Lecture 7: Equilibrium Physcis of Bioregulation

20251023

Lecturer: Fangzhou Xiao   *Scribe: Yiqiao Deng, Shaocong Fang, Chengqian Li*

Contents

1   What is Ergodicity

Ergodicity is crucial for equilibrium. Ergodicity means that, given sufficient time, a system can evolve into any possible state. In a closed system with ergodicity, after evolving for a sufficiently long time, it will inevitably reach an equilibrium state, which unifies temporal equilibrium and spatial equilibrium. Here, temporal equilibrium refers to the average value of the trajectory of a single point in the system after evolving for a sufficiently long time, while spatial equilibrium refers to the average value formed by the superposition of all states of all points in the system at a specific moment. In an ergodic system, the equilibrium state ensures that the temporal equilibrium equals the spatial equilibrium. Such ergodic systems provide us with a simplified method for solving their equilibrium states. We only need to simulate a single point, starting from a certain initial state, and let this point evolve for a sufficiently long time. The temporal average obtained from its trajectory will then equal the equilibrium state of the system. For example, consider the velocity distribution of ideal gas molecules at a certain temperature. We only need to simulate one molecule, let it evolve under the system's conditions for a sufficiently long time, and the distribution of the velocities it has assumed will correspond to the velocity distribution of all the ideal gas molecules in the system—that is, the equilibrium state of the system.

Furthermore, if a system is not ergodic, then its equilibrium state is neither unique nor determinate, which means we cannot discuss its equilibrium in a definitive manner. For example, a system loses ergodicity if one of its state variables can go extinct. By "extinct," we mean that once this state variable reaches zero, it becomes trapped at zero forever. Consider a state variable x, where dx/dt depends only on x, and when x equals zero, $dx/dt = 0$. In this case, once the system reaches $x = 0$, x will be permanently trapped at that point. Therefore, even if the system possesses several stable points (or equilibrium states) where $x \neq 0$, once x reaches zero, it can never access those non-zero stable points, even over an infinite amount of time. This system thus loses ergodicity. Consequently, the

equilibrium of such a system is indeterminate. Depending on the initial conditions (i.e., the value of x at $t = 0$), the system might sometimes reach a non-zero stable point, and other times become trapped at $x = 0$. In this scenario, the stable points and equilibrium states are not uniquely determined.

2 Similarities and differences between Steady State and Equilibrium

Next, we will discuss the similarities and differences between steady state and equilibrium. First, we need to clarify how steady state and equilibrium are described mathematically.

2.1 Mathematical Definitions

Steady state refers to a situation where all state variables in a system do not change with time, i.e.,

$$\frac{dx_i}{dt} = 0 \quad \text{for all state variables } x_i. \tag{1}$$

Equilibrium (specifically, detailed balance) refers to a condition where the transitions between different state variables are balanced. That is, for any state x_i, the total rate of transitions into x_i from all directly connected states x_j equals the total rate of transitions out of x_i to all directly connected states x_k. Mathematically,

$$\sum_j v_{ji} = \sum_k v_{ik}, \tag{2}$$

where v_{ji} denotes the rate from state x_j to x_i, and v_{ik} denotes the rate from state x_i to x_k. This condition is known as **detailed balance**.

2.2 Are Steady State and Equilibrium Equivalent?

Now, is steady state equivalent to equilibrium? For linear (chain) networks, yes; but for cyclic (ring) networks, not necessarily. We will discuss two examples separately.

2.2.1 Example 1: Linear Network $x_1 \leftrightarrow x_2 \leftrightarrow x_3$

In a linear network, the endpoints are x_1 and x_3. If the system is in a steady state, then for the endpoints, we have:

For x_1:

$$\frac{dx_1}{dt} = v_{21} - v_{12} = 0 \quad \Rightarrow \quad v_{21} = v_{12}. \tag{3}$$

For x_2:

$$\frac{dx_2}{dt} = (v_{12} + v_{32}) - (v_{21} + v_{23}) = (v_{12} - v_{21}) + (v_{32} - v_{23}) = 0. \tag{4}$$

Since $v_{12} - v_{21} = 0$ from the steady state condition of x_1, it follows that $v_{32} - v_{23} = 0$. Thus, the steady state condition propagates along the chain, ensuring that detailed balance holds at every node. Therefore, in a linear network, steady state implies equilibrium.

2.2.2 Example 2: Cyclic Network $x_1 \leftrightarrow x_2 \leftrightarrow x_3 \leftrightarrow x_1$

In a cyclic network, there are no endpoints. Consider the steady state condition for x_1:

$$\frac{dx_1}{dt} = (v_{21} + v_{31}) - (v_{12} + v_{13}) = (v_{21} - v_{12}) + (v_{31} - v_{13}) = 0. \tag{5}$$

This equation can be satisfied without requiring $v_{21} - v_{12} = 0$ and $v_{31} - v_{13} = 0$ individually. For example, if $v_{21} - v_{12} < 0$ and $v_{31} - v_{13} > 0$, their sum can still be zero. This corresponds to a net flow in the cycle: for instance, $x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_1$, with more flow from x_3 to x_1 than from x_1 to x_3, and more flow from x_1 to x_2 than from x_2 to x_1. Thus, steady state does not necessarily imply equilibrium in a cyclic network.

2.3 Energy Consumption and Biological Implications

Achieving a steady state without equilibrium in a cyclic network requires energy input. This is due to the second law of thermodynamics: in a closed system, a net flow would increase entropy, so maintaining a non-equilibrium steady state requires external energy. Biological systems often exploit such non-equilibrium steady states in cyclic networks to perform functions.

Two classic examples are the **Citric Acid Cycle** (Krebs cycle) and the **Calvin-Benson Cycle**:

- In the **Citric Acid Cycle**, metabolites remain at approximately constant concentrations under stable conditions (steady state), but the cycle is not at equilibrium. It continuously consumes acetyl-CoA and produces CO_2, NADH, and $FADH_2$. If it were at equilibrium, no net production of these energy carriers would occur, preventing aerobic respiration and ATP synthesis.

- In the **Calvin-Benson Cycle**, metabolites are also in steady state, but the cycle consumes CO_2 and produces organic carbon for biomass synthesis. Equilibrium would halt carbon fixation. This cycle is driven by energy from light reactions, via NADPH.

The ability of these cycles to maintain steady-state concentrations despite external perturbations reflects their **robustness**.

3 Consider Equilibrium under Macro and Micro View

In the macroscopic perspective, equilibrium is described using equilibrium constants. In the microscopic perspective, equilibrium is described using statistical mechanics. We will now consider how to connect these two approaches.

We begin by outlining the tools we will use. In the macroscopic perspective, we will use the van't Hoff equation, which relates the Gibbs free energy change to the equilibrium constant. Its mathematical form is:

$$\Delta G = -kT \ln K_{eq} \tag{6}$$

where ΔG is the Gibbs free energy change, k is the Boltzmann constant, T is the temperature, and K_{eq} is the equilibrium constant for the reaction.

In the microscopic perspective, statistical mechanics uses the Boltzmann distribution as a fundamental postulate. The Boltzmann distribution is suitable for describing any non-quantum (or classical) system in thermodynamic equilibrium. The Boltzmann distribution states that the probability P_i of a system being in state i is:

$$P_i = \frac{1}{Z} e^{-E_i/kT} \tag{7}$$

where k is the Boltzmann constant, T is the temperature, and Z is the partition function. Here, the partition function serves primarily as a normalization factor to ensure that the probabilities sum to one. In my view, the Boltzmann distribution essentially tells us that as the energy of a state increases, its likelihood of existence decreases exponentially.

Macroscopic Thermodynamic Starting Point

Thermodynamics provides the relationship between the standard Gibbs free energy change $\Delta G°$ and the equilibrium constant K_{eq} (van 't Hoff isotherm):

$$\Delta G° = -RT \ln K_{eq} \tag{8}$$

For the reaction $A \rightleftharpoons B$, the equilibrium constant is defined as:

$$K_{eq} = \frac{[B]}{[A]} \tag{9}$$

Statistical Mechanics Interpretation

In the framework of statistical mechanics, the concentration ratio can be interpreted as the probability ratio of the system being in different macroscopic states. For a randomly selected molecule:

$$\frac{[B]}{[A]} = \frac{P_B}{P_A} \tag{10}$$

where P_A and P_B represent the probabilities of the molecule being in state A and B, respectively.

According to the Boltzmann distribution, the probability of the system being in a particular macroscopic state is proportional to its partition function:

$$P_A \propto Z_A, \quad P_B \propto Z_B \tag{11}$$

Therefore, the probability ratio equals the partition function ratio:

$$\frac{P_B}{P_A} = \frac{Z_B}{Z_A} \tag{12}$$

Microscopic Nature of the Partition Function

The microscopic definition of the partition function Z is the sum of Boltzmann factors over all possible microstates. For macroscopic state B:

$$Z_B = \sum_{\text{all microstates } i \text{ belonging to B}} e^{-E_i^B/kT} \tag{13}$$

We can group microstates by energy. Assuming that at energy $E_{B,j}$, state B has $\Omega_B(E_{B,j})$ degenerate microstates, then:

$$Z_B = \sum_j \Omega_B(E_{B,j}) \cdot e^{-E_{B,j}/kT} \tag{14}$$

Similarly, for state A:

$$Z_A = \sum_j \Omega_A(E_{A,j}) \cdot e^{-E_{A,j}/kT} \tag{15}$$

Complete Derivation Chain

Combining equations (8) through (15), we obtain the complete derivation:

$$
\begin{aligned}
\Delta G^\circ &= -RT \ln K_{eq} \quad \text{(Macroscopic thermodynamic relation)} \\
&= -RT \ln\left(\frac{[B]}{[A]}\right) \quad \text{(Equilibrium constant definition)} \\
&= -RT \ln\left(\frac{P_B}{P_A}\right) \quad \text{(Statistical interpretation)} \\
&= -RT \ln\left(\frac{Z_B}{Z_A}\right) \quad \text{(Boltzmann distribution)} \\
&= -RT \ln\left(\frac{\sum_j \Omega_B(E_{B,j}) \cdot e^{-E_{B,j}/kT}}{\sum_j \Omega_A(E_{A,j}) \cdot e^{-E_{A,j}/kT}}\right) \quad \text{(Microscopic nature of partition functions)}
\end{aligned}
$$

Final Conclusion

We thus arrive at the microscopic statistical expression for the equilibrium constant:

$$K_{eq} = \frac{Z_B}{Z_A} = \frac{\sum_j \Omega_B(E_{B,j}) \cdot e^{-E_{B,j}/kT}}{\sum_j \Omega_A(E_{A,j}) \cdot e^{-E_{A,j}/kT}} \tag{16}$$

This result reveals the microscopic nature of the macroscopic equilibrium constant K_{eq}: it is determined by the ratio of statistical weights of competing states A and B, which in turn depend on their respective energy level structures (E) and state degeneracies (Ω).

4 Biological Examples: Enzyme Allosteric Regulation, Gene Regulation

Based on lecture notes (pages 7-14), this section applies equilibrium physical theory from previous parts to specific biological processes, demonstrating how to analyze single molecule states and directly calculate probability distributions using energies and weights.

4.1 Enzyme Catalytic Reaction: Michaelis-Menten Model

4.1.1 Model Framework

Consider enzyme E catalyzing substrate S to product P:

$$E + S \rightleftharpoons ES \tag{17}$$
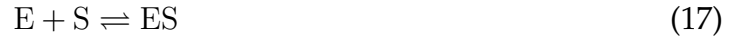
Using a lattice model with one enzyme and multiple substrate sites.

4.1.2 States and Weights

- **Free state** (E + S): weight $w_{\text{free}} \propto N_s e^{-E_f}$

- **Bound state** (ES): weight $w_{\text{bound}} \propto e^{-E_b}$

where N_s is the number of substrate molecules, E_f and E_b are energies of free and bound states.

4.1.3 Binding Probability Derivation

Binding probability:

$$p_{\text{bound}} = \frac{w_{\text{bound}}}{w_{\text{free}} + w_{\text{bound}}} = \frac{N_s e^{-\Delta E}}{1 + N_s e^{-\Delta E}} \tag{18}$$

where $\Delta E = E_f - E_b$.

4.1.4 Concentration Parameters

Define:

- Solution volume V, site volume Ω

- Substrate concentration $C_s = N_s/V$

- Dissociation constant $K_d = C_0 e^{\Delta E}$ (C_0 is standard concentration, 1M)

Rewriting binding probability:

$$p_{\text{bound}} = \frac{C_s/K_d}{1 + C_s/K_d} \tag{19}$$

4.1.5 Reaction Rate

Reaction rate is given by:

$$v = kE_{tot}p_{bound} = \frac{kE_{tot}C_s}{K_m + C_s} \tag{20}$$

where $K_m \approx K_d$, yielding the classical Michaelis-Menten equation.

4.1.6 Physical Interpretation

- Energy difference $\Delta E < 0$ indicates favorable binding
- Multiplicity N_s represents entropy contribution
- Chemical potential: $\mu \approx E_f + kT \ln(N_s/\Omega)$
- Free energy change: $\Delta G = \Delta E - T\Delta S$

4.2 Allosteric Regulation: MWC Model

4.2.1 Model Framework

MWC (Monod-Wyman-Changeux) model describes enzyme allosteric regulation:

- Two conformations: Active (A) and Inactive (I)
- Each conformation can bind substrate

4.2.2 States and Weights

Considering substrate concentration C_s, bound state weights are multiplied by factor C_s/C_0.

4.2.3 Active Probability

Probability of active conformation:

$$p_{active} = \frac{e^{-E_A}\left(1 + \frac{C_s}{K_A}\right)}{e^{-E_A}\left(1 + \frac{C_s}{K_A}\right) + e^{-E_I}\left(1 + \frac{C_s}{K_I}\right)} \tag{21}$$

where $K_A = C_0 e^{E_A - E_{ab}}$, $K_I = C_0 e^{E_I - E_{ib}}$.

4.2.4 Cooperativity Effects

When $K_A \ll K_I$ (tighter binding in active state), p_{active} changes significantly with C_s.

Introducing intersite interaction energy $E_{int} < 0$, binding probability for dimer system:

$$p_2 \propto \left(\frac{C_s}{K}\right)^2 e^{-E_{int}} \tag{22}$$

producing steeper transition curves (Hill coefficient $n > 1$).

4.2.5 Biological Significance

- Allostery enables enzymes to respond to substrate concentration changes

- Implements biological switching behavior

- Example: Hemoglobin oxygen binding curve

4.3 Gene Regulation: Lac Operon Model

4.3.1 Model Framework

Consider Lac operon transcriptional regulation with repressor and RNA polymerase competing for promoter binding.

4.3.2 States and Weights

- **No repressor state** (RNA polymerase bound):

$$w_{\text{pol}} \propto e^{-\Delta E_p}\frac{P}{N_{NS}} \tag{23}$$

- **With repressor state** (repressor bound):

$$w_{\text{rep}} \propto e^{-\Delta E_r}\frac{R}{N_{NS}} \tag{24}$$

where:

- ΔE_p, ΔE_r: binding energy differences

- P, R: RNA polymerase and repressor numbers

- N_{NS}: number of non-specific sites (multiplicity)

4.3.3 Transcription Probability

Transcription probability (RNA polymerase bound):

$$p_{\text{bound}} = \frac{e^{-\Delta E_p}P/N_{NS}}{e^{-\Delta E_p}P/N_{NS} + e^{-\Delta E_r}R/N_{NS} + \cdots} \tag{25}$$

4.3.4 Allosteric Extension

In homework problem 1, LacI repressor has active/inactive forms:

- Active form binds DNA, preventing transcription

- Inactive form does not bind DNA

- Inducer concentration c affects active form probability $p_{\text{active}}(c)$

Deriving fold-change formula:

$$\text{fold-change} = \frac{1}{1 + p_{\text{active}}(c)\frac{R}{N_{NS}}e^{-\Delta E_r}} \tag{26}$$

4.3.5 Model Advantages

- Parameters (ΔE_p, ΔE_r) can be measured from independent experiments

- No need for detailed kinetic mechanisms

- Provides quantitative predictive power

5 Beyond Equilibrium: Markov Chain

5.1 Motivation and Overall Logic

While equilibrium statistical mechanics provides powerful tools for analyzing biological systems at steady state, many crucial biological processes operate **beyond equilibrium** due to continuous energy input and driving forces. Biological systems are not closed–they exchange energy and matter with their environment, leading to behaviors that cannot be captured by equilibrium descriptions alone.

- **Key insight**: Biological systems utilize equilibrium in most components but are **sparsely driven** at critical control points

- **Examples**: Protein degradation with precise timing, phosphorylation cascades, metabolic networks with sustained fluxes

- **Challenge**: When detailed balance is broken, we need new tools to analyze steady states and dynamics

The Markov chain framework provides the mathematical foundation for analyzing these non-equilibrium systems, bridging the gap between equilibrium statistics and non-equilibrium dynamics.

5.2 Markov Chain Fundamentals

5.2.1 Definition and Derivation from Chemical Master Equation

For a continuous-time Markov chain describing molecular state transitions:

$$\frac{d\mathbf{P}}{dt} = \mathbf{Q}\mathbf{P} \tag{27}$$

$$Q_{ij} = \text{transition rate from state } j \text{ to state } i \quad (i \neq j) \tag{28}$$

$$Q_{jj} = -\sum_{i \neq j} Q_{ij} \quad \text{(conservation of probability)} \tag{29}$$

Properties:

- Probability conservation: $\mathbf{1}^\top \mathbf{P} = 1$

- Column sum zero: $\mathbf{1}^\top \mathbf{Q} = \mathbf{0}$

- Steady state: $\mathbf{Q\Pi} = \mathbf{0}$ (find null space of $\mathbf{Q}$)

- Ergodicity: Irreducible chain $\Rightarrow$ unique steady state $\mathbf{\Pi}$

Detailed Balance Condition:

$$Q_{ij}\Pi_j = Q_{ji}\Pi_i \quad \text{iff no cyclic fluxes (equilibrium)} \tag{30}$$

5.2.2 Biological Example: Phosphorylation Cascades

Consider kinase-phosphatase cycles with states representing phosphorylation levels $(0, 1, 2, \dots)$. The Markov chain captures stochastic phosphorylation/dephosphorylation events:

Process	Transition	Rate
Phosphorylation	$S_n \to S_{n+1}$	$k_n[K]$
Dephosphorylation	$S_n \to S_{n-1}$	$\gamma_n[P]$

5.3 Handling Non-Markovian Systems

Many biological processes exhibit memory effects, violating the Markov assumption. We can **embed** these systems into Markov chains by extending the state space:

$$X(t) \to (X(t), X(t - \Delta t), X(t - 2\Delta t), \dots) \tag{31}$$

Application: Protein degradation with precise timing control, where current degradation probability depends on previous states.

5.4 Dynamic Analysis: Hitting Times

5.4.1 Definition and Computation

The hitting time τ_{ji} represents the expected time to first reach state i starting from state j:

$$\tau_{ji} = \frac{1}{|q_{jj}|} + \sum_{k \neq i} \frac{q_{jk}}{|q_{jj}|} \tau_{ki} \tag{32}$$

Matrix formulation: Remove row/column i from $\mathbf{Q}$ to obtain $\mathbf{Q}^{(i)}$, then solve:

$$\mathbf{Q}^{(i)} \boldsymbol{\tau}^{(i)} = -\mathbf{1} \tag{33}$$

where $\boldsymbol{\tau}^{(i)}$ is the vector of hitting times to state i from all other states.

5.4.2 Biological Application

In phosphorylation cascades, hitting times quantify:

- Signal propagation speed from unphosphorylated to fully phosphorylated state
- Response time to external stimuli
- Temporal precision of signaling pathways

5.5 Steady State Distribution and Finite State Projection

5.5.1 Steady State Calculation

For finite-state systems, solve $\mathbf{Q\Pi} = \mathbf{0}$ subject to $\sum_i \Pi_i = 1$.

Homework Example: Binding reaction $E + S \rightleftharpoons C$ yields steady state:

$$\Pi \propto \frac{K_d^{-N_C}}{N_E! N_S! N_C!} \quad \text{(Poisson-like distribution)} \tag{34}$$

At high molecular numbers, the most probable state approaches the deterministic equilibrium.

5.5.2 Finite State Projection

For systems with large or infinite state spaces, project the Chemical Master Equation onto a finite Markov chain:

$$\text{Infinite CME} \rightarrow \text{Finite Markov Chain} \rightarrow \text{Approximate } \mathbf{\Pi} \text{ and dynamics} \tag{35}$$

Application: Metabolic cascades with non-equilibrium fluxes that cannot be captured by equilibrium thermodynamics.

5.6 Why We Use Markov Chains for Simulation and Analysis

1. **Unified Framework**: Markov chains provide a consistent mathematical structure for analyzing both equilibrium and non-equilibrium systems

2. **Computational Tractability**: The matrix $\mathbf{Q}$ enables efficient numerical computation of steady states and dynamic properties

3. **Connection to Physical Laws**: Markov chains naturally emerge from the Chemical Master Equation, which derives from fundamental stochastic reaction kinetics

4. **Bridge Between Scales**: Captures molecular stochasticity while connecting to macroscopic observables through ergodicity

5. **Experimental Validation**: Hitting times and steady state distributions provide testable predictions for single-molecule and bulk experiments

5.7   Key Takeaways

- Markov chains extend equilibrium statistical mechanics to driven biological systems

- The framework captures essential non-equilibrium features: cyclic fluxes, broken detailed balance, and energy dissipation

- Both steady-state properties (distributions) and dynamic properties (hitting times) are accessible

- Finite state projection makes experimentally relevant computations feasible

- Biological implementation requires careful consideration of Markovian assumptions and potential state space extensions

Final Insight: "It takes a lot of driving to not look like equilibrium" – weak driving forces often yield behaviors indistinguishable from equilibrium, but biological systems exploit strong, sparse driving at critical control points.

Table 1 MWC Model State Weights

State	Energy	Weight
Active free (EA)	E_A	e^{-E_A}
Active bound (EA·S)	$E_A + E_{ab}$	$e^{-(E_A + E_{ab})}$
Inactive free (EI)	E_I	e^{-E_I}
Inactive bound (EI·S)	$E_I + E_{ib}$	$e^{-(E_I + E_{ib})}$