. [1]; Matsuno et al. [24]; Bockmayr and Courtois [3]). Bockmayr and Courtois used hybrid constraint programming methods to model an alternative splicing regulation model. This implementation is very useful under circumstances where detailed knowledge about the model is unavailable. Meanwhile, Alur et al. used CHARON, a formal description language of hybrid system which combines ODE with "mode switching" mechanism to model the quorum sensing phenomenon in *Vibrio fischeri*, a marine bacterium that involves the Lux regulon. A Hybrid Petri Net [24] approach has been employed to model a hybrid system using ODEs and discrete events. This method has been used to model the growth pathway control of λ phage.

Hybrid algorithms aim to close the gap between macroscopic and mesoscopic scales of the system. In particular, the relevance of hybrid modeling has been proved necessary to capture the behavior of a real biological system. Moreover, hybrid algorithms have substantially cut down the computational cost of large scale modeling and simulation. One major drawback here is that by introducing additional numerical treatment to the algorithms, more parameters have to be defined and the accuracy of the solutions is dependent on the accuracy of parameters. Mostly, the simulations result in solutions of highly tuned parameters. Although these hybrid approaches show significant improvements in the computational cost, there are still lots of computational issues to be resolved before it can be applied to a realistic problem. Some of the issues are:

- accuracy of results,
- consistency of system parameters between different levels of abstraction,
- highly non-linear system,
- methodology to separate the systems into different subsystems, dynamic switching between different mathematical formalisms.

XVI. Conclusions

We end this paper by noticing that the biochemical approach to understand biological processes is essentially one of simulation. A biochemist typically prepares a cell-free extract that can mediate a well-described physiological process. Once the extract is fractioned to purify the components that catalyze individual reactions, the physiological process in reconstructed in vitro. The validity of this approach is measured by how closely the in vitro reconstructed process matches physiological observations.

Similarly, the validity of a model in its conceptual framework is measured by how closely its simulation matches physiological observations. Unfortunately, often controlled experiments cannot be performed on the system to validate our model (for example, how can the model be validated if only a single historical dataset exists ?). The validation becomes difficult also when the model is stochastic, i. e. it has random elements. However, whatever the nature of the model is, in general validation ensures that the model meets its intended requirements in terms of the methods employed and the results obtained. The ultimate goal of model validation is to make the model useful in the sense that the model addresses the right problem, provides accurate information about the system being modeled, and to makes the model actually used [23], [29].

References

[1] R. Alur, C. Belta, F. Ivancic, V. Kumar, M. Mintz, G. Pappas, H. Rubin, and J. Schug. Hybrid modeling and simulation of biomolecular networks. In *Hybrid System. Computation and Control, 4th International Workshop, HSCC, Rome Italy*, 2001.

[2] T. M. Bartol and J. R. Stiles. *M-cell, http://www.MCell.cnl.salk.edu*, 2002.

[3] A. Bockmayr and A. Courtois. Using hybrid concurrent programming to model dynamics biological systems. In *18th International Conference on Logic Programming, ICLP02*, pages 85–99. Springer, LNCS 2401, July 2002.

[4] J. Elf, A. Doncic, and M. Eherenberg. Mesoscopic reaction-diffusion in intracellular signaling. In *Proceedings of SPIE 5110*, pages 114–124, 2003.

[5] P. Erdi and G. Barna. Self-organisation in neural systems. some illustrations. *Lecture Notes in Bioinformatics*, 71, 1993.

[6] B. Ermentrout. *Simulating, analyzing, and animating dynamical systems: a guide to XPPAUT for researchers and students*. 1st edition. SIAM New York, 2002.

[7] M. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A*, 104, 2000.

[8] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical species. *J.Comp. Physics*, 22:403–434, 1976.

[9] D. T. GIllespie. Exact stochastic simulation of coupled chemical reactions. *The J. of Physical Chemistry*, 81(25), 1977.

[10] D. T. Gillespie. *Markov Processes*. Academic Press, 1992.

[11] D. T. GIllespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.

[12] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115:1716–1733, 2001.

[13] E. L. Haseltine and J. B. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.*, 117:6959–6969, 2002.

[14] J. Hasty and F. Issacs. Designer gene networks: toward fundamental cellular control. *CHAOS*, 11(1):207–220, 2001.

[15] W. Horsthemke and L. Hanson. Non equilibrium chemical instabilities in continuous flow stirred tank reactors: the effect of stirring. *J. Chem. Phys.*, 81, 1993.

[16] P. S. Jöberg. Numerical solution of the Fokker-Planck approximation of the chemical master equation. Master's thesis, Dept. of Information Technology, Uppsala University, 2005.

[17] T. R. Kiehl, R. M. Mattheyses, and M. K. Simmons. Hybrid simulation of cellular behavior. *Bioinformatics*, 20:316–322, 2004.

[18] P. Lecca. Simulating the cellular passive transport of glucose using a time-dependent extension of gillespie algorithm for stochastic π - calculus. *Int. Journal of Data Mining and Bioinformatics*, 1(4), 2006.

[19] P. Lecca. A time-dependent extension of gillespie algorithm for biochemical stochastic π-calculus. In *SAC ACM '06*, 2006.

[20] P. Lecca and L. Demattè. Stochastic simulation of reaction diffusion systems. *In. Journal of Medical and Biological Engineering*, 1(4):211–231, 2008.

[21] P. Lecca, L. Demattè, and C. Priami. Modeling and simulating reaction-diffusion systems with state-dependent diffusion coefficients. In *Int. Conference on Bioinformatics and Biomedicine 2008*, volume 34. World Academy of Science, Engineering and Technology.

[22] T. Lu., L. Tsimring D. Volfson, and J. Hasty. Cellular growth and division in the gillespie algorithm. *Syst. Biol.*, 1, 2004.

[23] C. M. maclan. Model verification and validation. In *Workshop on "Threat Anticipation: Social Science Methods and Models", The University of Chicago and Argonne National Laboratory*.

[24] H. Matsuno, A. Doj, M. Nagasaki, and S. Miyano. Hybrid petri net representations of gene regulatory networks. In *Pac. Symp. Biocomput.*, pages (5) –333–349, 2000.

[25] D. A. McQuarrie. Stochastic approach to chemical kinetics. *J. Appl. Prob.*, 4:413–478, 1967.

[26] W. Kolch O. Wolkenhauer, M. Ullah and K. Cho. Modelling and simulation of intracellular dynamics: Choosing an appropriate framework. *IEEE Transaction on Nano-Bioscience, Special Issue molecular and subcellular system biology*, 2004.

[27] J. Puchalka and A. M. Kierzek. Binding the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks. *Biophys. J.*, 86:1357–1372, 2004.

[28] F. Hayot R. Bundschuh and C. Javaprakash. Fluctuations of slow variables in generic netwroks. *Biophys. J.*, 84:1606–1615, 2003.

[29] R. G. Sargent. Simulation model verification and validation. In *Proceedings of the 23rd conference on Winter simulation*, pages 37–47. IEEE Computer Society Washington, DC, USA.

[30] T. S. Shimizu. *The spatial organisation of cell signaling pathways - a computer based study*. PhD thesis, University of Cambridge, UK, 2002.

[31] A. B. Stundzia and C. J. Lumsden. Stochastic simulation of coupled reaction-diffusion processes. *J. Comput. Phys.*, 127:196–207, 1996.

[32] N. G. van Kampfen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, 1992.

[33] D. J. Wilkinson. *Stochastic Modeling for System Biology*. Chapman & Hall, 2006.