

Mathematics of Planet Earth 4

Arianna Bianchi
Thomas Hillen
Mark A. Lewis
Yingfei Yi *Editors*

The Dynamics of Biological Systems



 Springer

Chapter 6

Stochastic Population Kinetics and Its Underlying Mathematicothermodynamics



Hong Qian

Abstract Based on differential calculus, classical mechanics represents the natural world in terms of featureless point masses and their movements. Chemistry studies molecules each of which has a large number of internal degrees of freedom in terms of atoms, electrons, etc.; the behavior of even a single biomolecule like a protein is often so complex that the foundation of chemical kinetics is essentially based on stochastic mathematics. Stochastic population kinetics is a more powerful and more realistic representation of the biological world. This chapter introduces this new mathematical modeling paradigm and shows the existence of a hidden thermodynamic structure underlying any stochastic nonlinear kinetic description of a multi-population biological system. The mathematicothermodynamics presented here is a generalization of J. W. Gibbs' chemical thermodynamics for equilibrium chemical reaction systems, as heterogeneous matters.

6.1 Introduction

Françis Jacob (1920–2013), one of the leading molecular biologists of the twentieth century, stated in his book “The Possible and the Actual” [13] that Western art had radically changed since the Renaissance from “symbolizing” to “represent” the real world. One can in fact view pure versus applied mathematics as a change from the former to the latter. The ultimate goal of mathematical science is to quantitatively represent the real world in terms of mathematics.

Currently there is a sharp contrast between the mathematical models, or theories, in physics and in biology. While we take Newton's equation of motion as almost the “Truth” under the appropriate conditions, one does not have such a level of confidence for the mathematical models in biology.

H. Qian (✉)

Department of Applied Mathematics, University of Washington, Seattle, WA, USA
e-mail: qian@amath.washington.edu

© Springer Nature Switzerland AG 2019

A. Bianchi et al. (eds.), *The Dynamics of Biological Systems*, Mathematics of Planet Earth 4, https://doi.org/10.1007/978-3-030-22583-4_6

149

In Newtonian mechanics, the natural world is represented by point masses and described by their movements. Each point mass, e.g., a Newtonian particle, has a unique position and velocity. The natural world according to chemistry, however, consists of “identical” molecules made of atoms. While each individual molecule has intrinsic stochasticity, e.g., a molecular individualism [4] due to the atomic motions within, population wise molecules follow statistical rate laws in their syntheses (birth), degradations (death), spatial diffusion (migration), state transitions (character switching), and interactions. Such a formal reaction kinetic system in a small volume V , such as biochemical reaction kinetics in a single cell, can be rigorously treated in terms of an integer-valued, continuous-time Markov process describing its nonlinear behavior, counting the molecules and their reactions, one at a time.

Population dynamics in biology has long been described in terms of nonlinear differential equations [17]. Many of the equations are remarkably similar to the kinetic equation for chemical reactions. In this chapter, we shall introduce in a rigorous fashion the rate law of rare events in term of exponential waiting time and the Poisson process. We shall show that the type of differential equations for population dynamics has a mathematical foundation in the theory of probability and Markov processes.

After introducing the stochastic mathematical representation of population kinetics, in the Sect. 6.9 of the chapter, we present a recently discovered universal mathematical structure that is inherent in any Markov population kinetics. This structure has a remarkable resemblance to the theory of thermodynamics, first developed in the nineteenth century by physicists dealing with heat—the stochastic motions of atoms and molecules. To distinguish the mathematical structure in the stochastic population kinetics from the subject from physics, we coined the term *mathematicothermodynamics*, within which we axiomatically introduce notions such as closed systems, open-driven systems, entropy production, free energy dissipation, etc. We shall derive two “laws”: The first is concerned with the balance of a free energy like function, and the second is concerned with certain monotonicity in the dynamics.

Finally, phase transition in physics, conformational transition in biochemistry, and phenotypic switching in cell biology are all nonlinear phenomena intrinsically related to multi-stability and saddle-node bifurcation, in the limits of time $t \rightarrow \infty$ and system’s size $V \rightarrow \infty$ [12, 26].

6.2 Probability and Stochastic Processes: A New Language for Population Dynamics

There are fundamentally two types of mathematical modeling: (a) representing scientific data in terms of mathematical formula or equations and (b) describing a system’s behavior (natural or engineered, physical or biological, electronic,

chemical, economical, social, ...) based on existing, established formula and equations. For lack of better terminology, we shall call the former ***data-driven modeling*** and the latter ***mechanistically derived modeling***. Note, according to Karl Popper (1902–1994) and his philosophy of science, the only legitimate scientific activity is falsifying a hypothesis: that requires first to formulate a hypothesis, which sometime is just looking for patterns in the data (e.g., numerical hypothesis) and sometime is proposing a mechanism (e.g., modeling); and (b) to derive rigorous predictions from a hypothesis, which is a form of logical, or mathematical, deduction.

Let us revisit some of the key notions already discussed, or widely used, in many of the other chapters—but let us try to be critical. In Chap. 1 Hillen and Lewis introduced the growth rate through a limiting process: if a population grows two person every 100 days, then it is “equivalent” to one person every 50 days, and half a person every 25 days. In fact, the growth rate is

$$r = \lim_{\Delta t \rightarrow 0} \frac{P(t + \Delta t) - P(t)}{\Delta t}.$$

Instantaneous rate (fluxion) is one of the most important concepts of Newton’s calculus! But does this make sense to quantify population growth? A half of a person, one tenth of a person? Clearly this theory cannot be true when the Δt is too small: *population change* cannot have non-integer numbers.

Second, has anyone ever seen such a regular population growth with exactly two person in the first 100 days, and another two in the next 100 days? I am sure some of you will say “that is just an average”.

Indeed, *discreteness* and *probability* are two fundamental issues in any population dynamics. Both have been ignored in the differential equation-based description of population dynamics. We shall start discussing population kinetics anew below. Most of the materials are taken from [1, 19, 20, 22, 23, 28, 31].

6.2.1 Brief Review of Elementary Probabilities

A ***random variable*** X taking a continuous real value has a probability density function (pdf) $f_X(x)$:

$$\int_{-\infty}^{\infty} f_X(x) dx = 1, \quad f_X(x) \geq 0. \quad (6.1)$$

The meaning of the $f_X(x)$ is this: for infinitesimal dx , the probability of observing $X \in (x, x + dx]$ is $f_X(x)dx$:

$$\Pr\{x < X \leq x + dx\} = f_X(x)dx. \quad (6.2)$$

Then, the cumulative probability distribution of X is defined as

$$F_X(x) = \Pr\{X \leq x\} = \int_{-\infty}^x f_X(z)dz, \text{ and } f_X(x) = \frac{dF_X(x)}{dx}. \quad (6.3)$$

The mean (or expected value) and variance of the random variable X then are

$$\langle X \rangle = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx, \quad (6.4)$$

$$\text{Var}[X] = \mathbb{E}[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx, \quad (6.5)$$

in which we have denoted $\mathbb{E}[X]$ by μ . Two most important examples of random variables taking real values are “exponential” and “normal”, also called Gaussian. The former has the standard form

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0, \quad \lambda > 0, \quad (6.6)$$

with mean and variance being λ^{-1} and λ^{-2} ; the latter has a standard form

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad (6.7)$$

with mean μ and variance σ^2 .

Gaussian normal distribution is widely discussed; including in popular press [11]. It is understood as a consequence of the *central limit theorem*. It is a statistical law emerging from a large collection of identical, independent parts. In the following sections, we shall show that for dynamical processes involving populations, there is a much less known, but equally if not more important statistical law: exponentially distributed time between “rare events”. In stochastic modeling of population dynamics, one’s primary focus is not the random number of individuals at a particular time; rather it is the random time of the next event that changes the number of individuals by one.

The best known discrete, integer-valued random variables are Bernoulli, binomial, Poisson, and geometric [30].

6.2.2 Radioactive Decay and Exponential Time

Let us revisit the simplest differential equation

$$\frac{dy}{dt} = -\lambda y, \quad (6.8)$$

where $\lambda > 0$. This equation has been introduced as a mathematical model for the remaining fraction of a radioactive material at time t

$$\frac{y(t)}{y(0)} = e^{-\lambda t}. \quad (6.9)$$

If all the atomic nuclei are *identical and independent*, then

$$\Pr\{\text{a nucleus remaining radioactive at time } t\} = e^{-\lambda t}. \quad (6.10)$$

However, if T is the random time at which the event of radioactive decay occurs, then

$$\Pr\{\text{a nucleus remaining radioactive at time } t\} = \Pr\{T \geq t\}. \quad (6.11)$$

T is a non-negative real-valued random variable with cumulative probability distribution $F_T(t) = \Pr\{T \leq t\} = 1 - e^{-\lambda t}$ and probability density function $f_T(t) = dF(t)/dt = \lambda e^{-\lambda t}$.

What types of problems, or more precisely “scenarios” and “mechanisms”, will give rise to this exponentially distributed waiting time? Why is it so universal? A good understanding of these questions will provide the reader a deeper understanding of the mathematical foundation of population dynamics, as emergent statistical laws, in terms of seemingly random behavior of a large population of individuals [15].

6.2.2.1 Rare Event

Let T be the random time at which a certain event occurs. If the occurrence of such an event is independent in time intervals $[t_1, t_2]$ and $[t_2, t_3]$, and if its occurrence is uniform in time (e.g., the system and its environment are stationary), then

$$\text{Prob. of no event occurring in } [0, t + \Delta t] = \quad (6.12)$$

$$\text{Prob. of no event occurring in } [0, t] \times \text{Prob. of no event occurring in } [t, t + \Delta t].$$

That is,

$$\Pr\{T > t + \Delta t\} = \Pr\{T > t\} \times \text{Prob. of no event occurring in } [t, t + \Delta t].$$

Now if the probability of one such event occurring in the time interval $[t, t + \Delta t]$ is proportional to Δt , and the probability of more than one events is $\propto o(\Delta t)$, then

$$\Pr\{T > t + \Delta t\} = \Pr\{T > t\} \times (1 - \lambda \Delta t + o(\Delta t)). \quad (6.13)$$

Then,

$$\frac{d}{dt}\Pr\{T > t\} = -\lambda\Pr\{T > t\}, \implies F_T(t) = e^{-\lambda t}. \quad (6.14)$$

Example: The waiting time for the first shopper coming in a store in the morning on a regular day.

6.2.2.2 Memoryless

One of the most important, in fact defining, properties of exponential distributed waiting time is

$$\frac{\Pr\{T \geq t + \tau\}}{\Pr\{T \geq t\}} = \frac{e^{-\lambda(t+\tau)}}{e^{-\lambda t}} = e^{-\lambda\tau}. \quad (6.15)$$

Example: You and your lazy brother doing experiments to observe the mean time of an exponentially distributed event. Even though your brother starts counting time a whole hour later than you, his resulting statistics will be exactly the same as yours!

More interestingly, the more individuals in a population, the faster the next event to occur. In mathematical terms: if all $T_k \sim \lambda_k e^{-\lambda_k t}$ and they are independently distributed, then $T^* = \min(T_1, T_2, \dots, T_n)$ also has an exponential distribution

$$\begin{aligned} \Pr\{T^* > t\} &= \Pr\{T_1 > t, \dots, T_n > t\} \\ &= \Pr\{T_1 > t\} \times \Pr\{T_2 > t\} \times \dots \times \Pr\{T_n > t\} = e^{-\mu t}, \end{aligned} \quad (6.16)$$

where $\mu = \lambda_1 + \lambda_2 + \dots + \lambda_n$. Thus, $f_{T^*}(t) = \mu e^{-\mu t}$.

6.2.2.3 Minimal Time of a Set of Non-Exponential i.i.d. Random Times

Now consider a set of random times $\{T_k\}$. They are *identical, independently distributed* (i.i.d.) random times with pdf $f_T(t)$ and cumulative probability distribution $F_T(t)$. Then $T^* = \min(T_1, T_2, \dots, T_n)$ has its distribution

$$\Pr\{T^* > t\} = \left(1 - F_T(t)\right)^n. \quad (6.17)$$

Now, introducing scaled $\hat{T}^* = nT^*$ and considering n to be very large, its distribution is

$$\Pr\{\hat{T}^* > t\} = \left(1 - F_T\left(\frac{t}{n}\right)\right)^n \simeq \left(-\frac{F_T'(0)}{n}t + O(n^{-2})\right)^n \rightarrow e^{-F_T'(0)t}. \quad (6.18)$$

Therefore, if $F'_T(0) = f_T(0)$ is finite, one obtains an exponentially distributed time.

We note the mathematical condition $f_T(0) > 0$: in an application, this implies that the time scale involved in the mechanism for the occurrence of an event is several orders of magnitude faster than the time scale in question.

6.2.3 Known Mechanisms That Yield an Exponential Distribution

In the previous section, we have derived the exponentially distributed waiting time based on some very elementary assumptions concerning (1) time homogeneous and (2) independent. Furthermore, in Sect. 6.2.2.3, we have shown that for non-exponential T , as long as $f_T(0) \neq 0$, the minimum of a large collection of i.i.d. T 's will be exponential. This is a strong argument for why one can use, on an appropriate time scale, the equations like (6.8) to model population dynamics.

6.2.3.1 Khinchin's Theorem

Let us consider a house that uses n light bulbs. One bought a large box of new light bulbs, and let us assume all the bulbs having identical, independently distributed life time X with pdf $f_X(x)$. For each light-bulb socket, one puts on a new bulb when the old one is burnt. The time sequence $0, T_1, T_2, \dots, T_k, \dots$ is called a *renewal process*, in which $T_k = \sum_{\ell=1}^k X^{(\ell)}$, where the $X^{(\ell)}$ with different ℓ are i.i.d. random variables drawn from the distribution $f_X(x)$. Now for the entire house, there are n identical, independent renewal processes. The time sequence of bulb changing form a *superposition* of the n renewal processes [3], as illustrated in Fig. 6.1.

For a single renewal process with renewal time distribution $f_X(x)$, the corresponding counting process, e.g., the number of renewals occurred before time t , N_t , has the distribution

$$\Pr \{N_t \geq k\} = \Pr \{T_k \leq t\} = F_{T_k}(t) = \int_0^t f_{T_k}(x)dx. \tag{6.19}$$

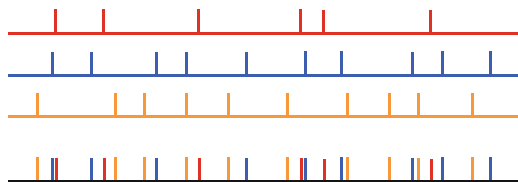


Fig. 6.1 If the red, orange, and blue point processes represent the renewal events of light bulbs for 3 different sockets, then the fourth row is the combined point process for all the bulb changes. It is the superposition of the three individual processes. With more and sockets, a statistical law emerges

Therefore,

$$\Pr \{N_t = k\} = F_{T_k}(t) - F_{T_{k+1}}(t). \quad (6.20)$$

Now if one randomly picks a time t , and let T_t^* be the waiting time for the next renewal, T_t^* is known as residual time in renewal theory. Its distribution is different from $f_X(x)$. In fact, one has

$$\begin{aligned} \Pr\{T_t^* \leq s\} &= \sum_{\ell=0}^{\infty} \Pr \{N_t = \ell\} \Pr \{T_{\ell+1} \leq t + s\} \\ &= \sum_{\ell=0}^{\infty} \left(F_{T_\ell}(t) - F_{T_{\ell+1}}(t) \right) F_{T_{\ell+1}}(t + s). \end{aligned} \quad (6.21)$$

Therefore, the probability density function for the stationary T_t^* is

$$f_{T_t^*}(s) = \frac{d}{ds} \Pr\{T_t^* \leq s\}. \quad (6.22)$$

$$f_{T_t^*}(0) = \sum_{\ell=1}^{\infty} \left(F_{T_\ell}(t) - F_{T_{\ell+1}}(t) \right) f_{T_\ell}(t) \neq 0. \quad (6.23)$$

Applying the result in Sect. 6.2.2.3, we then have the following theorem, which can be found in [3].

Theorem *If $T_k^{(1)}, T_k^{(2)}, \dots, T_k^{(n)}$ are n i.i.d. renewal processes with waiting time distribution $f_X(x)$, then the superposition of the n renewal processes has an exponential waiting time for the next event in the limit of $n \rightarrow \infty$, with rate parameter $n\mathbb{E}^{-1}[X]$.*

6.2.3.2 Kramers' Theory and Saddle-Crossing as a Rare Event

We have discussed the minimal time of a large collection of i.i.d. waiting times, and we have discussed superposition of renewal processes. We now turn to a third mechanism: the emergence of discrete chemical reactions from a description of atoms continuously moving in a molecule in an aqueous solution.

From a classical mechanics stand point, a molecule is a collection of atoms. For a protein with N number of atoms, a Newtonian mechanical description of its dynamics has $6N$ degrees of freedom, without even considering the atoms in the solvent, which is at least an order of magnitude more. This is what one observes from a molecular dynamics (MD) simulation. It is very complicated.

However, any such mechanical system has a potential energy function (its gradient is called a force field of MD simulations). Treating the solvent as a viscous

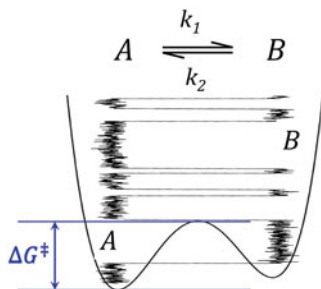


Fig. 6.2 The mathematical description of a chemical reaction of a single molecule. It is an emergent statistical law of a large number of discrete, stochastic reactions. $k_1 \propto e^{-\Delta G^\ddagger/k_B T}$, where the ΔG^\ddagger is called *activation energy*. Similarly, k_2 has its own activation barrier height. According to this description, the ratio k_1/k_2 becomes independent of the barrier

medium with frictional coefficient η , the dynamics of a protein is over damped and spends most of the time at the bottom of an “energy well”, as illustrated in Fig. 6.2. However, since the solvent is not truly continuous, but rather corpuscular, the collisions with the solvent molecules constitute a random force. Therefore, the dynamics can be described by a stochastic differential equation like

$$dY(t) = b(Y)dt + AdB(t), \quad (6.24)$$

in which $b(y) = -\eta^{-1}\nabla_y U(y)$, and $A = \sqrt{2\eta^{-1}k_B T}$.

With the presence of random forcing term $B(t)$, $Y(t)$ will once a while move against the deterministic force field and even cross the barrier (a saddle point in a high-dimensional space). But this is a rare event. This randomly perturbed nonlinear dynamical systems thus behaves, on a very long time scale, as $A \rightleftharpoons B$, with only two parameters k_1 and k_2 . The rate constants are related to the height of the barrier. H. A. Kramers first worked out the mathematical theory for this type of problems in 1940. The idea is not limited to chemical reactions; it is applicable to any nonlinear dynamics with random perturbations [7].

With one line of mathematics from Kramers, $k \propto e^{-\Delta G^\ddagger/k_B T}$ (Fig. 6.2), all the detailed atomic motions are deemed irrelevant—only two parameters, called forward and backward rate constants, are useful to a chemist. Furthermore, the theory shows that the transition from $A \rightarrow B$ spends most of the time in the waiting; the actual transition event is instantaneous! Indeed, one can mathematically prove in the limit of $\Delta G^\ddagger/k_B T \rightarrow \infty$, the waiting time distribution asymptotically approaches to exponential. From a molecular biological function perspective, the notion of discrete conformational states and the events of transitions among them are fundamental.

6.2.4 Population Growth

We have discussed $\frac{dy}{dt} = -\lambda y$ with positive λ : radioactive decay. And it does not seem that a similar discussion can be applied to $\frac{dx}{dt} = rx$ with a positive r , the other half of a population dynamics.

The answer turns out to be simple but profound: one should treat the birth as an event! The waiting time for the next birth is expected to be exponential. Furthermore, the rate is expected to be proportional to the number of individuals currently in the population (Exercise 1.2), say $X(t)$. Therefore, *on average the growth is 1 additional person in $(r\mathbb{E}[X])^{-1}$ time:*

$$\frac{d}{dt}\mathbb{E}[X(t)] = r\mathbb{E}[X(t)]. \quad (6.25)$$

Death is an event, birth is an event, state transition is an event. Most biological dynamics is about counting the populations, and about biological events that lead to changing populations. Stochasticity is in the timings of the various events. This is why J. D. Murray stated in [17] that continuous growth models for a species at time t have the universal conservation equation:

$$\frac{dY}{dt} = \text{births} - \text{deaths} + \text{migration}, \quad (6.26)$$

where $Y(t)$ is the population density.

6.2.5 Discrete State Continuous Time Markov (Q) Processes

Discrete state continuous time Markov processes are sometime called quasi Markovian, or Q-processes, a terminology first introduced in Arne Jensen's 1954 book *A Distribution Model, Applicable to Economics* and then by David Freedman in his 1971 book *Markov Chains*. In terms of the probability of state k at time t , $p_k(t)$, one has

$$p_k(t + dt) - p_k(t) = \left(\sum_{\ell=1}^N p_\ell(t) q_{\ell k} \right) dt, \quad (6.27)$$

where $q_{\ell k}dt$ is the transition probability from state ℓ to k within the infinitesimal time interval dt . Eq. (6.27) is called a *master equation*. Its fundamental solution is $\mathbf{P}(t) = e^{\mathbf{Q}t}$, where the \mathbf{Q} matrix has off-diagonal elements $q_{ij} \geq 0$ and

$$q_{ii} = - \sum_{j \neq i} q_{ij}. \quad (6.28)$$

Therefore, \mathbf{Q} has each and every row sums to zero. It is often referred to as infinitesimal transition rate matrix. It is easy to show that in this case, the sum

$$\sum_{k=1}^N p_k(t)$$

is independent of time t . The total probability is conserved over time. Note several important differences between Eqs. (6.26) and (6.27): The former is an equation for population *density* $Y(t)$ while the latter is an equation for the probability of population *size* $p_k(t) \equiv \Pr\{N(t) = k\}$; the right-hand side of former usually is a nonlinear function of Y while the latter is necessarily linear. The dimension of the latter ODE system, however, is much higher than the former.

6.2.5.1 Kolmogorov Forward and Backward Equations

In matrix form, Eq. (6.27) can be expressed as $\frac{d}{dt} p = p\mathbf{Q}$, where $p = (p_1, \dots, p_N)$ is a row vector. This equation is called *Kolmogorov forward equation*. Note strictly speaking the forward equation is not about the probability distribution (a vector), but about the transition probability matrix (fundamental solution) $\mathbf{P}(t)$ with initial value $\mathbf{P}(0) = \mathbf{I}$. More interestingly,

$$\frac{d}{dt} \mathbf{P} = \mathbf{P}\mathbf{Q} = \left(e^{\mathbf{Q}t}\right) \mathbf{Q} = \mathbf{Q}\mathbf{P}. \quad (6.29)$$

This is a different differential equation:

$$\frac{du_k}{dt} = \sum_{\ell=1}^N q_{k\ell} u_{\ell}, \quad (6.30)$$

which is called *Kolmogorov backward equation*. If $\{\pi_k\}$ is a stationary probability distribution, e.g., the solution to

$$\sum_{\ell=1}^N \pi_{\ell} q_{\ell k} = 0, \quad k = 1, 2, \dots, N,$$

then the solution to the backward equation, $u_k(t)$ has the important property of

$$\sum_{k=1}^N u_k(t) \pi_k$$

being independent of time t , e.g., it is a conserved quantity.

The solutions to the Kolmogorov forward and backward equations also have another important property. Let $p_k(t)$ and $q_k(t)$ be two solutions to a forward equation with different initial distributions $p_k(0)$ and $q_k(0)$. Then

$$\frac{d}{dt} \sum_{k=1}^N p_k(t) \ln \left(\frac{p_k(t)}{q_k(t)} \right) \leq 0. \quad (6.31)$$

One special case of this, which is widely known, is the choice of $q_k(t) = \pi_k$, if $\pi_k > 0 \forall k$.

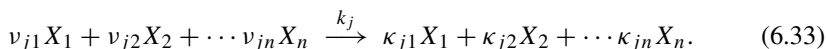
Similarly, two positive solutions to a Kolmogorov backward equation, $u_k(t)$ and $v_k(t)$ with different initial conditions $u_k(0)$ and $v_k(0)$, respectively, have

$$\frac{d}{dt} \sum_{k=1}^N \left(\pi_k u_k(t) \right) \ln \left(\frac{u_k(t)}{v_k(t)} \right) \leq 0. \quad (6.32)$$

One special case of this is when choosing $v_k(t) \equiv 1$. The quantity in Eq. (6.32) is called an H -function; the quantity in Eq. (6.31) is called relative entropy, or Kullback–Leibler divergence in information theory, or free energy in physical chemistry. These results have a deep implication for the second law of thermodynamics.

6.3 Theory of Chemical and Biochemical Reaction Systems

A general representation for complex chemical reaction systems is



$1 \leq j \leq m$. There are n species and m reactions. $(v_{ji} - \kappa_{ji})$ are called *stoichiometric coefficients*, they relate a species i to the reaction j . In a broader sense, a “reaction” is just a type of “events”.

6.3.1 Differential Equation and Nonlinear Dynamics

Because of the conservation of matter,

$$\frac{dx_i}{dt} = \sum_{j=1}^m (\kappa_{ji} - v_{ji}) \hat{\phi}_j(\mathbf{x}) \quad (6.34)$$

where x_i is the concentration of chemical species X_i , $1 \leq i \leq n$, and

$$\hat{\varphi}_j(\mathbf{x}) = k_j x_1^{v_{j1}} x_2^{v_{j2}} \cdots x_n^{v_{jn}} \quad (6.35)$$

is called the instantaneous flux of the j th reaction. $\mathbf{x} = (x_1, x_2, \dots, x_n)$. Eq. (6.34) is called rate equations, and Eq. (6.35) is called *the law of mass action* (LMA).

6.3.2 Delbrück-Gillespie Process (DGP)

Let us now consider probabilistically the discrete, individual events of the m possible reactions in Eq. (6.33), one at a time. The DGP assumes that the j th reaction occurs following an exponentially distributed waiting time, with rate parameter

$$\varphi_j(\mathbf{X}) = k_j V \prod_{\ell=1}^n \left(\frac{X_\ell!}{(X_\ell - v_{j\ell})! V^{v_{j\ell}}} \right), \quad (6.36)$$

when the molecular numbers of i th chemical species being X_i . Note $\varphi_j(\mathbf{X})$ has the dimension of $[\text{time}]^{-1}$. Clearly, the first reaction that occurs also follows an exponential time, with the rate being the sum of the rates of the m reactions:

$$\sum_{j=1}^m \varphi_j(\mathbf{X}). \quad (6.37)$$

Among the i.i.d. T_1, T_2, \dots, T_n , all exponentially distributed with respective rate parameters $\lambda_1, \lambda_2, \dots, \lambda_n$, the probability of the smallest one being T_k is

$$\begin{aligned} \Pr\{T^* = T_k\} &= \Pr\{T_k \leq \min(T_1, \dots, T_{k-1}, T_{k+1}, \dots, T_n)\} \\ &= \frac{\lambda_k}{\lambda_1 + \dots + \lambda_n}. \end{aligned} \quad (6.38)$$

More importantly,

$$\begin{aligned} &\Pr\{T^* = T_k, T^* \geq t\} \\ &= \Pr\{T_1 \geq T_k, \dots, T_{k-1} \geq T_k, T_k \geq t, T_{k+1} \geq T_k, T_n \geq T_k, \} \\ &= \int_t^\infty \lambda_k e^{-\lambda_k t} \prod_{\ell=1, \ell \neq k}^n \left(\int_{T_k}^\infty \lambda_\ell e^{-\lambda_\ell t_\ell} dt_\ell \right) \end{aligned}$$

$$\begin{aligned}
&= \int_t^\infty \lambda_k e^{-\lambda_k t_k} \prod_{\ell=1, \ell \neq k}^n \left(\int_{t_k}^\infty \lambda_\ell e^{-\lambda_\ell t_\ell} dt_\ell \right) \\
&= \left(\frac{\lambda_k}{\lambda_1 + \dots + \lambda_n} \right) e^{-(\lambda_1 + \dots + \lambda_n)t}.
\end{aligned} \tag{6.39}$$

This means the following important fact: the minimal time among $\{T_k\}$ gives two random variables: $T^* \equiv \min_k \{T_k\}$ and $k^* \equiv \arg \min_k \{T_k\}$; the minimal time T^* and the identity k^* are statistically independent.

6.3.3 Integral Representations with Random Time Change

6.3.3.1 Poisson Process

A standard Poisson process $Y(t)$ is an integer-valued, continuous-time Markov process with distribution

$$\Pr\{Y(t) = k\} = \frac{t^k}{k!} e^{-t}. \tag{6.40}$$

A Poisson process has both a *point process* representation, T_1, T_2, \dots, T_n , and a *counting process* representation $Y(t)$. The former is a positive real-valued, discrete-time Markov process with independent increments, and $T_{i+1} - T_i$ is exponentially distributed with rate 1.

6.3.3.2 Random Time Changed Poisson Representation

In terms of Poisson processes, the stochastic trajectory of a DGP representing the integer number of the molecule X_i at time t ,

$$X_i(t) = X_i(0) + \sum_{j=1}^m (\kappa_{ji} - \nu_{ji}) Y_j \left(\int_0^t \varphi_j(\mathbf{X}(t)) dt \right) \tag{6.41}$$

in which $\varphi_j(\mathbf{X})$ is given in (6.36). We have abused the notation X_i as both the symbol of a type of molecule, as in Eq. (6.33), and its number in the reaction system.

We see that in the limit of $\mathbf{X} \rightarrow \infty$ and $V \rightarrow \infty$,

$$\varphi_j(\mathbf{X}) \rightarrow k_j V \prod_{\ell=1}^n \left(\frac{X_\ell}{V} \right)^{\nu_{j\ell}} = k_j V \prod_{\ell=1}^n x_\ell^{\nu_{j\ell}} = V \hat{\varphi}_j(\mathbf{x}). \tag{6.42}$$

$\varphi_j(\mathbf{X})$ is also called the *propensity* of the j th reaction.

6.3.4 Birth-and-Death Process with State-Dependent Transition Rates

6.3.4.1 One-Dimensional System

Consider the stochastic population kinetics of a single species. Let $p_n(t)$ be the probability of having n individuals in the population at time t . Then $p_n(t)$ satisfies the master equation

$$\frac{dp_n(t)}{dt} = p_{n-1}u_{n-1} - p_n(u_n + w_n) + p_{n+1}w_{n+1}, \quad (6.43)$$

in which u_k and w_k are the birth rate and death rate of the population with exactly k individuals. The stationary distribution to Eq. (6.43) can be obtained:

$$\frac{p_n^{ss}}{p_{n-1}^{ss}} = \frac{u_{n-1}}{w_n}. \quad (6.44)$$

Therefore,

$$p_n^{ss} = p_0^{ss} \prod_{k=1}^n \left(\frac{u_{k-1}}{w_k} \right), \quad (6.45)$$

in which p_0^{ss} is to be determined by normalization.

Eq. (6.43) is the DGP corresponding to the nonlinear population dynamics of a single species with birth and death rates $\hat{u}(x)$ and $\hat{w}(x)$, with $x(t) \equiv \frac{X(t)}{V}$,

$$\frac{dx}{dt} = \hat{u}(x) - \hat{w}(x), \quad (6.46)$$

where,

$$\hat{u}(x) = \lim_{V \rightarrow \infty} \frac{u_x V}{V}, \quad \hat{w}(x) = \lim_{V \rightarrow \infty} \frac{w_x V}{V}. \quad (6.47)$$

It is easy to verify that the peaks and troughs of stationary probability distribution p_n^{ss} correspond nicely with the stable and unstable fixed points of Eq. (6.47). For the rest of this chapter, this correspondence should be kept in mind.

6.4 Using Mathematics to Articulate a Fundamental Idea in Biology

I want to use the following example to illustrate how to use mathematics, not only as a tool for computation and for modeling, but also for representing fundamental ideas.

Consider a population with many subpopulations $\mathbf{x} = (x_1, x_2, \dots, x_n)$, all $x_i \geq 0$. In the absence of migration, if we denote per capita growth rate $r_i = b_i - d_i$, then

$$\frac{dx_i}{dt} = x_i r_i. \quad (6.48)$$

For simplicity, we shall assume that both per capita birth rate b_i and death rate d_i are constants. Then the per capita growth rate for the entire population, which is also the mean per capita growth rate,

$$\bar{r} = \frac{\sum_{i=1}^n \frac{dx_i}{dt}}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n x_i r_i}{\sum_{i=1}^n x_i}, \quad x_i \geq 0. \quad (6.49)$$

Then,

$$\frac{d\bar{r}(\mathbf{x})}{dt} = \left[\frac{\sum_{i=1}^n x_i r_i^2}{\sum_{i=1}^n x_i} - \left(\frac{\sum_{i=1}^n x_i r_i}{\sum_{i=1}^n x_i} \right)^2 \right]. \quad (6.50)$$

We note that the term inside $[\dots]$ on the right-hand side is never negative:

$$\frac{\sum_{i=1}^n x_i r_i^2}{\sum_{i=1}^n x_i} - \left(\frac{\sum_{i=1}^n x_i r_i}{\sum_{i=1}^n x_i} \right)^2 = \frac{\sum_{i=1}^n x_i (r_i - \bar{r})^2}{\sum_{i=1}^n x_i} \geq 0. \quad (6.51)$$

In fact, it is exactly the variance of r_i among the different subpopulations. Therefore, it is always positive if there are variations among r_i . This mathematical result is a part of the ideas of both Adam Smith, on economics, and Charles Darwin, on the natural selection. In fact, the term $[\dots]$ in Eq. (6.50) has been identified by R. A. Fisher, the British statistician and evolutionary biologist, as the “growth of fitness due to natural selection” [6]. Here is a quote from Smith’s *magnum opus* “An Inquiry into the Nature and Causes of the Wealth of Nations” (1776):

As every individual, therefore, endeavours as much as he can both to employ his capital in the support of domestic industry, and so to direct that industry that its produce may be of the greatest value; every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it. By preferring the support of domestic to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse for the society that it was no part of it. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it. I have never known much good done by those who affected to trade for the public good. It is an affectation, indeed, not very common among merchants, and very few words need be employed in dissuading them from it.

6.5 Ecological Dynamics and Nonlinear Chemical Reactions: Two Examples

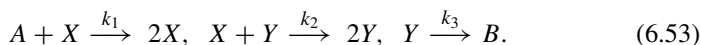
6.5.1 Predator and Prey System

Let $z(t)$ be the population density of a predator at time t and $x(t)$ be the population density of a prey at the same time. Then the simplest predator-prey dynamics is [17]

$$\begin{cases} \frac{dx}{dt} = \alpha x - \beta xz, \\ \frac{dz}{dt} = -\gamma z + \delta xz. \end{cases} \quad (6.52)$$

The detailed analysis of the nonlinear dynamics can be found in many textbooks on mathematical biology or differential equations [17].

Let us now consider the following chemical reaction system:



Then according to the LMA, the concentrations of X and Y , with fixed concentrations of A and B being a and b :

$$\frac{dx}{dt} = k_1ax - k_2xy, \quad \frac{dy}{dt} = -k_3y + k_2xy. \quad (6.54)$$

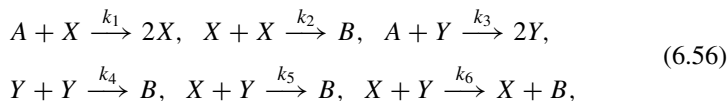
Therefore, we see that dynamics of an ecological predator-prey system is remarkable similar to that of a chemical reaction system with autocatalysis [16]: the first reaction in (6.53) requires an existing X serving as a catalyst for the reaction $A \rightarrow X$. A species that appears on the both sides of a chemical reaction is called a *catalyst*.

6.5.2 A Competition Model

Let us now consider another widely studied ecological dynamics with competition [17]:

$$\begin{cases} \frac{dN_1}{dt} = r_1N_1 - a_1N_1^2 - b_{21}N_1N_2, \\ \frac{dN_2}{dt} = r_2N_2 - a_2N_2^2 - b_{12}N_2N_1. \end{cases} \quad (6.55)$$

Can one “design” a system of chemical reactions that yields an identical system of differential equation? Without loss of generality, let us assume that $b_{12} > b_{21}$.



which, according to the LMA,

$$\begin{cases} \frac{dx}{dt} = (k_1 a)x - k_2 x^2 - k_5 xy, \\ \frac{dy}{dt} = (k_3 a)y - k_4 y^2 - (k_5 + k_6)xy. \end{cases} \quad (6.57)$$

If we identify x, y with N_1, N_2 , and

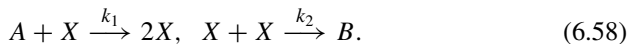
$$(k_1 a) \leftrightarrow r_1, \quad k_2 \leftrightarrow a_1, \quad k_5 \leftrightarrow b_{21}, \quad (k_3 a) \leftrightarrow r_2, \quad k_4 \leftrightarrow a_2, \quad (k_5 + k_6) \leftrightarrow b_{12},$$

then (6.57) is the same as (6.55). Note that the last reaction, $X + Y \rightarrow X + B$, is introduced to represent $b_{12} > b_{21}$.

A close inspection of the system of chemical reactions in (6.56) indicates that the overall reaction is $2A \rightarrow B$. Since each and every reaction is irreversible, there can be no chemical equilibrium. Rather, the system eventually reaches a *nonequilibrium steady state* in which there is a continuous, overall chemical flux converting $2A$ to B .

6.5.3 Logistic Model and Keizer's Paradox

We now turn to studying some issues more in-depth. Let us now consider a much simpler chemical reaction system,



It is easy to see that the ODE according to the LMA,

$$\frac{dx}{dt} = r \left(1 - \frac{x}{K}\right) x, \quad r = k_1 a, \quad K = \frac{r}{k_2}, \quad (6.59)$$

is the celebrated *logistic equation* in population dynamics. In the ecological context, r is known as the per capita growth rate in the absence of intra-species competition; and K is known as *carrying capacity*.

The DGP stochastic model has a *chemical master equation* (CME) for the probability of n X molecules in a reaction volume of V :

$$\frac{dp_n(t)}{dt} = u_{n-1}p_{n-1} - (u_n + w_n)p_n + w_{n+1}p_{n+1}, \quad (6.60a)$$

in which the state-dependent birth and death rates are

$$u_n = rn, \quad w_n = \frac{k_2n(n-1)}{V}. \quad (6.60b)$$

Then, according to Eq. (6.45),

$$p_0^{eq} = 1 \text{ and } p_n^{eq} = 0, \quad n \geq 1, \quad (6.61)$$

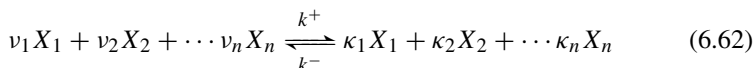
since $u_0 = 0$! In other words, according to this theory, the stationary probability distribution is “population extinction with probability 1”. But the ODE in (6.59) says that the stable steady state is $x = K$, with $x = 0$ being a unstable steady state which is not “relevant”.

This seeming disagreement between the deterministic ODE in (6.59) and stochastic dynamics described by (6.60) is known as *Keizer's paradox*. We refer the readers to [33] for the resolution of the paradox.

6.6 Chemical Thermodynamics and Entropy Production

6.6.1 Classical Chemical Thermodynamics

A single *reversible* chemical reaction



is said to be in a chemical equilibrium when

$$\frac{\hat{\phi}_k^+(\mathbf{x}^{eq})}{\hat{\phi}_k^-(\mathbf{x}^{eq})} = 1 \Leftrightarrow \left(\frac{x_1^{v_1} x_2^{v_2} \cdots x_n^{v_n}}{x_1^{\kappa_1} x_2^{\kappa_2} \cdots x_n^{\kappa_n}} \right)^{eq} = \frac{k^-}{k^+}. \quad (6.63)$$

(k^-/k^+) is known as the *equilibrium constant* of the reaction. The ratio on the lhs is a constant independent of the total amount participating species.

Chemical thermodynamics introduces the notions of chemical energy and chemical potential: for ideal solutions chemical species i has a chemical potential

$$\mu_i = \mu_i^0 + k_B T \ln x_i. \quad (6.64)$$

in which μ_i^o is determined by the atomic structure of a molecule, e.g., internal energy. k_B is Boltzmann's constant, and T is temperature in Kelvin. Then the Gibbs free energy of the lhs of (6.62) is the sum of the chemical potential

$$G = \sum_{i=1}^n v_i \left(\mu_i^o + k_B T \ln x_i \right). \quad (6.65)$$

When the reaction reaches its equilibrium, one has the total chemical potentials being equal on both sides:

$$\sum_{i=1}^n (v_i - \kappa_i) \left(\mu_i^o + k_B T \ln x_i^{eq} \right) = 0. \quad (6.66)$$

This implies

$$\prod_{i=1}^n (x_i^{eq})^{v_i - \kappa_i} = e^{-\frac{(v_i - \kappa_i) \mu_i^o}{k_B T}} = \frac{k^-}{k^+}, \quad (6.67)$$

or

$$\Delta G^o = \left(\sum_{i=1}^n \kappa_i \mu_i^o \right) - \left(\sum_{i=1}^n v_i \mu_i^o \right) = k_B T \ln \left(\frac{k^-}{k^+} \right). \quad (6.68)$$

This is a very well-known formula that can be found in every college chemistry textbook.

6.6.2 Mass-Action Kinetics

Following Eqs. (6.34) and (6.35), we have

$$\begin{aligned} \frac{dx_j}{dt} &= \sum_{j=1}^m (\kappa_{ji} - v_{ji}) \left(\hat{\phi}_j^+ - \hat{\phi}_j^- \right) \\ &= \sum_{j=1}^m (\kappa_{ji} - v_{ji}) \hat{\phi}_j^- \left\{ \exp \left[\sum_{\ell=1}^n (\kappa_{j\ell} - v_{j\ell}) \ln \left(\frac{x_\ell}{x_\ell^{eq}} \right) \right] - 1 \right\} \\ &= \sum_{j=1}^m (\kappa_{ji} - v_{ji}) \hat{\phi}_j^+ \left\{ 1 - \exp \left[\sum_{\ell=1}^n (v_{j\ell} - \kappa_{j\ell}) \ln \left(\frac{x_\ell}{x_\ell^{eq}} \right) \right] \right\}. \quad (6.69) \end{aligned}$$

Equation (6.69) shows that when $x_\ell = x_\ell^{eq}$, the term $[\dots] = 0$ and the term $\{\dots\} = 0$ as well, for every j . Therefore, the kinetic equation in (6.69) is consistent with the chemical equilibrium according to thermodynamics, e.g., Eqs. (6.66) and (6.67). Interestingly, recent work has shown that both macroscopic kinetics as in (6.69) and equilibrium thermodynamics in Sect. 6.6.1 are consequences of a stochastic kinetic description of a reaction system [10].

6.6.3 Stochastic Chemical Kinetics

We now apply the above formalism to a nonlinear chemical reaction in a small volume V with small number of molecules, n_A , n_B , and n_C numbers of A , B , and C :



We note that the $n_A + n_C$ and $n_B + n_C$ do not change in the reaction. Hence we can denote $n_A + n_C = n_A^o$ and $n_B + n_C = n_B^o$ as the total amount of A and B , including those in C , at the initial time. Now if we use n_C as the non-negative integer-valued random variable to describe the stochastic chemical kinetics, this simple nonlinear chemical reaction, according to DGP, is a one-dimensional birth-and-death process, with state-dependent birth and death rates $u_n = k^+ n_A n_B$ and $w_n = k^- n_C$. Then, according to Eq. (6.45), we have an equilibrium distribution $p^{eq}(m) = \Pr\{n_C^{eq} = m\}$:

$$\frac{p^{eq}(m+1)}{p^{eq}(m)} = \frac{k^+(n_A^o - m)(n_B^o - m)}{k^-(m+1)V}, \quad (6.71)$$

in which $n_A^o = n_A(0) + n_C(0)$ and $n_B^o = n_B(0) + n_C(0)$. Therefore,

$$p^{eq}(m) = \frac{\Xi^{-1} n_A^o! n_B^o!}{m!(n_A^o - m)!(n_B^o - m)!} \left(\frac{k^+}{k^- V} \right)^m, \quad (6.72)$$

where Ξ is a normalization factor

$$\Xi(\lambda) = \sum_{m=0}^{\min(n_A^o, n_B^o)} \frac{n_A^o! n_B^o! \lambda^m}{m!(n_A^o - m)!(n_B^o - m)!}, \quad \lambda = \left(\frac{k^+}{k^- V} \right). \quad (6.73)$$

More importantly, by noting $n_A + n_B + n_C = n_A^o + n_B^o - n_C$,

$$\begin{aligned} & -\ln p^{eq}(n_C) \\ &= -\ln \left[\frac{\lambda^{n_C}}{n_C!(n_A^o - n_C)!(n_B^o - n_C)!} \right] + \text{const.} \end{aligned}$$

$$\begin{aligned}
&= n_A \ln \left(\frac{n_A}{V} \right) - n_A + n_B \ln \left(\frac{n_B}{V} \right) - n_B + n_C \ln \left(\frac{n_C}{V} \right) - n_C - n_C \ln \left(\frac{k^+}{k^-} \right) \\
&= n_A \ln x_A + n_B \ln x_B + n_C \ln x_C + n_C \left(\frac{\mu_C^o - \mu_A^o - \mu_B^o}{k_B T} \right) - (n_A + n_B + n_C) \\
&= \sum_{\sigma=A,B,C} n_\sigma \left(\frac{\mu_\sigma^o}{k_B T} + \ln x_\sigma - 1 \right). \tag{6.74}
\end{aligned}$$

This agrees with Eq. (6.65).

In classical chemical kinetics, for a given $\mathbf{x}(t)$, the Ideal function of the chemical reaction system is

$$G^{eq}[\mathbf{x}(t)] = \sum_{\sigma=1}^n x_\sigma \left(\mu_\sigma^o + k_B T \ln x_\sigma - k_B T \right). \tag{6.75}$$

Then, following Eq. (6.34), assuming each and every reaction is reversible with rate constants k_j^+ and k_j^- ,

$$\begin{aligned}
\frac{d}{dt} G^{eq}[\mathbf{x}(t)] &= \sum_{i=1}^n \frac{dx_i}{dt} \left(\mu_i^o + k_B T \ln x_i \right) \\
&= k_B T \sum_{i=1}^n \sum_{j=1}^m \ln \left(\frac{x_i}{x_i^{eq}} \right) (\kappa_{ji} - \nu_{ji}) \left(k_j^+ \prod_{\ell=1}^n x_\ell^{\nu_{j\ell}} - k_j^- \prod_{\ell=1}^n x_\ell^{\kappa_{j\ell}} \right) \\
&= -k_B T \sum_{j=1}^m \left\{ \sum_{i=1}^n \ln \left(\frac{x_i}{x_i^{eq}} \right)^{\nu_{ji} - \kappa_{ji}} \right\} (\hat{\varphi}_j^+ - \hat{\varphi}_j^-) \\
&= -k_B T \sum_{j=1}^m (\hat{\varphi}_j^+ - \hat{\varphi}_j^-) \ln \left(\frac{\hat{\varphi}_j^+}{\hat{\varphi}_j^-} \right) \tag{6.76}
\end{aligned}$$

$$\leq 0. \tag{6.77}$$

The minus-log stationary probability distribution is a Lyapunov function for the dynamics. The rhs of Eq. (6.76) is known as *entropy production rate*.

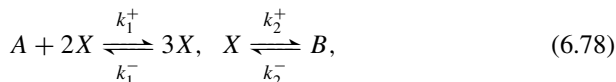
6.6.4 Nonequilibrium Steady-State and Driven Chemical Systems

If a chemical reaction system reaches its chemical equilibrium, then each and every reaction in the system is in *detailed balance* with zero net flux. This puts

a very strong condition on the dynamics. When a chemical reaction system has a sustained source and sink with different chemical potentials, it cannot reach a chemical equilibrium. Rather, it reaches a *nonequilibrium steady state* (NESS).

Let us consider the following two examples, the Schlögl model for bistability [34] and Schnakenberg model for nonlinear oscillation [17, 25, 35].

6.6.4.1 Schlögl Model



in which the concentrations (or chemical potentials) of A and B are sustained by an external agent. This reaction is known as *Schlögl model*, whose dynamics can be described by the differential equation

$$\frac{dx}{dt} = k_1^+ ax^2 - k_1^- x^3 - k_2^+ x + k_2^- b = f(x), \quad (6.79)$$

which is a third-order polynomial. It can exhibit bistability and saddle-node bifurcation phenomenon. All of them only occur under driven condition, when $\mu_A \neq \mu_B$. Note in the chemical equilibrium: $\mu_A = \mu_A^0 + k_B T \ln a = \mu_B^0 + k_B T \ln b$, and

$$\left(\frac{b}{a}\right)^{eq} = \frac{k_1^+ k_2^+}{k_1^- k_2^-}. \quad (6.80)$$

Differential equation (6.79), with its parameters $ak_1^+ k_2^+ = bk_1^- k_2^-$, has the right-hand-side

$$\begin{aligned} f(x) &= k_1^+ ax^2 - k_1^- x^3 - k_2^+ x + k_2^- b \\ &= k_1^+ ax^2 - k_1^- x^3 - k_2^+ x + \frac{ak_1^+ k_2^+}{k_1^-} \\ &= \left(x^2 + \frac{k_2^+}{k_1^-}\right) \left(ak_1^+ - k_1^- x\right). \end{aligned} \quad (6.81)$$

Therefore, the $f(x)$ has a unique fixed point at $x = \frac{ak_1^+}{k_1^-}$, the chemical equilibrium. In general, system (6.78) can exhibit chemical bistability; but this is only possible when A and B have a sufficiently large chemical potential difference, e.g., a *chemostat*.

More interestingly, when a and b satisfying (6.80), the DGP of the number of X , $n_X(t)$, is again a one-dimensional birth-and-death process, with

$$\begin{aligned}
 u_n &= \frac{k_1^+ a n(n-1)}{V} + k_2^- b V = \frac{k_1^+ a}{V} \left(n(n-1) + \frac{k_2^+ V^2}{k_1^-} \right), \quad (6.82) \\
 w_{n+1} &= \frac{k_1^-(n+1)n(n-1)}{V^2} + k_2^+(n+1) \\
 &= \frac{k_1^-(n+1)}{V^2} \left(n(n-1) + \frac{k_2^+ V^2}{k_1^-} \right).
 \end{aligned}$$

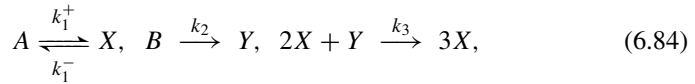
Therefore, the stationary distribution, according to Eq. (6.45),

$$p_n^{eq} = C \prod_{\ell=0}^{n-1} \frac{k_1^+ a / V}{k_1^- (\ell+1) / V^2} = \frac{\lambda^n}{n!} e^{-\lambda}, \quad \lambda = \left(\frac{k_1^+ a V}{k_1^-} \right). \quad (6.83)$$

This is a Poisson distribution, with expected value being $\mathbb{E}[n_X^{eq}] = \lambda$. Therefore, the expected concentration is $(k_1^+ a / k_1^-)$.

6.6.4.2 Schnakenberg Model

Similarly,



is known as *Schnakenberg model*, whose dynamics follow

$$\begin{cases} \frac{dx}{dt} = k_1^+ a - k_1^- x - k_3 x^2 y = f(x, y), \\ \frac{dz}{dt} = k_2 b - k_3 x^2 y = g(x, y). \end{cases} \quad (6.85)$$

This system can exhibit limit cycle oscillation and Hopf bifurcation. In terms of the DGP, it exhibits a rotational diffusion. We refer the readers to [25, 35] for an in-depth analysis of the model.

6.7 The Law of Large Numbers—Kurtz's Theorem

6.7.1 Diffusion Approximation and Kramers–Moyal Expansion

Starting with the master equation in (6.43), let us consider a partial differential equation (PDE) for a continuous density function $f(x, t)dx = p_{Vx}(t)$ where $x = \frac{n}{V}$, $dx = \frac{1}{V}$, then

$$\begin{aligned}
 \frac{\partial f(x, t)}{\partial t} &= V \frac{dp_{Vx}(t)}{dt} \\
 &= \frac{1}{dx} \left(f(x - dx, t) \hat{u}(x - dx) - f(x, t) (\hat{u}(x) + \hat{w}(x)) \right. \\
 &\quad \left. + f(x + dx, t) \hat{w}(x + dx) \right) \\
 &= \frac{\partial}{\partial x} \left(f(x + dx/2, t) \hat{w}(x + dx/2) - f(x - dx/2, t) \hat{u}(x - dx/2) \right) \\
 &\approx \frac{\partial}{\partial x} \left\{ \frac{\partial}{\partial x} \left(\frac{\hat{w}(x) + \hat{u}(x)}{2V} \right) f(x, t) - (\hat{u}(x) - \hat{w}(x)) f(x, t) \right\} + \dots
 \end{aligned} \tag{6.86}$$

in which

$$V^{-1}u_{Vx} = \hat{u}(x), \quad V^{-1}w_{Vx} = \hat{w}(x), \tag{6.87}$$

as $V \rightarrow \infty$.

6.7.2 Nonlinear Differential Equation, Law of Mass Action

Therefore, in the limit of $V \rightarrow \infty$,

$$\frac{\partial f(x, t)}{\partial t} = - \frac{\partial}{\partial x} (\hat{u}(x) - \hat{w}(x)) f(x, t), \tag{6.88}$$

which corresponds to the ordinary differential equation

$$\frac{dx}{dt} = \hat{u}(x) - \hat{w}(x), \tag{6.89}$$

that defines the characteristic lines of (6.88).

6.7.3 Central Limit Theorem, a Time-Inhomogeneous Gaussian Process

Now consider the process

$$Y(t) = \frac{X(t) - Vx(t)}{\sqrt{V}}, \quad (6.90)$$

which characterizes the “deviation” of $\frac{X(t)}{V}$ from $x(t)$. In the limit of $V \rightarrow \infty$, its pdf $f_Y(y, t)$ satisfies a linear PDE with time-varying coefficients

$$\begin{aligned} \frac{\partial f_Y(y, t)}{\partial t} = & \frac{\partial}{\partial y} \left\{ \frac{\partial}{\partial y} \left(\frac{\hat{w}(x(t)) + \hat{u}(x(t))}{2} \right) f_Y(y, t) \right. \\ & \left. - \left(\hat{u}'(x(t)) - \hat{w}'(x(t)) \right) y f_Y(y, t) \right\}. \end{aligned} \quad (6.91)$$

Therefore, $Y(t)$ is a continuous time, real-valued, time-inhomogeneous Markov process. Note the PDE (6.91) is very different from PDE (6.86). They are known in physics literature as the Kramers–Moyal expansion and van Kampen’s Ω -expansion, respectively [32]. The former is not related to the central limit theorem.

6.7.4 Diffusion’s Dilemma

Truncating the Eq. (6.86) after the second order, it has a stationary distribution

$$-\ln \hat{f}_Y^{st}(y) = 2V \int \left(\frac{\hat{w}(x) - \hat{u}(x)}{\hat{u}(x) + \hat{w}(x)} \right) dx. \quad (6.92)$$

On the other hand, the stationary solution given in (6.45),

$$p_n^{ss} = p_0^{ss} \prod_{k=1}^n \left(\frac{u_{k-1}}{w_k} \right),$$

in the limit of $V \rightarrow \infty$ with $V^{-1}u_{Vx} = \hat{u}(x)$, $V^{-1}w_{Vx} = \hat{w}(x)$, and $V^{-1} = dx$, yields

$$-\ln p_{Vx}^{ss} = -\sum_{k=1}^n \ln \left(\frac{u_{k-1}}{w_k} \right) + C \leftrightarrow -\ln f^{ss}(x) = V \int \ln \left(\frac{\hat{w}(x)}{\hat{u}(x)} \right) dx. \quad (6.93)$$

Is it possible Eqs. (6.92) and (6.93) are actually the same? We notice that both have identical local extrema:

$$\frac{d}{dx} \left(-\ln f_Y^{st}(x) \right) = 2V \left(\frac{\hat{w}(x) - \hat{u}(x)}{\hat{w}(x) + \hat{u}(x)} \right) = 0 \implies \hat{w}(x) = \hat{u}(x). \quad (6.94)$$

In fact, the curvature at a local extremum is identical:

$$\begin{aligned} \left[\frac{d^2}{dx^2} \left(-\ln f_Y^{st}(x) \right) \right]_{\hat{u}=\hat{w}} &= 2V \left(\frac{\hat{w}'(x) - \hat{u}'(x)}{\hat{w}(x) + \hat{u}(x)} \right) = V \left(\frac{\hat{w}'(x) - \hat{u}'(x)}{\hat{u}(x)} \right) \\ &= \left[\frac{d^2}{dx^2} \left(-\ln f^{ss}_S(x) \right) \right]_{\hat{u}=\hat{w}}. \end{aligned} \quad (6.95)$$

However, it can be shown, via an example, that the global minimum can be different [20, 37]! This implies that Kramers–Moyal expansion is not a valid approximation for stochastic kinetics with multiple stability. Continuous time, real-valued Markov processes are also called *diffusion processes*. The above result illustrates that there is no globally valid diffusion approximation for stochastic population kinetics in general.

6.8 The Logic of the Mechanical Theory of Heat and Nonequilibrium Thermodynamics

In order to present some rather recent results in Sect. 6.9 and put those results into a proper context, let us first revisit the celebrated work of L. Boltzmann on the *mechanical theory of heat* [8], and the generally accepted macroscopic nonequilibrium thermodynamics presented in the classic treatise of de Groot and Mazur [5]. The readers will recognize the logical threads of both theories in Sect. 6.9, as well as the finding of a missing link between the above two theories.

Boltzmann’s theory is based on the general Hamiltonian dynamics and starts with a definition of an entropy function $S = -k_B \ln \Omega(E)$. Section 6.9 will be based on the general Markov dynamics and starts with a definition of an entropy function according to Shannon [29]. Note that Boltzmann’s entropy is a static quantity, the entropy in Sect. 6.9, Eq. (6.110) below, is a function of time.

De Groot and Mazur’s theory is based on continuity equations for mass and energy, relating time changes of the density of these quantities to transport processes in three-dimensional space, and identifies entropy productions as “transport flux \times driving force”, *à la* Onsager [18]. Section 6.9 is based on a continuity equation for the probability in the state space, relating time change of probability to its transport, and also identifies the entropy production as “probability flux \times chemical potential difference”.

There is a missing link between Boltzmann's theory and the nonequilibrium thermodynamics. In addition to the continuity equations, the de Groot-Mazur approach also requires the *entropy balance equation* [5],

$$\frac{dS}{dt} = e_p + J_S, \quad (6.96)$$

as one of its fundamental premises, where e_p is the entropy production rate and J_S is the rate of entropy supplied to a system by its surroundings. The second law of thermodynamics, e.g., Clausius inequality, dictates that $e_p \geq 0$. Unfortunately, Boltzmann's mechanical theory of heat is not able to derive an equation like (6.96) from a Hamiltonian dynamics without resorting to additional assumptions based on a *stosszahlansatz*.¹ As one will see from Sect. 6.9, however, Markov dynamics is able to provide nicely an equation like (6.96). A stochastic dynamic approach to nonequilibrium thermodynamics is able to fill this logic gap, as was first demonstrated by Bergmann and Lebowitz in 1955 [2].

6.8.1 Boltzmann's Mechanical Theory of Heat

The entire world, as long as one is interested in phenomena that are at not too small a scale (e.g., quantum) and not too close to the speed of light (e.g., relativity), follows the Newtonian mechanics which can be represented mathematically in terms of a Hamiltonian system

$$\frac{dx}{dt} = \frac{\partial H(x, y)}{\partial y}, \quad \frac{dy}{dt} = -\frac{\partial H(x, y)}{\partial x}. \quad (6.97)$$

One of the most important result concerning the Eq. (6.97) is the dynamics invariance of $H(x(t), y(t))$:

$$\frac{d}{dt} H(x(t), y(t)) = \frac{\partial H}{\partial x} \left(\frac{dx}{dt} \right) + \frac{\partial H}{\partial y} \left(\frac{dy}{dt} \right) = 0. \quad (6.98)$$

¹In the phase space, the Hamiltonian system has a Liouville equation

$$\frac{\partial u(x, y, t)}{\partial t} = -\left(\frac{\partial H}{\partial y} \right) \frac{\partial u}{\partial x} + \left(\frac{\partial H}{\partial x} \right) \frac{\partial u}{\partial y}.$$

It is easy to show that

$$\frac{d}{dt} \iint u(x, y, t) \ln u(x, y, t) dx dy = 0.$$

Therefore, the information-entropy like quantity is time invariant under a deterministic diffeomorphism [36].

Now, let us assume that the Hamiltonian function contains also several parameters, $H(x, y, V, N)$ where V is the box size of a mechanical system and N is the number of particles in the box, then the next question which an applied mathematician might ask, but interestingly which has not been extensively studied, is this: “What is the long-time behavior of the system as a function of V , N , and other parameters?”

A Hamiltonian system, however, is fundamentally different from the earlier systems we have studied, which have attractive fixed point(s). In fact, it is clear that the long-time behavior is a function of the initial condition $H(x(0), y(0)) = E$. Helmholtz and Boltzmann (1884) realized that a “thermodynamic equilibrium state” of a mechanical system is *not a single point in the phase space, but rather, it is an entire invariant manifold* defined by the level set $H(x, y, V, N) = E$. It was Boltzmann’s ingenuity to realize that one can define

$$\begin{aligned} S(E, V, N) &= k_B \ln \{ \text{phase volume contained by the surface } H(x, y) = E \} \\ &= k_B \ln \int_{H(x,y) \leq E} dx dy. \end{aligned} \quad (6.99)$$

Since $S(E)$ is monotonic, one has an implicit function $E = E(S, V)$. Then

$$\begin{aligned} dE &= \left(\frac{\partial E}{\partial S} \right)_{V,N} dS + \left(\frac{\partial E}{\partial V} \right)_{S,N} dV + \left(\frac{\partial E}{\partial N} \right)_{S,V} dN \\ &= T dS - p dV + \mu dN. \end{aligned} \quad (6.100)$$

What is the significance of Eq. (6.100)? First, it is completely based on the fact that a Hamiltonian system has a *conservation of mechanical energy* H . Furthermore, however, this conservation of energy is valid not only for a single Hamiltonian system on a single invariant torus, but also the Hamiltonian system with multiple level sets, and even among an entire class of Hamiltonian systems with varying V and N , and other parameters. It becomes a universally valid equation, known as *the First Law of Thermodynamics*. Note, according to this theory, thermodynamic quantities like T , p , μ are mathematically defined via Eq. (6.100). They are emergent phenomena.

T and p have mechanical interpretations, though not perfect, as mean kinetic energy and mean momentum transfer to a wall. μ , however, has no interpretation in terms of classical motion; rather, it has an interpretation in terms of Brownian motion:

$$\frac{\partial \rho(x, t)}{\partial t} = D \frac{\partial^2 \rho(x, t)}{\partial x^2} = -\frac{1}{\eta} \frac{\partial (\hat{F} \rho)}{\partial x}, \quad (6.101)$$

where

$$\hat{F} = -\frac{\partial \mu}{\partial x}, \quad \text{and} \quad \mu = D \eta \ln \rho(x, t) = k_B T \ln \rho(x, t). \quad (6.102)$$

\hat{F} is known as *entropic force* in chemistry, and μ is known as chemical potential.

6.8.2 Classical Macroscopic Nonequilibrium Thermodynamics

Equation (6.100) is valid only when the entire torus $H(x, y) = E$ is visited in the long time limit; this is known as *ergodicity*. In other words, with time t in mind, the equation is valid only when the dS and dV are very slowly changing. What happens if the changes are not slow? Then, the *Second Law of Thermodynamics* states that

$$T dS \geq dQ = dE - dW, \quad (6.103)$$

in which dQ is the amount of heat that flows into the system, and dW is the amount of work done to the system. Both are path dependent, as indicated by the d . Eq. (6.103) is known as the Clausius inequality. The notion of *entropy production* is introduced to account for the inequality:

$$\frac{dS}{dt} = e_p - \frac{h_d}{T}, \quad e_p \geq 0, \quad (6.104)$$

in which e_p is called entropy production, which is never negative. $h_d = -dQ/dt$ is called heat dissipation. In general, neither e_p nor h_d is a time derivative. Eq. (6.104) is known as an *entropy balance equation*.

6.8.2.1 Local Equilibrium Assumption and Classical Derivation of Entropy Production

If one assumes that Eq. (6.100) is valid locally in space and time, then one has

$$\frac{\partial s(x, t)}{\partial t} = \frac{1}{T} \frac{\partial u(x, t)}{\partial t} - \sum_{i=1}^n \mu_i \frac{\partial c_i(x, t)}{\partial t}, \quad (6.105)$$

in which we have assumed incompressibility $dV = 0$. $s(x, t)$, $u(x, t)$, and $c_i(x, t)$ are entropy density, energy density, and concentration of the i th species.

Realizing that both energy and particles follow continuity equation in space-time, one has

$$\frac{\partial u(x, t)}{\partial t} = -\frac{\partial J_u(x, t)}{\partial x}, \quad \frac{\partial c_i(x, t)}{\partial t} = -\frac{\partial J_i(x, t)}{\partial x}. \quad (6.106)$$

Then, substituting these into Eq. (6.105), and use a certain amount of physical intuition, one arrives at

$$\frac{\partial s(x, t)}{\partial t} = e_p(x, t) + J_S(x, t) \quad (6.107a)$$

where entropy production rate per unit volume

$$e_p(x, t) = J_u \frac{\partial}{\partial x} \left(\frac{1}{T} \right) - \sum_{i=1}^n J_i \frac{\partial}{\partial x} \left(\frac{\mu_i}{T} \right) - \sum_{j=1}^m \frac{\Delta \mu_j \hat{\phi}_j}{T}, \quad (6.107b)$$

and entropy flux

$$J_S(x, t) = \frac{\partial}{\partial x} \left(\frac{J_u}{T} - \sum_{i=1}^n \frac{\mu_i J_i}{T} \right). \quad (6.107c)$$

According to Onsager's theory [18], each term in the entropy production e_p is a

$$\text{transport flux} \times \text{driving force} \quad (6.108)$$

which should be non-negative. The theory of nonequilibrium thermodynamics concerns with transport processes of various kinds: diffusion, heat, charge, chemical, etc. More information on the various transport fluxes can only be obtained, phenomenologically, from engineering.

6.9 Mathematicothermodynamics of Markov Dynamics

We now consider discrete-state Markov system with stochastic dynamics in terms of "continuity equation for probability in state space", e.g., Chapman–Kolmogorov equation, or master equation

$$\frac{dp_i(t)}{dt} = \sum_{j=1}^N (p_j q_{ji} - p_i q_{ij}), \quad (6.109)$$

in which q_{ij} are the infinitesimal transition probability rate given in (6.27).

We shall now follow the same logic steps of Boltzmann, illustrated in Sect. 6.8.1, to develop a "thermodynamic theory" based on the general dynamics by introducing the notion of entropy. Eq. (6.109) replaces the Hamiltonian system (6.97), and in the place of Boltzmann's celebrated $S = k_B \ln \Omega(E)$ will be the Gibbs-Shannon entropy:

$$S(t) = - \sum_{i=1}^N p_i(t) \ln p_i(t). \quad (6.110)$$

Then, one has

$$\frac{dS}{dt} = e_p + J_S, \quad (6.111a)$$

where

$$e_p(t) = \frac{1}{2} \sum_{i,j=1}^N \left(p_i(t)q_{ij} - p_j(t)q_{ji} \right) \ln \left(\frac{p_i(t)q_{ij}}{p_j(t)q_{ji}} \right), \quad (6.111b)$$

$$J_S(t) = \frac{1}{2} \sum_{i,j=1}^N \left(p_i(t)q_{ij} - p_j(t)q_{ji} \right) \ln \left(\frac{q_{ji}}{q_{ij}} \right). \quad (6.111c)$$

It is immediately obvious that $e_p \geq 0$ since for every pair of i, j in Eq. (6.111b), the term has the form of $(a - b) \ln(a/b) \geq 0$. We also note the resemblance of (6.111b) to Eq. (6.76).

Therefore, we have derived an entropy balance equation based on Markov dynamics, without the assumption of local equilibrium. Equations (6.111b) and (6.111c) further give explicit expressions, in terms of the $\{p_i(t)\}$, for the entropy flux J_S the non-negative entropy production e_p . As we shall show below, there is a complete nonequilibrium thermodynamics on the mesoscopic scale, in state space. This theory is effectively an isothermal theory with the “temperature” being 1.

6.9.1 Non-Decreasing Entropy in Systems with Uniform Stationary Distribution

If the master Eq. (6.109) has a stationary distribution $p_n^{ss} = 1 \forall n$, then

$$\sum_{j=1}^N (q_{ji} - q_{ij}) = \sum_{j=1}^N q_{ji} = 0, \quad \forall i.$$

In this case,

$$\begin{aligned} \frac{dS}{dt} &= - \sum_{i=1}^N \left(\frac{dp_i(t)}{dt} \right) \ln p_i = - \sum_{i,j=1}^N (p_j q_{ji} - p_i q_{ij}) \ln p_i \\ &= \sum_{i,j=1}^N p_i q_{ij} \ln \left(\frac{p_i}{p_j} \right) \geq \sum_{i,j=1}^N p_i q_{ij} \left(\frac{p_j}{p_i} - 1 \right) \\ &= \sum_{j=1}^N p_j \left(\sum_{i=1}^N q_{ij} \right) = 0. \end{aligned} \quad (6.112)$$

We therefore have a “theorem” stating that if the stationary probability distribution is uniform, then the entropy S is non-decreasing function of time.

6.9.2 Q -Processes with Detailed Balance

If a Q process has a stationary distribution $p_i^{ss} q_{ij} = p_j^{ss} q_{ji}$, known as *detailed balance*, then

$$\begin{aligned}
 J_S(t) &= \frac{1}{2} \sum_{i,j=1}^N \left(p_i(t) q_{ij} - p_j(t) q_{ji} \right) \ln \left(\frac{q_{ji}}{q_{ij}} \right) \\
 &= \frac{1}{2} \sum_{i,j=1}^N \left(p_i(t) q_{ij} - p_j(t) q_{ji} \right) \ln \left(\frac{p_i^{ss}}{p_j^{ss}} \right) \\
 &= \sum_{i,j=1}^N \left(p_j(t) q_{ji} - p_i(t) q_{ij} \right) \ln p_j^{ss} = - \sum_{j=1}^N \frac{dp_j(t)}{dt} \ln p_j^{ss} \\
 &= \frac{d}{dt} \left(\sum_{j=1}^N p_j(t) \left(- \ln p_j^{ss} \right) \right) = \frac{1}{T} \frac{d\bar{E}}{dt}, \tag{6.113}
 \end{aligned}$$

in which

$$\bar{E} = \sum_{j=1}^N p_j(t) E_j, \tag{6.114}$$

should be identified as the mean energy, with $E_j = -T \ln p_j^{ss}$ as the “energy” of the state i according to Boltzmann’s law. Then, Eq. (6.11a) becomes

$$\frac{d}{dt} \left(\frac{\bar{E}}{T} - S \right) = -e_p \leq 0. \tag{6.115}$$

$F = \bar{E} - TS$ is known as the “free energy” of a thermodynamic system. It is expected to monotonically decrease with time in an isothermal system approaching to equilibrium. In an equilibrium steady state, the free energy reaches its minimum.

6.9.3 Monotonicity of F Change in General Q -Processes

Encouraged by the above results, let us consider the Kullback–Leibler divergence, also known as relative entropy:

$$F(t) = \sum_{i=1}^N p_i(t) \left(-\ln p_i^{ss} + \ln p_i(t) \right) = \sum_{i=1}^N p_i(t) \ln \left(\frac{p_i(t)}{p_i^{ss}} \right) \geq 0. \quad (6.116)$$

One can actually show that $dF/dt \leq 0$ for general Q -process without the detailed balance:

$$\begin{aligned} \frac{dF(t)}{dt} &= \sum_{i=1}^N \left(\frac{dp_i(t)}{dt} \right) \ln \left(\frac{p_i(t)}{p_i^{ss}} \right) = \sum_{i,j=1}^N (p_j q_{ji} - p_i q_{ij}) \ln \left(\frac{p_i(t)}{p_i^{ss}} \right) \\ &= \sum_{i,j=1}^N p_j q_{ji} \ln \left(\frac{p_i(t) p_j^{ss}}{p_i^{ss} p_j(t)} \right) \leq \sum_{i,j=1}^N p_j q_{ji} \left(\frac{p_i(t) p_j^{ss}}{p_i^{ss} p_j(t)} - 1 \right) \\ &= \sum_{i=1}^N \frac{p_i}{p_i^{ss}} \left(\sum_{j=1}^N (p_j^{ss} q_{ji} - p_i^{ss} q_{ij}) \right) = 0. \end{aligned} \quad (6.117)$$

6.9.4 F Balance Equation of Markov Dynamics

More interestingly, we have a new, balance equation for the $F(t)$:

$$\frac{dF(t)}{dt} = E_{in}(t) - e_p(t), \quad (6.118a)$$

where $e_p(t) \geq 0$ is given in (6.111b), and

$$E_{in}(t) = \frac{1}{2} \sum_{i,j=1}^N (p_i(t) q_{ij} - p_j(t) q_{ji}) \ln \left(\frac{p_i^{ss} q_{ij}}{p_j^{ss} q_{ji}} \right) \geq 0. \quad (6.118b)$$

See [9] for the proof of this inequality. Both $E_{in}(t)$ and $e_p(t)$ are non-negative which means that Eq. (6.118a) can be interpreted as “the $F(t)$ has a source and a sink”, its change equals to an input $E_{in}(t)$, a source term, and dissipation $e_p(t)$, a sink term. There is a mesoscopic conservation of the quantity F . Equation (6.118a) is more meaningful than the Eq. (6.111a), in which J_S does not have a definitive sign.

The balance Eq. (6.118a) and the monotonicity of $dF/dt \leq 0$ have remarkable resemblance to the first and the second laws of thermodynamics. But they are really a part of a mathematical structure of any stochastic Markov dynamics.

To emphasize this mathematical nature, we call all the results in this section, collectively, *mathematicothermodynamics* [9, 10, 21, 24].

6.9.5 Driven System and Cycle Decomposition

The entropy production given in (6.111b) can be written as

$$e_p = \sum_{\text{all edge } ij}^N (\varphi_{ij} - \varphi_{ji}) \ln \left(\frac{\varphi_{ij}}{\varphi_{ji}} \right), \quad (6.119)$$

where $\varphi_{ij} = p_i(t)q_{ij}$ is the one-way probability flux from state i to j . It can be proven that, in a stationary Q-process, the above expression can be expressed also as [14]

$$e_p = \sum_{\text{all cycles } \Gamma}^N (\varphi_{\Gamma}^+ - \varphi_{\Gamma}^-) \ln \left(\frac{\varphi_{\Gamma}^+}{\varphi_{\Gamma}^-} \right), \quad (6.120)$$

in which φ_{Γ}^{\pm} is the number of Γ cycle completed in a unit time, in the forward and backward direction. Most importantly, for cycle $\Gamma = (i_0, i_1, \dots, i_n, i_0)$

$$\frac{\varphi_{\Gamma}^+}{\varphi_{\Gamma}^-} = \frac{q_{i_0 i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} q_{i_n i_0}}{q_{i_1 i_0} q_{i_2 i_1} \cdots q_{i_n i_{n-1}} q_{i_0 i_n}}, \quad (6.121)$$

which is independent of the probabilities! Therefore, $\ln(\varphi_{\Gamma}^+/\varphi_{\Gamma}^-)$ can and should be understood as the entropy production per cycle, and the term $(\varphi_{\Gamma}^+ - \varphi_{\Gamma}^-)$ is simply a kinematic term that counts the number of cycle completed along a trajectory. All the nonequilibrium thermodynamics is contained in the (6.121); it is about kinetic cycles [27]. If a Markov process is detail balanced, then its entropy production is zero on each and every kinetic cycle.

It is well known since the work of A. N. Kolmogorov that the quantity in (6.121) equals unity for each and every cycle if and only if the Markov process is detailed balanced. Therefore, the mathematical notion of *detailed balance* provides a fitting description of a non-driven kinetic system whose steady state is an equilibrium. For a driven kinetic system, at least one of the cycles in the state space Γ has unbalanced circulation: $\varphi_{\Gamma}^+ \neq \varphi_{\Gamma}^-$.

6.9.6 Macroscopic Thermodynamics in the Kurtz Limit

For a DGP with N species and M reactions, the F function introduced in Sect. 6.9.4 is a functional of the probability distribution $p_V(\mathbf{n}, t)$ which is itself a function of the reaction system's volume V . Then one naturally asks what its macroscopic limit is as $V \rightarrow \infty$ as in the Kurtz limit? It can be shown that [10]

$$\begin{aligned} \lim_{V \rightarrow \infty} \frac{F[p_V(\mathbf{n}, t)]}{V} &= \lim_{V \rightarrow \infty} \frac{1}{V} \sum_{\mathbf{n}} p_V(\mathbf{n}, t) \ln \left[\frac{p_V(\mathbf{n}, t)}{p_V^{ss}(\mathbf{n})} \right] \\ &= - \lim_{V \rightarrow \infty} \frac{1}{V} \sum_{\mathbf{n}} p_V(\mathbf{n}, t) \ln p_V^{ss}(\mathbf{n}) \\ &= G^{ss}[\mathbf{x}(t)], \end{aligned} \quad (6.122)$$

in which $\mathbf{n} = (n_1, n_2, \dots, n_N)$, n_k is the number of molecules of the k th species, $\mathbf{x} = (x_1, \dots, x_N)$ is the corresponding number density $\mathbf{x} = \frac{\mathbf{n}}{V}$. The Kurtz theorem in Sect. 6.7 states that the stochastic trajectory of a DGP, $\mathbf{n}_V(t)$,

$$\lim_{V \rightarrow \infty} \frac{\mathbf{n}_V(t)}{V} = \mathbf{x}(t), \quad (6.123)$$

where $\mathbf{x}(t)$ is the solution to the deterministic, nonlinear rate equation (e.g., Eq. (6.89)). Most interestingly, according to the large deviation principle from the theory of probability, when the steady state probability $p_V^{ss}(\mathbf{n})$ converges to a Dirac- δ function, its tail probability has an asymptotic expression

$$- \lim_{V \rightarrow \infty} \frac{\ln p_V^{ss}(\mathbf{n})}{V} = - \lim_{V \rightarrow \infty} \frac{\ln p_V^{ss}(V\mathbf{x})}{V} = G^{ss}(\mathbf{x}). \quad (6.124)$$

This steady state large deviation rate function $G^{ss}(\mathbf{x})$ can be identified as a generalized Gibbs function for nonequilibrium chemical reaction systems. It can be shown that

$$\frac{d}{dt} G^{ss}[\mathbf{x}(t)] = \left(\frac{d\mathbf{x}(t)}{dt} \right) \cdot \nabla_{\mathbf{x}} G^{ss}(\mathbf{x}) \leq 0. \quad (6.125)$$

This is a generalization of the inequality in Eq. (6.77). See [10] for the proof.

6.10 Summary and Conclusion

This chapter presents a new modeling paradigm for biological systems and processes that consist of multiple populations of individuals, each with an infinite many internal degrees of freedom. The individuals are grouped into subpopulations

and mathematically represented by their statistical behaviors in terms of birth, death, migration, and state switching. We show that the population kinetics in terms of nonlinear ordinary differential equations (ODEs) widely employed in mathematical biology is fundamentally a stochastic kinetic theory. This stochastic population kinetic representation of biological reality can be introduced quite rigorously, thus it provides one with confidence in the conclusions drawn from mathematical analysis. We called this formalism *Delbrück-Gillespie process*. In the large population limit, T. G. Kurtz's theorem, a law of large numbers, yields a system of nonlinear rate equations that is consistent with the traditional ODEs. In Sect. 6.9, very recent results on mesoscopic nonequilibrium thermodynamics and its corresponding macroscopic nonequilibrium thermodynamics are presented. Together the three parts, (1) stochastic kinetics in terms of DGP, (2) deterministic nonlinear dynamics in terms of ODEs, and (3) the mathematicothermodynamics, provide a comprehensive mathematical theory for a wide range of biological systems and processes from biochemistry to ecology.

6.11 Exercises: Simple and Challenging

6.11.1 Simple Exercises

1. Compute the expected value and the variance of an exponentially distributed random variable X with rate λ .
2. Let X_1, \dots, X_n be n i.i.d. exponential random variables with rate λ . Let $X^* = \min\{X_1, X_2, \dots, X_n\}$. Show that $f_{T^*}(t) = n\lambda e^{-n\lambda t}$.
3. If a set of n i.i.d. random times all with distribution $f_T(t)$, $f_T(0) = 0$ but $f'_T(0) \neq 0$, what is the distribution for $T^* = \min\{T_1, T_2, \dots, T_n\}$ in the limit of $n \rightarrow \infty$?

6.11.2 More Challenging Exercises

4. Consider a population consisting of identical and independent individual organisms, each with an exponentially distributed time for giving "birth", with rate λ , and going "death", with rate μ .
 - (i) Now when the population has exactly n individuals, what is the probability distribution for the waiting time to the next birth? What is the probability distribution for the waiting time to the next death? What is the probability distribution for the waiting time to the next birth or death event?

- (ii) Let $p_n(t)$ be the probability of having exactly n individuals in the population at time t :

$$\sum_{n=0}^{\infty} p_n(t) = 1.$$

What system of differential equations should $p_n(t)$ satisfy?

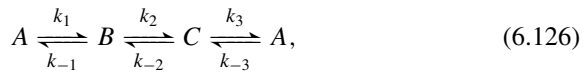
- (iii) The mean population at time t is defined as

$$\langle n \rangle(t) = \sum_{n=0}^{\infty} n p_n(t).$$

Based on the system of differential equations you obtained in (ii), show that

$$\frac{d}{dt} \langle n \rangle = (\lambda - \mu) \langle n \rangle.$$

5. The 3-state Markov system,



has been widely used in biochemistry to model the conformational changes of a single protein molecule undergoing through its three different states A , B , and C . For example, A is non-active, B is partially active, and C is fully active.

- (a) The probabilities for the states, $\mathbf{p} = (p_A, p_B, p_C)$, satisfies a differential equation

$$\frac{d}{dt} \mathbf{p}(t) = \mathbf{p}(t) \mathbf{Q},$$

where \mathbf{Q} is a 3×3 matrix. Write the \mathbf{Q} out in terms of the k 's. Show that the sum of each and every row is zero. Discuss in probabilistic terms, what is the meaning of this result?

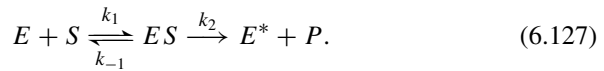
- (b) Compute the steady state probabilities p_A^{ss} , p_B^{ss} , and p_C^{ss} , and show that, in the steady state, the net (probabilistic) flux from state A to B ,

$$J_{A \rightarrow B}^{ss} = k_1 p_A^{ss} - k_{-1} p_B^{ss},$$

is the same as the net flux from state $B \rightarrow$ state C , and also the net flux from $C \rightarrow A$. Since they are all the same, it is called the steady state flux J^{ss} of the biochemical reaction cycle in (6.126).

- (c) What is the condition, in terms of all the k 's, for $J^{ss} = 0$?

6. Consider a single enzyme E in the sea of substrate molecule S . The Michaelis–Menten kinetics is



Because there is only a single enzyme molecule working, the concentration of S can be assumed as always constant, at the value c_S .

Write the differential equations for the probability of the enzyme being in state E , ES , and E^* : $p_E(t)$, $p_{ES}(t)$, and $p_{E^*}(t)$.

Given initial condition $p_E(0) = 1$, $p_{ES}(0) = 0$, and $p_{E^*}(0) = 0$, try to solve $p_{E^*}(t)$.

It is clear that the time for the enzyme to move from state E to E^* is stochastic. Let T be the random time. What is the probability distribution for T , $f_T(t)$? How is it related to $p_{E^*}(t)$?

Compute expected value $\mathbb{E}[T]$. Compare your result with the Michaelis–Menten formula.

References

1. D.A. Beard, H. Qian, *Chemical Biophysics: Quantitative Analysis of Cellular Systems*, in Texts in Biomedical Engineering (Cambridge University Press, London, 2008)
2. P.G. Bergmann, J.L. Lebowitz, New approach to nonequilibrium processes. *Phys. Rev.* **99**, 578–587 (1955)
3. R.D. Cox, *Renewal Theory* (Methuen & Company, New York, 1970)
4. P.G. de Gennes, Molecular individualism. *Science* **276**, 1999–2000 (1997)
5. S.R. de Groot, P. Mazur, *Nonequilibrium Thermodynamics* (Dover, New York, 1984)
6. A.W.F. Edwards, The fundamental theorem of natural selection. *Biol. Rev.* **69**, 443–474 (1994)
7. M.I. Freidlin, A.D. Wentzell, *Random Perturbations of Dynamical Systems*, 3rd ed. (Springer, New York, 2012)
8. G. Gallavotti, *Statistical Mechanics: A Short Treatise* (Springer, Berlin, 1999)
9. H. Ge, H. Qian, The physical origins of entropy production, free energy dissipation and their mathematical representations. *Phys. Rev. E* **81**, 051133 (2010)
10. H. Ge, H. Qian, Mesoscopic kinetic basis of macroscopic chemical thermodynamics: a mathematical theory. *Phys. Rev. E* **94**, 052150 (2016)
11. R.J. Herrnstein, C. Murray, *Bell Curve: Intelligence and Class Structure in American Life* (Free Press, New York, 1996)
12. S. Huang, F. Li, J.X. Zhou, H. Qian, Processes on the emergent landscapes of biochemical reaction networks and heterogeneous cell population dynamics: differentiations in living matters (review). *J. R. Soc. Interface* **14**, 20170097 (2017)
13. F. Jacob, *Possible and Actual—Jessie and John Danz Lectures* (University Washington Press, Seattle, 1994)
14. D.-Q. Jiang, M. Qian, M.-P. Qian, *Mathematical Theory of Nonequilibrium Steady States*. LNM, vol. 1833 (Springer, New York, 2004)
15. S.D. Levitt, S.J. Dubner, *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* (William Morrow, New York, 2005)

16. A.J. Lotka, Analytical note on certain rhythmic relations in organic systems. Proc. Natl. Acad. Sci. USA **6**, 410–415 (1920)
17. J.D. Murray, *Mathematical Biology I: An Introduction*, 3rd ed. (Springer, New York, 2002)
18. L. Onsager, Reciprocal relations in irreversible processes, I. Phys. Rev. **37**, 405–426 (1931)
19. H. Qian, Cellular biology in terms of stochastic nonlinear biochemical dynamics: emergent properties, isogenetic variations and chemical system inheritability. J. Stat. Phys. **141**, 990–1013 (2010)
20. H. Qian, Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems—an analytical theory (invited article). Nonlinearity **24**, R19–R49 (2011)
21. H. Qian, A decomposition of irreversible diffusion processes without detailed balance. J. Math. Phys. **54**, 053302 (2013)
22. H. Qian, *Nonlinear Stochastic Dynamics of Complex Systems, I: A Chemical Reaction Kinetic Perspective with Mesoscopic Nonequilibrium Thermodynamics* (2016). ArXiv:1605.08070
23. H. Qian, L.M. Bishop, The chemical master equation approach to nonequilibrium steady-state of open biochemical systems: linear single-molecule enzyme kinetics and nonlinear biochemical reaction networks (Review). Int. J. Mol. Sci. **11**, 3472–3500 (2010)
24. H. Qian, M. Qian, X. Tang, Thermodynamics of the general diffusion process: time-reversibility and entropy production. J. Stat. Phys. **107**, 1129–1141 (2002)
25. H. Qian, S. Saffarian, E.L. Elson, Concentration fluctuations in a mesoscopic oscillating chemical reaction system. Proc. Natl. Acad. Sci. USA **99**, 10376–10381 (2002)
26. H. Qian, P. Ao, Y. Tu, J. Wang, A framework towards understanding mesoscopic phenomena: Emergent unpredictability, symmetry breaking and dynamics across scales. Chem. Phys. Lett. **665**, 153–161 (2016)
27. H. Qian, S. Kjelstrup, A.B. Kolomeisky, D. Bedeaux, Entropy production in mesoscopic stochastic thermodynamics—Nonequilibrium kinetic cycles driven by chemical potentials, temperatures, and mechanical forces (Topical review). J. Phys. Condens. Matter **28**, 153004 (2016)
28. D.B. Saakian, H. Qian, *Nonlinear Stochastic Dynamics of Complex Systems, III: Nonequilibrium Thermodynamics of Self-Replication Kinetics* (2016). ArXiv:1606.02391
29. C.E. Shannon, W. Weaver, *The Mathematical Theory of Communication* (University Illinois Press, Chicago, 1963)
30. H.M. Taylor, S. Karlin, *An Introduction to Stochastic Modeling*, 3rd ed. (Academic Press, New York, 1998)
31. L.F. Thompson, H. Qian, *Nonlinear Stochastic Dynamics of Complex Systems, II: Potential of Entropic Force in Markov Systems with Nonequilibrium Steady State, Generalized Gibbs Function and Criticality* (2016). ArXiv:1605.08071
32. N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, revised and enlarged ed. (Elsevier, Amsterdam, 1992)
33. M. Vellela, H. Qian, A quasi-stationary analysis of a stochastic chemical reaction: Keizer’s paradox. Bull. Math. Biol. **69**, 1727–1746 (2007)
34. M. Vellela, H. Qian, Stochastic dynamics and nonequilibrium thermodynamics of a bistable chemical system: the Schlögl model revisited. J. R. Soc. Interface **6**, 925–940 (2009)
35. M. Vellela, H. Qian, On Poincaré–Hill cycle map of rotational random walk: locating stochastic limit cycle in reversible Schnakenberg model. Proc. R. Soc. A Math. Phys. Eng. **466**, 771–788 (2010)
36. J. Voigt, Stochastic operators, information, and entropy. Commun. Math. Phys. **81**, 31–38 (1981)
37. D. Zhou, H. Qian, Fixation, transient landscape and diffusion’s dilemma in stochastic evolutionary game dynamics. Phys. Rev. E **84**, 031907 (2011)