

Stochastic Energetics

based on Ken Sekimoto (Iwanami, 2004)

1. Fluctuation and Its Description¹

1 Fluctuation.

Let O be an observed result, and $\langle \rangle$ an appropriate averaging (local spatial average, time average over some window, etc). Then, $\delta O = O - \langle O \rangle$ is the *fluctuation* of O . The nature of δO depends on how O is observed (e.g., resolving power of the observer) and the average $\langle \rangle$. δO usually exhibits a non-anticipatory property.

Random (or stochastic) variable. Since we always discuss a collection of samples when we discuss fluctuations, we need a parameter ω to specify each sample. This parameter is called the *probability parameter* and the totality Ω of these parameter values is called the *probability space* (for the problem or phenomenon under study). A *random variable* is a map from Ω to some set; if it is a real random variable it is a map to \mathbf{R} . A *stochastic process* is a map from Ω to a set of functions on time; if you start observing a sample process you pick up $\omega \in W$, and $f_\omega(t)$ is what you observe for this fixed ω .

[Here, probability measure and various expectation values must be introduced.] An expectation value of a random variable $n(\omega)$ is defined by

$$\langle n \rangle = \int n(\omega) d\mu(\omega). \quad (1)$$

If we observe an observable O many times, we can describe such an experiment by a string of observation outcomes $n_1, n_2, \dots, n_k \dots$. We may consider a stochastic process such that k is regarded as time; that is, it is a map from Ω to the set of the totality of the semi-infinite sequences of observation results. A sample path $\{n_\omega(k)\}_{k=1}^\infty$ describes one such trial run. If there is no memory effect in this sequence, and if the sequence is statistically uniform (that is, if we collect all $n_\omega(k)$, the set is k independent) we say it is an iid sequence.

Law of large numbers and refinements.

For iid sequence we have (the strong) LLN:
for almost all $\omega \in \Omega$

$$\frac{1}{N} \sum_{k=1}^N n_\omega(k) \rightarrow \langle n \rangle. \quad (2)$$

This implies the weak law of LLN

$$P \left(\left| \frac{1}{N} \sum_{k=1}^N n_\omega(k) - \langle n \rangle \right| > \epsilon \right) \rightarrow 0 \quad (3)$$

for any $\epsilon > 0$.

If the fluctuation is stationary (uniform in time) and ‘very erratic’, we could expect LLN.

¹Very different from the original.

There are two refinements of LLN that answers the following questions:

- (1) When we fix N how (3) behaves as a function of ϵ ?
- (2) When we fix ϵ how (3) behaves as a function of N ?

Central limit theorem: This answers (1). It is clear that if ϵ is appropriately scaled, for large N , the distribution becomes to sharp (almost an atomic measure). If we choose $\epsilon = \zeta/\sqrt{N}$, then asymptotically for large N

$$P\left(\left|\frac{1}{\sqrt{N}}\sum_{k=1}^N(n_{\omega}(k) - \langle n \rangle)\right| > \zeta\right) \rightarrow 1 - \frac{1}{\sqrt{2\pi V}}e^{-\zeta^2/2V} \quad (4)$$

in law. This is the *central limit theorem*, where V is the variance $V = \langle n^2 \rangle - \langle n \rangle^2$.

Large deviation principle: This answers (2). The decay rate of the LHS of (3) is exponential:

$$P\left(\left|\frac{1}{N}\sum_{k=1}^N n_{\omega}(k) - \langle n \rangle\right| > \epsilon\right) \sim e^{-NJ(\epsilon)}. \quad (5)$$

Here, the decay rate $J(\epsilon)$ vanishes if $\epsilon = 0$ (weak law of large numbers). For physicists the following form may be more appealing:

$$P\left(\frac{1}{N}\sum_{k=1}^N n_{\omega}(k) \sim x\right) \sim e^{-NI(x)}, \quad (6)$$

where I is called the *rate function*: it is nonnegative, convex, its level sets are compact, and $x = \langle n \rangle$ is its unique zero. This is called *large deviation principle*. For iid variables with finite variance I is approximated with a quadratic function around $x = \langle n \rangle$, so the central limit theorem is recovered.²

7

Random thermal force.

We idealize the thermal force $\hat{\xi}(t) = \xi(t, \omega)$ as a Gaussian white noise: its average³ is zero, and

$$\langle \xi(t, \omega)\xi(t', \omega) \rangle = 2b\delta(t - t'), \quad (7)$$

where b is a positive constant.

Free Brownian motion.

The motion of a mass point under the force $\hat{\xi}$ is called the free Brownian motion. This is realized by a small particle suspended in water, if we ignore hydrodynamic effects. It is governed by

8

$$\frac{dp}{dt} = -\gamma p + \xi(t), \quad \frac{dx}{dt} = \frac{p}{m}. \quad (8)$$

9

Here, γ is a positive constant called the *viscous friction constant*.⁴

(a) If we demand that the ensemble average of the kinetic energy is $k_B T/2$ (equipartition of energy), $b = 2k_B T\gamma$ is required. This can be shown with the explicit solution of (8).

²However, LDP and CLD have nothing to do with each other logically.

³Of course this is the average over Ω .

⁴We suppress ω ; actually, for each ω we have a realization of a Brownian motion.

(b) If we do not pay attention to the very short time behavior (time scale much shorter than γ/m), then we may ignore the inertial effect:

$$0 = -\gamma \frac{dx}{dt} + \xi(t). \quad (9)$$

Since γ/m is assumed to be very large, p is slaved to ξ as $p = \xi m/\gamma$, we obtain the above.⁵ If (9) is averaged over time Δt , we obtain

$$0 = -\gamma \frac{x(t + \Delta t) - x(t)}{\Delta t} + \xi_{\Delta t}(t), \quad (10)$$

where

$$\xi_{\Delta t}(t) = \frac{1}{\Delta t} \int_t^{t+\Delta t} \xi(s) ds. \quad (11)$$

Notice that $\xi_{\Delta t}(t)$ is the same as $\xi(t)$ in law so long as we do not ask any question for the time scale shorter than Δt . That is, (9) is invariant under time coarse-graining.

Diffusion due to free Brownian motion.

From we see

$$\langle |x(t) - x(0)|^2 \rangle = 2Dt \quad (12)$$

with

$$D = k_B T / \gamma, \quad (13)$$

which is called the *diffusion constant*. This is *Einstein's relation*. γ is usually proportional to the particle size, so large particles experience larger forces $\propto \sqrt{\gamma}$. However, the effect of friction wins and the velocity is $\propto \sqrt{k_B T / \gamma}$.

The displacement $X = x(t) - x(0)$ obeys

$$P(X) = \frac{1}{\sqrt{4\pi Dt}} e^{-X^2/4Dt}. \quad (14)$$

This can be shown easily from the generating function (Fourier transform of P)

$$\langle e^{-ikX} \rangle = \left\langle e^{-ik \int_0^t \xi(s) ds} \right\rangle = e^{-k^2 Dt}. \quad (15)$$

This also shows that P obeys the *diffusion equation*

$$\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial X^2}. \quad (16)$$

Langevin equation.

If we have some external force (obtained from a potential $U(x, a)$, where a is a parameter that can control it externally):

$$\frac{dp}{dt} = -\frac{\partial U(x, a)}{\partial x} - \gamma \frac{p}{m} + \xi(t), \quad \frac{dx}{dt} = \frac{p}{m}. \quad (17)$$

⁵A more honest derivation is to use the singular perturbation.

Just in the free case, under the overdamped condition, we can use

$$0 = -\frac{\partial U(x, a)}{\partial x} - \gamma \frac{dx}{dt} + \xi(t). \quad (18)$$

These equations are often called *langevin equations*.

Langevin equations are ‘derived’ from equations of motion of mechanics through reduction. An artificial but solvable example is as follows:

The Brownian particle is in an external potential U_0 with mass unity, to which numerous tiny particles are connected with harmonic springs: Let (X, P) be the canonical coordinates of the Brownian particle and (x_j, p_j) that for the j -th tiny particle.

$$\frac{dX}{dt} = P, \quad \frac{dP}{dt} = -\frac{\partial U_0}{\partial X} + \gamma_j \sum_{j=1}^N \left(x_j - \frac{\gamma_j}{m\omega_j^2} X \right), \quad (19)$$

$$\frac{dx_j}{dt} = p_j, \quad \frac{dp_j}{dt} = -m\omega_j^2 x_j + \gamma_j X. \quad (20)$$

13 For any X x_j can be solved formally; thus we may get rid of all the tiny particle coordinates
 14 except for their initial conditions. If we average the initial conditions assuming the equilibrium. This is left for your exercise. state, we can derive the Langevin equation.

Projection operator method.

15 A standard formal strategy to derive the Langevin equation from the equation of motion is
 16 to use a projection operator technique. The projection \mathcal{P} averages out microscopic details.
 17 With the aid of the Kawasaki identity, we obtain the following form:

$$\frac{d}{dt}A = v(A) - \int_0^t \gamma(t-s) \frac{\partial U(A(s))}{\partial A^*(s)} ds + f(t) \quad (21)$$

Here, $v(A)$ is called the *streaming term*, γ is the memory kernel that is the time correlation function of f . Notice that there is no approximation, since f contains all the microscopic dynamics. If we replace this f with a stochastic process with a correlation function γ , (21) becomes a generalized Langevin equation; this replacement is supposedly correct in law, but no one can demonstrate this.

If the correlation time is short, then we may replace $\gamma(t) = \gamma\delta(t)$. This implies that the derivative of A at t is governed by A at t ; no past information is needed. This approximation is called *Markovian approximation*.

Stochastic Calculus.

19 The Gaussian white noise is an idealization of fast changing non-anticipatory noise. The first mathematical surprise was that such an object could be rationalized at all. This was done by Wiener: He demonstrated that the ‘integral’ of the Gaussian white noise $B(t)$, symbolically,⁶

$$B(t, \omega) = \int_0^t \xi(s, \omega) ds \quad (22)$$

may be defined.⁷ The *Wiener process* $B(t)$ is a Gaussian process with
 (1) $\langle B(t) \rangle = 0$,

⁶Again, the probability parameter ω may not be systematically explicitly written.

⁷It is possible to define ξ directly as a map from Ω to a space of generalized functions of time, but Schwartz came later. See Hida.

(2) $\langle B(t)B(s) \rangle = \min(t, s)$.⁸

Show $\langle (B(t) - B(s))^2 \rangle = |t - s|$. B is almost surely continuous, and its total variation is not bounded. Therefore, calculus of functions of B is crippled. Thus, the idealization of ξ as $W \equiv dB/dt$ can be taken only after all the calculation is over within the usual calculus.

20

Itô introduced the *Itô integral* as the forward Riemann-Stieltjes sum and the *stochastic differential equation* (SDE; Langevin equation in the mathematical sense)

$$dx = a(x, t)dt + b(x, t)dB \quad (23)$$

in interpreted as a short-hand notation of the following integral equation

$$x(t) = x(0) + \int_0^t a(x(s), s)ds + \int_0^t b(x(s), s)dB(s), \quad (24)$$

where the last term is defined as an Itô integral. For a C_1 function f

$$\int_0^t f(B(s))dB(s) = \int_{B(0)}^{B(t)} f(x)dx - \frac{1}{2}f'(B(s))ds. \quad (25)$$

Therefore, obviously the usual chain rule does not work. Instead we need the *Itô lemma*:

$$df(x(t), t) = \frac{\partial}{\partial x}f(x(t), t)dx(t) + \frac{\partial}{\partial t}f(x(t), t)dt + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(x(t), t)d\langle x(t)^2 \rangle \quad (26)$$

21

Here, for x governed by (23), $d\langle x(t)^2 \rangle = b(x(t), t)^2dt$. (26) is intuitively obtained by Taylor-expanding f with only the order dt terms reserved.

(25) implies that the integration and the white-noise limit are not commutative. Schematically, limit first + Itô calculus = Ordinary calculus + limit. This implies that the ordinary Stieltjes integral must be replaced by

$$\int_{x(0)}^{x(t)} f(x(s))dx(s) \rightarrow \int_0^t f(x(s))dx(s) + \frac{1}{2}\int_0^t f'(x(s))dx(s)^2, \quad (27)$$

if we wish to take the white noise limit first. Therefore, it is convenient to introduce \circ product

$$f(x(t)) \circ dx(t) = f(x(t))dx(t) + \frac{1}{2}f'(x(t))dx(t)^2. \quad (28)$$

The integral

$$\int_0^t f(x(s)) \circ dx(s) = \int_0^t f(x(s))dx(s) + \frac{1}{2}\int_0^t f'(x(s))dx(s)^2. \quad (29)$$

is called the Stratonovich integral.⁹

Physical Langevin equation and \circ

In physics everything is of bounded variation, so the ordinary Langevin equation written in the form (23) must be interpreted in terms of the \circ product, IF it is written down physically (through physics reasoning):

$$dx = a(x, t)dt + b(x, t) \circ dB. \quad (30)$$

⁸ $B(0) = 0$ is often included in the definition, but here we omit this constraint.

⁹but justified through Itô integral, and is more restricted than the Itô integral.

That is, as a respectable SDE, we must write¹⁰

$$dx = \left[a(x, t) + \frac{1}{2}b_x(x(t), t)b(x(t), t) \right] dt + b(x, t)dB. \quad (31)$$

24

Level of Langevin descriptions

As is long known, the coarse-graining along the time axis does not alter the proper SDE, but the physical Langevin equation (or SDE in the Stratonovich form) do. Actually, the Itô formula can be physically obtained by introducing shorter microscopic time scale and the longer observation time scale and averaging-out the fast time scale changes.

For example, if the Langevin equation is physically given by

$$\frac{dp}{dt} = -\frac{\partial U(x, a)}{\partial x} - \gamma(x)\frac{p}{m} + \sqrt{2k_B T(x)\gamma(x)}\xi, \quad \frac{dx}{dt} = \frac{p}{m}, \quad (32)$$

this should be interpreted as a Stratonovich SDE; however, p does not appear in the noise amplitude, so we can understand it as a standard SDE as well:

$$dp = -\left(\frac{\partial U(x, a)}{\partial x} + \gamma(x)\frac{p}{m} \right) dt + \sqrt{2k_B T(x)\gamma(x)}dB \quad dx = \frac{p}{m}dt. \quad (33)$$

We can eliminate p assuming that we pay attention only to the time scale much longer than $m/\gamma(x)$. Then, pdt/m behaves like dx , so γ times pdt/m must be interpreted as the \circ product.¹¹

$$0 = -\gamma(x) \circ dx - \frac{\partial U(x, a)}{\partial x}dt + \sqrt{2k_B T(x)\gamma(x)}dB. \quad (34)$$

Since $\gamma \circ dx = \gamma dx + (1/2)\gamma'(dx)^2 = \gamma dx + k_B T(x)[\gamma'(x)/\gamma(x)]dt$,

$$dx = -\frac{1}{\gamma(x)}\frac{\partial U}{\partial x}dt - k_B T(x)\frac{\gamma'(x)}{\gamma(x)^2}dt + \sqrt{\frac{2k_B T(x)}{\gamma(x)}}dB. \quad (35)$$

In other words,

$$dx = \left[-\frac{1}{\gamma(x)}\frac{\partial U}{\partial x} + k_B T(x)\frac{d}{dx}\frac{1}{\gamma(x)} \right] dt + \sqrt{\frac{2k_B T(x)}{\gamma(x)}}dB. \quad (36)$$

26

If we wish to use the \circ product, (38) reads

$$\gamma \circ dx = -\frac{\partial U(x, a)}{\partial x}dt - \frac{1}{2}\frac{d}{dx}\sqrt{2k_B T(x)\gamma(x)}dx dB + \sqrt{2k_B T(x)\gamma(x)}dB, \quad (37)$$

$$= -\frac{\partial U(x, a)}{\partial x}dt - \frac{1}{2\gamma(x)}\frac{d}{dx}k_B T(x)\gamma(x)dt + \sqrt{2k_B T(x)\gamma(x)}dB. \quad (38)$$

That is, the corresponding ‘physical’ equation reads

$$\gamma(x)\frac{dx}{dt} = -\frac{\partial U}{\partial x} - \frac{1}{2\gamma(x)}\frac{d}{dx}k_B T(x)\gamma(x) + \sqrt{2k_B T(x)\gamma(x)}\xi(t). \quad (39)$$

Its true interpretation is (38) or (36).

Relation between long-time average and equilibrium statistical mechanics.

Ideally, equilibrium statistical mechanics allows infinitely long time scale. However, in practice, we could classify variables needed to describe the system into two categories: (1) characteristic times are sufficiently longer than the observation time; (2) characteristic times are sufficiently shorter than the observation time. That this dichotomy is clear is the condition for equilibrium statistical mechanics to be useful in practice.

In this case we can change the mode of observation maintaining the dichotomy (*quasiequilibrium processes*). In this case it is expected that the time average over the observation time scale may be calculated accurately with an appropriate canonical distribution parameterized by the variables in category (1).

Evolution of probability distribution.

Since $P(x, t) = \langle \delta(x(t) - x) \rangle$, where the average is over the initial condition and the noise (i.e., ω), to derive the equation of motion for P , we have only to write down the time derivative of $\delta(x(t) - x)$. With the aid of Itô's lemma (26), for (36)

$$d\delta(x(t) - x) = -\frac{d}{dx}\delta(x(t) - x)dx + \frac{1}{2}(dx)^2\frac{d^2}{dx^2}\delta(x(t) - x), \quad (40)$$

$$= \frac{dt}{\gamma(x(t))} \frac{\partial U}{\partial x(t)} \frac{d}{dx}\delta(x(t) - x) + dt \frac{k_B T(x(t))}{\gamma(x(t))} \frac{d^2}{dx^2}\delta(x(t) - x) + \text{the term } \propto dB. \quad (41)$$

Therefore, after taking the expectation value, we obtain

$$dP(x, t) = dt \left[\frac{\partial}{\partial x} \left(\frac{1}{\gamma(x)} \frac{\partial U}{\partial x} P(x, t) - k_B T(x) \frac{\partial}{\partial x} \frac{1}{\gamma(x)} P(x, t) \right) + \frac{\partial^2}{\partial x^2} \frac{k_B T(x)}{\gamma(x)} P(x, t) \right]. \quad (42)$$

That is,

$$\frac{\partial}{\partial t} P(x, t) = \frac{\partial}{\partial x} \frac{1}{\gamma(x)} \left(\frac{\partial U}{\partial x} + \frac{\partial}{\partial x} k_B T(x) \right) P(x, t). \quad (43)$$

This type of equation is called the *Smolchowski equation*. We can of course derive an equation for the simultaneous distribution of x and p . This equation is called *Kramers equation* or *Fokker-Planck equation*. Generally speaking, all such equations are loosely called the Fokker-Planck equation.

Some general properties of (43)

(1) It is linear, so the superposition principle applies; in particular Green's function method may be used.

(2) The probability is conserved, because (43) has the following local conservation law form:

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x} J. \quad (44)$$

(3) *H-theorem*: The Kullback-Leibler entropy (\rightarrow p77)

$$K(P|P_{eq}) = \int dx P(x, t) \log \frac{P(x, t)}{P_{eq}(x)}, \quad (45)$$

where P_{eq} is the solution to $J = 0$, is a Lyapunov function for P_{eq} . Remark. P_{eq} has the

¹⁰This is essentially the Wong-Zakai theorem.

¹¹However, the last term is the standard stochastic differential, not the \circ product.

usual form of the Boltzmann factor, if T is uniform; if T is position dependent, P_{eq} depends on T and U nonlocally.

2. Macroscopic Thermodynamics

Thermodynamic laws of macroscopic systems

37

Several laws are known to hold universally for any macroscopic system in equilibrium.¹² Here, the word, ‘macroscopic’ and ‘equilibrium’ are primitive concepts of the logical system of thermodynamics. Conventionally, we have four laws. They are, informally:

O: (a) An isolated system¹³ left alone eventually reaches a state (= *equilibrium state*) in which no further time evolution is possible under the same isolated condition. (b) Thermal equilibrium relation is an equivalence relation.

I: The state of a system in equilibrium may be specified by its energy E and several extensive¹⁴ work coordinates that can be (operationally) defined directly in terms of mechanical and electromagnetic quantities. There is a process in which any change of E may be computed with the aid of energy conservation of electromechanical works W (called *adiabatic process*): $\Delta E = W$. The space spanned by E and the work coordinates is called the *thermodynamic state space*.

II: There are processes such that $Q \equiv \Delta E - W$ is not zero. Q is called *heat*. No process is possible, the sole result of which is that heat is imported from outside and work is done to it (the *Kelvin-Planck principle*).

IV: Thermodynamic quantities are extensive or intensive.

The above ordering is not logical, because the fourth law is much more fundamental than the first and second laws. It automatically excludes small systems and extremely large systems. A better informal exposition of equilibrium thermodynamics is:

Thermodynamic states may be characterized by extensive parameters alone (Oa+IV, i.e., there is a state space); one of them is energy. There are adiabatic processes that the energy change can be solely due to mechanical work (I), but there are also non-adiabatic processes. The deficit $\Delta E - W$ is called heat. No process is possible, the sole result of which is that heat is imported from outside and work is done to it (II). This allows us to define temperature and entropy. Thermal contact establishes thermal equilibrium; Thermal equilibrium relation is an equivalence relation (Ob). Entropy is bounded from below (III).

38

In the fluctuating world, however long we wait, the system never reaches any time-independent state.

Gibbs relation

Any process that can be represented by a continuous path in the state space is called a quasi-equilibrium process. If E is differentiable, we may write

$$dE = TdS + \sum X_i dx_i, \quad (46)$$

¹²‘Universal’ implies that there is an equivalence relation \sim among objects for whose corresponding equivalence class the universal properties are perfect: *sim* \iff identity of universal properties. Whether universality is useful or not depends on how small or large the equivalence class is.

¹³System = any part of the world; isolated system = a system that has no interaction with the rest of the world.

¹⁴a homogeneous function of degree one of the amount of matters .

39

where x_i are work coordinates. This is called the *Gibbs relation*. Here, the independent variables are S and x_i . Legendre transformation allows us to replace some of them with their conjugate variables T or X_i :

$$F = \inf_{\{x_k\}} \{E - \sum x_k X_k\}. \quad (47)$$

Is extensivity meaningful in the fluctuating world? If we consider an ensemble of non-interacting clones of the systems, then thermodynamics holds for the ensemble average.

Since dF is a complete its second order partial derivatives allow symmetry wrt the order of two partial differentiations. These are the *Maxwell relations*.

Since E is a homogeneous function of degree 1 of extensive variables $E = \sum x_i X_i + ST$, so $\sum x_i dX_i + S dt = 0$ (the *Gibbs-Duhem relation*).

42

The origin of the energy or entropy scale may be shifted by constant amounts without affecting the thermodynamic relations.

Free energy as effective potential

43

If a system is placed in a constant temperature environment at T , then $A = \inf_S \{E - TS\}$ depends only on the work coordinates, because

$$dA = -SdT + \sum X_i dx_i. \quad (48)$$

44

That is, we can change A ‘freely’ at our will.

The energy balance in this case can be observed in three ways. We consider the whole world consists of agent + system+ environment:

45

* Agent’s point of view: the external agent does work W and this energy is stored in the system’s potential energy as ΔA .

* System’s point of view: System get energy ΔE and $W - \Delta E = \Delta A - \Delta E$ is transferred to the environment.

* System+Environment’s point of view (the combined system is regarded adiabatic): This obtains total energy W from the agent and ΔE is stored in the system, and $-T\Delta S$ is given to the environment, where ΔS is the entropy increase of the environment.

If the reversible work W_{21} needed to change the system from state 1 and state 2 under constant temperature T is known as a function of T , then the required energy ΔE_{21} may be obtained (this is Gibbs-Helmholtz):

$$\Delta E = \frac{\partial W/T}{\partial 1/T} = W - T \frac{\partial W}{\partial T}. \quad (49)$$

Here, the partial derivatives keep the initial and terminal state work coordinates.

If we assume that the environment is always in equilibrium, then

$$T(\Delta S + \Delta S_{env}) \geq 0. \quad (50)$$

Since $\Delta E = W - \Delta S_{env}$,

$$W = \Delta E + \Delta S_{env} \leq \Delta E - T\Delta S = \Delta A. \quad (51)$$

That is, the work the agent must do cannot be smaller than ΔA . If we rewrite this as (the *minimum work principle*)

$$-W \geq -\Delta A, \quad (52)$$

this implies that the work the system can do cannot be larger than ΔA (the *maximum work principle*).

Open systems

47

An open system is a system that can exchange heat and particles with its environment. The energy balance reads

$$\Delta E = T\Delta S + W + \sum \mu_\alpha \Delta N_\alpha. \quad (53)$$

However, generally, chemical and thermal effects cannot be separated esp when we must pay attention to the particular nature of chemical substances.

The natural thermodynamic potential for open systems is

$$J = \inf_{S, N_\alpha} \{E - ST - \sum \mu_\alpha N_\alpha\}. \quad (54)$$

We can formulate the principle minimum or maximum work for open systems in terms of J just as before.

Mechanochemical system

49

Such a system consists of the reactant system, the product system and the motor. We assume that everything happens under constant pressure. The internal energy is replaced by enthalpy, and the Helmholtz free energy is replaced by the Gibbs free energy. If we write work other than due to volume change, the energy balance reads

$$\Delta H = W + Q. \quad (55)$$

If no work is done, $\Delta H = Q$ as is well known. Since $Q \leq \Delta S$,

$$W = \Delta H - Q \geq \Delta G. \quad (56)$$

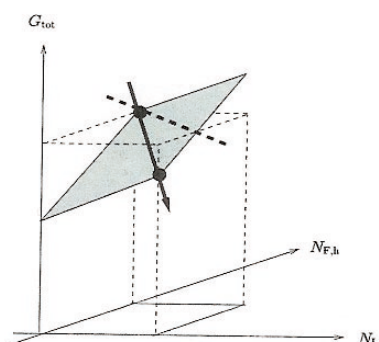
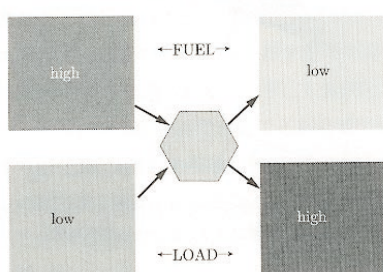
50

This is the principle of minimum work. If we do not take out the max work Notice that if we do not take out maximum work (under a quasiequilibrium condition), we have more heat to produce, but the heat produced by the process is $\delta H - \Delta G +$ this excess heat.

51

Conjugate chemical system

If the total change of the Gibbs free energy is negative, then the coupled reactions can proceed spontaneously. This way we could use the decrease of G due to one reaction to pump another chemical against the chemical potential gradient (active transport, ATP synthesis, etc).



Efficiency of engines

52 Here, the efficiency is defined by the ratio of the actual work and the maximal work.

55 If a system is not macroscopic, then we must pay due attention to the 'costs' required to connect/disconnect the system to various reservoirs. Also in general adiabatic processes are not conceivable in the usual sense of the word.

Enthalpy-entropy compensation

56 If two states of a system have similar G , i.e., $H_1 - H_2 \sim T(S_1 - S_2)$ but the difference of enthalpies (entropies) is significant, we say there is an *enthalpy-entropy compensation* (or H-S compensation) between these two states.

Notice that this is natural if the two states are distinguishable (as gas and liquid) but can coexist or can change reversibly. Thus, for protein conformations this is also natural. In this case the number of coordinated water molecules could change, so even the analogy with water phase transition may not be meaningless.

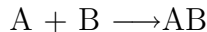
3. Chemical Reactions

62 Timescale: quantum chemistry \sim 1fs; protein conformation change: \sim 1ms-min.

Molecules are, strictly speaking, ephemeral beings, so they are nonequilibrium states of atom assemblies. However, they are quite different from glasses which are also nonequilibrium phenomena, because the dynamical time scale (or the time scale of the phenomena relevant to them) is much shorter than their lifespan.

Macroscopic reaction theory

Let us take



as an example.

$$\frac{d[A]}{dt} = k[A][B], \quad (57)$$

65 where k is called the *rate constant*. Here $[X]$ implies the number density of X.. If there is a reverse reaction, we write

$$\frac{d[AB]}{dt} = k[A][B] - k'[AB]. \quad (58)$$

A system is said to be closed wrt the reaction, if the system is closed wrt to A, B and AB molecules. In such a system in the long run there would be no time change (reach an equilibrium state),¹⁵ so

$$[AB]/[A][B] = k'/k. \quad (59)$$

66 This is called the *law of mass action*. This must be consistent with thermodynamics. There should not be any change of Gibbs free energy, so introducing the expression of the chemical potential for X

$$\mu_X = \mu_X^\circ + k_B T \log [X], \quad (60)$$

we have

$$\mu_{AB} - \mu_A - \mu_B = 0. \quad (61)$$

That is,

$$K \equiv k'/k = \exp[(\mu_{AB}^\circ - \mu_A^\circ - \mu_B^\circ)/k_B T]. \quad (62)$$

¹⁵That there is no time dependence does not imply equilibrium, if the system is not close.

K is called the *equilibrium constant*.

Discrete stochastic process

To describe chemical reaction at the mesoscopic level, we assume

- 73 (1) Transition states cannot be seen; reaction occurs instantaneously,
 (2) Markovian approximation: the future is completely determined by the state at present.
 Thus the change from state S_i to S_j is described by the transition probability $w_{i \rightarrow j}$. The probability flux $J_{i \rightarrow j}$ from S_i to S_j may be written as

$$J_{i \rightarrow j} = P_i w_{i \rightarrow j} - P_j w_{j \rightarrow i}. \quad (63)$$

74 The probability P_i of the state S_i obeys the *master equation*

$$\frac{dP_i}{dt} = - \sum_j J_{i \rightarrow j}. \quad (64)$$

If P_i are time independent, we say the system is in a steady state; if $J_{i \rightarrow j} = 0$ for all pairs i, j , we say the system is in a *detailed balanced state*.¹⁶

- 75 In equilibrium the system must be in a detailed balanced state; if not, there must be a circular balance somewhere, which may be used to drive an engine that violates the second law. Thus, using $P_i w_{i \rightarrow j} = P_j w_{j \rightarrow i}$ one by one, we can determine the equilibrium P_i .

For the detailed balance to hold, $w_{i \rightarrow j}$ must satisfy stringent conditions. Let us write $\log w_{i \rightarrow j} = Q_{ij}$ and $\log P_i = e_i$. Then,

$$Q_{ij} + e_i = Q_{ji} + e_j. \quad (65)$$

If we introduce $\Delta_{ij} = Q_{ij} + e_i$, then this must be symmetric. This is the constraint for these set of equations to have a set of solutions. Therefore, a necessary and sufficient condition for the reaction system described by the master equation to have an equilibrium state is that the transition probability may be written in terms of a symmetric matrix Δ_{ij} and a set of numbers $\{F_i\}$ and an overall constant $1/\tau$ as

$$w_{i \rightarrow j} = \frac{1}{\tau} e^{-\beta(\Delta_{ij} - F_i)}. \quad (66)$$

We can write

$$P_i = Z^{-1} e^{-\beta F_i}. \quad (67)$$

- 76 Example: 1D nearest neighbor hopping: states $\{S_{-n}, \dots, S_m\}$ ($n, m > 0$). The rate is nonzero between adjacent states such that $\Delta_{j,j+1} = \Delta_L$ for $-n \leq j < 0$ and $\Delta_{j,j+1} = \Delta_H$ for $0 \leq j < m$. Even if $\Delta_L \neq \Delta_H$, $P_i = \text{const}$. Explain why. \square

Nonequilibrium steady state

If the condition (66) is not satisfied, the steady state of the Markov process corresponds to a nonequilibrium steady state.

- 77 The state is characterized by the minimum of the Kullback-Leibler entropy (wrt the steady state) See below.

¹⁶'State' means about the ensemble, of course, since we are discussing probabilities.

If P_i obeys the above master equations, since it is linear $P_i(t + dt)$ may be written a linear combination of $\{P_i(t)\}$:

$$P_i(t + dt) = \sum_j \Pi_{ij} P_j(t). \quad (68)$$

This is a Markov chain (Π is a stochastic matrix). Let P and Q be stochastic vectors of the same dimension.

$$K(P, Q) = \sum_i P_i \log(P_i/Q_i) \quad (69)$$

is called the *Kullback-Leibler entropy*, and is positive definite: since $x \log x$ is convex,

$$\sum_i P_i \log(P_i/Q_i) = \sum_i Q_i (P_i/Q_i) \log(P_i/Q_i) \geq \left[\sum_i Q_i (P_i/Q_i) \right] \log \left[\sum_i Q_i (P_i/Q_i) \right] = 0. \quad (70)$$

For any stochastic matrix Π , $K(P, Q) \leq K(\Pi P, \Pi Q)$. That is, if we assume $\Pi Q = Q$, then it is a Lyapunov function for the Markov chain defined by Π . This can be shown as, with the aid of $\sum_i \Pi_{ij} = 1$,

$$K(P, Q) - K(\Pi P, \Pi Q) = \sum_j P_j \log \frac{P_j}{Q_j} - \sum_i \left(\sum_j \Pi_{ij} P_j \log \frac{\sum_j \Pi_{ij} P_j}{\sum_j \Pi_{ij} Q_j} \right), \quad (71)$$

$$= \sum_{ij} \Pi_{ij} P_j \left(\log \frac{\Pi_{ij} P_j}{\Pi_{ij} Q_j} - \log \frac{\sum_j \Pi_{ij} P_j}{\sum_j \Pi_{ij} Q_j} \right), \quad (72)$$

$$= \sum_i \left(\sum_j \Pi_{ij} P_j \right) \left[\sum_j \frac{\Pi_{ij} P_j}{\sum_j \Pi_{ij} P_j} \left(\log \frac{\Pi_{ij} P_j}{\Pi_{ij} Q_j} - \log \frac{\sum_j \Pi_{ij} P_j}{\sum_j \Pi_{ij} Q_j} \right) \right]. \quad (73)$$

78 The quantity in [] is just the Kullback-Leibler entropy $K(\Pi_{ij} P_j / \sum_j \Pi_{ij} P_j, \Pi_{ij} Q_j / \sum_j \Pi_{ij} Q_j) \geq 0$.

Thus, we have demonstrated that $K(P, P_e)$, where P_e is the stationary state of the markov chain, is a Lyapunov function for the fixed point P_e . The relation between K and nonequilibrium thermodynamics was extensively studied by Schlögl in the 70s.

79 Notice that the sample paths of any continuous Markov process can be realized (almost surely) as solutions of a SDE. For a Langevin equation, the transition rate may be defined as

$$\langle \delta(x - x(t)) \delta(x' - x(t + dt)) \rangle = P_e(x) [\delta(x - x') + w(x \rightarrow x') dt] + o[dt], \quad (74)$$

where the average is over the initial equilibrium condition at time t and ω . If the detailed balance

$$P_e(x) w(x \rightarrow x') = P_e(x') w(x' \rightarrow x) \quad (75)$$

holds, then

$$\langle \delta(x - x(t)) \delta(x' - x(t + dt)) \rangle = \langle \delta(x' - x(t)) \delta(x - x(t + dt)) \rangle. \quad (76)$$

That is, there is no absolute direction of time. This is natural in equilibrium. A condition for this to hold is the fluctuation-dissipation relation (\rightarrow p9 or the Einstein relation \rightarrow p10).¹⁷

Stochastic description of chemical reactions

Let us take the example $A + B \rightarrow AB$. $k[A][B]$ is the rate of production of AB per unit volume. If N_X is the number of X in the unit volume we are considering, then $[X]$ should be replaced with N_X . If we assume that each reaction is N_X stochastically independent, then $kN_A N_B$ must be the transition rate. Therefore,

$$w(N_{AB} \rightarrow N_{AB} + 1) = kN_A N_B \quad (77)$$

k must contain sort of an activation factor. Since during the time dt/dP one reaction occurs, so this must be equal to $1/w(N_{AB} \rightarrow N_{AB} + 1)$. Therefore, on the average, we recover the macroequation:

$$\frac{dN_{AB}}{dt} \times \frac{dt}{dP} = 1 \Rightarrow \frac{dN_{AB}}{dt} = k[A][B]. \quad (78)$$

The reverse reaction has

$$w(N_{AB} + 1 \rightarrow N_{AB}) = k'(N_{AB} + 1). \quad (79)$$

Here, the master equation, etc., are written, it is better to utilize explicitly the Poisson process, and the corresponding stochastic differential equations. This will be rewritten later.

¹⁷ADD MORE HERE.