Chapter 2
Typicality and Ergodicity

2.1 Typicality and Ergodicity in Classical Statistical Mechanics

The PEPP is the postulate that all the possible microscopic states appear in the same probability. Since all the probabilities for the states are severely restricted by this postulate, it is counter-intuitive and thus many people have tried to justify it microscopically. In many textbooks, ergodicity of the state is often referred for the justification of PEPP. However, it is not physically realistic and cannot be the justification, although ergodic theorem itself has brought fruitful mathematics. Instead of the ergodicity, I introduce the typicality for the interpretation of PEPP. In this section, I review these points in a classical setup.

2.1.1 Ergodicity and Its Problems

When we employ a microcanonical ensemble in statistical mechanics, an equilibrium value is obtained by microcanonical ensemble average,

$$
\langle A \rangle_{mc} \equiv \frac{\int_{E-\Delta E < E} A (\{p_i\}, \{q_i\}) \, d\Gamma}{W} \quad (2.1)
$$

where $(\{p_i\}, \{q_i\})$ is a set of positions $p_i$ and momentums $q_i$ for all particles ($i$ is an index of particles), $A (\{p_i\}, \{q_i\})$ is a value of a physical variable $A$ for $(\{p_i\}, \{q_i\})$, $E$ is energy, $(E-\Delta E, E]$ is an energy shell, $d\Gamma \equiv \Pi_i d\pi_i d\eta_i$, and $W$ is the number of states

$$
W \equiv \int_{E-\Delta E < E} d\Gamma. \quad (2.2)
$$
The averaged value over time is

$$\langle A \rangle_{\text{time}} \equiv \lim_{T \to \infty} \frac{1}{T} \int_0^T A \left( \{ p_i(t) \}_i, \{ q_i(t) \}_i \right) \, dt \quad (2.3)$$

where \( \{ p_i(t) \}_i, \{ q_i(t) \}_i \) is a set of positions and momentums at time \( t \). Ergodicity is the property that the microcanonical ensemble average value of every physical variable \( A \) is equal to the time average value of it, i.e., for \( \forall A \),

$$\langle A \rangle_{\text{mc}} = \langle A \rangle_{\text{time}}. \quad (2.4)$$

In Newtonian mechanics, the ergodicity is proved for some models. However, even if Eq. (2.4) holds, the time for the state \( \{ p_i(t) \}_i, \{ q_i(t) \}_i \) to evolve over all the possible states \( T_{\text{evolve}} \) is usually too long to measure. \( T_{\text{evolve}} \) easily exceeds the lifetime of the universe. Namely, we cannot wait for \( T_{\text{evolve}} \) coming, even if the ergodicity holds. On the other hand, we do not need to measure the physical quantities for a long time \( \sim T_{\text{evolve}} \) to know the equilibrium values in reality. For example, when we measure the temperature of a glass of water for a minute, we can obtain the equilibrium value of it. Therefore, the ergodicity will not be the reason why the ensemble average gives the correct equilibrium value. In addition to this, \( T_{\text{evolve}} \) increases exponentially as the size of the system increases. In contrast to this, statistical mechanics is an asymptotic theory whose prediction gets closer to thermodynamics as the system size increases. Hence, the ergodicity seems to have nothing to do with the foundation of statistical mechanics.

### 2.1.2 Typicality in Weak Sense

In Sect. 2.1.1, we saw that ergodicity cannot be used for the basis of statistical mechanics. Alternatively, we can easily prove that “weak typicality” among states can be substituted for the PEPP.

The weak typicality is stated as follows: Let a macroscopic physical variable \( A \) be extensive, that is, when a system gets \( x \)-times larger, a value of \( A \) also gets \( x \)-times larger. For example, energy, magnetization, number of particle are the extensive variables. I say that the weak typicality for \( A \) holds if \( \forall \varepsilon > 0 \), \( ^3 \)system size \( V \), such that

$$\frac{\sqrt{\Delta A^2}}{V} < \varepsilon \quad (2.5)$$

where

$$\Delta A^2 \equiv \int \frac{A(\{ p_i(t) \}_i, \{ q_i(t) \}_i) - \langle A \rangle_{\text{ens}}^2}{W} \, dW. \quad (2.6)$$
Since the number of particles increases linearly as $V$ increases, $\sqrt{\Delta A^2}$ also increases as $V$ does. However, the weak typicality postulates that $\sqrt{\Delta A^2} = o(V)$. It means that if I prepare a state for sufficiently large $V$ randomly among all possible states,

$$a((p_i(t))_i, (q_i(t))_i) \equiv \frac{A((p_i(t))_i, (q_i(t))_i)}{V}$$

rarely deviates from its ensemble average. Therefore, I do not need to assume the PEPP to obtain the thermodynamic predictions from Newtonian mechanics. Since the PEPP restricts all the realizing probability of the possible states, it is a stronger assumption than the weak typicality. If I only assume the weak typicality instead of PEPP, $a((p_i(t))_i, (q_i(t))_i)$ takes the value which is equal to the ensemble average within a negligible error.

Inversely, I can also show that the weak typicality is proved from PEPP. I assume that PEPP, the Boltzmann’s entropy formula, and the convexity of entropy, i.e.,

$$\frac{\partial^2 S(E, A, V)}{\partial A^2} \leq 0,$$

where $S(E, A)$ is entropy and I omit variables except for energy $E$, some physical quantity $A$, and the volume $V$. By the definition of the Boltzmann’s formula, I get

$$W(E, A, V) = \exp[S(E, A, V)] = \exp[V \{s(u, a; V)\}]$$

where $s(u, a; V) = S(E, A, V)/V$, $u \equiv E/V$, and $a \equiv A/V$. $s(E, A; V)$ take the maximum at $a^*$ where

$$\frac{\partial s(u, a; V)}{\partial a} \bigg|_{a=a^*} = 0.$$

Expanding the rhs of Eq. (2.10) around $a^*$, I get

$$W(E, A, V) = \exp[V \{s(u, a^*; V) + s''(u, a^*; V)(a - a^*)^2 + O((a - a^*)^3)\}]$$

where

$$s''(u, a^*; V) \equiv \frac{\partial^2 s(u, a; V)}{\partial a^2} \bigg|_{a=a^*} < 0$$

and $s''(u, a^*; V)$ is $\Theta(V^0)$. Therefore, $W(E, A, V)$ behaves like a Gaussian distribution with a peak at

$$a = a^* \quad (\Leftrightarrow A = V a^*)$$

and the width

$$\Delta a^2 = \Theta\left(\frac{1}{V}\right) \quad (\Leftrightarrow \Delta A^2 = \Theta(V)).$$
When one specifies the value of energy but does not specify that of \( a \), \( a \) may take various values. However, since \( \lim_{V \to \infty} \Delta a^2 = 0 \) in Eq. (2.15), not only the ensemble average takes \( a^* \) but also the values of \( a \) of almost all the possible states take \( a^* \) with errors vanishing to zero with increasing \( V \). This is the proof of the weak typicality for \( a \) derived from the above three assumptions, PEPP, the Boltzmann’s entropy formula, and the convexity of entropy. Since the Boltzmann’s entropy formula and the convexity of entropy are always assumed in a usual situation, the PEPP automatically implies the weak typicality.

In summary, the weak typicality is sufficient for obtaining thermodynamic predictions. Although the PEPP is the sufficient condition for the weak typicality, it is not the necessary condition for it. Therefore, one may expect that the weak typicality is more fundamental than the PEPP in statistical mechanics. However, this discussion of the weak typicality focuses only on the extensive variables. We cannot apply the above discussion to other variables; for example, fluctuations, one-particle observables like its position, and intensive variables are excluded. I do not know whether this discussion is extended to these variables. Thus, the justification of statistical mechanics using the weak typicality in this form is not completed.

### 2.2 Typicality and Ergodicity in Quantum Statistical Mechanics

In Sect. 2.1, I have explained that neither ergodicity nor PEPP is necessary to obtain thermodynamic predictions of some extensive variables in classical system, but the weak typicality is sufficient. For quantum systems, I will show that ergodicity generally breaks down [1–5], but typicality is proved in a stronger sense than what I have explained in classical statistical mechanics [6–11]. I call it “strong typicality.” The strong typicality is stated as follows: First, the expectation value \( \langle \psi | \hat{A} | \psi \rangle \) for any “mechanical variables” (defined in this section) is close to the microcanonical ensemble average under a natural probability measure of the coefficients \( \{c_n\}_n \). Second, the difference between \( \langle \psi | \hat{A} | \psi \rangle \) and the ensemble average is of the order \( \exp[-\Theta(-V)] \), which immediately vanishes to zero as \( V \) increases. It will turn out that the strong typicality can substitute for the PEPP as a principle of statistical mechanics.

#### 2.2.1 Ergodicity in Quantum Mechanics

Let me consider an arbitrary quantum pure state \( |\psi_{erg}\rangle \) in an energy shell \( (E - \delta E, E] \). It can be written in a general form,

\[
|\psi_{erg}\rangle = \sum_{n \in \mathcal{S}} c_n |n\rangle. \quad \mathcal{S} = \{n | u_n \in (E - \delta E, E)\}
\]  

(2.16)
where \( \{c_n\}_n \) is a set of coefficients and \( |n\rangle \) is an energy eigenstate, \( \hat{H}|n\rangle = u_n|n\rangle \). The time evolution of \( |\psi\rangle \) is

\[
|\psi_{\text{erg}}(t)\rangle \equiv \sum_n c_n e^{-i\frac{\hbar}{\epsilon}u_n t}|n\rangle,
\]

(2.17)

and the expectation value of \( |\psi(t)\rangle \) for an operator \( \hat{A} \) is

\[
\langle \psi_{\text{erg}}(t)|\hat{A}|\psi_{\text{erg}}(t)\rangle = \sum_{n,m} c_n^* c_m e^{i\frac{\hbar}{\epsilon}(u_n - u_m)t} \langle n|\hat{A}|m\rangle.
\]

(2.18)

Then, the time average of the expectation value is

\[
\langle \hat{A} \rangle_{\text{time}} \equiv \lim_{T \to \infty} \frac{1}{T} \int_0^T \langle \psi_{\text{erg}}(t)|\hat{A}|\psi_{\text{erg}}(t)\rangle dt
\]

\[
= \sum_{n,m} c_n^* c_m \langle n|\hat{A}|m\rangle \lim_{T \to \infty} \frac{1}{T} \int_0^T e^{i\frac{\hbar}{\epsilon}(u_n - u_m)t} dt
\]

\[
= \sum_{n,m} c_n^* c_m \langle n|\hat{A}|m\rangle \delta_{n,m}
\]

(2.19)

\[
= \sum_n |c_n|^2 \langle n|\hat{A}|n\rangle
\]

(2.20)

On the other hand, the microcanonical ensemble average is

\[
\langle \hat{A} \rangle_{\text{mc}} \equiv \frac{1}{d} \sum_n \langle n|\hat{A}|n\rangle.
\]

(2.21)

where \( d = \text{dim}\mathcal{H} \). Therefore,

\[
\langle \hat{A} \rangle_{\text{time}} \neq \langle \hat{A} \rangle_{\text{mc}},
\]

(2.22)

holds for all observable \( \hat{A} \), if and only if

\[
|c_n|^2 = \frac{1}{d}
\]

(2.23)

for \( \forall n \). Except for the case of Eq. (2.24), the ergodicity does not hold in quantum statistical mechanics. As I have emphasized, ergodicity is too strict and unnecessary statement. The failure of the ergodicity is because the ergodicity tries to prove Eq. (2.23) for all observables, which include arbitrary \( N \)-body correlations. For the foundation of statistical mechanics, therefore, we have to think seriously what kind of operators we look at. This leads to the understanding of the strong typicality.
### 2.2.2 Strong Typicality in Quantum System

We examine the typicality in quantum systems. The discussion in Sect. 2.2.2 mostly follows Refs. [8]. We consider a state \(|\psi_E\rangle\) in an energy shell;

\[ |\psi_E\rangle \equiv \sum_{n \in S} c_n |n\rangle, \quad S = \{ n \mid u_n \in (E - \delta E, E) \} \quad (2.25) \]

Here, \(\{c_n\}_n\) is a set of coefficients labeled by \(n\), \(|n\rangle\) and \(u_n\) are energy eigenstates and corresponding energy eigenvalues, respectively. Since the question “Is it typical?” depends on a measure. We introduce the probability measure \(P(\{c_n\}_n)\) into \(\{c_n\}_n\).

Although \(|\psi_E\rangle\) looks the same as \(|\psi_{\text{erg}}\rangle\), \(\{c_n\}_n\) in \(|\psi_E\rangle\) is random variables in this case. In order to make the measure be natural for quantum systems, we impose two conditions to \(P(\{c_n\}_n)\): (1) Invariance—The measure should be invariant under arbitrary unitary transformations. (2) Normalization—\(\{c_n\}_n\) should be normalized; \(\sum_i |c_n|^2 = 1\). Therefore, \(P(\{c_n\}_n)\) should be

\[ P(\{c_n\}_n) = \frac{\Gamma(d - 1/2)}{2\pi^{d-1}} \delta \left( \sum_n |c_n|^2 - 1 \right), \quad (2.26) \]

where \(d = \text{dim} S\).

Under this measure [13], the variance of \(\langle \psi_E | \hat{A} | \psi_E \rangle - \langle \hat{A} \rangle_{\text{MC}}\) is calculated as

\[ \left( \langle \psi_E | \hat{A} | \psi_E \rangle - \langle \hat{A} \rangle_{\text{MC}} \right)^2 = \frac{\langle \Delta \hat{A}^2 \rangle_{\text{MC}} + \langle \hat{A} \rangle_{\text{MC}}^2}{d + 1}, \quad (2.27) \]

where \(\cdots\) denotes the random average of \(\cdots\), and \(\langle \Delta \hat{A}^2 \rangle_{\text{MC}} \equiv \langle (\hat{A} - \langle \hat{A} \rangle_{\text{MC}})^2 \rangle_{\text{MC}}\). In Eq. (2.27), the denominator of the rhs is \(\exp[\Theta(V)]\), and the numerator is the order of polynomial of \(V\) for any mechanical variable (defined in Sect. 2.2.3). Hence,

\[ \text{the rhs of Eq. (2.27) = } \exp[-\Theta(V)]. \quad (2.28) \]

This result means that almost all \(|\psi_E\rangle\)'s give correct results which are close to the microcanonical ensemble average. In this sense, the typicality holds. Moreover, the error shown in Eq. (2.27) is exponentially small as \(V\) increases. This error is much smaller than \(\Theta(V)\), which is the error of the weak typicality for classical systems in Sect. 2.1.2. Hence, I call this typicality a “strong typicality.” The strong typicality is the most important property for statistical mechanics based on pure quantum states.

I note that the measure given in Eq. (2.26) is simple and natural but not necessary. I expect that the typicality will hold for other reasonable probability measures of \(\{c_n\}_n\). I also note a possible interpretation of \(|\psi_E\rangle\). In the ensemble formulation of quantum statistical mechanics, equilibrium state corresponding to microcanonical ensemble is
\[
\hat{\rho}_{MC} \equiv \frac{1}{d} \sum_{n \in \mathcal{S}} |n\rangle\langle n|.
\]  

(2.29)

Then, I can show that

\[
\hat{\rho}_{MC} = |\psi_E\rangle\langle \psi_E|.
\]  

(2.30)

Hence, \(|\psi_E\rangle\) can be interpreted as a single realization of a pure quantum state from \(\hat{\rho}_{MC}\) of the form of Eq. (2.30).

### 2.2.3 Mechanical Variables and Genuine Thermodynamic Variables

In a macroscopic system, we do not measure all of the possible observable, but measure several macroscopic observables such as magnetization and correlation functions. I call these macroscopic observables as “mechanical variables (MVs)” and define them in a rigorous way [12].

**Def.**) Mechanical Variables

A mechanical variable is a quantum-mechanical observable which is low-degree polynomials of local operators and satisfies the condition that

\[
\langle \hat{A}^2 \rangle_{\text{ens}}^{\beta, V} \leq K(\beta) V^{2m} \text{ for all } \beta, V.
\]  

(2.31)

Here, a function \(K(\beta)\) and a constant \(m\) are positive and independent of \(\hat{A}\) and \(V\). I make \(\hat{A}\) dimensionless by dividing an appropriate unit. Since I have fixed \(m\), the number of independent MVs \(N_{\text{MV}}\) is \(\Theta(V^m)\).

Returning to the main result, we can rewrite Eq. (2.27) by using Markov’s inequality as

\[
P \left( \left| \langle \psi_E | \hat{A} | \psi_E \rangle - \langle \hat{A} \rangle_{\text{MC}} \right| \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} \frac{\langle \Delta \hat{A}^2 \rangle_{\text{MC}} + \langle \hat{A} \rangle_{\text{MC}}^2}{d + 1},
\]  

(2.32)

where \(P(\cdots)\) denotes the probability of event \(\cdots\). Therefore, the probability that all expectation values of the MVs for \(|\psi_E\rangle\) do not deviate from the corresponding ensemble averages is estimated as follows:

\[
P \left( \sum_{\hat{A} \in \text{MV}} \left| \langle \psi_E | \hat{A} | \psi_E \rangle - \langle \hat{A} \rangle_{\text{MC}} \right| \geq \varepsilon \right) \leq \frac{N_{\text{MV}}^2}{\varepsilon^2} \frac{1}{d + 1} \max_{\hat{A} \in \text{MV}} \left( \langle \Delta \hat{A}^2 \rangle_{\text{MC}} + \langle \hat{A} \rangle_{\text{MC}}^2 \right).
\]  

(2.33)
Since \( N_{\text{MV}} \) is \( \Theta(V^m) \), the rhs is still \( \exp[\Theta(-V)] \). Hence, a single realization of \( |\psi_E\rangle \) for sufficiently large \( V \) correctly gives all equilibrium values of the MVs simultaneously.

The number of mechanical variables \( N_{\text{MV}} \) is exponentially smaller than \( d \), which is the dimension of the Hilbert space in the energy shell. Thus, even when we measure all the MVs, we cannot identify a pure quantum state but almost all \( |\psi_E\rangle \) return the same value for all the MVs.

However, entropy and temperature are not included in this discussion, because they cannot be represented as a MV. If one stick to represent these variables as quantum mechanical observables, they become \( N \)-body operators in general, while low-degree polynomials of local operators are the only MVs. In contrast to quantum mechanics, any equilibrium values for macroscopic variables including entropy and temperature are obtained from any of the thermodynamic functions in thermodynamics. I call these macroscopic variables except for the MVs “genuine thermodynamic variables (GTVs)” [12].

**Def.** Genuine Thermodynamic Variables

Genuine thermodynamic variables are observables which are not MVs but can be derived from entropy.

For example, (inverse) temperature is the derivative of entropy with respect to energy. Free energy is obtained by Legendre transformation of entropy. Hence, they are GTVs.

### 2.2.4 Canonical Typicality

Using the results in Sect. 2.2.2, it is revealed that the subsystem of the \( |\psi_E\rangle \) is almost identical to the Gibbs state in that subsystem [6, 7, 14–18]. Let the total system be a composite of system (\( S \)) and environment (\( E \)), and the interaction between \( S \) and \( E \) be negligible. The size of the system is \( v \), and the size of the environment is \( V \). Here, we consider the case \( v \ll V \). We assume the Hamiltonian

\[
\hat{H} = \hat{H}_S + \hat{H}_E,
\]

where the support of \( \hat{H}_S \) is on \( S \) and that of \( \hat{H}_E \) is on \( E \). We prepare \( |\psi_E\rangle \) of the total system. The reduced density matrix of \( |\psi_E\rangle \) in \( S \) is

\[
\hat{\rho}_S = \text{Tr}_B[|\psi_E\rangle\langle\psi_E|].
\]

Next, let \( \hat{A}_S \) be an arbitrary observable in \( S \). \( N_S \), which is the number of \( \hat{A}_S \), scales as \( O(v^\nu) \) while the total dimension of Hilbert space \( d \) in Eq. (2.33) scales as \( O(V^\nu) \). Here, \( \nu \) is a constant, e.g., \( \nu = 2 \) for the spin-1/2 model and \( \nu = 4 \) for the Hubbard model. Hence, using Inequality (2.33), we get
Since the rhs is \(\exp[\Theta(-V)]\), we show that \(|\psi_E\rangle\) gives equilibrium value for all the observable in \(S\). I note that \(\hat{A}_S\) includes \(v\)-body correlations, which is a microscopic quantities for the system \(S\). Since \(v \ll V\), the \(v\)-body correlations are low-degree polynomial for the total system. Therefore, \(|\psi_E\rangle\), which is the pure state of the total system can give all equilibrium value for observables in \(S\).

When two quantum states give the same expectation value for all the possible observable in the subsystem \(S\), these state are identical in \(S\). Namely, \(\hat{\rho}_S \simeq e^{-\beta \hat{H}_S} / \text{Tr}[e^{-\beta \hat{H}_S}]\) where \(\beta\) is temperature such that

\[
\langle \hat{H}_E \rangle^{\text{can}}_{\beta, V} = E. 
\]

The expectation value of any observable \(\hat{A}\) in \(S\) gives

\[
\langle \psi_E | \hat{A} | \psi_E \rangle = \langle \hat{A} \rangle^{\text{can}}_{\beta, V}
\]

with the exponentially small error. This is so-called “canonical typicality” [7].

### 2.3 Summary and Discussions of this Chapter

In this section, we saw that the ergodicity is physically nonsense for the justification of the PEPP. Moreover, it breaks down in quantum system. Thus, we saw the typicality among states. Even in classical system, we can show the weak typicality, which is enough to obtain thermodynamic predictions for some extensive variables. However, it may not be sufficient for the foundation of statistical mechanics. In quantum system, the strong typicality was shown. Owning to the exponentially large Hilbert space, all equilibrium values of the MVs, which include even a position of a particle and higher order fluctuation, can be predicted from a single pure state in the strong typicality. The strong typicality can be another principle of statistical mechanics by substituting the PEPP. It will be also expected that the exponentially small error enable us to apply such a pure state to practical applications.

However, the GTVs such as entropy are excluded from the MVs. From any of the thermodynamic functions, all the thermodynamic predictions are derived. Since the genuine thermodynamic variables are related closely to the number of states through Boltzmann entropy formula, it may seem to be difficult to obtain them from a single pure quantum state. In addition to this, we do not know whether there exist pure quantum states which are specified by intensive variables such as temperature and chemical potential. For practical applications, it is also preferred that such pure quantum states can be generated easily, but we do not know how. I will answer these questions positively from now on. Here, I summarize the important points in this chapter:
Lesson:
As far as we look at MVs, we can neither distinguish different realizations of $|\psi_E\rangle$'s, nor do $|\psi_E\rangle$ from $\hat{\rho}_{MC}$.

Finding Of Previous Works:
The probabilistic error of $|\psi_E\rangle$ is $\exp[\Theta(-V)]$, which is much smaller than that of the weak typicality, $O\left(\frac{1}{\sqrt{V}}\right)$.

Questions:
Can we obtain the genuine thermodynamic variables from a single realization of a pure quantum state?
By specifying the intensive variables, can we generate pure states which are regarded as equilibrium states?
How to construct these states?

References

5. Z.X. Gong, L.-M. Duan, arXiv:1109.4696
13. N. Ullah, Nucl. Phys. 58, 65 (1964)
Formulation of Statistical Mechanics Based on Thermal Pure Quantum States
Sugiura, S.
2017, XII, 73 p. 4 illus. in color., Hardcover