Game-Theoretic Approach to Thermodynamics

Ken Hiura

Abstract

To understand the fluctuating world, we introduce probabilistic assumptions on microscopic systems and study their statistical properties. It is known that for macroscopic systems in equilibrium and in the relaxation to equilibrium, the principle of equal a priori probabilities is useful to understand their equilibrium fluctuations and irreversible relaxation behavior. In addition, a statistical thermodynamical theory, called stochastic thermodynamics, for mesoscopic systems that are weakly coupled to equilibrium environments has been developed over the past decades. In recent years, it was found that a game-theoretic approach to stochastic thermodynamics based on martingale theory sheds light on novel properties in non-equilibrium systems including the stopping times and extreme value statistics. In this thesis, inspired by thesis works, we investigate problems in statistical thermodynamics from the viewpoint of martingale theory.

Martingale theory is useful for studying two properties. First, as suggested by previous works, martingale theory provides a useful tool to examine stopping time statistics through several mathematical properties, e.g., the information inequality at stopping times. In Chapter 2, by using this property, we derive a fundamental bound on first passage times for accumulated currents in Markov jump processes. Second, the martingale processes are used to characterize the randomness of the microscopic realizations through the condition that we cannot make much money in a fair betting game. In Chapter 3, via a simple dynamical model, we argue that random initial microstates characterized by martingale processes exhibit irreversible behavior macroscopically. Moreover, in Chapter 4, we consider a repeated work extraction from a small heat engine and find that the accumulation of the extracted work is a martingale process. By using martingale theory and information-theoretic techniques, we show that the second law of thermodynamics in this situation leads to the Gibbs distribution.

Publication List

This thesis is based on the following three publications.

[1] Ken Hiura and Shin-ichi Sasa, Microscopic Reversibility and Macroscopic Irreversibility: From the Viewpoint of Algorithmic Randomness, J. Stat. Phys. **177**, 725 (2019).

[2] Ken Hiura and Shin-ichi Sasa, Kinetic uncertainty relation on first passage time for accumulated current, Phys. Rev. E **103**, L050103 (2021).

[3] Ken Hiura, Gibbs Distribution From Sequentially Predictive Form of the Second Law, J. Stat. Phys. **185**, 2 (2021)

The following two publications are related to the thesis but are not included.

[4] Ken Hiura and Shin-ichi Sasa, How Does Pressure Fluctuate in Equilibrium?, J. Stat. Phys. **173**, 285 (2018).

[4] Ken Hiura, Macroscopic Dynamical Fluctuations in Kac Ring Model, J. Stat. Phys. **186** (2022).

Contents

Chapter 1

Introduction

1.1 Probability in Physics

1.1.1 Statistical Mechanics for Large Systems

One of the most important achievements of probability theory in physics is statistical mechanics for macroscopic equilibrium systems. The purpose of equilibrium statistical mechanics is to connect the microscopic dynamical properties of the large system with its macroscopic behavior in equilibrium through probabilistic considerations. A fundamental assumption of equilibrium statistical mechanics is the principle of equal a priori probabilities: Consider a system with a microscopic Hamiltonian H_N, where the subscript N denotes the system size. The statistical properties of the equilibrium state specified by its energy E are then described by the microcanonical measure,

$$\mu_{N;E}(d\omega) = \frac{1}{W_{N;E}}\delta(E - H_N(\omega))d\omega, \tag{1.1.1}$$

where ω is a microscopic state of the system and $W_{N;E} = \int \delta(E - H_N(\omega))d\omega$ is the normalization factor. This probabilistic assumption allows us to obtain two important information on the equilibrium state:

- First, by considering the consistency of the microcanonical form (1.1.1) with the variational principle in thermodynamics, we can identify the logarithm of the normalization factor $\ln W_{N;E}$ with the entropy function up to the additive constant: $S_N(E) = k_{\mathrm{B}} \ln W_{N;E} + (\text{const.})$. Since the entropy function completely determines the most probable behavior of the equilibrium state, we can obtain all the typical macroscopic properties from the microscopic Hamiltonian through this relationship, called the Boltzmann formula.

- Second, the probability distribution (1.1.1) provides useful expressions for macroscopic fluctuations. As an example, we consider an isolated system with the total energy E_{tot}. We suppose that this system is separated into the left-hand side and the right-hand side

by a diathermal wall and the position of the wall is fixed. By applying the principle of equal a priori probabilities to this composite system, we find that the probability that the energy of the left-hand side E deviates from the equilibrium value Ne^* tends to zero in the thermodynamic limit: For any $\epsilon > 0$,

$$\lim_{N\to\infty} \mu_{N;E_{\rm tot}} \left(\left| \frac{E}{N} - e^* \right| \geq \epsilon \right) = 0. \tag{1.1.2}$$

Here, the equilibrium value e^* is determined by the variational formula: $e^* := \mathrm{argmax}_e s(e)$ for $s(e) := \lim_{N\to\infty}(S_N(E) + S_N(E_{\rm tot} - E))/N$. The law of large numbers (1.1.2) guarantees the reproducibility of thermodynamic observations, i.e., the outcomes in the instantaneous measurements of the energy for different samples give the same result with high probability. Moreover, we can obtain more detailed information on the deviation probability from the principle of equal a priori probabilities. The deviation probability decays exponentially in the system size N and the decay rate is expressed in terms of the entropy function:

$$\mu_{N;E_{\rm tot}} \left(\frac{E}{N} \approx e \right) \asymp \exp\left[N k_{\rm B}^{-1} (s(e) - s(e^*)) \right]. \tag{1.1.3}$$

Thus, the entropy function characterizes instantaneous fluctuations for macroscopic observables as well as the most probable values. The expression (1.1.3) was found by Einstein, and it is regarded as the large deviation principle for static fluctuations of equilibrium states.

As well as equilibrium states, the idea of the principle of equal a priori probabilities is useful when we consider the irreversibility paradox. Let us imagine a macroscopic simple fluid in an adiabatic container. According to the zeroth law of thermodynamics, any thermodynamically isolated macroscopic system reaches a macroscopically stationary state, called the equilibrium state, after a sufficiently long time. In this relaxation process, the spatiotemporal change of coarse-grained density fields of conserved quantities is believed to be described universally by deterministic and irreversible hydrodynamic equations. This asymmetry in the direction of time is referred to as macroscopic irreversibility. In contrast, the macroscopic system microscopically consists of many interacting molecules. If the system is microscopically isolated, the time evolution of the constituent molecules is described by deterministic and reversible equations such as the classical Hamiltonian equations or quantum Schrödinger equations. This apparent inconsistency between macroscopic irreversibility and microscopic reversibility is called the irreversibility paradox. To resolve this paradox, we have to formulate the macroscopic irreversible laws within microscopic reversible dynamical systems [Leb93, Bri96].

A crucial idea is again the principle of equal a priori probabilities and the associated law of large numbers. Suppose that initial microscopic states are sampled from the microcanonical measure corresponding to a given non-equilibrium macroscopic state. The validity of the deterministic macroscopic law is then formulated as a result of the law of large numbers [LPS88], i.e., there is a set of microscopic states satisfying the macroscopic equations with

probability approaching one in a macroscopic limit. This is why the hydrodynamic equations describe even a single experimental result with high accuracy. Although to prove the law of large numbers for a given microscopic dynamics and initial probability measure is not an easy task in general, the above scenario is believed to be valid for a wide class of models and it is proved rigorously for specific models. Thus, the probabilistic assumption for microscopic states and the appropriate choice of the measure describing the macroscopic system play essential roles in statistical mechanics.

1.1.2 Thermodynamics for Small Systems

In contrast to macroscopic systems, small systems such as biomolecular motors are strongly affected by thermal fluctuations. Therefore, whether or not the thermodynamic structure can be found in small systems is a non-trivial problem. Nevertheless, a thermodynamic theory for mesoscopic systems that are weakly coupled to equilibrium environments has been developed over the past decades. For example, we can decompose the fluctuating energy transfer into two contributions, work and heat, similarly to macroscopic thermodynamics, and moreover we can verify the second law of thermodynamics for averaged quantities under this decomposition. Thanks to this framework, we can discuss the energy balance at nanoscales and a thermodynamic efficiency for molecular motors. This theory is now called stochastic thermodynamics.

A remarkable achievement in this field is the discovery of the fluctuation relations governing the fluctuation of the entropy production in generic stochastic systems. While the second law of thermodynamics claims only the positivity of the averaged entropy production, the fluctuation relation reveals a symmetry in *fluctuations* of the entropy production. The fluctuation relations can be applied to systems far from equilibrium and present a unifying picture of existing results in non-equilibrium statistical mechanics such as the second law of thermodynamics and the linear response theory. Moreover, the theoretical ideas and tools found in the course of these developments provided many interesting results such as steady state thermodynamics for small systems and information thermodynamic.

We must emphasize that the theory of stochastic thermodynamics has evolved with technological advances that allow us to observe and control micro- or nanoscale fluctuations. While experimental techniques that can access nanoscale fluctuations validate stochastic thermodynamics, the theoretical side provides novel methods to extract useful information on thermodynamic properties of small systems. For example, Jarzynski and Crooks relations can be used to estimate the free energy difference of a single molecule from nonequilibrium measurements of work fluctuations [LDS+02, CRJ+05]. Another application of stochastic thermodynamics is the experimental estimation of heat dissipation. While the energy dissipation into the heat bath is difficult to measure directly in experiments, the Harada-Sasa equality [HS05, HS06], which connects the heat dissipation with the violation of the fluctuation-dissipation relation, enables us to estimate it from measurements of the response and autocorrelation function [ATM18].

Finally, we mention a novel inequality that has been intensively studied recently and

called the thermodynamic uncertainty relation (TUR). A typical form of TUR is as follows:

$$\frac{(\mathbb{E}[J_t])^2}{\mathrm{Var}[J_t]} \leq \mathbb{E}[\Sigma_t], \tag{1.1.4}$$

where J_t and Σ_t are a time-integrated current and the entropy production up to the time t, respectively. As well as the fluctuation relations, this type of inequality is universally valid for various non-equilibrium stochastic systems under various situations, particularly, Markov jump processes and overdamped Langevin systems. This inequality is important in two respects. First, this inequality gives a tighter lower bound on the entropy production than the second law of thermodynamics if $\mathbb{E}J_t \neq 0$. Second, this inequality expresses a trade-off relation between the precision of the time-integrated current and the mean entropy production: The smaller the dissipation of the system, the larger the uncertainty of the current. In other words, this inequality provides a novel thermodynamic efficiency $\eta_{\mathrm{TUR}} \coloneqq (\mathbb{E}[J_t])^2/\mathrm{Var}[J_t]\mathbb{E}[\Sigma_t] \leq 1$, which may be useful in quantifying the efficiency of biomolecular motors.

1.2 Thermodynamics meets Gambling

1.2.1 Stochastic Thermodynamics with Martingale

In a typical situation in stochastic thermodynamics, we observe a stochastic system in a certain observation time interval and investigate the statistics of an observable at the fixed time. Here, we address a complementary problem, in which we exchange the roles of the observable and time, and study the statistics of the random time at which the observable first reaches a fixed threshold. Such random times are called first passage times. The distributions of first passage times, or more generally stopping times, are extensively studied in various fields such as the theory of stochastic processes [Red01, vK07], reaction rate theory [HTB90], biology [IBZ16, Ewe04], statistical estimations [GMS97], and finance [BBDG18]. Moreover, in nonequilibrium physics, the universal natures of the first-passage-time statistics for thermodynamically relevant quantities were found recently, including the fluctuation relations at stopping times [NRJ17, MSM+21], the universality of the asymptotic behavior of the first-passage-time distributions [SD16, Pan18, SMG+19], and several tradeoff relations concerning the first passage time [GH17, Gar17, FE20]. In particular, the thermodynamic constraint on the precision of the first passage times, which is a topic of this thesis, may be useful in measuring the efficiency of the biological clocks beyond merely theoretical interests [MCH19] (see also [BS15, BS16, BS17]).

Among these studies, several interesting results including the integral fluctuation relations at stopping times are obtained from applications of martingale theory in probability theory to stochastic thermodynamics. Martingale theory also plays essential roles in topics we consider in this thesis. Let us give a concise explanation on martingale theory below.

Martingale

To understand a motivation to introduce the concept of "martingale", let us consider a simple situation. A gambler visits a casino with money $M_0 > 0$ in his pocket and plays the following gambling with a dealer. First, the gambler predicts the outcome of coin tossing and bet his money on two outcomes, head and tail. After his betting, the dealer tosses a coin. We introduce a variable x_1 that takes 1 and 0 if the outcome of the first coin tossing is head and tail, respectively. Let q_1 and q_0 be the amount of money the gambler bet on head and tail, respectively. If $x_1 = i$ ($i \in \{0, 1\}$), the payoff is assumed to be $\pi_i q_i$. Then, the gambler's wealth $M_1(x_1)$ after the first coin tossing is given by

$$M_1(0) = M_0 - (q_0 + q_1) + \pi_0 q_0, \ M_1(1) = M_0 - (q_0 + q_1) + \pi_1 q_1. \tag{1.2.1}$$

How should we choose odds (π_0, π_1) to realize "fair" gambling if the probability that $X_1 = 1$ is p? One reasonable definition is that we say that the gambling is fair if the gambler's wealth neither increase nor decrease on average. Since the average of the wealth is given by

$$pM_1(1) + (1-p)M_1(0) = M_0 + (p\pi_1 - 1)q_1 + ((1-p)\pi_0 - 1)q_0, \tag{1.2.2}$$

the fair odds must be $\pi_1 = 1/p$ and $\pi_0 = 1/(1-p)$. For instance, the coin is unbiased, i.e., $p = 1/2$, the fair odds are $\pi_1 = \pi_0 = 2$.

We now suppose that the gambler and the dealer repeat this fair gambling many times. We only allow the gambler's strategy to choose his betting money that depends causally on the past outcomes and is independent of the future outcomes of coin tossing. We can identify a strategy for the gambler with a sequence of functions $M_n : \{0, 1\}^n \to \mathbb{R}_{\geq 0}$ satisfying the condition

$$M_{n-1}(x_1, \ldots, x_{n-1}) = pM_n(x_1, \ldots, x_{n-1}, 1) + (1-p)M_n(x_1, \ldots, x_{n-1}, 0) \tag{1.2.3}$$

for any n and $x_1 \ldots x_{n-1} \in \{0, 1\}^{n-1}$. The value $M_n(x_1, \ldots, x_n)$ gives the gambler's wealth after n-th round of the gambling when the outcome up to the n-th round is $(x_1, \ldots, x_n) \in \{0, 1\}^n$. Therefore, Eq. (1.2.3) expresses the fairness condition similarly to Eq. (1.2.2) at $\pi_1 = 1/p$ and $\pi_0 = 1/(1-p)$. After extending the domain of M_n into $\Omega \coloneqq \{0, 1\}^{\mathbb{N}}$ in the trivial way, we say that the process $\{M_n : n \in \mathbb{N}\}$ is a *martingale* if it satisfies the condition (1.2.3) for all n and $\omega \in \Omega$.

The concept of "martingale" is useful to study two problems in which the gambler is interested.

- The first interest is when should the gambler stop betting. The gambler must decide whether or not he stops betting based on only the information on past outcomes of the coin tossing. If the stopping rule satisfies such a causality condition, the time at which the gambler stops betting is called a stopping time. More formally, we say that a random time $T : \Omega \to \mathbb{N} \cup \{\infty\}$ is a stopping time if

$$\forall n \in \mathbb{N} \cup \{\infty\}, \ \{\omega \in \Omega : T(\omega) \leq n\} \in \mathcal{F}_n, \tag{1.2.4}$$

where $\mathcal{F}_n = \sigma(\{\omega \in \Omega : \omega_k = i\} : 0 \le k \le n, i \in \{0,1\})$ for $n \in \mathbb{N}$ and $\mathcal{F}_\infty = \sigma(\cup_n \mathcal{F}_n)$. A typical example of stopping times is a first passage time:

$$\begin{aligned} T &= \inf\{n \ge 0 : M_n = m\} \\ &= \text{“the first time for the wealth to reach a threshold } m\text{”}, \end{aligned} \tag{1.2.5}$$

where $\inf(\emptyset) = \infty$. The following random time is *not* a stopping time because we must know the future in order to stop betting:

$$T = \inf\{n \ge 0 : n \le 10, M_n = \max_{0 \le k \le 10} M_k\}. \tag{1.2.6}$$

Indeed, $\{T = 0\} = \{M_1 \le M_0, \ldots, M_{10} \le M_0\} \notin \mathcal{F}_0$.

If the gambler stops betting at a fixed time, the average of his wealth neither increase nor decrease because of the fairness condition, i.e., $\mathbb{E}M_n = M_0$. How about when the gambler applies a fluctuating stopping time. According to Doob's optional stopping theorem, we have that

$$\mathbb{E}(M_T) = M_0, \tag{1.2.7}$$

if M is uniformly integrable, where $M_T \coloneqq \lim_{n\to\infty} M_{T\wedge n}$ with $T \wedge n \coloneqq \min\{T, n\}$. Due to this theorem, even if the gambler applies a stopping rule, his wealth cannot be greater than the initial value.

- The second interest is how the maximum value of the wealth is distributed. For example, the gambler may stop betting by the n-th round if his wealth exceeds $\lambda > 0$ times the initial wealth. The gambler is then interested in the probability that the maximum value of the wealth up to the n-th round becomes greater than the threshold value, i.e., $\{\sup_{k\le n} M_n \ge \lambda M_0\}$. According to Doob's martingale inequality, we have that

$$\mathbb{P}\left(\sup_{k\le n} M_k \ge \lambda M_0\right) \le \lambda^{-1}\mathbb{E}(M_n 1_{\{\sup_{k\le n} M_k \ge \lambda M_0\}}) \le \lambda^{-1} M_0^{-1}\mathbb{E}(M_n) = \lambda^{-1}. \tag{1.2.8}$$

  In contrast to Markov's inequality, Doob's inequality gives an upper bound on a tail probability of the *maximum* value. Since a martingale expresses a capital process in the fair game under a strategy, this inequality implies that whatever strategy the gambler applies, the probability that the gambler can quit betting while his wealth exceeds 100 times the initial wealth is smaller than 1%.

The above arguments can be extended to more general settings. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space (the parameter t may be discrete or continuous). A stochastic process M_t is said to be a $(\mathbb{P}, \{\mathcal{F}_t\})$-martingale (or simply, martingale) if M is adapted to $\{\mathcal{F}_t\}$, $\mathbb{E}[|M_t|] < \infty$ for any t, and

$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s, \text{ a.s. } \forall t \ge s. \tag{1.2.9}$$

Examples of martingale processes that are important in this thesis are likelihood ratio martingales. We now have two stochastic processes $\mathbb{P}$ and $\mathbb{Q}$ such that $\mathbb{Q}$ is absolutely continuous relative to $\mathbb{P}$ on every $\mathcal{F}_t$ and define a process M_t as a version of the likelihood ratio between $\mathbb{P}$ and $\mathbb{Q}$:

$$M_t \coloneqq \frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}}. \tag{1.2.10}$$

Then, we find that $\{M_t\}$ is a $(\mathbb{P}, \{\mathcal{F}_t\})$-martingale. Since the likelihood ratio function measures the distinguishability between two processes, it mainly plays two roles in stochastic thermodynamics.

- If we consider the original process and the time-reversed process, the likelihood ratio between them quantifies the time irreversibility of the original process. Moreover, it is identified with the entropy production characterizing the thermodynamic dissipation. The theoretical idea on fluctuation relations is based on this perspective. Recently, martingale property of the entropy production has been studied and several interesting results are obtained: the integral fluctuation relation at stopping times, the work relation at stopping times, a bound on the tail probability of the infimum of the entropy production, the detailed fluctuation relation associated with the first passage time distributions.

- If we consider the original process and the perturbed process, the likelihood ratio between them characterizes the response property of the system against the perturbation. Indeed, the several response properties in Markovian dynamics can be obtained through the Taylor expansion of the logarithmic likelihood ratio function with respect to the small response field. A recent important finding based on the second perspective is the fluctuation-response inequality, which provides a new tool to study universal bounds on the response and fluctuations. For instance, the thermodynamic uncertainty relation explained before is derivable from the fluctuation-response inequality. In this thesis, we employ this technique with the martingale theory to study a class of universal bounds on the statistics of the stopping times.

Thus, the likelihood processes play important roles in stochastic thermodynamics. However, the martingale aspect of the likelihood processes has been studied only recently. One of the purposes of this thesis is to develop stochastic thermodynamics from the viewpoint of the martingale.

1.2.2 Thermodynamic Uncertainty Relation at Stopping Time

We explain in more detail the second perspective on the likelihood ratio processes in the context of stochastic thermodynamics. We are interested in the response property of a stochastic dynamics, a Markov jump process or a diffusion process in many cases, against perturbations. We use $\mathbb{P}$ and $\mathbb{P}^\theta$ to denote the path probability measures for the original

dynamics and the perturbed dynamics, respectively. Here, the parameter $\theta \in \mathbb{R}$ quantifies the amplitude of the perturbation, e.g., the external driving force on a Brownian particle. We write the likelihood ratio between the original and perturbed dynamics as

$$p_t^\theta := \frac{d\mathbb{P}^\theta|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}}. \tag{1.2.11}$$

For an observable A_t that depends on trajectories up to the time t, we have an expression for the expectation of A_t under the perturbed dynamics in terms of the original dynamics:

$$\mathbb{E}_\theta[A_t] = \mathbb{E}[A_t p_t^\theta]. \tag{1.2.12}$$

The (linear response) fluctuation-response inequality provides an upper bound on the intensity of the linear response. Here, the response intensity is measured by the ratio of the squared response function to the variance in the original process:

$$R(A_t) := \frac{(\partial_\theta \mathbb{E}[A_t]|_{\theta=0})^2}{\mathrm{Var}[A_t]} = \frac{1}{\mathrm{Var}[A_t]} \left(\int A_t \partial_\theta p_t^\theta|_{\theta=0} dP \right)^2. \tag{1.2.13}$$

According to the fluctuation-response inequality, this intensity is bounded above by the *Fisher information* I_t:

$$R(A_t) \leq I_t(0) := \mathbb{E}[(\partial_\theta \ln p_t^\theta)^2]|_{\theta=0}. \tag{1.2.14}$$

We remark that while the left-hand side depends on the choice of the observable A_t, the right-hand side is dependent only on the likelihood ratio function. Therefore, the fluctuation-response inequality gives a *uniform* upper bound on the response intensity. This inequality is interesting and useful in two respects. First, although few results are available for non-Markov processes, this inequality can be applied in principle to general stochastic processes, not limited to diffusion processes and Markov jump processes. Second, the perturbations added to the system are not limited to physical ones. For example, by adding a drift perturbation that does not correspond to any physical operations, which is called a "virtual" perturbation in [DS20], we obtain the thermodynamic uncertainty relation from the fluctuation-response inequality.

It is known that even if we apply a stopping rule in the experiment, we can prove the fluctuation-response inequality by employing martingale theory. Therefore, it is natural to ask whether the extended inequalities may reveal novel bounds on fluctuations *at stopping times*, in a similar manner that the thermodynamic uncertainty relation provides a measure quantifying an efficiency of a fluctuating system in a fixed observation time. In Chapter 2, we study this issue. Specifically, we derive the *kinetic* bound on the precision of the first passage time for the time-integrated current and find that this kinetic bound can be relevant far from equilibrium in contrast to the *thermodynamic* bound.

1.2.3 Randomness from Gambling

Until now we have reviewed martingale theory as a useful tool to study the stopping time and extreme value statistics. However, the concept of "martingale" has also been studied in

another context as a tool to characterize randomness. To see this, we return to the gambling with coin tossing, i.e., stochastic processes on $\Omega = \{0,1\}^{\mathbb{N}}$. First, we review an important observation by Jean Ville [Vil39]. For a martingale process M_n, let us take $\lambda = 2^n$ and define

$$U_n := \{\omega \in \Omega : (\exists k \in \mathbb{N})(M_k(\omega) > 2^n M_0)\}. \tag{1.2.15}$$

The maximal inequality (1.2.8) implies $\mathbb{P}(U_n) \leq 2^{-n}$. Then, we conclude that

$$U := \bigcap_{n\in\mathbb{N}} U_n = \{\sup_n M_n = \infty\} \tag{1.2.16}$$

is a null set, i.e., $\mathbb{P}(U) = 0$. This means that the probability that the gambler make infinite profits is zero. This consequence is consistent with our intuition on fair gamblings. Conversely, for a given null set N, there is a non-negative martingale that is divergent on N. Let us prove this statement. For a null set N, we have an open subsets $\{U_n\}$ such that $N \subseteq \bigcap_n U_n$ and $\mathbb{P}(U_n) \leq 2^{-n}$. We now define a process $M_{n,k}$ as

$$M_{n,k}(\omega) := \frac{\mathbb{P}(U_n \cap C(\omega_1, \dots, \omega_k))}{\mathbb{P}(C(\omega_1, \dots, \omega_k))}, \tag{1.2.17}$$

where $C(\omega_1, \dots, \omega_k) := \{\sigma \in \Omega : \omega_i = \sigma_i, 1 \leq i \leq k\}$. We can see that this process is a non-negative martingale:

$$\begin{aligned}
&pM_{n,k+1}(\omega_1, \dots, \omega_k, 1, \dots) + (1-p)M_{n,k+1}(\omega_1, \dots, \omega_k, 0, \dots) \\
&= \sum_{i\in\{0,1\}} \frac{\mathbb{P}(C(\omega_1, \dots, \omega_k, i))}{\mathbb{P}(C(\omega 1, \dots, \omega_k))} \frac{\mathbb{P}(U_n \cap C(\omega_1, \dots, \omega_k, i))}{\mathbb{P}(C(\omega_1, \dots, \omega_k, i))} \\
&= \sum_{i\in\{0,1\}} \frac{\mathbb{P}(U_n \cap C(\omega_1, \dots, \omega_k, i))}{\mathbb{P}(C(\omega 1, \dots, \omega_k))} \\
&= \frac{\mathbb{P}(U_n \cap C(\omega_1, \dots, \omega_k))}{\mathbb{P}(C(\omega_1, \dots, \omega_k))} = M_{n,k}(\omega).
\end{aligned} \tag{1.2.18}$$

Since $M_{n,0} = \mathbb{P}(U_n) \leq 2^{-n}$, M_n defined by $M_n = \sum_{k=0}^{\infty} M_{n,k}$ is finite and the process $\{M_n : n \in \mathbb{N}\}$ defines a non-negative martingale. Moreover, since $M_{n,k}(\omega) = 1$ for $\omega \in N$ by definition, we obtain that $\sup_n M_n = \infty$ on N.

The above correspondence between null sets and martingale processes suggests that the concept of "martingale" is useful to define "random sequences". Approximately speaking, a sequence $\omega \in \Omega$ is called *random* if it satisfies no exceptional properties. For instance, consider an experiment where we toss an unbiased coin many times. A sequence in which 0 and 1 appear at a rate of 1/3 and 2/3, respectively, is not random under this experiment because its frequencies of 0 and 1 are exceptional. Exceptional properties are mathematically defined as null sets. The correspondence between null sets and non-negative martingales motivates the following definition of randomness: A sequence ω is random if $\sup_n M_n(\omega) < \infty$ for any non-negative martingales M. In other words, if we bet on the future bits by utilizing the

knowledge of the previous bits of the random sequence, there is no betting strategy by which we are able to make much money.

Since a singleton $\{\omega\}$ is a null set, there is a betting strategy = non-negative martingale such that $\lim_{n\to\infty} M_n(\omega) = \infty$. Therefore, unfortunately, the above definition of randomness turns out to be meaningless. We have to restrict the class of martingales to provide a meaningful definition of randomness. Following the celebrated idea of Levin and Schnorr, we restrict martingales to ones that can be performed feasibly in an algorithmic manner. They specified such a feasible martingale as an effective martingale, more precisely, a lower semi-computable martingale. Then, a sequence ω is called random if $\sup_n M_n(\omega) < \infty$ for any effective martingales. Remarkably, this randomness, which characterizes random sequences from the viewpoint of gambling, is equivalent to other notions of randomness characterized by typicalness and incompressibility conditions. As a result, this randomness, called *algorithmic randomness*, is regarded as a natural notion of randomness.

Is the characterization of randomness through martingale theory relevant to statistical physics? In the next two subsections, we introduce two problems, thermodynamic irreversibility and the second law of thermodynamics, and argue the relevancy of randomness in statistical physics.

1.2.4 Thermodynamic Irreversibility and Randomness

In Chapter 3, we study thermodynamic irreversibility from the viewpoint of algorithmic randomness. In subsection 1.1.1, we explain that the law of large numbers gives a clear account of the emergence of macroscopic laws from microscopic dynamics. However, we should notice that it refers only to the probability that the macroscopic law is satisfied and does not tell us which microscopic states among all realizable states obey the macroscopic law. When considering the reversibility paradox, one finds that this fact becomes problematic. Loschmidt pointed out that if a microscopic trajectory satisfying the microscopic equation of motion obeys the macroscopic law, the time-reversed trajectory is also a solution of the same equation due to the microscopic reversibility, but violates the macroscopic law due to the irreversibility of that law [Los76]. Thus, the apparent inconsistency between the macroscopic irreversibility and the microscopic reversibility is relevant to individual trajectories. We note that the recurrence paradox posed by Zermelo also refers to a single trajectory [Zer96], but this paradox is resolved by considering the thermodynamic limit first. To resolve the reversibility paradox, it is desirable to have a more direct formulation studying individual microscopic states in the thermodynamic limit. In particular, we need a criterion to determine whether a given microscopic state belongs to a set characterized by typical macroscopic properties.

To this end, in Chapter 3, we apply the theory of algorithmic randomness to statistical physics on the basis of its characteristics that the algorithmically random sequences satisfy statistical properties such as the law of large numbers and of the iterated logarithm. The

strong law of large numbers in probability theory states that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}x(i)=\frac{1}{2}\ \ \textit{for almost all } x \text{ with respect to } \mathbb{P}. \tag{1.2.19}$$

In contrast, the counterpart in the theory of algorithmic randomness states that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}x(i)=\frac{1}{2}\ \ \textit{for random } x \text{ with respect to } \mathbb{P}. \tag{1.2.20}$$

Although the former statement refers to only the probability that the law of large numbers is satisfied, the latter refers to individual sequences. In the context of statistical physics, the statistical properties of random sequences imply that the probability-theoretic statement,

> "*Almost all* microscopic states with respect to a probability measure obey a macroscopic law,"

can be replaced by the pointwise one,

> "*Random* microscopic states with respect to a probability measure obey a macroscopic law."

Since this statement is expressed at the level of individual states, the notion of algorithmic randomness may be useful to discuss the foundations of statistical mechanics beyond probability-theoretic statements.

The formulation with the notion of randomness provides a new perspective on the reversibility paradox. From a measure-theoretic point of view, the microscopic reversibility is consistent with our experience since the time-reversed state of a typical state with respect to a probability measure violates the macroscopic law, but has only an extremely small probability with respect to the same measure. In contrast, from a viewpoint of algorithmic randomness, the microscopic reversibility implies that the time-reversed state of a random state is not random because it is contained in a null set involved with a violation of the macroscopic law. This fact has an implication in the relation between the randomness of a physical state and the ease of preparation of that state. If we can specify a description of a binary sequence completely in an algorithmic manner, the sequence is not random because we can easily predict the sequence by using the algorithmic description. In contrast, a sequence we generate by a stochastic device such as tossing a coin many times is algorithmically random. Now, when we prepare a state of a physical system, we cannot avoid a certain source of noise. Although the relation between a stochastic device and noises in preparation of physical states is not obvious, we may say that it is difficult to experimentally prepare the time-reversed state of a random state, which is nonrandom, because we have to avoid all sources of noises to prepare a nonrandom state. In this manner, the theory of algorithmic randomness clarifies a conceptually new aspect of the reversibility paradox. In Chapter 3, we show a part of the results that the theory of algorithmic randomness reveals.

1.2.5 Game-Theoretic Thermodynamics

The role of martingales in the study of thermodynamic irreversibility is only of secondary importance. Indeed, in Chapter 3, we analyze it based on the measure-theoretic typicality instead of martingale theory. Therefore, the connection of thermodynamic concepts with martingales is indirect. In Chapter 4, we find that a process of the extracted work in repeated cyclic operations is identified with a martingale process. The combination of this identification and Ville's theorem reveals a novel relationship between the Gibbs distribution and the second law of thermodynamics.

The problem we consider in Chapter 4 is formulated as follows. A fundamental assumption of equilibrium statistical mechanics is that microscopic states are randomly sampled according to the Gibbs distribution for the Hamiltonian. While we can obtain statistical properties of observables, such as means and variances, from this probabilistic assumption, the assumption is also consistent with the second law of thermodynamics. In fact, we cannot extract a strictly positive amount of work through any cyclic process *on average* if the initial probability distribution is Gibbsian [PW78, Len78, GP80, Dan81, Jar97, Cro98]. This result can be regarded as a derivation of the second law of thermodynamics from statistical mechanics. It is natural to ask whether the second law conversely characterizes the Gibbs distribution or not. This question has been traditionally studied in terms of passivity [PW78, Len78, GP80, Dan81]. These studies showed that the initial probability distribution is Gibbsian if and only if any number of copies of the identical state satisfy the second law of thermodynamics. In this approach, we crucially assume that the system is described by a probability distribution on the phase space. In contrast, the question we consider in Chapter 4 is how the probabilistic description based on the Gibbs distribution emerges from the second law of thermodynamics, particularly the absence of the perpetual motion machine of the second kind.

A similar question was posed in the context of probability theory. In measure-theoretic probability theory, we assign a real value in $[0, 1]$ to each event under certain compatibility conditions such as additivity. Although measure-theoretic probability theory is a useful tool to analyze the random behavior of phenomena in nature, it does not provide a characterization of randomness itself. As an alternative approach, Shafer and Vovk proposed game-theoretic probability theory [SV01, SV19] whose fundamental idea is Ville's theorem. In a gambling, they say that the gambling is fair if the gambler never become infinitely rich in the limit as the gamble continues and that the sequence of events obtained from such gambling is random. Based on these ideas, they proved, say, the law of large numbers in terms of the gambling, without using measure-theoretic concepts. If we think of the work extraction as a certain gambling between an agent and the nature, the second law of thermodynamics corresponds to the fairness condition in the gambling. If this reasoning is true, we can expect that randomness of the microscopic states in an equilibrium system is characterized through the second law of thermodynamics without using the probabilistic assumption in statistical mechanics. The purpose of Chapter 4 is to validate this idea and to answer in the affirmative based on the identification of the work processes with martingale processes and the game-theoretic probability theory.

Chapter 2

Kinetic Uncertainty Relation on Fist Passage Time

2.1 Introduction

2.1.1 Summary of results

In this chapter, we study the kinetic uncertainty relations (KURs) on the first passage times for time-integrated currents. Whereas the thermodynamic uncertainty relation gives a bound on the precision of an observable in terms of entropy production, the KUR [TB19] gives a bound in terms of the time-symmetric dynamical activity [Mae20]. Garrahan [Gar17] obtained a kinetic bound on the FPT for a stationary continuous-time Markov chain,

$$\frac{\mathbb{E}[\tau]^2}{\mathrm{Var}[\tau]} \leq n\mathbb{E}[\tau], \tag{2.1.1}$$

in the large threshold limit. Here $\mathbb{E}[\tau]$ denotes the mean time at which the time-integrated current first reaches a threshold, $\mathrm{Var}[\tau]$ the variance of τ, and n the mean number of jumps per time in the stationary state. This trade-off relation implies that the smaller the activity of the stochastic system, the larger the uncertainty of the time to reach the threshold. The inequality (2.1.1) was derived via the large deviation theory and verified only in the large threshold limit. The main purpose of this chapter is to derive the KUR on the first passage time that is valid for any *finite* threshold and to simplify the derivation based on the technique recently developed in [DS20, TB19].

2.1.2 Outline of the chapter

The remainder of the chapter is organized as follows. In Section 2.2, we review previous studies and present main results of this chapter. In Section 2.3, we applies our results to two examples and find that the kinetic bounds are relevant far from equilibrium. In Section 2.4, we review the information inequalities at stopping times, which is the key ingredient in the

derivation of our results, and present a proof of our main result. In Section 2.5, we end this chapter with concluding remarks.

2.2 Main Result

2.2.1 Setup and previous results

Setup

We consider a time-homogeneous and continuous-time Markov chain on a directed multigraph $G = (S, E)$. Here S is a discrete state space and E is the set of directed edges between two states. Let $k_e(x, y)$ be the transition rate from the state x to y via the edge $e \in E$ and $\lambda(x) \coloneqq \sum_e \sum_{y(\neq x)} k_e(x, y)$ the escape rate from x, where the summation is taken over edges starting from x. For a fixed time $t \geq 0$, $X_{[0,t]} = (X_s)_{s\in[0,t]}$ denotes a single trajectory of the system and is characterized by the discrete-time sequence $(x_0, t_0 = 0; x_1, t_1, e_1; \ldots; x_{N_t}, t_{N_t}, e_{N_t})$, which indicates that the total number of jumps for trajectory $X_{[0,t]}$ over $[0, t]$ is N_t, and the transition from x_{i-1} to $x_i (\neq x_{i-1})$ occurs via the edge e_i at times t_i for $i = 1, \ldots, N_t$. We focus on a time-integrated current $J_t \coloneqq J(X_{[0,t]})$ defined as

$$J(X_{[0,t]}) = \sum_{i=1}^{N_t} g_{e_i}(x_{i-1}, x_i), \tag{2.2.1}$$

where $g_e(x, y)$ weights the contribution of the transition from x to y via the edge e. The class of observables of this form includes many important physical quantities. Here we address two significant examples, the number of jumps and the fluctuating entropy production. The total number of jumps via edge f is obtained by taking $g_e = \delta_{e,f}$. This quantity measures how active the system is on edge f and is called dynamical activity. Next, we consider two edges, f connecting from x to y, and b connecting from y to x, with $k_f(x, y) \neq 0$ and $k_b(y, x) \neq 0$. We assume that these edges are in contact with the same heat bath, and require that the entropy per the Boltzmann constant k_B produced in the heat bath during the transition $x \to y$ is given by $\ln(k_f(x, y)/k_b(y, x))$. The fluctuating entropy production associated with the paired edges (f, b) is then obtained by taking $g_e = \ln(k_f(x, y)/k_b(y, x))[\delta_{e,f} - \delta_{e,b}]$. The requirement we impose here is called the local detailed balance condition.

Thermodynamic and kinetic uncertainty relations

If the system satisfies the local detailed balance condition, the system starts from the steady state, and we consider the time-asymmetric currents, i.e., $g(x, y) = -g(y, x)$, the precision of the time-integrated current is bounded above by the entropy production

$$\frac{\mathbb{E}[J_t]^2}{\mathrm{Var}[J_t]} \leq \mathbb{E}[\Sigma_t], \tag{2.2.2}$$

where Σ_t is the total entropy production up to the time t. This inequality, called the thermodynamic uncertainty relation (TUR), is regarded as a trade-off relation between the precision

of the current and the dissipation. We have another trade-off relation called the kinetic uncertainty relation (KUR). The KUR claims that

$$\frac{\mathbb{E}[J_t]^2}{\mathrm{Var}[J_t]} \leq \mathbb{E}[N_t]. \tag{2.2.3}$$

Here, N_t denotes the total number of jumps up to the time t and therefore the right-hand side of the KUR quantifies a dynamical activity of the system. Although the KUR is valid for any observables in contrast to the TUR, we consider only time-integrated currents for simplicity. Numerical simulation for simple systems shows that while the TUR is tight near equilibrium, the KUR gives a better bound in the regime far from equilibrium.

First passage statistics

In the above situation, we fix an observation time t and study the statistics of time-integrated currents J_t. In this chapter, instead of studying the statistics at a fixed time, we consider the statistics of the random time at which the time-integrated current first reaches a fixed threshold. The first passage time (FPT) τ for the time-integrated current J_t is defined as

$$\tau \coloneqq \inf\{t \geq 0 : J_t = J_{\mathrm{th}}\}, \tag{2.2.4}$$

where J_{th} denotes the threshold value. The FPT is obviously a stochastic variable and accompanies fluctuations. The first time for the system to reach the specific state z can be represented in this form by taking $g_e(x, y) = \delta_{y,z}$ and $J_{\mathrm{th}} \in (0, 1)$. Methods to analyze the statistics of the FPTs in the class are well established [Red01, vK07]. Our concern here is the precision of the FPT quantified by the ratio of the squared mean FPT to the variance, $\mathbb{E}[\tau]^2/\mathrm{Var}[\tau]$.

Main result

Garrahan [Gar17] derived the kinetic uncertainty relation on the first passage time (2.1.1) via the large deviation theory. By using a similar method, Gingrich and Horowitz obtained the thermodynamic uncertainty relation on the first passage time. Their argument is based on the asymptotic behavior in the *large threshold limit* $J \to \infty$. Therefore, the same method cannot be straightforwardly applied to first passage times with *finite thresholds*. Since there are problems that concern the first passage times with finite thresholds, e.g., the escape rate from metastable states and searching problem, it is desired to develop a method to study trade-off relations on first passage times with finite thresholds.

Our main result is the kinetic uncertainty relation that is valid for any finite threshold. We suppose that the mean and variance of τ are finite. We find that the precision of the FPT is bounded from above by the mean dynamical activity, which is quantified by the mean number of jumps:

$$\frac{\mathbb{E}_{x_0}[\tau]^2}{\mathrm{Var}_{x_0}[\tau]} \leq \mathbb{E}_{x_0}[N_\tau], \tag{2.2.5}$$

where N_τ is the total number of jumps up to the first passage time τ and $\mathbb{E}_{x_0}[\cdot]$ denotes the expectation value conditioned on the initial configuration $X_0 = x_0$. This activity bound (2.2.5) is the main result of this chapter. The inequality (2.2.5) implies that the reduction in the number of jumps up to the time for the system to passage the threshold inevitably accompanies the worsening of the optimal precision of the FPT.

We make several remarks on our result.

- First, our result holds for any finite threshold $J_{\rm th}$ in contrast to the inequality (2.1.1) in [Gar17]. For a sufficiently large threshold $J_{\rm th}$ and ergodic Markov process, we expect that $\mathbb{E}[N_\tau]$ nearly equals $n\mathbb{E}[\tau]$ because in that situation the jump number per time is well approximated over a sufficiently large time interval by the stationary value n. Hence the inequality (2.1.1) is recovered in this asymptotic limit from our result. Although the precision of the FPT for the integrated current is also bounded from above by the mean entropy production, the thermodynamic bound of the form in [GH17] is guaranteed only in this asymptotic limit. We illustrate the violation of the TUR with examples in the next section.

- Second, we should notice that the KUR (2.2.5) can be applied to the case that the system starts from a specific initial state x_0. In addition, the inequality (2.2.5) is still valid if we replace the conditional expectation $\mathbb{E}_{x_0}$ by the expectation value $\mathbb{E}_\rho$ with respect to the arbitrary initial distribution ρ.

- Third, the thermodynamic bound is tighter than the activity bound around equilibrium because in the regime close to equilibrium the mean entropy production tends to zero but the mean number of jumps remains finite. In contrast, even when the mean entropy production goes to infinity in the regime far from equilibrium, the mean activity up to the first passage time may be finite due to a nonequilibrium force driving the system to the threshold. In that case, the KUR provides a tighter bound than the TUR.

- Forth, the KUR is applicable even if the Markov chain is not ergodic or does not satisfy the reversibility condition, i.e., $k(x,y) \neq 0$ iff $k(y,x) \neq 0$. Examples of stochastic processes violating the reversibility condition are models including absorbing states in the population dynamics and stochastic resetting systems [EMS20].

- Fifth, we note that the mean and variance of the FPT may diverge; in such circumstances, the KUR (2.2.5) may be violated. For instance, when the accumulated current has a positive drift, the FPT for a negative threshold takes infinite value with positive probability. However, the modified KUR still holds in the following form,

$$\frac{\widetilde{\mathbb{E}}_{x_0}[\tau]^2}{\widetilde{\mathrm{Var}}_{x_0}[\tau]} \leq \widetilde{\mathbb{E}}_{x_0}[N_\tau]. \tag{2.2.6}$$

  We note that if the probability that the first passage time is infinite is positive, the modified probability distribution is unnormalized, $\widetilde{\mathbb{E}}[1] = \mathrm{Prob}(\tau < \infty) < 1$.

2.3 Examples

Before going to the derivation of main result, we examine the KUR (2.2.5) in two paradigmatic examples.

2.3.1 Biased Random Walk

The first example is the biased random walk X_t on $\mathbb{Z}$ starting from $X_0 = 0$. The transition rates between neighboring sites are set to $k_\pm := k(x, x \pm 1) = ae^{\pm\epsilon/2}$ and other transitions do not occur. Here $a > 0$ and $\epsilon > 0$ are positive constants. Suppose that this system describes a colloid under an external driving force f in a channel having a periodic structure of length l and filled with water in equilibrium at temperature T (Fig. 2.1(a)). According to the local detailed balance condition, $\epsilon = fl/k_{\mathrm{B}}T = \ln(k_+/k_-)$ is the entropy per k_{B} produced in the water by the one forward jump. We consider the random time $\tau_x = \inf\{t \geq 0 : X_t = x\}$ at which the colloid first reaches the site $x > 0$.

The entropy production along the path $X_{[0,t]}$ is given by $\Sigma_t := \epsilon X_t$ and therefore the stationary entropy production is $\sigma := \mathbb{E}[\Sigma_t]/t = 2\epsilon \sinh(\epsilon/2)$. We easily find that the precision of τ_x is given by $\mathbb{E}[\tau_x]^2/\mathrm{Var}[\tau_x] = x \tanh(\epsilon/2)$ and the TUR [GH17], $\mathbb{E}[\tau_x]^2/\mathrm{Var}[\tau_x] \leq \sigma\mathbb{E}[\tau_x]/2$, is directly verified for any thresholds [GRJ20]. In addition, the mean dynamical activity is given by $\mathbb{E}[N_{\tau_x}] = x \coth(\epsilon/2)$ and is in agreement with the KUR, $\mathbb{E}[\tau_x]^2/\mathrm{Var}[\tau_x] \leq \mathbb{E}[N_{\tau_x}]$.

In Fig. 2.1(b), we see that while the TUR is tighter near the equilibrium $\epsilon \lesssim 1$, the KUR becomes relevant as the nonequilibrium driving force increases. We remark that the TUR may be violated in general for finite thresholds. As an example, we consider a random walk with a reflecting boundary condition at the origin, i.e., $k(0, -1) = 0$ and a precision of the FPT for the threshold $x = 1$. The distribution of τ_1 is the exponential distribution with the decay rate k_+ and therefore $\mathbb{E}[\tau_1]^2/\mathrm{Var}[\tau_1] = 1 = \mathbb{E}[N_{\tau_1}]$ for any ϵ, whereas the upper bound of the TUR, $\sigma\mathbb{E}[\tau_1] = \epsilon/2$, becomes less than 1 for sufficiently small ϵ.

2.3.2 Two Level System

The second example is a two-level system in contact with two heat baths at different temperatures (Fig. 2.2(a)). The lower (resp. higher) energy level is coded by 0 (resp. 1) and the energy gap is set to $\Delta > 0$. The transition rate from $x \in \{0, 1\}$ to $y(\neq x)$ associated with the heat bath at the inverse temperature β_e is given by $k_e(x, y) > 0$ for $e \in \{h, c\}$. We assume that $\beta_h < \beta_c$ and define $\overline{\beta} := (\beta_c + \beta_h)/2$. The local detailed balance condition imposes $k_e(1, 0)/k_e(0, 1) = e^{\beta_e \Delta}$ for each bath e. We observe the heat produced in the cold bath $e = c$ and measure the accumulated heat current per $\beta_c\Delta$ from the system into the cold bath $J_t := \sum_{i=1}^{N_t} [\delta_{x_{i-1},1}\delta_{x_i,0} - \delta_{x_{i-1},0}\delta_{x_i,1}]\delta_{e_i,c}$. Our interest is the variation in precision of the first passage time $\tau_m = \inf\{t \geq 0 : J_t = m\}$ $(m \in \mathbb{Z})$ along with the temperature difference, which quantifies a distance from equilibrium in this model. To measure the degree by which temperatures differ, we introduce a dimensionless parameter $\epsilon := (\beta_c - \beta_h)/\overline{\beta} \in [0, 2]$. We

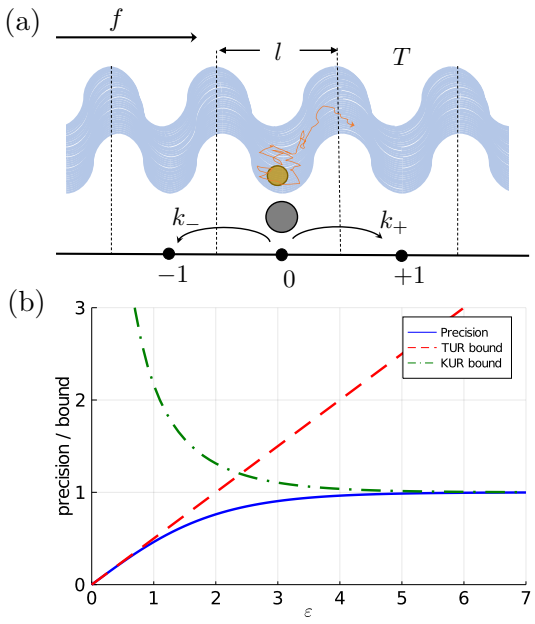

Figure 2.1: (a) Schematics of a colloid being driven by an external force f in a periodic channel filled with water at temperature T and 1D biased random walk describing the dynamics of the colloid. (b) Plots showing the precision $\mathbb{E}[\tau_x]^2/(x\mathrm{Var}[\tau_x])$ (blue solid line), the TUR bound $\sigma\mathbb{E}[\tau_x]/(2x)$ (red dash line), and the KUR bound $\mathbb{E}[N_{\tau_x}]/x$ (green dash-dot line) as functions of ϵ. The dimensionless parameter $\epsilon = fl/k_{\mathrm{B}}T$ measures a distance from equilibrium in this model.

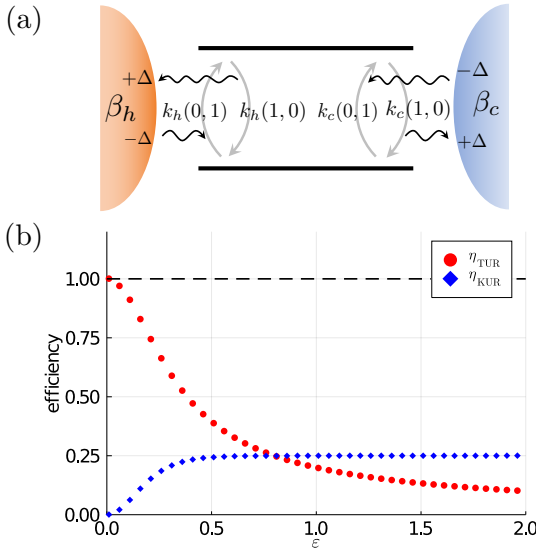

Figure 2.2: (a) Schematic of two-level system. (b) Efficiencies associated with the TUR (red circles) and KUR (blue diamonds) obtained from the Monte Carlo simulations for 10^7 samples for various temperature differences. We set $X_0 = 0$, $k_c(0,1) = 1$, $k_c(1,0) = e^{\beta_c \Delta}$, $k_h(0,1) = 1$, $k_h(1,0) = e^{\beta_h \Delta}$, $\overline{\beta}\Delta = 10$, $m = 1$ in our simulation.

define the efficiencies associated with the TUR and KUR as

$$\eta_{\mathrm{TUR}} := \frac{2}{\sigma} \frac{\mathbb{E}[\tau_m]}{\mathrm{Var}[\tau_m]}, \quad \eta_{\mathrm{KUR}} := \frac{1}{\mathbb{E}[N_{\tau_m}]} \frac{\mathbb{E}[\tau_m]^2}{\mathrm{Var}[\tau_m]} \leq 1, \tag{2.3.1}$$

respectively. From the plot of the efficiencies obtained by Monte Carlo simulations (Fig. 2.2(b)), we see that although they are actually reversed as the temperature difference increases, the efficiency associated with the KUR is always far from optimal.

2.4 Derivation based on Information Inequalities

2.4.1 Preliminaries

The key ingredient of the derivation of our result is the information inequality at stopping times, which provides a uniform bound on the ratio of the response to the fluctuation. In this section, we review the information inequality at stopping times. The readers who are not interested in technical details can skip this section.

Re-weighting formula at stopping time

We first derive the re-weighting formula, which is useful to study the response property of the stochastic process. We consider a family of stochastic processes over time interval $[0, \infty)$ with path probability measures P_θ parametrized by a real parameter θ. We can imagine that

P_0 and P_θ ($\theta \neq 0$) are the reference process and the perturbed process, respectively, and the parameter θ represents the amplitude of the perturbation, e.g., the external field. This situation is mathematically modeled by the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P_\theta)$. Here Ω denotes the space of stochastic trajectories over $[0, \infty)$, $\mathcal{F}$ the collection of all events that can occur in this system, $\mathcal{F}_t (\subset \mathcal{F})$ the collection of events that can occur up to fixed time t, and $P_\theta : \mathcal{F} \to [0, 1]$ the map assigning events in $\mathcal{F}$ into probabilities. $\mathcal{F}$ and $\{\mathcal{F}_t\}_{t\geq 0}$ are called the σ-algebra and filtration, respectively. We use P_θ^t to denote the probability distribution that assigns probabilities to events in $\mathcal{F}_t$. More formally, P_θ^t is defined as the restriction of P_θ to $\mathcal{F}_t$, i.e., $P_\theta^t \coloneqq P_\theta|_{\mathcal{F}_t}$. We assume that for any θ, θ' and $t \geq 0$, P_θ^t are mutually absolutely continuous with respect to $P_{\theta'}^t$. Under this assumption, we define the likelihood ratio function $p_\theta^t \coloneqq dP_\theta^t/dP_0^t$ for each $t \geq 0$, which characterizes the response property of the reference system $\theta = 0$ against the perturbation. Indeed, the expectation value with respect to the perturbed process of an observable A_t that depends only on the trajectory up to the time t is written in terms of the expectation with respect to the original process and the likelihood ratio function:

$$\mathbb{E}_\theta[A_t] = \int A_t dP_\theta = \int A_t \frac{dP_\theta^t}{dP_0^t} dP_0^t = \mathbb{E}_0[A_t p_\theta^t], \tag{2.4.1}$$

where $\mathbb{E}_\theta$ denotes the expectation operator with respect to the process P_θ. Thus, we can obtain the information on the response property of the original dynamics by studying the re-weighting factor p_θ^t.

The purpose of this section is to find a fundamental limitation on the statistics at *stopping times*. A stopping time is a random time when the stochastic process satisfies a certain condition first time. Formally, a random variable τ is said to be a stopping time if $\{\tau \leq t\} \in \mathcal{F}_t$ for each $t \geq 0$, which means that the value of τ is smaller than t is determined by the information about the trajectory up to the time t. Typical examples of stopping times are first passage times, e.g.,

$$\tau = \inf\{t \geq 0 : X_t \in R\}, \tag{2.4.2}$$

where R is a region. We define the σ-algebra $\mathcal{F}_\tau$ generated by the stopping time τ as

$$E \in \mathcal{F}_\tau \Leftrightarrow (\forall t \geq 0)(E \cap \{\tau \leq t\} \in \mathcal{F}_t). \tag{2.4.3}$$

This means that $\mathcal{F}_\tau$ is the collection of events of which we can determine the happening before the stopping time τ. As an example of events in $\mathcal{F}_\tau$, let us consider whether an observable at the time τ is in a certain range. For an observable A_t that depends on the trajectory up to time t, we define the observable A_τ at the stopping time τ as

$$A_\tau(\omega) \coloneqq A_{\tau(\omega)}(\omega) \tag{2.4.4}$$

for trajectories $\omega \in \Omega$ such that $\tau(\omega) < \infty$. If $\tau(\omega) = \infty$, there is no canonical way to define $A_\tau(\omega)$. In this chapter, we always study events on the set $\{\tau < \infty\}$ and therefore this ambiguity does not matter. Since whether $A_\tau \in [a, a + da)$ or not is determined by

the information on the trajectory up to the stopping time τ, $\{A_\tau \in [a, a+da)\} \in \mathcal{F}_\tau$. The probability distribution P_θ^τ that assigns probabilities to events in $\mathcal{F}_\tau$ is defined as the restriction of P_θ to $\mathcal{F}_\tau$, i.e., $P_\theta^\tau \coloneqq P_\theta|_{\mathcal{F}_\tau}$. The expectation value of an observable A_τ at the stopping time is obtained via this distribution P_θ^τ: $\mathbb{E}_\theta[A_\tau] = \int A_\tau dP_\theta^\tau$ if $P_\theta(\tau < \infty) = 1$.

Can we obtain the re-weighting formula (2.4.1) when the experiment is terminated at the random time τ? Whether the random variable $p_\theta^\tau(\omega) \coloneqq p_\theta^{\tau(\omega)}(\omega)$ is related to the Radon-Nikodym derivative of P_θ^τ with respect to P_0^τ is a non-trivial problem because these two objects are defined in different ways. However, we can prove that

$$P_\theta(E \cap \{\tau < \infty\}) = \int_E p_\theta^\tau 1_{\{\tau<\infty\}} dP_0 \tag{2.4.5}$$

for any events $E \in \mathcal{F}_\tau$. By defining the modified probability distribution $\widetilde{P}_\theta[E] = P_\theta[E \cap \{\tau < \infty\}]$ and the corresponding expectation $\widetilde{\mathbb{E}}_\theta$, we obtain the re-weighting formula for the modified distributions,

$$\widetilde{\mathbb{E}}_\theta[A_\tau] = \widetilde{\mathbb{E}}_0[A_\tau p_\theta^\tau] \tag{2.4.6}$$

from (2.4.5). We remark that $\widetilde{\mathbb{E}}_\theta[1] = P_\theta(\tau < \infty)$ may be less than 1, i.e., the distribution $\widetilde{P}_\theta$ may be unnormalized. If $P_\theta(\tau < \infty) = 1$ for any θ, we have that P_θ^τ is mutually absolutely continuous with respect to $P_{\theta'}^\tau$ and the Radon-Nikodym derivative is actually related to p_θ^τ:

$$p_\theta^\tau = \frac{dP_\theta^\tau}{dP_0^\tau}. \tag{2.4.7}$$

In that case, $\widetilde{P}_\theta$ and $\widetilde{\mathbb{E}}_\theta$ yield the normalized probability distribution and the usual expectation operator.

A proof of (2.4.5) is as follows. First, we consider a *truncated* stopping time $\tau \wedge n \coloneqq \min\{\tau, n\}$ for a fixed natural number n. Then, we have $P_0^{\tau\wedge n} \ll P_\theta^{\tau\wedge n}$ since $\mathcal{F}_{\tau\wedge n} \subseteq \mathcal{F}_n$ and $P_0^n \ll P_\theta^n$. Therefore, we obtain that

$$\frac{dP_\theta^{\tau\wedge n}}{dP_0^{\tau\wedge n}} = \mathbb{E}_0\left[\frac{dP_\theta^n}{dP_0^n}\middle|\mathcal{F}_{\tau\wedge n}\right] = \mathbb{E}_0\left[p_\theta^n|\mathcal{F}_{\tau\wedge n}\right]. \tag{2.4.8}$$

Because $\tau \wedge n$ is bounded by definition and the process $(p_\theta^n)_{n\in\mathbb{N}}$ is a $\{\mathcal{F}_n\}$-martingale with respect to $\mathbb{P}_0$, we obtain from the optional stopping theorem that

$$p_\theta^{\tau\wedge n} = \mathbb{E}_0\left[p_\theta^n|\mathcal{F}_{\tau\wedge n}\right]. \tag{2.4.9}$$

We now fix a natural number $k \in \mathbb{N}$ and choose $E \in \mathcal{F}_{\tau\wedge k}$. Since $E \cap \{\tau \le m\} \in \mathcal{F}_{\tau\wedge m}$ for any $m \ge k$ and $p_\theta^\tau = p_\theta^{\tau\wedge m}$ on the truncated set $\{\tau \le m\}$, we obtain from Eqs. (2.4.8) and (2.4.9) that

$$P_\theta(E \cap \{\tau \le m\}) = \int_{E\cap\{\tau\le m\}} dP_\theta = \int_{E\cap\{\tau\le m\}} p_\theta^{\tau\wedge m} dP_0 = \int_{E\cap\{\tau\le m\}} p_\theta^\tau dP_0. \tag{2.4.10}$$

By taking $m \to \infty$, we have from the monotone convergence theorem for $1_{E\cap\{\tau\le m\}}p_\theta^\tau \uparrow 1_{E\cap\{\tau<\infty\}}p_\theta^\tau$ that

$$P_\theta(E\cap\{\tau<\infty\}) = \int_{E\cap\{\tau<\infty\}} p_\theta^\tau dP_0. \tag{2.4.11}$$

Since

$$\{\tau<\infty\}\cap\mathcal{F}_\tau = \{\tau<\infty\}\cap\sigma\left(\bigcup_{k=1}^\infty \mathcal{F}_{\tau\wedge k}\right), \tag{2.4.12}$$

Eq. (2.4.11) is true for $E\in\mathcal{F}_\tau$ as well as elements in $\mathcal{F}_{\tau\wedge k}$ for each $k\in\mathbb{N}$.

Information inequality at stopping time

Based on the re-weighting formula (2.4.5), we explain the information inequalities at stopping times. For a stopping time τ, if $P_\theta(\tau<\infty)=1$ for any θ, the following inequality holds:

$$\frac{(\partial_\theta\mathbb{E}_\theta[A_\tau])^2}{\mathrm{Var}_\theta[A_\tau]} \le I_\tau(\theta), \tag{2.4.13}$$

where $I_\tau(\theta) = \mathbb{E}_\theta[-\partial_\theta^2\ln p_\theta^\tau]$ is called the *Fisher information* of the family $\{P_\theta\}_\theta$. This inequality is called *information inequality* or *Cramér-Rao inequality*. The linear response property of the observable at the stopping time τ against the perturbation $\theta\to\theta+d\theta$ is captured by the derivative $\partial_\theta\mathbb{E}_\theta[A_\tau]$. Therefore, the relative amplitude of the linear response against the fluctuation is quantified by the left-hand side of Eq. (2.4.13). Since the Fisher information is independent of the choice of the observable, it gives a *uniform* upper bound on the ratio of the linear response to the fluctuation.

We prove a modified form of Eq. (2.4.13):

$$\frac{(\partial_\theta\widetilde{\mathbb{E}}_\theta[A_\tau])^2}{\widetilde{\mathrm{Var}}_\theta[A_\tau]} \le \widetilde{I}_\tau(\theta), \tag{2.4.14}$$

where we have defined the modified Fisher information $\widetilde{I}_\tau(\theta) \coloneqq \widetilde{\mathbb{E}}_\theta[-\partial_\theta^2\ln p_\theta^\tau]$. A crucial assumption here is that $\widetilde{\mathbb{E}}_\theta[1]$ is independent of θ, i.e.,

$$\partial_\theta\widetilde{\mathbb{E}}_\theta[1] = \partial_\theta P_\theta(\tau<\infty) = 0. \tag{2.4.15}$$

This assumption means that the probabilities that the stopping time diverges have the same value over all parameters. Under this assumption and several regularity conditions, we obtain the inequality,

$$\begin{aligned}
\partial_\theta\widetilde{\mathbb{E}}_\theta(A_\tau) &= \widetilde{\mathbb{E}}_0[p_\theta^\tau(\partial_\theta\ln p_\theta^\tau)A_\tau] \\
&= \widetilde{\mathbb{E}}_\theta[(\partial_\theta\ln p_\theta^\tau)A_\tau] \\
&= \widetilde{\mathbb{E}}_\theta[(\partial_\theta\ln p_\theta^\tau)(A_\tau-\widetilde{\mathbb{E}}_\theta[A_\tau])] \\
&\le \sqrt{\widetilde{\mathbb{E}}_\theta[(\partial_\theta\ln p_\theta^\tau)^2]}\sqrt{\widetilde{\mathrm{Var}}_\theta[A_\tau]},
\end{aligned} \tag{2.4.16}$$

where we have used (2.4.6) and $\partial_\theta p_\theta^\tau = p_\theta^\tau \partial_\theta \ln p_\theta^\tau$ in the first line, (2.4.6) again in the second line, $\widetilde{\mathbb{E}}_\theta[\partial_\theta \ln p_\theta^\tau] = \widetilde{\mathbb{E}}_0[\partial_\theta p_\theta^\tau] = \partial_\theta \widetilde{\mathbb{E}}_\theta[1]$ and the assumption (2.4.15) in the third line, and the Cauchy-Schwarz inequality in the fourth line. By using a similar argument, we find that

$$\widetilde{\mathbb{E}}_\theta[(\partial_\theta \ln p_\theta^\tau)^2] = \widetilde{\mathbb{E}}_\theta[-\partial_\theta^2 \ln p_\theta^\tau]. \tag{2.4.17}$$

Hence, we have the information inequality at the stopping time, Eq. (2.4.14). If the stopping time τ is finite with probability one for any parameter θ, the above inequality simplifies to the form (2.4.13).

2.4.2 Application of information inequalities: the kinetic uncertainty relation

Before giving a proof of (2.2.5), we explain the idea behind our derivation via the derivation of the kinetic uncertainty relation at a fixed time [TB19]. The essential step of the derivation is to find a perturbation, or equivalently, a family of modified processes $\{P_\theta : \theta\}$ that satisfies two conditions:

$$\partial_\theta \mathbb{E}_\theta[J_t]|_{\theta=0} = \mathbb{E}[J_t], \tag{2.4.18}$$

and

$$I_t(0) = \mathbb{E}[N_t]. \tag{2.4.19}$$

If we find such a family, by applying the information inequality (2.4.13) with $\tau = t$ and $\theta = 0$, we obtain the KUR (2.2.3).

According to Ref. [TB19], the time rescaled process satisfies the above two conditions. To see this, we define the modified processes with transition rates $k_{e,\theta} := (1+\theta)k_e(x,y)$, where θ is the real parameter. The escape rates for the modified process are given by $\lambda_\theta(x) = (1+\theta)\lambda(x)$. This modification of the rate clearly corresponds to a rescaling of time in the stochastic process. The distribution $P_0 = P$ corresponds to the original stochastic process. Since the process $\{Y_t^\theta := X_{(1+\theta)t} : t \geq 0\}$ has the distribution P_θ, we have that

$$\mathcal{J}_\theta(t) := \mathbb{E}_\theta[J_t] = \mathbb{E}[J_{(1+\theta)t}] =: \mathcal{J}((1+\theta)t). \tag{2.4.20}$$

Therefore, we obtain $\partial_\theta \mathbb{E}_\theta[J_t]|_{\theta=0} = t d_t \mathcal{J}(t)$. If the system starts from the steady state, this equation becomes $\partial_\theta \mathbb{E}_\theta[J_t]|_{\theta=0} = \mathcal{J}(t)$ because $\mathcal{J}(t) = t j_{\rm ss}$, where $j_{\rm ss}$ is the averaged current in the steady state.

Let us calculate the Fisher information for these modified processes. The logarithm of the likelihood ratio function $p_\theta^t = dP_\theta^t / dP_0^t$ is given by

$$\begin{aligned} \ln p_\theta^t &= \sum_{i=1}^{N_t} \ln \frac{k_{e_i,\theta}(x_{i-1}, x_i)}{k_{e_i}(x_{i-1}, x_i)} - \int_0^t [\lambda_\theta(X_s) - \lambda(X_s)] ds \\ &= N_t \ln(1+\theta) - \theta \int_0^t \lambda(X_s) ds \end{aligned} \tag{2.4.21}$$

Therefore, we obtain that $-\partial_\theta^2 \ln p_\theta^t = (1+\theta)^{-2} N_t$ and $I_t(\theta) = (1+\theta)^{-2}\mathbb{E}_\theta[N_\tau]$. Consequently, by applying the information inequality to the time-rescaled processes, we obtain the KUR at a fixed time (2.2.3).

2.4.3 Proof of (2.2.5)

We present a derivation of (2.2.6) via the information inequality at stopping times (2.4.14). For a time-integrated current J_t of the form (2.2.1), we consider the first passage time $\tau \coloneqq \inf\{t \geq 0 : J_t > J_{\rm th}\}$ at which the accumulated current J_t first reaches a certain threshold $J_{\rm th}$. Similarly to the above derivation, let us find a family of modified processes that satisfies

$$\partial_\theta \widetilde{\mathbb{E}}[\tau]|_{\theta=0} = \widetilde{\mathbb{E}}[\tau] \tag{2.4.22}$$

and

$$\widetilde{I}_\tau(0) = \widetilde{\mathbb{E}}[N_\tau]. \tag{2.4.23}$$

Recall that the Fisher information at a fixed time for the time-rescaled processes is the dynamical activity. Therefore, we use the time-rescaled processes as modified processes. From (2.4.21), the Fisher information at the stopping time for the time-rescaled processes is given by $\widetilde{I}_\tau(\theta) = (1+\theta)^{-2}\widetilde{\mathbb{E}}_\theta[N_\tau]$. The second condition is satisfied.

We check that the time-rescaled processes satisfies the first condition. Since the observable J_t has the property $J[(Y_s \coloneqq X_{cs})_{s\in[0,c^{-1}t]}] = J_t$ for any $c > 0$, we have that

$$\tau[(X_{cs})_{s\in[0,\infty)}] = c^{-1}\tau[(X_s)_{s\in[0,\infty)}]. \tag{2.4.24}$$

This means that the first passage time for the time-integrated current scales by factor c^{-1} because of the rescaling of time for the stochastic trajectories $X_s \mapsto X_{cs}$. Therefore, the cumulative distribution function of τ with respect to P_θ and P_0 are related through

$$\begin{aligned}
P_{\theta,x_0}(\tau \leq t) &= \sum_{n=1}^{\infty} P_{\theta,x_0}(\tau \leq t \wedge N_t = n) \\
&= \sum_{n=1}^{\infty} \sum_{(x_1,e_1:\cdots:x_n,e_n)} \int_{0<t_1<\cdots<t_n<t} dt_1 \dots dt_n \; 1_{\{\tau\leq t\}}[(x_s)_{s\in[0,t]}] \\
&\quad \times e^{-\lambda_\theta(x_0)\cdot t_1} k_{e_1,\theta}(x_0,x_1) e^{-\lambda_\theta(x_1)\cdot(t_2-t_1)} \dots k_{e_n,\theta}(x_{n-1},x_n) e^{-\lambda_\theta(x_n)\cdot(t-t_n)} \\
&= \sum_{n=1}^{\infty} \sum_{(x_1,e_1:\cdots:x_n,e_n)} \int_{0<t_1<\cdots<t_n<t} dt_1 \dots dt_n \; 1_{\{\tau\leq t\}}[(x_s)_{s\in[0,t]}] \\
&\quad \times (1+\theta)^n e^{-\lambda(x_0)\cdot(1+\theta)t_1} k_{e_1}(x_0,x_1) e^{-\lambda(x_1)\cdot(1+\theta)(t_2-t_1)} \dots k_{e_n}(x_{n-1},x_n) e^{-\lambda(x_n)\cdot(1+\theta)(t-t_n)} \\
&= \sum_{n=1}^{\infty} \sum_{(x_1,e_1:\cdots:x_n,e_n)} \int_{0<t_1'<\cdots<t_n'<(1+\theta)t} dt_1' \dots dt_n' \; 1_{\{\tau\leq(1+\theta)t\}}[(x_{(1+\theta)^{-1}s})_{s\in[0,(1+\theta)t]}] \\
&\quad \times e^{-\lambda(x_0)\cdot t_1'} k_{e_1}(x_0,x_1) e^{-\lambda(x_1)\cdot(t_2'-t_1')} \dots k_{e_n}(x_{n-1},x_n) e^{-\lambda(x_n)\cdot((1+\theta)t-t_n')} \\
&= P_{x_0}(\tau \leq (1+\theta)t).
\end{aligned}$$

This relation implies that the unnormalized probability density function $f_\theta(t)$ of τ satisfies $f_\theta(t) = (1+\theta)f_0((1+\theta)t)$. Hence, we have that $\widetilde{\mathbb{E}}_{\theta,x_0}[\tau] = (1+\theta)^{-1}\widetilde{\mathbb{E}}_{x_0}[\tau]$ and $\partial_\theta \widetilde{\mathbb{E}}_{\theta,x_0}[\tau] = -(1+\theta)^{-2}\widetilde{\mathbb{E}}_{x_0}[\tau]$.

The rescaling of time with finite factor $(1+\theta)$ does not change the probability that the first passage time diverges, i.e., $\widetilde{\mathbb{E}}_\theta[1]$ is independent of θ. Therefore, we can apply the information inequality (2.4.14) to this modified processes. By taking $\theta = 0$ in (2.4.14), we have the KUR (2.2.6). If $P(\tau < \infty) = 1$, the KUR (2.2.5) is recovered.

2.5 Discussion

We have derived the kinetic bound (2.2.5) on the first passage time for time-integrated current that is valid for any finite threshold. In contrast to TURs, KUR may be relevant for a system far from equilibrium. An interesting challenge is to apply our result to biological systems such as circadian clocks, molecular motors and enzyme kinetics, and measure the efficiency of these systems from the perspective of the precision of the first passage time.

Refs. [GH17, Gar17] use the connection between the rate functions for current statistics and first-passage-time statistics to derive the same type of inequality. Nevertheless, our derivation is based on the idea given in [DS20, TB19] that finite-time TUR and KUR are obtained from the information inequalities for virtually perturbed systems. This method significantly simplifies the derivation and extends the range of applicability. Indeed, recently, the uncertainty relation on the precision of the first passage time in Markovian open quantum systems was derived based on the same technique. The extension of the uncertainty relation on the first passage time to diffusion processes will be studied in future work.

Finally, we make a remark on the derivation of the thermodynamic uncertainty relation from the information inequality. In Ref. [DS20], Dechant and Sasa derived the finite-time TUR on a time-integrated current from the information inequality. Specifically, they found a family of stationary continuous-time Markov chains $\{P_\theta\}$ for which $\mathbb{E}_\theta[J_t] = (1+\theta)\mathbb{E}[J_t]$ and $I_t(\theta) \leq \mathbb{E}[\Sigma_t]/2$, where Σ_t is the total entropy production over $[0,t]$. Although one may expect that the mean first passage time for J_t has rescaling property $\mathbb{E}_\theta[\tau] = (1+\theta)^{-1}\mathbb{E}[\tau]$ for this family and the finite-threshold TUR on the FPT can be derived from (2.4.13), it is not true. This is because the statistics of the first passage time depends on the transition probability. Although the perturbation considered in [DS20] corresponds to the time-rescaling of the single-time probability distribution and current, it does not have the same property at the level of the transition probability. Nevertheless, TUR-type bounds were derived for first passage times to reach a specific state based on more complicated perturbations [PRR21]. The upper bounds on precisions of first passage times contains kinetic contributions associated with absorbing states as well as the entropic contributions, which is consistent with the conclusion in the subsection 2.3.1. More recently, Shiraishi [Shi21] found the optimal upper bound on the precision of the time-integrated currents that is derived from the information inequality and presented a unifying picture of the thermodynamic and kinetic uncertainty relation. Whether the same picture can be applied to the precision of the first passage time remains an open problem.

Appendix 2.A Mean dynamical activity in random walk

The random walker X_t is described in the form,

$$X_t = \sum_{i=1}^{N_t} Z_i. \tag{2.A.1}$$

Here N_t is the number of jumps over $[0, t]$ and $(Z_1, Z_2, \dots)$ are random variables independently and identically distributed according to $P(Z_i = \pm 1) = k_\pm/(k_+ + k_-)$. Let us consider the first passage time $\tau_x = \inf\{t \geq 0 : X_t = x\}$. Because condition $X_t = x$ is equivalent to $\sum_{i=1}^{N_t} Z_i = x$, N_{τ_x} is a first passage time at which the discrete-time stochastic process $(Z_1, Z_2, \dots)$ first reaches the threshold x. If $x > 0$, the expectation value of N_{τ_x} is finite. By applying the Wald identity [GMS97, Theorem 2.4.4], we obtain $\mathbb{E}[X_{\tau_x}] = \mathbb{E}[N_{\tau_x}] \cdot \mathbb{E}[Z_i]$. Hence,

$$\mathbb{E}[N_{\tau_x}] = \frac{\mathbb{E}[X_{\tau_x}]}{\mathbb{E}[Z_i]} = \frac{x}{\tanh(\epsilon/2)}. \tag{2.A.2}$$

Chapter 3

Macroscopic Law for Algorithmic Random States

3.1 Introduction

3.1.1 Summary of the results

In this chapter, we demonstrate with the aid of a pedagogical model how the emergence of macroscopic irreversible laws from reversible microscopic dynamics is formulated in terms of algorithmic randomness. We expect that the following results hold true for a wide class of models although we investigate a specific model in this chapter.

We study a variant of the Kac ring model [Kac59, GO09, MNS09], which consists of two kinds of degrees of freedom, particles with spin $2x(i) - 1 \in \{-1, 1\}$ ($i \in \mathbb{Z}$) on a one-dimensional infinite lattice $\mathbb{Z}$ and scatterers $y(i) \in \{0, 1\}$ located between particles. At each discrete time step, a particle at site $i \in \mathbb{Z}$ moves to site $i + 1$. Then, the bit $x(i)$ is flipped if the scatterer at site i is present, $y(i) = 1$, and it remains its value if absent, $y(i) = 0$. This evolution rule $\varphi : \{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}} \to \{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$ defines a discrete-time, deterministic and reversible dynamical system on $\{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$. If we choose a set of macroscopic variables $m = (m_0, m_1)$ as the average magnetization m_0^N and the fraction of scatterers m_1^N over $2N+1$ sites around the origin, the system exhibits deterministic and irreversible behavior in the sense of the law of large numbers: *for almost all* (x, y) *with respect to* $\mu_{(1+m_0)/2} \times \mu_{m_1}$,

$$\lim_{N\to\infty} (m_i^N \circ \varphi^t)(x, y) = \Phi_i^t(m) \text{ for all } i \in \{0, 1\} \text{ and } t \in \{0, \dots, T\}, \tag{3.1.1}$$

where μ_p is the Bernoulli measure on $\{0, 1\}^{\mathbb{Z}}$ with parameter $p \in [0, 1]$. $\varphi^t(x, y)$ is the microscopic state at time t starting from a microstate (x, y). $\Phi(m) = ((1-2m_1)m_0, m_1)$ represents a macroscopic law in the model. Our main claim of this chapter is that algorithmically random microstates with respect to the initial probability measure satisfy the macroscopic law in the thermodynamic limit. That is to say, *for Martin-Löf random* (x, y) *with respect to* $\mu_{(1+m_0)/2} \times \mu_{m_1}$,

$$\lim_{N\to\infty} (m_i^N \circ \varphi^t)(x, y) = \Phi_i^t(m) \text{ for all } i \in \{0, 1\} \text{ and } t \in \{0, \dots, T\}. \tag{3.1.2}$$

This result implies the zeroth law of thermodynamics for individual random microstates. Thus, the notion of algorithmic randomness opens the possibility of formulating macroscopic properties such as hydrodynamic equations and the zeroth law of thermodynamics at the level of individual microscopic states.

To quantify the irreversibility for individual trajectories, we define a quantity called the irreversible information loss as the logarithm of the ratio of probabilities at time t of a microscopic state $\varphi^t(x, y)$ and the time-reversed one $(\pi \circ \varphi^t)(x, y)$ [SK00]. We prove that the irreversible information loss is positive for any random state, which implies the difficulty of realizing the time-reversed state of a random state in a measure-theoretic sense. The randomness notion sheds light on another aspect of the reversibility paradox. We show that the time-reversed state of a random microstate is not random as well as it violates the macroscopic law.

3.1.2 Previous studies

There are a few works that have applied the theory of algorithmic randomness to statistical physics. The basic idea of such previous studies is to employ the Kolmogorov complexity for a microscopic state, which is the shortest program length outputting the state, and to present a formulation combining the Shannon entropy with the Kolmogorov complexity [Ben82]. For example, a new definition of entropy for microstates was proposed to provide some insight into Maxwell's demon problems [Zur89b, Zur89a, Cav93]. It should be noted that the Kolmogorov complexity is independent of the probability measure, while the Martin-Löf randomness is defined for a probability measure. The most important relation between the two concepts is that an infinite sequence is Martin-Löf random with respect to a probability measure μ if and only if the Kolmogorov complexity of the sequence is not smaller than the optimal compression length under the probability measure, $-\log \mu$, calculated from Shannon information theory. See Theorem 3.2.4.6 for the precise statement. Therefore, the difference of the optimal compression length from the Kolmogorov complexity, which is referred to as the randomness deficiency [Gác, Gác94], is the most important quantity to identify the Martin-Löf randomness. By using the randomness deficiency, we can express our statement as "The randomness deficiency for an initial state diverges if the macroscopic behavior does not obey a macroscopic law." As far as we know, no statement using the Kolmogorov complexity of initial microstates has been addressed for describing the macroscopic irreversibility.

3.1.3 Outline of the chapter

The remainder of the chapter is organized as follows.

In section 3.2, we review the theory of algorithmic randomness. To explain it in a self-contained manner, we include a brief review of computability theory and measure theory on the binary Cantor space. In section 3.3, we first introduce a variant of the Kac ring model. We prove the law of large numbers in a measure-theoretic sense. With this in mind, we provide the pointwise version of the law of large numbers on the basis of algorithmic randomness. In section 3.4, we define the Shannon and Boltzmann entropies. The pointwise law of large

numbers leads to a pointwise version of the zeroth law of thermodynamics. In section 3.5, we investigate the consequence of microscopic reversibility. We define a quantity called irreversible information loss quantifying the asymmetry between a microscopic trajectory and the time-reversed one, and prove the positivity of this quantity for random states. By using the reversibility property of microscopic dynamics, we construct a probability measure with respect to which the Boltzmann entropy decreases along the typical macroscopic trajectory. Similarly, we prove the nonrandomness of time-reversed states. In section 3.6, we conclude with open problems and related topics.

3.1.4 Notations

We use the following notations throughout this chapter.

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Q}_{\geq 0}$, and $\mathbb{R}_{\geq 0}$ denote the set of natural numbers, integers, rational numbers, real numbers, nonnegative rational and real numbers, respectively. Let $\{0,1\}^{\mathbb{N}}$ denote the set of all infinite binary sequences, which is identified the set of all functions from $\mathbb{N}$ to $\{0,1\}$, $\{0,1\}^{<\mathbb{N}}$ the set of finite binary strings including the empty string $\Box$, $|\sigma|$ the length of a string $\sigma \in \{0,1\}^{<\mathbb{N}}$, and $\sigma\tau$ the concatenation of finite string σ and finite or infinite string τ. A subset of natural numbers $A \subseteq \mathbb{N}$ is identified with its characteristic function $\chi_A \in \{0,1\}^{\mathbb{N}}$. For a finite string σ and finite or infinite string τ, we let $\sigma \sqsubseteq \tau$ denote that σ is a prefix of τ. For a finite or infinite string x, $x(n)$ denotes the n-th element of x and $x \upharpoonright n$ or $x(0:n-1)$ its first n bits $x(0)x(1)\dots x(n-1)$.

3.2 Preliminaries

In this section, we review the algorithmic theory of randomness. Since this theory is based on computability theory, we also provide a brief review of computability theory. We hope that the chapter will be read by theoretical physicists who are unfamiliar with computability theory. This section includes only a minimal set of concepts necessary for reading this chapter and omits proofs of theorems. For more details of topics and proofs of theorems, see [LV08, Nie09, DH10, Gác] for the theory of algorithmic randomness and [Coo04, Odi99a, Odi99b] for the computability theory.

3.2.1 Computability theory

A function from a subset $A \subseteq \{0,1\}^{<\mathbb{N}}$ to $\{0,1\}^{<\mathbb{N}}$ is called a *partial function* on $\{0,1\}^{<\mathbb{N}}$ and is denoted by $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$. The subset A is called the domain of f and is denoted by $\mathrm{dom}(f)$. The range of f is denoted by $\mathrm{range}(f)$. If $A = \{0,1\}^{<\mathbb{N}}$, f is called *total* and is denoted by $f : \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$. A central concept of computability theory is the following.

Definition 3.2.1.1 (computable function)**.** A partial function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ is *computable* if there exists a Turing machine M such that M computes f.

Informally, each Turing machine M represents a computer program. A partial function f is computable if there is a program or algorithm such that for any input string $\sigma \in \{0,1\}^{<\mathbb{N}}$, it either outputs $f(\sigma)$ if $f(\sigma)$ is defined, or it outputs nothing if $f(\sigma)$ is not defined. If we choose a coding function from $\{0,1\}^{<\mathbb{N}}$ to a countable object such as natural numbers, finite tuples of natural numbers, integers and rational numbers, we can extend the notion of computability of functions on $\{0,1\}^{<\mathbb{N}}$ to functions on the object. For instance, we can represent a natural number $n \in \mathbb{N}$ as a binary string $\beta(n) \in \{0,1\}^{<\mathbb{N}}$ by using the binary expansion. A partial function f on $\mathbb{N}$ is called computable if there is a partial computable function $g :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ with $g \circ \beta = \beta \circ f$. Similarly, a function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \mathbb{Q}$ is computable if there exists computable functions $\delta \times p \times q :\subseteq \{0,1\}^{<\mathbb{N}} \times \{0,1\}^{<\mathbb{N}} \times \{0,1\}^{<\mathbb{N}} \to \{0,1\} \times \mathbb{N} \times \mathbb{N}\backslash\{0\}$ such that $f(\sigma) = (-1)^{\delta(\sigma)} p(\sigma)/q(\sigma)$ for any $\sigma \in \{0,1\}^{<\mathbb{N}}$. All functions implemented in modern computers such as addition, multiplication, subtraction, division, and bounded summation are computable.

A set $A \subseteq \{0,1\}^{<\mathbb{N}}$ is computable if its characteristic function $\chi_A : \{0,1\}^{<\mathbb{N}} \to \{0,1\}$ is computable. For instance, the set of all prime numbers is computable because there is an algorithm that decides whether a given natural number is a prime number or not. To formulate the notion of algorithmic randomness, we use a weaker notion of the computability of sets.

Definition 3.2.1.2 (computably enumerable)**.** A set $A \subseteq \{0,1\}^{<\mathbb{N}}$ is *computably enumerable* (*c.e.*) if there exists a partial computable function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ such that $A = \mathrm{range}(f)$.

This means that there exists an algorithm enumerating or listing all the members of the set. For example, for a polynomial $p(y_1, y_2)$ with integer coefficients, $D = \{x \in \mathbb{N} : \exists y_1, y_2 \in \mathbb{N}\ p(y_1, y_2) = x\}$ may not be computable but is computably enumerable. It is easy to prove that $A \subseteq \{0,1\}^{<\mathbb{N}}$ is computable if and only if both A and $\bar{A}$ are c.e. In particular, if A is computable, then A is c.e. Computable enumerability is a properly weaker notion than computability because there is a set that is computably enumerable but not computable. Examples of such sets are the halting problem of Turing machines and Hilbert's tenth problem.

We also define uniformly computable enumerability of sequences of sets.

Definition 3.2.1.3 (uniformly c.e.)**.** A sequence $(A_n)_{n\in\mathbb{N}}$ of sets $A_n \subseteq \{0,1\}^{<\mathbb{N}}$ is *computably enumerable uniformly in* n if there exists a partial computable function $f :\subseteq \{0,1\}^{<\mathbb{N}} \times \mathbb{N} \to \{0,1\}^{<\mathbb{N}}$ such that $A_n = \mathrm{range}(f(\cdot, n))$ for all $n \in \mathbb{N}$.

A Turing machine is a special-purpose machine in the sense that the machine computes one computable function. Since we can code a program by a natural number in a computable manner, we can construct a *universal Turing machine*, which is a model of present-day computers. This is why we can now implement any program by using only one computer.

Theorem 3.2.1.4 (universal Turing machine)**.** There is a partial computable function of two variables g such that $g(e, x) = f_e(x)$ for any input x and any partial computable function f_e indexed by a natural number e.

There is no coding that maps from finite strings to real numbers because the set of all real numbers is uncountable. Therefore, we say that a real number is computable if there exists a sequence of rationals approximating the real number from below and above in a computable way.

Definition 3.2.1.5 (computable real, computable real-valued function)**.** A real number $x \in \mathbb{R}$ is *lower semicomputable* if the set $\{q \in \mathbb{Q} : q < x\}$ is computably enumerable. x is *upper semicomputable* if $-x$ is lower semicomputable. x is *computable* if it is both lower and upper semicomputable. Similarly, a real-valued function $f : \{0,1\}^{<\mathbb{N}} \to \mathbb{R}$ is *lower semicomputable* if the set $\{(\sigma, q) \in \{0,1\}^{<\mathbb{N}} \times \mathbb{Q} : q < f(\sigma)\}$ is computably enumerable. f is *upper semicomputable* if $-f$ is lower semicomputable. f is *computable* if it is both lower and upper semicomputable.

3.2.2 Topology and measure theory in Cantor space

We review the basics of topology and measure theory on the set of infinite binary sequences. For a finite string $\sigma \in 2^{<\mathbb{N}}$, we use $[\sigma]$ to denote the cylinder set, that is, the set of all infinite binary sequences whose prefix is σ, $[\sigma] = \{\sigma\tau : \tau \in \{0,1\}^{\mathbb{N}}\}$. For $S \subseteq \{0,1\}^{<\mathbb{N}}$, we let $[S] = \bigcup_{\sigma \in S} [\sigma]$.

Definition 3.2.2.1 (c.e. open)**.** The Cantor space is $\{0,1\}^{\mathbb{N}}$ equipped with the product topology of a countable number of copies of the discrete topological space $\{0,1\}$. The Cantor space has a countable basis of cylinder sets $\{[\sigma] : \sigma \in \{0,1\}^{<\mathbb{N}}\}$. A subset $A \subseteq \{0,1\}^{\mathbb{N}}$ is *open* if it is the union of a subset of cylinder sets, that is,

$$A = [S] = \bigcup_{\sigma \in S} [\sigma] \tag{3.2.1}$$

for some subset of strings $S \subseteq \{0,1\}^{<\mathbb{N}}$. If there exists a computably enumerable set S such that $A = [S]$, then A is called *c.e. open*. A sequence $(A_n)_{n \in \mathbb{N}}$ of sets $A_n \subseteq \{0,1\}^{\mathbb{N}}$ is *c.e. open uniformly in n* if there exists a sequence $(S_n)_{n \in \mathbb{N}}$ of c.e. sets uniformly in n such that $A_n = [S_n]$ for all n.

Let $(\{0,1\}^{\mathbb{N}}, \mathcal{B})$ be the measurable space, where $\mathcal{B}$ is the Borel σ-algebra. It is known that a probability measure on $(\{0,1\}^{\mathbb{N}}, \mathcal{B})$ can be constructed from a premeasure on $\{[\sigma] : \sigma \in \{0,1\}^{<\mathbb{N}}\}$ with the aid of the Carathéodory's extension theorem. In the following, μ_ρ denotes the induced probability measure from a premeasure ρ and is identified with the premeasure.

Definition 3.2.2.2 (premeasure, computable measure)**.** A *probability premeasure* is a function $\rho : \{0,1\}^{<\mathbb{N}} \to \mathbb{R}_{\geq 0}$ such that $\rho(\square) = 1$ and $\rho(\sigma 0) + \rho(\sigma 1) = \rho(\sigma)$ for all $\sigma \in \{0,1\}^{<\mathbb{N}}$. The induced probability measure μ_ρ is *computable* if ρ is computable as a real-valued function.

Example 3.2.2.3.

(1) A premeasure $\rho(\sigma) = 2^{-|\sigma|}$ for $\sigma \in \{0,1\}^{<\mathbb{N}}$ induces the *uniform measure* or the *Lebesgue measure* λ.

(2) Let p be a real number such that $p \in (0,1)$. We set $\rho(1) = p$, $\rho(0) = 1-p$, and define a probability premeasure $\rho_p : \{0,1\}^{<\mathbb{N}} \to \mathbb{R}_{\geq 0}$ by

$$\rho_p(\sigma) = \prod_{i=0}^{|\sigma|-1} \rho(\sigma(i)). \tag{3.2.2}$$

We call the induced measure μ_{ρ_p} the *Bernoulli measure* of parameter $\bar{p}$, which is denoted simply by μ_p. The Bernoulli measure μ_p of parameter p is computable if and only if p is a computable real. We note that the Bernoulli measure with $p = 1/2$ is the uniform measure λ.

(3) Let $x \in \{0,1\}^{\mathbb{N}}$ be a sequence. The *Dirac measure* δ_x concentrated on x is induced by the premeasure

$$\rho_x(\sigma) = \begin{cases} 1 & \text{if } \sigma \sqsubseteq x \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.3}$$

For any measurable set $A \subseteq \{0,1\}^{\mathbb{N}}$,

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases} \tag{3.2.4}$$

We use the first Borel-Cantelli lemma to prove the strong form of the law of large numbers. Additionally, Martin-Löf randomness has an alternative characterization in terms of the effective version of the Borel-Cantelli lemma (see Definition 3.2.3.7).

Lemma 3.2.2.4 (first Borel-Cantelli lemma)**.** Let $(C_n)_{n\in\mathbb{N}}$ be a sequence of measurable sets. If $\sum_{n=0}^{\infty} \mu(C_n) < \infty$, then

$$\mu(\{x : x \in C_n \text{ for infinitely many } n\}) = \mu\left(\bigcap_{n=0}^{\infty} \bigcup_{k\geq n}^{\infty} C_k\right) = 0 \tag{3.2.5}$$

3.2.3 Martin-Löf randomness

In the probability theory, any realization of a stochastic process is assumed to occur randomly. In an n times fair coin tossing experiment, a realization $0^n = 00\ldots0$ (n zeros) has the same probability 2^{-n} as any other realization. There is no difference between all realizations in this sense. Nevertheless, we believe that the relative frequencies of heads and tails approach asymptotically to $1/2$ as $n \to \infty$ under this experiment. This belief is represented by the strong law of large numbers in the probability theory. Although $0^{\mathbb{N}} = 00\ldots$ (infinitely many zeros) is a realizable outcome, it is not *random* in that it does not satisfy the law of large numbers. Thus, it is possible to distinguish between random sequences and nonrandom ones according to the statistical laws that have the probability one. In other words, the

notion of random sequences generated by a stochastic process is defined as ones having typical properties, or equivalently, having no exceptional properties. However, it is not clear what class of typical properties or exceptional properties we should choose to define random sequences. For instance, although $(01)^{\mathbb{N}} = 010101\ldots$ satisfies the law of large numbers, our intuition tells us that it is not a typical sequence generated by a fair coin tossing and therefore should not be random. Even if we require that the law of large numbers should hold for subsequences selected from a whole sequence by countable rules, there is a sequence satisfying the requirement but violating the law of the iterated logarithm, which is known as Ville's counterexample [Vil39] (see also Theorem 6.5.1 in [DH10]). Hence, just the law of large numbers is not enough to characterize randomness. One naive idea is to consider all exceptional properties. We then define a set describing an exceptional property.

Definition 3.2.3.1 (null set)**.** A set $N \subseteq \{0,1\}^{\mathbb{N}}$ is a *null set* with respect to a probability measure μ if there is a sequence $(U_n)_{n\in\mathbb{N}}$ of open sets such that $N \subseteq \bigcap_{n\in\mathbb{N}} U_n$ and $\mu(U_n) \leq 2^{-n}$.

Example 3.2.3.2.

(1) For $x \in \{0,1\}^{\mathbb{N}}$, the one-element set $\{x\}$ is a null set with respect to λ. Indeed, $\{x\} = \bigcap_{n\in\mathbb{N}} U_n$, where $U_n = [x \upharpoonright n]$ with $\lambda(U_n) = 2^{-n}$.

(2) $N = \{x \in \{0,1\}^{\mathbb{N}} : x(2n) = 1 \text{ for all } n \in \mathbb{N}\}$ is a null set with respect to λ. Indeed, $N = \bigcap_{n\in\mathbb{N}} U_n$, where $U_n = \{x \in \{0,1\}^{\mathbb{N}} : x(2k) = 1,\ 0 \leq k < n\}$ with $\lambda(U_n) = 2^{-n}$.

Example 3.2.3.2 (1) shows that the naive idea fails because there is no sequence not contained in all null sets. To obtain a meaningful definition of random sequences, we have to restrict the class of null sets. The definition must satisfy the following two requirements at least.

1. The set of random sequences is typical in measure-theoretic sense, that is, it has probability one.

2. Sequences generated by some simple rule such as $(01)^{\mathbb{N}}$ are not random with respect to λ.

We should note that a countable union of null sets is also a null set.

Proposition 3.2.3.3. Let $(N_e)_{e\in\mathbb{N}}$ be a sequence of null sets with respect to a probability measure μ. The countable union $\bigcup_{e\in\mathbb{N}} N_e$ of the sequence is a null set.

Therefore, if we choose a countable family of null sets to define random sequences, the first condition is automatically satisfied. We should also impose some computability conditions on the null sets if we interpret the generation by simple rules as listing elements in an algorithmic manner. The above argument motivates the following definition.

Definition 3.2.3.4 (Martin-Löf random [ML66])**.** Let μ be a computable probability measure on $\{0,1\}^{\mathbb{N}}$. A *Martin-Löf test* with respect to the measure μ (ML μ-test) is a sequence $(U_n)_{n\in\mathbb{N}}$ of c.e. open sets uniformly in n such that $\mu(U_n) \leq 2^{-n}$ for all $n \in \mathbb{N}$. A set $N \subseteq \{0,1\}^{\mathbb{N}}$ is called a Martin-Löf null set with respect to μ (ML μ-null set) if there is a Martin-Löf test $(U_n)_{n\in\mathbb{N}}$ such that $N \subseteq \bigcap_{n\in\mathbb{N}} U_n$. A sequence $x \in \{0,1\}^{\mathbb{N}}$ is Martin-Löf random with respect to μ (ML μ-random) if $\{x\}$ is not a Martin-Löf null set. MLR_μ denotes the set of ML μ-random sequences.

Martin-Löf randomness satisfies the first requirement. Indeed, there are only countably many ML tests because there are only countably many c.e. sets. Since the union of all ML μ-null sets is a null set with respect to μ from Proposition 3.2.3.3, μ-almost every sequence is ML μ-random.

Theorem 3.2.3.5. $\mu(\mathsf{MLR}_\mu) = 1$.

We present an interpretation of Definition 3.2.3.4 from the viewpoint of the hypothesis testing. In a casino, a gambler Skeptic and a dealer Reality are about to play a game using an unbiased coin. However, the coin is prepared by Reality and Skeptic doubts whether or not the coin is really unbiased. Skeptic needs to make sure of the unbiasedness of the coin in some way. One simple way to check this is to toss the coin many times and to test whether or not the generated sequence $\omega \in \{0,1\}^{\mathbb{N}}$ is random. To this end, we apply the idea in hypothesis testing. Skeptic prepare a subset $C \subseteq \{0,1\}^{\mathbb{N}}$ consisting of the sequences that are clearly non-random. As an example, we here consider the set of sequences whose $2k$-th bits are 1,

$$C = \{\eta \in \{0,1\}^{\mathbb{N}} : (\forall k \in \mathbb{N})(\eta(2k) = 1)\}. \tag{3.2.6}$$

We remark that this subset consists of non-random sequences although it does not cover whole non-random sequences. On the one hand, since Reality claims that the coin is unbiased, the null hypothesis H_0 is expressed as

$$H_0 : \omega \notin C.$$

On the other hand, since Skeptic doubts the claim H_0, the alternative hypothesis H_1 is expressed as

$$H_1 : \omega \in C.$$

To judge H_0 or H_1, Skeptic observes the sequence ω up to the $2(n-1)$-th bits. If there is a bit $k \in \{0,\dots,n-1\}$ such that $\omega(2k) = 0$, Skeptic accept H_0 and judge $\omega \notin C$. If $\omega(2k) = 1$ for all $k \in \{0,\dots,n-1\}$, Skeptic reject H_0 and judge $\omega \in C$. This means that Skeptic take

$$C_n = \{\eta \in \{0,1\}^{\mathbb{N}} : (\forall k \in \{0,\dots,n-1\})(\eta(2k) = 1)\} \tag{3.2.7}$$

as a rejection region. In this testing procedure, although the probability of the type II error is zero, the probability of the type I error is given by $\lambda(C_n) = 2^{-n}$. Therefore, if n is sufficiently

large, the type I error probability is extremely small. Definition 3.2.3.4 applies the idea that a sequence of rejection regions whose probabilities decay to zero defines a test for random sequences.

As we mentioned above, a single test $\{U_n\}$ does not cover whole non-random sequences. To determine whether a given sequence is random or not, we have to implement a countable number of ML tests. However, the existence of a universal Turing machine implies that the union of all ML μ-tests is also a ML μ-test. Such a test is called universal Martin-Löf μ-test.

Theorem 3.2.3.6 (universal Martin-Löf test)**.** There exists a Martin-Löf test $\{U_n\}_{n\in\mathbb{N}}$ with respect to μ such that for any ML μ-test $\{V_n\}_{n\in\mathbb{N}}$, $\bigcap_{n\in\mathbb{N}} V_n \subseteq \bigcap_{n\in\mathbb{N}} U_n$.

Proof. See Fact 3.2.4 in [Nie09] or Theorem 6.2.5 in [DH10] □

ML randomness also satisfies the second requirement. If $x \in \{0,1\}^{\mathbb{N}}$ is computable, then x is not ML λ-random because $(U_n)_{n\in\mathbb{N}} = ([x \upharpoonright n])_{n\in\mathbb{N}}$ is a ML λ-test. We remark that even $z = 0^{\mathbb{N}}$ is ML random with respect to the Dirac measure δ_z concentrated on z.

The notion of randomness can be extended to objects in $\{0,1\}^{\mathbb{Z}}$ and $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$. We fix a bijective coding $\iota : \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{N}}$ in the following. We assign a two-sided infinite binary sequence $x = \ldots x(-1)x(0)x(1)\cdots \in \{0,1\}^{\mathbb{Z}}$ to a one-sided infinite binary sequence

$$\iota(x) = x(0)x(-1)x(1)x(-2)x(2)\cdots \in \{0,1\}^{\mathbb{N}}. \tag{3.2.8}$$

We let μ be a computable probability measure on $\{0,1\}^{\mathbb{Z}}$. We say that $x \in \{0,1\}^{\mathbb{Z}}$ is Martin-Löf random with respect to μ if $\iota(x)$ is ML random with respect to $\mu \circ \iota^{-1}$. Similarly, we define a coding function from $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ to $\{0,1\}^{\mathbb{N}}$ as

$$\kappa(x,y) = x(0)y(0)x(1)y(1)\cdots \in \{0,1\}^{\mathbb{N}} \text{ for } (x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}. \tag{3.2.9}$$

For a computable probability measure μ on $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$, $(x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ is called Martin-Löf random with respect to μ if $\kappa(x,y)$ is ML random with respect to $\mu \circ \kappa^{-1}$.

There is an alternative characterization of Martin-Löf randomness in term of Solovay tests. We use this equivalence to prove the effective law of large numbers.

Definition 3.2.3.7 (Solovay random)**.** Let μ be a computable probability measure. A *Solovay test* is a sequence $(S_n)_{n\in\mathbb{N}}$ of c.e. open sets uniformly in n such that $\sum_n \mu(S_n) < \infty$. $x \in \{0,1\}^{\mathbb{N}}$ is *Solovay μ-random* if x is in only finitely many S_n.

Proposition 3.2.3.8. An element $x \in \{0,1\}^{\mathbb{N}}$ is Martin-Löf μ-random iff it is Solovay μ-random.

Proof. See Proposition 3.2.19 in [Nie09] or Theorem 6.2.8 in [DH10]. □

3.2.4 Robustness of Martin-Löf randomness

Conservation of ML randomness

The definition of Martin-Löf randomness refers to an underlying probability measure. Therefore, even if a sequence $x \in \{0,1\}^{\mathbb{N}}$ is ML μ-random, it may not be random with respect to another measure ν. A trivial example is a sequence that is λ-random but not random with respect to the Bernoulli measure $\mu_{1/3}$. However, ML randomness is a robust notion in the sense that it is preserved by simple transformations. Here a simple transformation means a computable function *from* $\{0,1\}^{\mathbb{N}}$ *to* $\{0,1\}^{\mathbb{N}}$, which is defined through a partial computable monotone function on $\{0,1\}^{<\mathbb{N}}$.

Definition 3.2.4.1. A partial function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ is *monotone* if

$$\sigma \sqsubseteq \tau \Rightarrow f(\sigma) \sqsubseteq f(\tau) \tag{3.2.10}$$

holds for all $\sigma, \tau \in \mathrm{dom}(f)$. For a partial monotone function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$, we define a partial function $\hat{f} :\subseteq \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ as

$$\hat{f}(x) = \begin{cases} \bigcup_{\sigma \sqsubseteq x} f(\sigma) & \text{if } \sup\{|\sigma| : \sigma \sqsubseteq x, \sigma \in \mathrm{dom}(f)\} = \infty \\ \text{undefined} & \text{otherwise} \end{cases} \tag{3.2.11}$$

for $x \in \{0,1\}^{\mathbb{N}}$. A partial function $F :\subseteq \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ is *computable* if there exists a partial computable monotone function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ with $F = \hat{f}$. If $\mathrm{dom}(F) = \{0,1\}^{\mathbb{N}}$, the function F is called *total*. Hereafter, we consider only total (computable) functions on $\{0,1\}^{\mathbb{N}}$.

We define an image measure μF^{-1} of μ under a function $F : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ as

$$(\mu F^{-1})(A) = \mu\left(F^{-1}(A)\right) \tag{3.2.12}$$

for any measurable set $A \in \mathcal{B}$. Then, computable functions on $\{0,1\}^{\mathbb{N}}$ preserve the computability of probability measures and the ML randomness.

Proposition 3.2.4.2. Let μ be a measure on $\{0,1\}^{\mathbb{N}}$ and $F : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ be a total computable function. If μ is computable, μF^{-1} is computable.

Proof. See Lemma 2.6 in [BP12]. □

Theorem 3.2.4.3 (conservation of ML randomness)**.** Let μ be a computable probability measure on $\{0,1\}^{\mathbb{N}}$ and $F : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ a total computable function. If $x \in \{0,1\}^{\mathbb{N}}$ is ML μ-random, then $F(x)$ is ML μF^{-1}-random.

Proof. See Theorem 3.2 in [BP12]. □

Other characterization of ML randomness

Martin-Löf randomness has robustness in another sense. It has other characterizations in terms of incompressibility and unpredictability.

Incompressibility approach Let us consider a sequence. If the sequence has a simple structure, it can be compressed into a shorter length one by using an algorithm. For instance, $(01)^{100000}$ can be transformed to the shorter program "output one hundred thousand 01s ". Conversely, if the sequence is "random", there is no simple description of it. This consideration leads to the idea of the Kolmogorov complexity [Sol64a, Sol64b, Kol65]. Here we use a prefix-free version of the Kolmogorov complexity for technical reasons.

Definition 3.2.4.4 (prefix-free computable function)**.** A set of strings $A \subseteq \{0,1\}^{<\mathbb{N}}$ is called *prefix-free* if for any two distinct elements σ and τ in A, σ is not a prefix of τ. A partial computable function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$ is *prefix-free* if $\mathrm{dom}(f)$ is prefix-free.

The Kolmogorov complexity of a string $\sigma \in \{0,1\}^{<\mathbb{N}}$ with respect to a prefix-free computable function f is defined as the length of a shortest program (string) $\tau \in \{0,1\}^{<\mathbb{N}}$ with $f(\tau) = \sigma$. That is,

$$K_f(\sigma) = \min\{|\tau| : f(\tau) = \sigma\}, \tag{3.2.13}$$

where the minimum is taken to be ∞ if the set after the "min" is empty. The Kolmogorov complexity of a string σ depends on the underlying prefix-free computable function f. However, there exists an optimal prefix-free computable function U in the sense that if for any partial computable function $f :\subseteq \{0,1\}^{<\mathbb{N}} \to \{0,1\}^{<\mathbb{N}}$, there is a positive constant $c_f < \infty$ such that for all $\sigma \in \{0,1\}^{<\mathbb{N}}$,

$$K_U(\sigma) \leq K_f(\sigma) + c_f. \tag{3.2.14}$$

Thus, if a string can hardly be compressed by an optimal function, then the string cannot be compressed by any computable function. In other words, the Kolmogorov complexity is an intrinsic property of strings.

Definition 3.2.4.5 (prefix-free Kolmogorov complexity)**.** We fix an optimal prefix-free computable function U and define the *prefix-free Kolmogorov complexity* $K(\sigma)$ of a string $\sigma \in \{0,1\}^{<\mathbb{N}}$ as $K(\sigma) = K_U(\sigma)$.

According to the following theorem, the Kolmogorov complexity provides a characterization of randomness in terms of incompressibility. That is to say, the measure-theoretic typicalness of a sequence is equivalent to the incompressibility of it in the sense of the Kolmogorov complexity.

Theorem 3.2.4.6 (equivalence between ML randomness and complexity randomness)**.** Let μ be a computable probability measure. A binary sequence $x \in \{0,1\}^{\mathbb{N}}$ is Martin-Löf random with respect to μ if and only if there exists a positive constant c such that for all n

$$K(x \upharpoonright n) > -\log \mu([x \upharpoonright n]) - c. \tag{3.2.15}$$

Proof. See Theorem 6.2.3 in [DH10] for the case of the uniform measure. The extension to an arbitrary computable measure is straightforward. □

Unpredictability approach There is another characterization by unpredictability. If a sequence is random, the knowledge of the first n bits of the sequence provides no useful information on the $(n+1)$-th bit. Therefore, if we bet on the future bits by utilizing the knowledge of the previous bits of the random sequence, there is no betting strategy by which we are able to make much money. A betting strategy is represented mathematically by a martingale, which is a crucial concept in the theory of stochastic process [Doo53]. See [Nie09, DH10] for more details in the context of the theory of algorithmic randomness.

3.3 Kac infinite chain model

We aim to understand how deterministic and irreversible macroscopic laws emerge from deterministic and reversible microscopic dynamics. The Kac ring model has often been used as an instructive model to demonstrate the macroscopic law as the law of large numbers [Kac59, GO09, MNS09]. The model is also suitable as an example of an application of the randomness notion because the dynamical system is defined on infinite binary sequences.

3.3.1 Model

Let us consider the one-dimensional lattice $\mathbb{Z}$. For each site $i \in \mathbb{Z}$, there is one particle having a spin variable $\eta(i) \in \{-1, 1\}$ and at most one scatterer. The occupation number of the scatterer at site i is denoted by $y(i) \in \{0, 1\}$. For convenience, we set $x(i) = (1+\eta(i))/2 \in \{0, 1\}$ for all $i \in \mathbb{Z}$ and think of them as dynamical variables. Then, a microscopic state of our model is represented by $(x, y) = (x(i), y(i))_{i\in\mathbb{Z}}$ and the state space is $\{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$. A discrete-time deterministic dynamical system on $\{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$ is defined by a function $\varphi : \{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}} \to \{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$ with

$$\varphi(x, y)(i) = (x(i-1) + y(i-1) - 2x(i-1)y(i-1), y(i)). \tag{3.3.1}$$

By using (3.3.1), we obtain the time evolution of the spin variables, $\eta(x, y)(i) = 2x(i) - 1$, as

$$(\eta \circ \varphi^t)(x, y)(i) = [1 - 2y(i-1)] \dots [1 - 2y(i-t)](2x(i-t) - 1). \tag{3.3.2}$$

Thus, this dynamical system has the following interpretation. Let us prepare an initial configuration of spins and scatterers. For each time step, the configuration of scatterers remains unchanged and the particle at site i jumps to the neighbor site $i+1$. Then, the spin $\eta(i)$ of the particle is flipped if the scatterer at site i is present, $y(i) = 1$, or it keeps its value if absent, $y(i) = 0$.

The dynamical system is deterministic and invertible. In fact, the map

$$\varphi^{-1}(x, y)(i) = (x(i+1) + y(i) - 2x(i+1)y(i), y(i)) \tag{3.3.3}$$

is the inverse of φ. Obviously, this dynamics corresponds to jumps of particles to the left site. We discuss the details of the microscopic reversibility in section 3.4.

The dynamical system on $\{0,1\}^{\mathbb{Z}}\times\{0,1\}^{\mathbb{Z}}$ can be regarded as that on $\{0,1\}^{\mathbb{N}}\times\{0,1\}^{\mathbb{N}}$ and $\{0,1\}^{\mathbb{N}}$ by the encoding function ι and κ. Hereafter, (x,y) (resp. φ) represents an element of $\{0,1\}^{\mathbb{Z}}\times\{0,1\}^{\mathbb{Z}}$, $\{0,1\}^{\mathbb{N}}\times\{0,1\}^{\mathbb{N}}$, or $\{0,1\}^{\mathbb{N}}$ (resp. the function on $\{0,1\}^{\mathbb{Z}}\times\{0,1\}^{\mathbb{Z}}$, $\{0,1\}^{\mathbb{N}}\times\{0,1\}^{\mathbb{N}}$, or $\{0,1\}^{\mathbb{N}}$) interchangeably. It is easy to show that if we think φ as a function from $\{0,1\}^{\mathbb{N}}$ to $\{0,1\}^{\mathbb{N}}$, φ is a total computable function on $\{0,1\}^{\mathbb{N}}$.

Remark 3.3.1.1. Our model is a variant of the Kac ring model [Kac59]. The original model is defined on the ring of size N. We use the infinite chain model in this chapter because the randomness notion in section 3.2 is sharply defined for infinite sequences. Therefore, Zermelo's recurrence paradox, which is a characteristic of finite systems, does not occur.

The model can be thought to be a dynamical system that consists of spin degrees of freedom with quenched scatterers because the configuration of scatterers does not change in time. In this chapter, we include the scatterers in state variables for simplicity. See Remark 3.3.3.2

3.3.2 Measure-theoretic approach

Let us imagine the situation we observe the system macroscopically. We introduce the following two macroscopic variables over $2N+1$ sites for a microscopic state (x,y):

$$m_0^N(x,y)=\frac{1}{2N+1}\sum_{i=-N}^{N}(2x(i)-1),\quad m_1^N(x,y)=\frac{1}{2N+1}\sum_{i=-N}^{N}y(i). \tag{3.3.4}$$

If we observe the time evolution of the macroscopic variables $m(t)=(m_0(t),m_1(t))$, the variables obey a macroscopic law and relax to the equilibrium values. In fact, at each time step t, we assume that the up or down spins are scattered at a rate $m_1(t)$ for sufficiently large N. Then, the fraction of the up or down spins changes from $(1\pm m_0(t))/2$ to $[1\pm(1-2m_1(t))m_0(t)]/2$. Therefore, the average magnetization changes from $m_0(t)$ to $(1-2m_1(t))m_0(t)$. Because the average density of the scatterers is constant, the macroscopic law has the form $m(t)=\Phi^t(m(0))$ with $\Phi(m)=([1-2m_1]m_0,m_1)$.

This "molecular chaos" argument provides the form of the macroscopic law that the system should obey *on average*. However, the hydrodynamic equations for fluids predict the macroscopic behavior of a single experiment, not just the ensemble average. The same holds true for this model. Suppose that initial microscopic states are sampled according to an initial probability measure corresponding to a nonequilibrium state. Then, the macroscopic law is understood as typical behavior with respect to the initial probability measure. This scenario is represented mathematically by the law of large numbers.

In statistical mechanics, if we have information on only the values of relevant macroscopic variables at the initial time, then one natural choice of an initial probability measure is the Gibbs measure corresponding to the initial macroscopic state [ZMR96, ZMR97]. In the case of the Kac infinite chain model, the relevant macroscopic variables are the average magnetization m_0 and the average density of scatterers m_1. Then, the Gibbs measure in this case is the product of the Bernoulli measures $\mu_{(1+m_0)/2}\times\mu_{m_1}$ on $\{0,1\}^{\mathbb{Z}}\times\{0,1\}^{\mathbb{Z}}$, where $m=(m_0,m_1)\in[-1,1]\times[0,1]$ is an initial nonequilibrium state.

Under the above settings, the weak and strong laws of large numbers hold. Although the facts are widely known, we give complete proofs of the theorems in the following. Henceforth, we write $\mu_m = \mu_{(1+m_0)/2} \times \mu_{m_1}$ for notational simplicity and $\mathbb{E}[X]$ denotes the expectation value of a random variable X with respect to μ_m. For instance,

$$\mathbb{E}[2x(i) - 1] = m_0, \quad \mathbb{E}[y(i)] = m_1 \quad \text{for } i \in \mathbb{Z}. \tag{3.3.5}$$

Theorem 3.3.2.1 (weak law of large numbers [Kac59, GO09, MNS09])**.** For any $T \in \mathbb{N}$ and any $\delta > 0$,

$$\lim_{N\to\infty} \mu_m \left(\bigcup_{t=0}^{T} \bigcup_{i\in\{0,1\}} \left\{ (x,y) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}} : |(m_i^N \circ \varphi^t)(x,y) - \Phi_i^t(m)| > \delta \right\} \right) = 0. \tag{3.3.6}$$

Proof. Fix $T \in \mathbb{N}$ and $\delta > 0$. By the subadditivity of measures, it is enough to show that for any $t \in \{0, \dots, T\}$ and any $i \in \{0, 1\}$,

$$\lim_{N\to\infty} \mu_m \left(|(m_i^N \circ \varphi^t)(x,y) - \Phi_i^t(m)| > \delta \right) = 0. \tag{3.3.7}$$

First, we show that

$$\mathbb{E}[(m_i^N \circ \varphi^t)(x,y)] = \Phi_i^t(m). \tag{3.3.8}$$

By using (3.3.2), (3.3.4), (3.3.5) and statistical independence of $x(i)$ and $y(j)$, we have

$$\begin{aligned}
\mathbb{E}[(m_0^N \circ \varphi^t)(x,y)] &= \frac{1}{2N+1} \sum_{i=-N}^{N} \mathbb{E}[(1-2y(i-1)) \dots (1-2y(i-t))(2x(i-t)-1)] \\
&= \frac{1}{2N+1} \sum_{i=-N}^{N} \mathbb{E}[1-2y(i-1)] \dots \mathbb{E}[1-2y(i-t)] \mathbb{E}[2x(i-t)-1] \\
&= (1-2m_1)^t m_0 = \Phi_0^t(m).
\end{aligned} \tag{3.3.9}$$

$\mathbb{E}[(m_1^N \circ \varphi^t)(x,y)] = m_1 = \Phi_1^t(m)$ is obvious. Next, we evaluate the second moments of $(m_i^N \circ \varphi^t)(x,y)$.

$$\begin{aligned}
\mathbb{E}\left[\left((m_0^N \circ \varphi^t)(x,y) \right)^2 \right] &= \frac{1}{(2N+1)^2} \sum_{i,j=-N}^{N} \mathbb{E}[(1-2y(i-1)) \dots (1-2y(i-t)) \\
&\quad \times (1-2y(j-1)) \dots (1-2y(j-t))] \cdot \mathbb{E}[(2x(i-t)-1)(2x(j-t)-1)] \\
&= \frac{1}{(2N+1)^2} \sum_{k=-2N}^{2N} (2N+1-|k|)\, \mathbb{E}[(1-2y(0)) \dots (1-2y(t-1)) \\
&\quad \times (1-2y(k)) \dots (1-2y(k+t-1))] \cdot \mathbb{E}[(2x(0)-1)(2x(k)-1)] \\
&= \frac{1}{(2N+1)^2} \left[(2N+1) + 2m_0^2 \sum_{k=1}^{2N} (2N+1-k)(1-2m_1)^{2\min\{t,k\}} \right].
\end{aligned} \tag{3.3.10}$$

We have used the translation invariance of μ_m and statistical independence of random variables at different sites. We take N such that $T \le 2N$. Then, we have

$$\sum_{k=1}^{2N}(2N+1-k)(1-2m_1)^{2\min\{t,k\}} = \sum_{k=1}^{t}(2N+1-k)(1-2m_1)^{2k} + \sum_{k=t+1}^{2N}(2N+1-k)(1-2m_1)^{2t}$$
$$\le (2N+1)t + \frac{1}{2}(1-2m_1)^{2t}(2N+1)^2. \tag{3.3.11}$$

Therefore,

$$\mathrm{Var}[(m_0^N \circ \varphi^t)(x,y)] \le \frac{1+2m_0^2 t}{2N+1} \tag{3.3.12}$$

for $t \in \{0,\dots,T\}$ and $T \le 2N$. Additionally, we obtain the variance of m_1^N,

$$\mathrm{Var}[(m_1^N \circ \varphi^t)(x,y)] = \frac{1-m_1^2}{2N+1}. \tag{3.3.13}$$

By using Chebyshev's inequality, we have

$$\mu_m\left(|(m_i^N \circ \varphi^t)(x,y) - \Phi_i^t(m)| > \delta\right) \le \frac{\mathrm{Var}[(m_i^N \circ \varphi^t)(x,y)]}{\delta^2} \le \frac{C}{\delta^2(2N+1)} \tag{3.3.14}$$

with a constant C independent of N, which implies (3.3.7). □ □

We have the strong form of the law of large numbers from the inequality (3.3.14) and the first Borel-Cantelli lemma (Theorem 3.2.2.4).

Theorem 3.3.2.2 (strong law of large numbers)**.** For any natural number $T \in \mathbb{N}$,

$$\mu_m\left(\lim_{N\to\infty}(m_i^N \circ \varphi^t)(x,y) = \Phi_i^t(m) \text{ for all } i \in \{0,1\} \text{ and } t \in \{0,\dots,T\}\right) = 1. \tag{3.3.15}$$

Proof. For $k \in \mathbb{N}_{>0}$, we set

$$C_{N,k} = \bigcup_{t=0}^{T}\bigcup_{i\in\{0,1\}}\left\{(x,y) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}} : |(m_i^N \circ \varphi^t)(x,y) - \Phi_i^t(m)| > \frac{1}{k}\right\}. \tag{3.3.16}$$

From (3.3.14), we have

$$\mu_m(C_{N^2,k}) \le \frac{2(T+1)Ck^2}{2N^2+1} \tag{3.3.17}$$

for N satisfying $T \le 2N$. Therefore,

$$\sum_{N=0}^{\infty}\mu_m(C_{N^2,k}) \le \sum_{N=0}^{\lceil T/2\rceil}\mu_m(C_{N^2,k}) + \sum_{N=\lceil T/2\rceil}^{\infty}\frac{2(T+1)Ck^2}{(2N^2+1)} < \infty. \tag{3.3.18}$$

By the first Borel-Cantelli lemma (Theorem 3.2.2.4) and Proposition 3.2.3.3, we have

$$\mu_m\left(\bigcup_{k\in\mathbb{N}_{>0}}\bigcap_{N\in\mathbb{N}}\bigcup_{M\geq N} C_{M^2,k}\right)=0. \tag{3.3.19}$$

This means that for μ_m-almost all configuration (x,y), the subsequence $((m_i^{N^2}\circ\varphi^t)(x,y))_{N\in\mathbb{N}}$ converges to $\Phi_i^t(m)$ for $t\in\{0,\dots,T\}$ and $i\in\{0,1\}$

Next, we show the convergence of the whole sequence. For any natural numbers L, M and a real number p with $p\leq(2L+1)/(2M+1)\leq 1$, we have

$$\begin{aligned}(m_0^L\circ\varphi^t)(x,y)-(m_0^M\circ\varphi^t)(x,y)&=\left(1-\frac{2L+1}{2M+1}\right)(m_0^L\circ\varphi^t)(x,y)\\&\quad-\frac{1}{2M+1}\left(\sum_{i=-M}^{-L-1}(2x(i)-1)+\sum_{i=L+1}^{M}(2x(i)-1)\right)\\&\leq 1-p+\frac{2(M-L)}{2M+1}\leq 2(1-p),\end{aligned} \tag{3.3.20}$$

where we have used $m_0^L\in[-1,1]$ and $2x(i)-1\in\{-1,1\}$. We consider a natural number K such that $N^2\leq K\leq(N+1)^2$. If we take $L=N^2$, $M=K$ and $p=p_N=(2N^2+1)/(2(N+1)^2+1)$ first and take $L=K$, $M=(N+1)^2$ and $p=p_N$ second, the inequality (3.3.20) gives

$$(m_0^{N^2}\circ\varphi^t)(x,y)-2(1-p_N)\leq(m_0^K\circ\varphi^t)(x,y)\leq(m_0^{(N+1)^2}\circ\varphi^t)(x,y)+2(1-p_N). \tag{3.3.21}$$

These inequalities also hold for $(m_1^K\circ\varphi^t)(x,y)$. Since $p_N\to 1$ as $N\to\infty$, for μ_m-almost all (x,y), the whole sequence $((m_i^N\circ\varphi^t)(x,y))_{N\in\mathbb{N}}$ converges to $\Phi_i^t(m)$ for $i\in\{0,1\}$ and $t\in\{0,\dots,T\}$. □ □

Remark 3.3.2.3. A similar analysis leads to the weak and strong law of large numbers for the microscopic dynamics φ^{-1} and positive integers $T>0$. This is a consequence of the microscopic reversibility and the statistical property of the initial measure. If an initial configuration has no correlation between sites, whether the microscopic dynamics is φ or φ^{-1}, which corresponds to the direction of movement of the spins, is irrelevant to the validity of the macroscopic relaxation.

3.3.3 Algorithmic randomness approach

We reformulate the law of large numbers associated with the macroscopic law as properties of individual microscopic states. The concept of algorithmic randomness introduced in section 3.2 helps us to do that. By using the randomness notion, we have the following theorem.

Theorem 3.3.3.1 (effective strong law of large numbers)**.** Let $m=(m_0,m_1)$ be computable reals and $T\in\mathbb{N}$. If $(x,y)\in\{0,1\}^{\mathbb{Z}}\times\{0,1\}^{\mathbb{Z}}$ is Martin-Löf random with respect to μ_m,

$$\lim_{N\to\infty}(m_i^N\circ\varphi^t)(x,y)=\Phi_i^t(m) \tag{3.3.22}$$

for all $i\in\{0,1\}$ and $t\in\{0,\dots,T\}$.

Proof. Fix $T \in \mathbb{N}$. Since $\Phi_i^t(m)$ are computable reals, $(C_{N^2,k})_{N\in\mathbb{N}}$ is c.e. open uniformly in N. By (3.3.18), it is a Solovay test. Therefore, if $(x, y) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ is ML random with respect to μ_m, then $(x, y) \in \bigcup_{N\in\mathbb{N}} \bigcap_{M\geq N} (C_{N^2,k})^c$. Because k is arbitrary, the subsequence $((m_i^N \circ \varphi^t)(x, y))_{N\in\mathbb{N}}$ converges to $\Phi_i^t(m)$ for any $i \in \{0,1\}$ and $t \in \{0,\dots,T\}$. The proof of the convergence of the whole sequence is the same as the proof of Theorem 3.3.2.2. □ □

According to Theorem 3.3.3.1, the algorithmic randomness of a microscopic state is a sufficient condition that the microstate obeys the macroscopic relaxation law. Since the set of all ML random microstates has measure one (see Theorem 3.2.3.5), this sufficient condition is not too strong from a viewpoint of measure-theoretic typicality. In particular, the strong law of large numbers (Theorem 3.3.2.2) follows from the effective law. We stress that the effective law of large numbers holds for a wide class of models. We discuss the generality of our result in **??**.

Remark 3.3.3.2. Van Lambalgen's theorem [vL87] implies that (x, y) is ML μ_m-random if and only if y is ML μ_{m_1}-random and x is ML $\mu_{(1+m_0)/2}$-random with *oracle* y. Therefore, Theorem 3.3.3.1 insists that for a given ML μ_{m_1}-random configuration of quenched scatterers y, $\mu_{(1+m_0)/2}$-random microstates with oracle y satisfy the macroscopic law. We note that for a $\mu_{(1+m_0)/2}$-random element x and μ_{m_1}-random element y, the pair (x, y) does not necessarily obey the macroscopic law. For instance, if $(1+m_0)/2 = m_1$ and x is $\mu_{(1+m_0)/2}$-random, (x, x) violates the law.

3.4 Entropy and the zeroth law of thermodynamics

Entropy is a fundamental concept in various fields such as thermodynamics, statistical physics, information theory and dynamical systems theory. Each type of entropy has a different role. We investigate the Boltzmann entropy quantifying irreversibility on transitions between macroscopic states.

3.4.1 Shannon entropy

Before considering the Boltzmann entropy, we review basic properties of the Shannon entropy for convenience, which is an information-theoretic quantity characterizing the optimal compression rate in the information source coding problem [CT12].

Remark 3.4.1.1. In this subsection, the configurations (x, y), the probability measures and the microscopic dynamics φ are regarded as ones defined on $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$, not on $\{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$, by the encoding function ι in section 3.2.3.

Definition 3.4.1.2 (Shannon entropy rate, self-entropy rate)**.** The *Shannon entropy rate* of the joint probability measure μ on $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ is defined as

$$\bar{H}(\mu) = \limsup_{n\to\infty} -\frac{1}{n} \sum_{(\sigma,\tau)\in 2^n\times 2^n} \mu([\sigma] \times [\tau]) \ln \mu([\sigma] \times [\tau]). \tag{3.4.1}$$

The self-entropy rate of $(x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ with respect to μ is defined as

$$\bar{H}_{\mu}(x,y) = \limsup_{n\to\infty} -\frac{1}{n} \ln \mu([x \upharpoonright n] \times [y \upharpoonright n]). \tag{3.4.2}$$

A straightforward calculation provides

$$\bar{H}(\mu_m) = h\left(\frac{1+m_0}{2}\right) + h(m_1), \tag{3.4.3}$$

where $h(p) = -p\ln p - (1-p)\ln(1-p)$ $(p \in [0,1])$ is the binary entropy function. For $x \in \{0,1\}^{\mathbb{N}}$ and $n \in \mathbb{N}$, set $N(x,n) = |\{i : x(i) = 1, 0 \le i \le n-1\}|$. If $(x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ satisfies

$$\lim_{n\to\infty} \frac{N(x,n)}{n} = p_x, \quad \lim_{n\to\infty} \frac{N(y,n)}{n} = p_y, \tag{3.4.4}$$

the self-entropy rate of (x,y) with respect to μ_m is given by

$$\bar{H}_{\mu_m}(x,y) = -p_x \ln\left(\frac{1+m_0}{2}\right) - (1-p_x)\ln\left(\frac{1-m_0}{2}\right) - p_y \ln m_1 - (1-p_y)\ln(1-m_1). \tag{3.4.5}$$

In particular, for any random element $(x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ with respect to μ_m,

$$\bar{H}_{\mu_m}(x,y) = \bar{H}(\mu_m), \tag{3.4.6}$$

because $p_x = (1+m_0)/2$ and $p_y = m_1$. This type of statement is referred to as the effective version of the *asymptotic equipartition property*.

For deterministic and reversible dynamical systems, the Shannon entropy of the probability measure describing the system does not provide useful information on irreversibility. If we define the probability measure at time t starting from the initial measure μ_m as $\mu_{m,t} = \mu_m \varphi^{-t}$, the Shannon entropy rate is invariant under the time evolution, that is, $\bar{H}(\mu_m) = \bar{H}(\mu_{m,t})$. This invariance remains true for random elements. In fact, since the initial segment of the first and second components of $\varphi^t(x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$, $x_t \upharpoonright n$ and $y_t \upharpoonright n$, depend only on $x \upharpoonright n+2t$ and $y \upharpoonright n+2t$ (the factor 2 comes from the way of encoding ι from $\{0,1\}^{\mathbb{Z}}$ to $\{0,1\}^{\mathbb{N}}$), the inclusion relation

$$[x \upharpoonright n+2t] \times [y \upharpoonright n+2t] \subseteq \varphi^{-t}([x_t \upharpoonright n] \times [y_t \upharpoonright n]) \subseteq [x \upharpoonright n-2t] \times [y \upharpoonright n-2t] \tag{3.4.7}$$

holds. Then,

$$\begin{aligned} -\frac{1}{n} \ln \mu_m([x \upharpoonright n+2t] \times [y \upharpoonright n+2t]) &\le -\frac{1}{n} \ln \mu_{m,t}([x_t \upharpoonright n] \times [y_t \upharpoonright n]) \\ &\le -\frac{1}{n} \ln \mu_m([x \upharpoonright n-2t] \times [y \upharpoonright n-2t]). \end{aligned} \tag{3.4.8}$$

For any random element $(x,y) \in \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$, the terms on the left- and right-hand sides converge to $\bar{H}(\mu_m)$. Therefore, the self-entropy rate of $\varphi^t(x,y)$ with respect to $\mu_{m,t}$ exists and equals that of (x,y) with respect to μ_m:

$$\bar{H}_{\mu_{m,t}}(\varphi^t(x,y)) = \lim_{n\to\infty} -\frac{1}{n} \ln \mu_{m,t}([x_t \upharpoonright n] \times [y_t \upharpoonright n]) = \bar{H}(\mu_m) = \bar{H}(\mu_{m,t}). \tag{3.4.9}$$

3.4.2 Boltzmann entropy and the zeroth law of thermodynamics

Because the Shannon entropy does not change in time in reversible dynamical systems, we need another quantity to characterize the macroscopic irreversibility. According to Boltzmann's idea, the asymmetry of the direction of time in the macroscopic behavior emerges from the large differences between the number of microstates consistent with macrostates. Since the number of microstates corresponding to a macrostate is proportional to the probability of the macrostate under the uniform measure λ and different macrostates usually have exponentially different probabilities, it is reasonable to introduce the rate function in the large deviation theory as the Boltzmann entropy.

Definition 3.4.2.1 (Boltzmann entropy)**.** The *Boltzmann entropy* of a macroscopic state $m = (m_0, m_1)$ is defined as

$$S_B(m) = \lim_{\delta \downarrow 0} \lim_{N \uparrow \infty} \frac{1}{2N+1} \ln \lambda \times \lambda \left(\{(x,y) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}} : |m_0^N(x) - m_0| \leq \delta, |m_1^N(x) - m_1| \leq \delta\} \right), \tag{3.4.10}$$

where λ is the uniform measure on $\{0,1\}^{\mathbb{Z}}$.

The following scenario is well-known [Leb93]: An initial microstate in a nonequilibrium macrostate with low Boltzmann entropy evolves *typically* toward macrostates with higher entropy and finally reaches the equilibrium state with the maximum entropy.

Although at first sight, it explains the macroscopic irreversibility qualitatively, it should be noted that we must suppose an initial probability measure in order to argue the *typical* macroscopic behavior. The above scenario is certainly true if we assume that initial microstates are chosen according to the microcanonical measure or the Gibbs measure. In fact, by Stirling's formula, we have

$$S_B(m) = -2 \ln 2 + \bar{H}(\mu_m). \tag{3.4.11}$$

If we prepare initial microstates according to the Gibbs measure μ_m, then the initial macrostate is m and the macrostate evolves according to the law Φ with probability one according to Theorem 3.3.2.2. Then, the Boltzmann entropy difference is typically positive:

$$S_B(\Phi^t(m)) - S_B(m) = \bar{H}(\mu_{\Phi^t(m)}) - \bar{H}(\mu_m) > 0 \text{ for } t > 0. \tag{3.4.12}$$

We can reformulate the argument from a viewpoint of randomness. If an initial state $(x,y) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ is ML μ_m-random, the macrostate at the initial time is m and at time t is $\Phi^t(m)$ (see Theorem 3.3.3.1). Then, the Boltzmann entropy increases over time. This is the zeroth law of thermodynamics for algorithmic random microstates.

3.5 Microscopic reversibility and anti-Boltzmann behavior

3.5.1 Microscopic reversibility

The microscopic dynamics $\varphi : \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ is invertible. This property is referred to as *microscopic reversibility*. Let us define a time-reversal transformation $\pi : \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ by

$$(\pi(x,y))(i) = (x(-i), y(-i-1)). \tag{3.5.1}$$

The time-reversal transformation is an involution $\pi^2 = 1$ and is totally computable. The microscopic reversibility is represented by $\pi \circ \varphi = \varphi^{-1} \circ \pi$. We note that $m_0^N(\pi(x,y)) = m_0^N(x,y)$ and $m_1^N(\pi(x,y)) = m_1^N(x, \Sigma^{-1}(y))$. Here, $\Sigma(x)(i) = x(i+1)$ is the shift map on $\{0,1\}^{\mathbb{Z}}$. In particular, the time-reversal transformation does not affect the macrostate.

3.5.2 Irreversible information loss

If there is a microscopic trajectory $(\varphi^t(x,y))_{t\in\{0,1,\dots,T\}}$ whose macroscopic trajectory is $(\Phi^t(m))_{t\in\{0,1,\dots,T\}}$, then the time-reversed one $((\varphi^t \circ \pi \circ \varphi^T)(x,y))_{t\in\{0,1,\dots,T\}} = ((\pi \circ \varphi^{T-t})(x,y))_{t\in\{0,1,\dots,T\}}$ is macroscopically observed as $(\Phi^{T-t}(m))_{t\in\{0,1,\dots,T\}}$. Loschmidt inquired how the above consequence of the microscopic reversibility is consistent with the macroscopic irreversibility. This question is called the reversibility paradox problem.

Sasa and Komatsu introduced the irreversible information loss quantifying the asymmetry between the trajectory $(\varphi^t(x,y))_{t\in\{0,1,\dots,T\}}$ and the time-reversed one $((\pi\circ\varphi^{T-t})(x,y))_{t\in\{0,1,\dots,T\}}$, and investigated the relation to the Boltzmann entropy change [SK00]. Following this idea, we define the rate of *irreversible information loss* as

$$I_{\mu_m,t}(x,y) = \limsup_{n\to\infty} -\frac{1}{n} \ln \frac{\mu_{m,t}(\pi([x_t \upharpoonright n] \times [y_t \upharpoonright n]))}{\mu_{m,t}([x_t \upharpoonright n] \times [y_t \upharpoonright n])}, \tag{3.5.2}$$

where the dynamical system is regarded as one on $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$. The positivity of the irreversible information loss of a microstate (x,y) implies the exponential difference between the probabilities at time t of the microstate $\varphi^t(x,y)$ and the time-reversed one $(\pi \circ \varphi^t)(x,y)$. Then, it explains how difficult it is to prepare the time-reversed state $(\pi \circ \varphi^t)(x,y)$ relative to the state $\varphi^t(x,y)$ in the measure-theoretic sense. We note that this argument is different from the standard one on the reversibility paradox indicating the practical impossibility of the time-reverse transformation.

We can easily calculate the above quantity for random states (x,y) as follows. The microscopic reversibility implies that $\varphi^{-t} \circ \pi \circ \varphi^t = \pi \circ \varphi^{2t}$. If (x,y) is ML μ_m-random, by Theorem 3.3.3.1, we have

$$\lim_{n\to\infty} \frac{N(x_{2t},n)}{n} = \frac{1+\Phi_0^{2t}(m)}{2}, \quad \lim_{n\to\infty} \frac{N(y_{2t},n)}{n} = \Phi_1^{2t}(m). \tag{3.5.3}$$

The same type of inclusion relation as (3.4.7) implies

$$\lim_{n\to\infty} -\frac{1}{n}\ln \mu_{m,t}(\pi([x_t \upharpoonright n]\times[y_t \upharpoonright n])) = -\frac{1+\Phi_0^{2t}(m)}{2}\ln\left(\frac{1+m_0}{2}\right) - \frac{1-\Phi_0^{2t}(m)}{2}\ln\left(\frac{1-m_0}{2}\right) - \Phi_1^{2t}(m)\ln m_1 - (1-\Phi_1^{2t}(m))\ln(1-m_1). \quad (3.5.4)$$

Therefore, the rate of the irreversible information loss of a random element (x, y) is given by

$$I_{\mu_m,t}(x,y) = \frac{m_0 - \Phi_0^{2t}(m)}{4}\ln\frac{1+m_0}{1-m_0} = \frac{1-(1-2m_1)^{2t}}{4} m_0 \ln\frac{1+m_0}{1-m_0}. \quad (3.5.5)$$

If $m_0 \neq 0$, $m_1 \neq 1/2$, and $t > 0$, then $I_{\mu_m,t}(x, y) > 0$. This result implies the measure-theoretic difficulty of preparing the time-reversed state relative to the random state. Moreover, by explicit calculation, we have

$$I_{\mu_m,t}(x,y) - (S_B(\Phi^t(m)) - S_B(m)) \geq 0. \quad (3.5.6)$$

The above equality holds for $m_0 = 0$, $m_1 \in \{0, 1/2\}$, or $t = 0$. That is, the degree of difficulty is greater than the Boltzmann entropy change in this case.

3.5.3 Violation of the macroscopic law and nonrandomness of time-reversed states

Suppose that an initial probability measure is μ_m and $(x, y) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ is ML μ_m-random. By the conservation of ML randomness (see Theorem 3.2.4.3), $\varphi^t(x, y)$ is ML random with respect to the probability measure at time t, $\mu_{m,t}$. That is, the randomness of the initial microstate is preserved under the dynamics. An intriguing question is whether the time-reversed state $(\pi \circ \varphi^t)(x, y)$ is ML $\mu_{m,t}$-random or not. The microscopic reversibility implies that the macroscopic evolution starting from the state $(\pi \circ \varphi^t)(x, y)$ does not obey the macroscopic law Φ. Therefore, $(\pi \circ \varphi^t)(x, y)$ is not $\mu_{m,t}$-random.

Theorem 3.5.3.1 (non-randomness of time-reversed state)**.** Let $m = (m_0, m_1)$ be computable reals and $m_0 \neq 0$, $m_1 \notin \{0, 1/2\}$. For any μ_m-random element $(x, y) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ and $t \in \mathbb{N}\backslash\{0\}$, $(\pi \circ \varphi^t)(x, y)$ is not ML $\mu_{m,t}$-random.

Proof. Assume $(\pi \circ \varphi^t)(x, y)$ is ML $\mu_{m,t}$-random for $t \in \mathbb{N}\backslash\{0\}$. By the conservation of ML randomness (Theorem 3.2.4.3), $(u, v) = (\varphi^{-t} \circ \pi \circ \varphi^t)(x, y) = (\pi \circ \varphi^{2t})(x, y)$ is ML μ_m-random. From Theorem 3.3.3.1, for $s \in \{0, \ldots, 2t\}$,

$$\begin{aligned} m_i^N(\varphi^s(u,v)) &= m_i^N((\varphi^s \circ \pi \circ \varphi^{2t})(x,y)) \\ &= m_i^N((\pi \circ \varphi^{2t-s})(x,y)) \\ &\to \Phi_i^{2t-s}(m) \ \text{ as } N \to \infty. \end{aligned} \quad (3.5.7)$$

Since $m_0 \neq 0$, $m_1 \notin \{0, 1/2\}$ and $t \neq 0$, there exists $s \in \{0, \ldots, 2t\}$ such that $\Phi_0^{2t-s}(m) \neq \Phi_0^s(m)$ (take $s(\neq t)$). This is a contradiction. □ □

This consideration leads to the following argument. $(u,v) = (\varphi^{-T} \circ \pi \circ \varphi^{T})(x,y) = (\pi \circ \varphi^{2T})(x,y)$ is ML $\mu_{m,2T}\pi$-random if and only if $(x,y) = (\pi \circ \varphi^{2T})(u,v)$ is ML μ_m-random. As in the proof of Theorem 3.5.3.1,

$$\lim_{N\to\infty} m_i^N(\varphi^t(u,v)) = \Phi_i^{2T-t}(m) \tag{3.5.8}$$

for $t \in \{0,\dots,2T\}$ and ML $\mu_{m,2T}\pi$-random element $(u,v) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$. Therefore, if we observe the macroscopic time evolution starting from a random microscopic state with respect to the *initial* probability measure $\mu_{m,2T}\pi$, the system exhibits the time-reversed behavior of the original macroscopic law Φ^t. In particular, the Boltzmann entropy along the typical macroscopic trajectory *decreases* monotonically:

$$S_B(\Phi^{2T-t}(m)) - S_B(\Phi^{2T}(m)) < 0 \quad \text{for} \quad 0 < t \le 2T. \tag{3.5.9}$$

Thus, typical macroscopic behavior depends on the choice of an initial probability measure. Even if an initial macroscopic state is given, the initial probability measure representing the state is not unique. Therefore, we have to demonstrate why we regard the Gibbs measure as important. See 3.6 for further discussion.

3.6 Concluding remarks

3.6.1 Remarks

Natural choice of initial measure We need to choose an initial probability measure to state a probabilistic law of large numbers for an irreversible macroscopic law. In this chapter, we have chosen the Gibbs measure because it works well in many examples in statistical physics. Then, the system evolves typically so that the entropy increases monotonically to equilibrium. In contrast, as shown in 3.5.3, if we choose another initial measure carefully, the entropy decreases along the typical macroscopic evolution with respect to the measure. To elucidate the origin of macroscopic irreversibility, we have to clarify the difference between these two measures and to demonstrate why the measures under which the entropy increases are realized in our world. The problem also occurs when we discuss the effective law of large numbers since the notion of randomness formalizes typical states *under a given probability measure*.

Nonrandom states satisfying the macroscopic law We have shown that random microstates satisfy the macroscopic law. However, the reverse is not generally true because the condition on the violation of the macroscopic law is just a part of Martin-Löf tests. That is to say, there are microscopic states satisfying the macroscopic law but not passing another ML test. Little is known about the physical meaning of such ML tests, and therefore also of nonrandom microstates satisfying the macroscopic law.

Relevant randomness test In the theory of algorithmic randomness, there are various classes of randomness according to the level of computability imposed on null sets in addition to the Martin-Löf randomness (see Chapter 7 of [DH10] for example). In any case, we take into account *all* effective null sets or corresponding statistical tests. However, all these tests are not necessarily realizable in physical experiments. Therefore, one may say that the theory of algorithmic randomness is unnecessary for the foundation of statistical physics. A critical problem here is to identify the class of null sets associated with macroscopic properties. To consider the problem, let us recall the argument in 3.2.3 motivating the definition of the Martin-Löf randomness. We have seen that the law of large numbers does not provide a sufficient characterization of randomness. Even if we add another law such as the law of the iterated logarithm to the requirement of randomness, we may find other probabilistic laws having probability one and the requirement may turn out to be insufficient. Avoiding these difficulties, the theory of algorithmic randomness considers *all* effective statistical laws and as a result clarifies a rich structure of randomness such as the equivalence between measure-theoretic typicalness, incompressibility and unpredictability. When we attempt to specify statistical laws involved with macroscopic properties, the above idea may be useful and there may be a deep connection between algorithmic randomness and statistical physics.

Applications to other models In this chapter, we have shown that the algorithmic randomness of microscopic states is a sufficient condition of macroscopic relaxation in the Kac chain model. We expect to extend the theorem to a wider class of models. To prove the effective law of large numbers, we need the upper bound on the probability of the sequence of sets involved with the violation of the macroscopic law that tends to zero in the thermodynamic limit and the computable enumerability of the sequence. The former condition follows from a purely measure-theoretic argument. As long as we focus on macroscopic properties, the latter condition is also expected to be satisfied. For instance, there are deterministic and reversible dynamical systems with particle conservation that exhibit diffusive behavior in the sense of the law of large numbers [Lef13, Lef15]. It is possible to extend our results to these models.

The models we refer above are cellular automata, that is, the dynamical systems on infinite lattices with local rules. The Martin-Löf randomness in 3.2.3 is defined on $\{0,1\}^{\mathbb{N}}$ and can be applied to only the dynamical systems on discrete state spaces. Recently, the notion of randomness has been generalized to computable metric spaces [Gác05, HR09] and applied to the dynamical system theory [GHR10, GHR11]. Applying the theory to statistical physics is an important problem.

Quantum randomness Another direction of future study is to generalize the notion of randomness to quantum systems. As in classical settings, the quantum Kolmogorov complexity of a quantum state is defined as the length of the shortest program outputting the description of the state [Gác01, Vit01]. The notion of Martin-Löf random quantum state and the relation to the quantum Kolmogorov complexity has been investigated only recently [NS19]. In either case, the algorithmic randomness theory of quantum systems has not yet been sufficiently studied compared to the case of classical systems. An example

of the application of quantum randomness is the typicality of thermal equilibrium states [Tas16, GLTZ06, PSW06, Sug07, Rei07]. Although there are various mathematical formulations of the typicality of thermal equilibrium in quantum systems, they all state that almost all quantum pure states in a Hilbert space spanned by a set of the energy eigenstates represent thermal equilibrium. With the analogy to the argument in classical systems, we expect that *random quantum states represent thermal equilibrium*. A more challenging theoretical issue in this context is the relation between the algorithmic randomness and the eigenstate thermalization hypothesis (ETH). The ETH insist that *all* the energy eigenstates in an energy shell represent thermal equilibrium with the energy [RDO08]. The ETH is regarded as a plausible sufficient condition of thermalization in isolated quantum systems. The thermodynamic structure such as the fluctuation theorem and second law of thermodynamics that has been studied on the basis of the Gibbs state is being re-examined for the energy eigenstates [IKT17, KIT19]. We anticipate that it is important to study the ETH from a viewpoint of algorithmic randomness.

Martingale and thermodynamics Finally, we comment on the unpredictability aspect of algorithmic randomness associated with thermodynamics and statistical physics. The proofs and arguments in this chapter are based on the measure-theoretic typicalness aspect of randomness. However, as explained in section 3.2.4, the randomness notion has several characterizations such as incompressibility and unpredictability. We should further develop these aspects of statistical physics. For instance, the notion of martingale, which captures the unpredictability aspect of randomness, has not been well studied in statistical physics. Recently, the martingale property of exponentiated entropy production in stochastic thermodynamics has been investigated [CG11, NRJ17]. A more challenging task is to investigate the fundamental assumption of statistical physic such as the principle of equal a priori probability from the viewpoint of the martingale property. For instance, the game-theoretic probability theory [SV01, SV19] provides a new formulation of limit theorems in probability theory such as the law of large numbers and the central limit theorem by utilizing only the betting game without the probabilistic structure. The idea in the game-theoretic probability theory that the probability emerges from the martingale property may be useful to study this problem. We will study this approach in Chapter 4.

Chapter 4

The Second Law implies the Gibbs Distribution

4.1 Introduction

4.1.1 Elementary Example: Single-Particle Ideal-Gas Engine

Before going to main results, we consider a single-particle ideal-gas engine as an elementary example to clarify the problem. A single particle is confined in a box of volume $V = L^3$ and in contact with a heat bath having a temperature T. An external agent attempts to extract work from the system. The agent inserts a barrier at the center of the box, $x = L/2$, and moves the barrier quasi-statically to $x = (1-\mu)L$, where $\mu \in (0,1)$ specifies the final position of the barrier. When the particle is on the left side (resp. right side), we set $\omega = 0$ (resp. $\omega = 1$). Since the volume of the region in which the particle is confined after the operation is $\mu^\omega(1-\mu)^{1-\omega}V$, the work extracted in this process is given by

$$W^\mu(\omega) = \int_{V/2}^{\mu^\omega(1-\mu)^{1-\omega}V} \frac{k_{\rm B}T}{V} dV = k_{\rm B}T \ln 2\mu^\omega(1-\mu)^{1-\omega}, \tag{4.1.1}$$

where $k_{\rm B}$ is the Boltzmann constant and we have used the equation of state $P(T,V,N) = Nk_{\rm B}T/V$ for the ideal gas. Finally, the agent removes the barrier and the system returns to the initial state such that the overall process becomes cyclic. According to the second law of thermodynamics, the mechanical work extracted by any cyclic operations is always non-positive. Eq. (4.1.1), however, becomes positive for some μ and ω, e.g., $\mu = 1/4$ and $\omega = 0$. In statistical mechanics and stochastic thermodynamics, this apparent inconsistency is considered to arise from the fluctuations in small systems. To resolve the inconsistency, we assume that ω is a random variable obeying the equilibrium distribution $P_{1/2}(\{\omega = 0\}) = P_{1/2}(\{\omega = 1\}) = 1/2$. Then, although the second law is violated with positive probability for $\mu \neq 1/2$, $P_{1/2}(W^\mu > 0) = 1/2 > 0$, due to the fluctuation, the *expectation value* of the extracted work W^μ is non-positive for any operations μ, $\mathbb{E}_{P_{1/2}}[W^\mu] = k_{\rm B}T \ln 4\mu(1-\mu) \leq 0$, and the second law of thermodynamics remains true *on average*. Moreover, the second law

conversely characterizes the equilibrium distribution. To prove this statement, we assume that the initial distribution of ω is given by $P_\rho(\{\omega = 0\}) = 1 - \rho$ and $P_\rho(\{\omega = 1\}) = \rho$, where $\rho \in (0, 1)$ quantifies the inhomogeneity of the particle distribution. The expectation value of the extracted work is given by $\mathbb{E}_{P_\rho}[W^\mu] = k_{\mathrm{B}}T(D(P_\rho\|P_{1/2}) - D(P_\rho\|P_\mu))$. Here $D(P_\rho\|P_\mu) \coloneqq \rho\ln(\rho/\mu) + (1-\rho)\ln[(1-\rho)/(1-\mu)]$ is the Kullback-Leibler divergence between the Bernoulli distributions P_ρ and P_μ. If the initial distribution is not uniform, i.e., $\rho \neq 1/2$, we have that $\mathbb{E}_{P_\rho}[W^\rho] = k_{\mathrm{B}}TD(P_\rho\|P_{1/2}) > 0$ by choosing $\mu = \rho$. Therefore, the condition that $\mathbb{E}_{P_\mu}[W^\mu] \leq 0$ for any $\mu \in (0, 1)$ implies $\rho = 1/2$. In summary, the non-positivity of the averaged work is equivalent to the equilibrium condition for the initial distribution in this example.

We stress again that the extracted work $W^\mu(\omega)$ is positive for some initial state ω and operation μ. Therefore, when we speak of the validity of the second law of thermodynamics in small systems, we have to consider a situation that we prepare microscopic states and extract work many times. Moreover, if the empirical frequency of the initial microscopic states is biased from the equilibrium distribution, the second law may be violated, as suggested by the argument in the previous paragraph. In this sense, the second law in small systems may require the stochastic behavior of the initial microscopic states in addition to the equilibrium condition for the probability distribution characterizing its stochasticity. In this chapter, we ask how can we formulate mathematically the emergence of stochasticity from the second law of thermodynamics. To answer the question, we investigate two issues in this subsection.

First, we provide a mathematical definition of "stochastic behavior of the initial microscopic states". In this chapter, we consider the situation that the agent repeats the cyclic operations infinitely many times and ask whether an infinite sequence $\omega_1\omega_2\cdots \in \{0,1\}^{\mathbb{N}_+}$ of initial microscopic states in the experiments is a random sequence or not with respect to the equilibrium distribution $P_{1/2}$. Here $\mathbb{N}_+ = \{1, 2, \dots\}$ denotes the set of positive integers and $\omega_n \in \{0, 1\}$ denotes the initial position of the particle in the n-th cycle. An example of non-random sequence is $00000\dots$, which corresponds to the situation that the particle is always in the left side of the box. Although the theory of algorithmic randomness [Nie09, DH10] provides a reasonable and rigorous definition of randomness for individual sequences, we pay attention to only the convergence of the empirical distribution and empirical mean in this chapter. For the single-particle ideal-gas engine, we regard an infinite sequence $\omega_1\omega_2\dots$ representing the positions of the particle as a random sequence if the sequence satisfies the strong law of large numbers (SLLN),

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\omega_i = \frac{1}{2}. \tag{4.1.2}$$

We discuss the relation to the algorithmic randomness in Section 4.5.

Second, we formulate the second law of thermodynamics without probability measures. Since our purpose is to clarify the emergence of the probabilistic description, we cannot suppose the underlying probability distribution at the starting point of the discussion and use the standard definition of the second law $\mathbb{E}[W^\mu] \leq 0$. To remove the probability distribution and expectation value from the description, we consider again the situation that the agent

repeats the cyclic operations infinitely many times (Fig. 4.1). In contrast to averaging the extracted work for a single cyclic operation, in repeated cyclic operations, we allow the agent to choose a different operation in each cycle. A critical problem here is to specify the information that the agent can use when he decides the cyclic operation in each cycle. In this chapter, we apply the *prequential scheme* [Daw84], where the agent determines the volume fraction $\mu_n \in (0,1)$ in the n-th cycle depending on only the past history $\omega_1, \dots, \omega_{n-1}$ of the initial positions up to the $(n-1)$-th cycle [1]. In other words, the agent predicts the position of the particle in the n-th cycle from the past results $\omega_1 \dots \omega_{n-1}$ and performs a cyclic operation based on the prediction. An assignment $\omega_1 \dots \omega_{n-1} \mapsto \mu_n$ for each $n = 1, 2, \dots$ represents a prediction scheme called a *strategy* for the agent and denoted by $\hat{\mu}$. For a given strategy $\hat{\mu}$ for the agent, the accumulation of the extracted work $W^{\hat{\mu}}$ is given by

$$W^{\hat{\mu}}(\omega_1 \dots \omega_n) \coloneqq \sum_{i=1}^{n} k_{\mathrm{B}} T \ln 2 \mu_t^{\omega_i} (1-\mu_t)^{1-\omega_i}. \tag{4.1.3}$$

We define the violation of the second law in terms of the asymptotic behavior of $W^{\hat{\mu}}$. We say that an infinite sequence $\omega_1 \omega_2 \cdots \in \{0,1\}^{\mathbb{N}_+}$ of initial positions of the particle violates the second law of thermodynamics under the strategy $\hat{\mu}$ if the total amount of the extracted work from the heat engine diverges to infinity, i.e.,

$$\lim_{n\to\infty} W^{\hat{\mu}}(\omega_1 \dots \omega_n) = \infty. \tag{4.1.4}$$

This means that the agent can extract work from such a sequence as much as he wants by repeating the cyclic operations sufficiently many times. For example, if the agent chooses $\mu_t = 1/4$ for any $t \in \mathbb{N}_+$, the infinite sequence $0000\dots$ violates the second law because $W^{\hat{\mu}}(\omega_1 \dots \omega_n) = n k_{\mathrm{B}} T \ln(3/2)$ diverges to infinity as $n \to \infty$. We adopt this definition as the second law in our study because it refers to no probability measure. We note that this definition is consistent with equilibrium statistical mechanics. In fact, if we assume that $\omega_1, \omega_2, \dots$ are independent and identically distributed random variables obeying the product distribution $P_{1/2}^{\otimes \mathbb{N}_+}(\{\omega_n = 0\}) = P_{1/2}^{\otimes \mathbb{N}_+}(\{\omega_n = 1\}) = 1/2$, the probability that the second law is violated is zero,

$$P_{1/2}^{\otimes \mathbb{N}_+}\left\{ \lim_{n\to\infty} W_n^{\hat{\mu}} = \infty \right\} = 0, \tag{4.1.5}$$

where we have defined a random variable $W_n^{\hat{\mu}}(\xi) \coloneqq W^{\hat{\mu}}(\omega_1 \dots \omega_n)$ for $\xi = \omega_1 \omega_2 \cdots \in \{0,1\}^{\mathbb{N}_+}$ [2].

[1] The protocol should not be confused with that in Szilard's engine [Szi64]. Imagine that positions of the particle are prepared independently and identically according to the uniform distribution $P_{1/2}$. The mutual information between the positions up to the $(n-1)$-th cycle and the position in the n-th cycle is zero due to the statistical independency. Therefore, even if the agent can use the information on the past history, the expectation value of the extracted work in the n-th cycle is always non-positive.

[2] We can prove a stronger statement $P\{\sup_n W_n^{\hat{\mu}} = \infty\} = 0$. First, we note that $P\{\sup_n W_n^{\hat{\mu}} = \infty\} = P\{\sup_n e^{\beta W_n^{\hat{\mu}}} = \infty\}$. $e^{\beta W_n^{\hat{\mu}}}$ is a positive martingale and bounded in L^1 because $\mathbb{E}_P(e^{\beta W_n^{\hat{\mu}}}) = 1$ for any n. We have from the martingale convergence theorem that $e^{\beta W_\infty^{\hat{\mu}}} \coloneqq \lim_{n\to\infty} e^{\beta W_n^{\hat{\mu}}}$ exists and is a finite non-negative value almost surely. Hence, $P\{\sup_n W_n^{\hat{\mu}} = \infty\} = 0$.

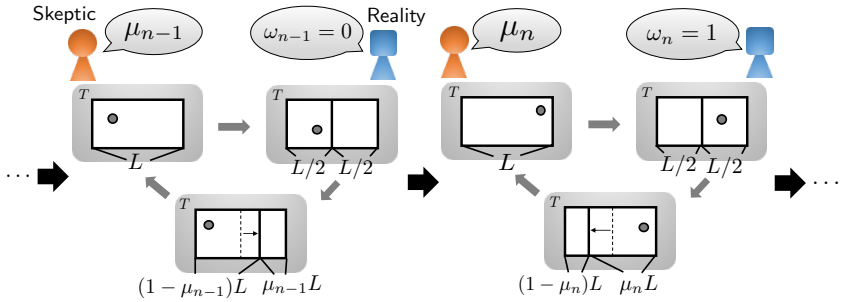

Figure 4.1: Schematic of the protocol. (1) An agent named Skeptic announces the position of the barrier $\mu_n \in (0,1)$ depending on the past history $\omega_1 \dots \omega_{n-1}$ of positions of the particle. (2) Another agent named Reality announces the position of the particle ω_n in the n-th cycle. (3) Skeptic moves the barrier from $x = L/2$ to $x = (1-\mu_n)L$ and remove it. (4) Go back to (1).

Based on the above two arguments, our question is formulated as follows. Instead of assuming that ω_n is a random variable and introducing a probability measure from the beginning, we ask which sequences can be realized under the second law of thermodynamics and what statistical properties such sequences have commonly. As we noted before, in this chapter, we consider the strong law of large numbers (4.1.2) as a relevant statistical property. The SLLN is formulated in probability theory as

$$P_{1/2}^{\otimes \mathbb{N}_+}\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\omega_i = \frac{1}{2}\right\} = 1. \tag{4.1.6}$$

This statement means that the relative frequencies of positions of particle becomes one half almost surely. However, we give a different formulation of the SLLN in our setting because no probability measure enters the setting. According to the excellent book "Probability and Finance: It's Only a Game!" by Shafer and Vovk [SV01], there exists a strategy $\hat{\mu}_{\mathrm{SLLN}}$ for the agent such that if an infinite sequence $\omega_1\omega_2\cdots \in \{0,1\}^{\mathbb{N}_+}$ of positions of the particle does *not* satisfy the SLLN (4.1.2), the second law of thermodynamics is then violated, i.e., (4.1.4) holds for the sequence. Equivalently, if the sequence retains the second law under the strategy $\hat{\mu}_{\mathrm{SLLN}}$, the positions of the particle necessarily obey the SLLN (4.1.2). This implies that even though no probability measure is assumed, the empirical distribution for positions of the particle must be consistent with the equilibrium distribution due to the constraint by the second law. From the viewpoint of statistical mechanics, the second law of thermodynamics is a consequence of the equilibrium distribution and equilibrium statistical mechanics gives a microscopic foundation for the second law. However, according to the theorem by Shafer and Vovk, the second law of thermodynamics (4.1.4) under some strategy requires that the sequence must be random in the sense that it satisfies the SLLN (4.1.2), and leads to the equilibrium distribution in this sense. Our purpose of the chapter is to investigate this novel aspect of the relationship between equilibrium statistical mechanics and the second law of thermodynamics for generic small systems.

4.1.2 Summary of the results

The argument in subsection 4.1.1 suggests that statistical properties of equilibrium states are characterized by the second law of thermodynamics. In this chapter, we extend the above example to a system with a generic Hamiltonian on a finite state space. First, we consider the situation that the external agent has an ability to prepare arbitrary Hamiltonians during cyclic operations. We show that there exists a strategy for the agent such that the empirical distribution for a sequence satisfying the second law of thermodynamics necessarily converges to the Gibbs distribution of the initial Hamiltonian (Theorem 4.2.3.1). As in the case of the single-particle ideal-gas engine, this result can be interpreted as a statement that the empirical statistics must be consistent with the assumption of statistical mechanics, i.e., the Gibbs distribution, due to the second law.

Second, we study the empirical statistics in the same manner when we restrict the ability of the agent. The assumption that the agent can prepare arbitrary Hamiltonians is too demanding for his ability because we control only a small number of parameters in the Hamiltonian, such as the magnetic field, in many physical situations. The main aim of the present chapter is to propose a protocol corresponding to such a restricted situation and to determine what statistical properties are observed in that situation. This restriction weakens the ability of the agent, and we thus expect that the statistical property the second law imposes also weakens. We find that the Gibbs distributions for Hamiltonians having a small number of parameters linearly coupled to conjugate variables form an exponential family and propose a new protocol where the agent has to construct his strategy by changing only these parameters. Our main contribution is that in the protocol there exists a strategy such that the empirical mean of the conjugate variable with respect to the control parameter for sequence satisfying the second law converges to the equilibrium value (Theorem 4.3.2.1). This result suggests that there is a hierarchy of statistical properties observed under the second law according to his ability.

4.1.3 Related Studies

Several studies shares the same mathematical structures and techniques with our work, although our motivation and formalism in this chapter is quite different from the usual one in statistical mechanics and stochastic thermodynamics. We review related studies in this subsection.

First, an important property of the exponential of the accumulation of the extracted work is martingality, which is a fundamental concept in the theory of stochastic processes [Wil91, Doo53]. The martingale property is useful to investigate the statistics at stopping times and extreme value statistics. In the context of nonequilibrium thermodynamics, the novel statistical properties of stochastic entropy production were recently studied [CG11, NRJ17, Ner20, MSM+21] based on this property of martingales and the fact that the exponentiated negative entropy production or its modification is a martingale. Although our results are also based on the fact that the exponential of the accumulation of the extracted

work is a martingale[3)], we use another property that was first found by Jean Ville [Vil39], the characterization of almost sure properties in terms of martingales. See subsection 4.2.4. As other interesting study concerning the martingale property in physical systems, see [MS20] for example.

Second, we use universal coding theory [Grü07] to construct a strategy for the agent. The prediction strategy in universal coding is useful for proving our main result because our problem is similar to the coding or prediction of the outcomes ω_n in the n-th cycle from the past sample data $\omega_1 \dots \omega_{n-1}$ where the performance of the prediction is measured by the log-loss function. The analogy between gambling and source coding problems was first discussed by Kelly [Kel56]. See also Chapter 6 of [CT12]. In addition, the analogy with information thermodynamics was pointed out in [VPM16, Ito16, TMMR20]. Specifically, Refs. [TMMR20] applies universal coding theory to information thermodynamics to construct an optimal work extraction protocol. A crucial difference of our work from this study is that we analyze the asymptotic behavior of extracted work for *individual* sequences [VPM16, KG11] and attempt to find statistical properties shared by sequences satisfying the second law.

Finally, we stress that our studies are based on an analogy between the work extraction in thermodynamics and betting in game-theoretic probability theory [SV01, SV19]. Game-theoretic probability theory is a mathematical formulation of probability theory alternative to the conventional measure-theoretic one. In game-theoretic probability theory, a gambler named Skeptic bets money on head or tail of a coin and a dealer named Reality choose the outcome. By repeating this gamble infinitely many times and imposing the duty that Reality must chooses a sequence of outcomes such that Skeptic cannot make infinitely much money, we study what statistical behavior is observed. For instance, as mentioned in subsection 4.1.1, it is possible to construct an explicit strategy such that Skeptic's capital grows infinitely as long as the sequence violates the law of large numbers. The work extraction we proposed in this chapter can be regarded as a game played between two players, Skeptic and Reality. The external agent who attempts to extract work from the heat engine as much as possible corresponds to the gambler Skeptic and the world who prepares the particle to retain the second law of thermodynamics corresponds to the dealer Reality. Although the first main result of this chapter in Section 4.2 is a straightforward extension of the theorem proved by Shafer and Vovk, the possibility of analogous analysis of thermodynamics and the novel aspect of the relationship between statistical mechanics and the second law are new findings of this chapter as long as the author knows.

[3)] In externally driven systems, the exponentiated negative entropy production in a time interval is not martingale in general [CG11, Ner20, MSM+21]. In this chapter, however, we do not consider the stochastic time evolution of the system explicitly and concentrate our interest on the sum of the extracted work obtained from statistically independent experiments. Therefore, the exponentiated negative entropy production in this chapter is indeed a martingale. See subsection 4.2.4.

4.1.4 Outline of the chapter

The remainder of the chapter is organized as follows. In Section 4.2, we formulate the work extraction in a similar manner to the single-particle ideal-gas engine in subsection 4.1.1 and give the first main result. We also discuss the mathematical backgrounds, Ville's theorem, behind our result. In Section 4.3, we propose another protocol where the ability of the agent is restricted, and prove the second main result. In Section 4.4, we discuss a game-theoretic interpretation of our protocols. We end our chapter with concluding remarks in Section 4.5.

4.1.5 Notations

This subsection summarize notations we use throughout this chapter. Since the author explain notations when they are first used, the readers can skip this subsection.

Let $\mathbb{N}_+ = \{1, 2, \dots\}$ be the set of positive integers. We use Ω to denote a finite set representing a microscopic state space. $\Omega^* \coloneqq \{\square\} \cup (\cup_{n=1}^{\infty} \Omega^n)$ denotes the set of finite strings over Ω, where $\square$ is the empty string. A string over Ω with length n is written as $\omega^n = \omega_1 \dots \omega_n \in \Omega^*$, $\omega_i \in \Omega$. We use $\xi = \omega_1 \omega_2 \cdots \in \Omega^{\mathbb{N}_+}$ to denote an infinite sequence on Ω. A real-valued function $H : \Omega \to \mathbb{R}$ defines a Hamiltonian on the state space Ω. For a positive real number $\beta > 0$, the Gibbs distribution for the Hamiltonian H at the inverse temperature β is defined as a probability distribution on Ω with density $g_{\beta H}(\omega) \coloneqq e^{-\beta(H(\omega) - F_\beta(H))}$ with respect to the counting measure, where $F_\beta(H) \coloneqq -\beta^{-1} \ln \sum_{\omega \in \Omega} e^{-\beta H(\omega)}$ is the free energy.

4.2 Analysis of Generic Hamiltonians

4.2.1 Setup

Let us consider a physical system whose state space is given by a finite set Ω. The thermodynamic property of the system in contact with the heat bath is described by a Hamiltonian $H : \Omega \to \mathbb{R}$ and an inverse temperature β of the bath. According to Kelvin's principle, which is one of the representation of the second law of thermodynamics, the positive amount of work cannot be extracted by any cyclic operations. Here an operation is said to be cyclic if the initial and final Hamiltonians coincide. This principle leads to the absence of the perpetual motion machine of the second kind. However, the second law of thermodynamics may be violated for some individual initial state and cyclic operation in the finite system. Therefore, the second law in small systems is usually formulated as a statement on the non-positivity of the *expectation value* of the extracted work assuming the underlying probability distribution. In this subsection, we review the usual formulation of the second law of thermodynamics in terms of the expectation value.

Let us consider the following type of cyclic process [EVdB11] to avoid taking the dynamical evolution of the system into consideration:

(P1) The agent quenches the Hamiltonian adiabatically from the initial Hamiltonian H to another one H'.

(P2) The agent equilibrates the system with the inverse temperature β.

(P3) The agent resets the system quasi-statically and isothermally

Let ω be an initial state of the system. We suppose that the microscopic state ω does not change during the adiabatic quenching process (P1). Under this assumption, the extracted work in the process (P1) is given by the decrease in internal energy $H(\omega) - H'(\omega)$. In the process (P2), the agent touches the system with the heat bath having the inverse temperature β and the system relaxes to the new equilibrium state for the quenched Hamiltonian H'. This equilibration process obviously requires no mechanical work. In the process (P3), the agent changes the Hamiltonian from the quenched one H' to the initial one H quasi-statically to make the whole process cyclic. Moreover, we crucially assume that the work extracted in the quasi-static isothermal process is equal to the decrease in free energy[4)]. This assumption implies that the extracted work in the process (P3) is given by $F_\beta(H') - F_\beta(H)$. Here the free energy for the Hamiltonian H at the inverse temperature β is defined as $F_\beta(H) \coloneqq -\beta^{-1} \ln \sum_{\omega\in\Omega} e^{-\beta H(\omega)}$. Therefore, the total amount of work extracted in this cyclic process is given by $W(\omega_1) \coloneqq H(\omega) - H'(\omega) + F_\beta(H') - F_\beta(H)$.

Now we suppose that the initial state is sampled according to an initial density ρ. The expectation value of the work extracted during the above cyclic process is given by

$$\mathbb{E}_\rho[W] = \sum_{\omega\in\Omega} \rho(\omega) \left[H(\omega) - H'(\omega) + F_\beta(H') - F_\beta(H)\right] = D(\rho\|g_{\beta H}) - D(\rho\|g_{\beta H'}). \tag{4.2.1}$$

Here $g_{\beta H}(\omega) \coloneqq e^{-\beta(H(\omega)-F_\beta(H))}$ is the Gibbs density function for the Hamiltonian with the inverse temperature β and $D(p\|q) \coloneqq \sum_{\omega\in\Omega} p(\omega) \ln(p(\omega)/q(\omega))$ is the Kullback-Leibler divergence between two densities p and q. From this expression, we have the following equivalence between the second law of thermodynamics and the Gibbs distribution.

Theorem 4.2.1.1. For a probability density ρ, $\mathbb{E}_\rho[W] \leq 0$ if and only if $\rho = g_{\beta H}$.

Proof. If the initial distribution is the Gibbs distribution for the initial Hamiltonian, the first term on the right-hand side of Eq. (4.2.1) vanishes and therefore the expectation value of the extracted work is non-positive for any quenched Hamiltonian H', i.e., $\mathbb{E}_\rho[W] = -D(\rho\|g_{\beta H'}) \leq 0$. Conversely, if $\rho \neq g_{\beta H}$, the value $\mathbb{E}_\rho[W]$ can be positive. Indeed, by choosing H' such that $\rho = g_{\beta H'}$, we have that $\mathbb{E}_\rho[W] = D(\rho\|g_{\beta H}) > 0$. □

Therefore, the second law of thermodynamics expressed in the form of the expectation value during the cyclic process (P1)-(P3) is equivalent to that the initial distribution is the Gibbs distribution for the initial Hamiltonian.

Finally, we remark that the single-particle ideal-gas engine in subsection 4.1.1 is formally considered to be an example of the above setup.

Example 4.2.1.2 (Single-particle ideal-gas engine)**.** Let $\Omega = \{0, 1\}$ be a state space. Each state $\omega \in \{0, 1\}$ codes the position of the particle in the box. The effective Hamiltonian for

[4)] We discuss the status of this assumption after Theorem 4.2.3.1 and Section 4.5.

the single-particle ideal gas is defined as $H_\mu(\omega) = -\beta^{-1}\ln \mu^\omega(1-\mu)^{1-\omega}V$, where V is the total volume of the box and the parameter $\mu \in (0,1)$ indicates the position of the barrier. The free energy for the Hamiltonian is given by $F_\beta(H_\mu) = -\beta^{-1}\ln V$. If the initial and quenched parameters are $1/2$ and μ, respectively, the extracted work is written as

$$W(\omega) = H_{1/2}(\omega) - H_\mu(\omega) + F_\beta(H_\mu) - F_\beta(H_{1/2}) = \beta^{-1}\ln 2\mu^\omega(1-\mu)^{1-\omega}, \tag{4.2.2}$$

which is identical to Eq. (4.1.1).

4.2.2 Prequential formulation

The purpose of the present chapter is to investigate the emergence of equilibrium statistical mechanics from the second law of thermodynamics without referring to any probability measure. Hence we have to remove the probability measure from the definition of the second law. In this subsection, as such a formulation, we give a prequential definition of the second law.

To remove the probability distribution, we consider the situation that the agent repeats cyclic processes (P1)-(P3) infinitely many times. First, the agent performs a cyclic process according to the protocol (P1)-(P3). Let ω_1 be an initial state and H_1 a quenched Hamiltonian in this process. After the first cycle, the agent determines a quenched Hamiltonian H_2 in the second cycle depending on the initial state ω_1 in the first cycle and performs the cyclic process (P1)-(P3) again. In general, we suppose that the agent chooses a quenched Hamiltonian in the n-th cycle depending on the past history $\omega_1 \dots \omega_{n-1} \in \Omega^{n-1}$ up to the $(n-1)$-th cycle. The assignment of a quenched Hamiltonian in the n-th cycle to each past history $\omega_1 \dots \omega_{n-1}$ specifies a strategy for the agent to extract work. Therefore, we call a function $\hat{H} : \Omega^* \to \mathbb{R}^\Omega$ *strategy* in this chapter. Here Ω^* denotes the set of finite strings over Ω including the empty string $\square$. For a strategy $\hat{H}$, $\hat{H}(\cdot|\square) : \Omega \to \mathbb{R}$ represents a quenched Hamiltonian in the first cycle and $\hat{H}(\cdot|\omega_1 \dots \omega_{n-1}) : \Omega \to \mathbb{R}$ represents a quenched Hamiltonian in the n-th cycle when the initial states up to the $(n-1)$-th cycle are $\omega_1 \dots \omega_{n-1}$. This scheme in which the agent decides his action based on the past history of outcomes is called *prequential* in statistics [Daw84] and *causal* or *nonanticipating* in information theory of gambling and portfolio theory [CT12].

The accumulation of the extracted work up to the n-th cycle is given by the sum of the extracted work in each cycle. For a strategy $\hat{H}$, we define the function $W^{\hat{H}} : \Omega^* \to \mathbb{R}$ as $W^{\hat{H}}(\square) = 0$ and

$$W^{\hat{H}}(\omega^n) \coloneqq \sum_{i=1}^{n} \left[H(\omega_i) - \hat{H}(\omega_i|\omega^{i-1}) + F_\beta(\hat{H}(\cdot|\omega^{i-1})) - F_\beta(H) \right] \tag{4.2.3}$$

for $\omega^n \coloneqq \omega_1 \dots \omega_n \in \Omega^n$. From the definition, $W^{\hat{H}}(\omega^n)$ gives the accumulation of the work up to the n-th cycle under the strategy $\hat{H}$ when the initial states up to the n-th cycle are $\omega_1 \dots \omega_n$.

Now we provide a prequential definition of the second law of thermodynamics.

Definition 4.2.2.1. Let $\xi = \omega_1\omega_2\cdots \in \Omega^{\mathbb{N}_+}$ be an infinite sequence over Ω and $\hat{H}$ be a strategy for the agent. We say that

(1) ξ *violates weakly* the second law of thermodynamics under the strategy $\hat{H}$ if

$$\sup_n W^{\hat{H}}(\omega_1\dots\omega_n) = \infty, \tag{4.2.4}$$

(2) ξ *violates* the second law of thermodynamics under the strategy $\hat{H}$ if

$$\lim_{n\to\infty} W^{\hat{H}}(\omega_1\dots\omega_n) = \infty. \tag{4.2.5}$$

Definition 4.2.2.1 is regarded as a definition of the perpetual motion machine of the second kind for individual sequences of states. Let ω_1 be an initial microscopic state of the engine in the first cycle. By performing the cyclic process (P1)-(P3) for the initial state, the agent extract work from the engine by $W^{\hat{H}}(\omega_1)$. The engine cannot be regarded as a perpetual motion machine of the second kind only because the value $W^{\hat{H}}(\omega_1)$ is positive. To say that the engine violates the second law, we require that for any given positive value $W_0 > 0$, the agent should extract an amount of work larger than W_0 by repeating cyclic processes as many times as he needs. Therefore, we define the second law for *infinite sequences of initial microscopic states of the engine* as indicated in Definition 4.2.2.1. We note that while in the above definition the violation of the second law depends on both a strategy the agent applies and an infinite sequences of microscopic states, the definition needs no underlying probability measure.

4.2.3 Convergence of empirical distribution to Gibbs distribution

Instead of introducing probability distributions, we consider which sequences satisfies the second law of thermodynamics and what statistical properties are shared among these sequences. In general, a statistical property is described by a subset of infinite sequences $E \subseteq \Omega^{\mathbb{N}_+}$. In this chapter, as a relevant statistical property, we focus on only the strong law of large numbers, i.e.,

$$\mathsf{SLLN} = \left\{ \xi = \omega_1\omega_2\cdots \in \Omega^{\mathbb{N}_+} : \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) = g_{\beta H}(\omega) \text{ for all } \omega \in \Omega \right\}. \tag{4.2.6}$$

Here the quantity $\sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i)/n$ is the empirical density for the string ω^n quantifying the relative frequency of ω in $\omega^n = \omega_1\dots\omega_n$. If the sequence $\xi = \omega_1\omega_2\dots$ are sampled identically and independently according to the Gibbs distribution, the SLLN (4.2.6) happens almost surely. In this sense, the SLLN (4.2.6) is a statistical property observed in equilibrium. Our purpose is to find a strategy $\hat{H}$ such that any sequences $\xi \in \Omega^{\mathbb{N}_+}$ satisfying the second law of thermodynamics under the strategy $\hat{H}$ in the sense of Definition 4.2.2.1 necessarily obey the SLLN, i.e., $\xi \in \mathsf{SLLN}$. The existence of such strategy implies that the empirical statistics of the sequences of initial states observed under the constraint by the second law must be consistent with equilibrium statistical mechanics. Based on the above reasoning, we provide the following theorem:

Theorem 4.2.3.1. Let $g_{\beta H}(\omega) = e^{-\beta H(\omega)+\beta F_\beta(H)}$ be the Gibbs density function for the initial Hamiltonian H at the inverse temperature β. There exists a strategy $\hat{H}$ such that if an infinite sequences $\xi \in \Omega^{\mathbb{N}_+}$ of initial states does not satisfy the SLLN, i.e., $\xi \notin \mathsf{SLLN}$, then the sequence ξ violates the second law of thermodynamics under the strategy.

This is the first main result of this chapter. We make several remarks on Theorem 4.2.3.1. First, we stress again that no probability measure for initial microscopic states are assumed. One may find that the Gibbs distribution is implicitly inserted in the definition of the extracted work (4.2.3) through the assumption that the work extracted in the quasi-static isothermal process is equal to the decrease in free energy. Although this reasoning is actually true as we will see in subsection 4.2.4, the assumption on the form of the extracted work does *not* immediately leads to the stochastic behavior of initial states and the content of Theorem 4.2.3.1 remains highly non-trivial. Second, we stress the difference from the argument in the end of subsection 4.2.1. There, if the initial distribution deviates from the Gibbs distribution for the initial Hamiltonian, we have to choose the quenched Hamiltonian depending on the deviation in order to extract a positive amount of work. In contrast, Theorem 4.2.3.1 claims the existence of a single *universal* strategy under which the second law is automatically violated for the sequence whose empirical statistics deviate from the Gibbs distribution even if we do not know the deviation.

We explain in subsection 4.2.4 the reason why the Gibbs distribution appears in Theorem 4.2.3.1 although no probability measure is assumed in our setting. We remark that Theorem 4.2.3.1 is just a straightforward extension of the case of single-particle ideal-gas engine in subsection 4.1.1 and it is nothing new mathematically. In addition to the proof in textbooks of Shafer and Vovk [SV01, SV19], there are several proofs of Theorem 4.2.3.1 such as Ref. [KT08] based on the maximum likelihood strategy and Ref. [KTT08] based on the Bayesian strategy. Nevertheless we prove Theorem 4.2.3.1 as a special case of Theorem 4.3.2.1.

Hereafter, we say that a strategy *forces* (resp. *weakly forces*) an event $E \subseteq \Omega^{\mathbb{N}_+}$ if infinite sequences over Ω that do not satisfy E violate (resp. weakly violate) the second of thermodynamics under strategy. According to this terminology, Theorem 4.2.3.1 claims the existence of a strategy that forces the strong law of large numbers for the empirical distribution.

Finally, we present a proof based on Refs. [SV01, SV19].

Proof of Theorem 4.2.3.1. First, we prove two lemmas.

Lemma 4.2.3.2. For any $\omega \in \Omega$ and any sufficiently small positive real $\epsilon > 0$, there exists a strategy that weakly forces

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) \leq g_{\beta H}(\omega) + \epsilon. \tag{4.2.7}$$

Similarly, for any $\omega \in \Omega$ and any sufficiently small positive real $\epsilon > 0$, there exists a strategy that weakly forces

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) \geq g_{\beta H}(\omega) - \epsilon. \tag{4.2.8}$$

Proof. Fix an element $\omega \in \Omega$. We define $R_\omega := \min\{1/g_{\beta H}(\omega), 1/(1-g_{\beta H(\omega)})\} \in (0,1)$ and choose a real number $\epsilon \in (0, R_\omega/2)$. Consider a strategy

$$\hat{H}(\omega_n|\omega_1 \dots \omega_{n-1}) = H^{+,\omega,\epsilon}(\omega_n) := H(\omega_n) - \beta^{-1}\ln[1+\epsilon(1_{\{\omega\}}(\omega_n) - g_{\beta H}(\omega))] \tag{4.2.9}$$

$$= \begin{cases} H(\omega) - \beta^{-1}\ln[1+\epsilon(1-g_{\beta H}(\omega))] & \text{if } \omega_n = \omega \\ H(\omega) - \beta^{-1}\ln[1-\epsilon g_{\beta H}(\omega)] & \text{if } \omega_n \neq \omega. \end{cases} \tag{4.2.10}$$

We remark that the free energy of $H^{+,\omega,\epsilon}$ is equal to $F_\beta(H)$ because

$$\begin{aligned} \sum_\eta e^{-\beta H^{+,\omega,\epsilon}(\eta)} &= \sum_\eta e^{-\beta H(\eta)}(1+\epsilon(1_{\{\omega\}}(\eta) - g_{\beta H}(\omega))) \\ &= \sum_\eta e^{-\beta H(\eta)} + \epsilon\left[e^{-\beta H(\omega)} - \left(\sum_\eta e^{-\beta H(\eta)}\right) g_{\beta H}(\omega)\right] = e^{-\beta F_\beta(H)}. \end{aligned}$$

Under this strategy, the accumulation of the extracted work is given by

$$\begin{aligned} W^{\hat{H}}(\omega^n) &= \sum_{i=1}^n [H(\omega_i) - H^{+,\omega,\epsilon}(\omega_i)] \\ &= \beta^{-1}\sum_{i=1}^n \ln[1+\epsilon(1_{\{\omega\}}(\omega_i) - g_{\beta H}(\omega))]. \end{aligned} \tag{4.2.11}$$

Since $\epsilon < R_\omega/2$, $\epsilon(1_{\{\omega\}}(\omega_i) - g_{\beta H}(\omega)) < 1/2$. By using the inequality $\ln(1+t) \geq t - t^2$ for $|t| \leq 1/2$, we have that

$$\begin{aligned} W^{\hat{H}}(\omega^n) &\geq \beta^{-1}\sum_{i=1}^n \left[\epsilon\left(1_{\{\omega\}}(\omega_i) - g_{\beta H}(\omega)\right) - \epsilon^2\left(1_{\{\omega\}}(\omega_i) - g_{\beta H}(\omega)\right)^2\right] \\ &\geq \beta^{-1}\sum_{i=1}^n \left[\epsilon\left(1_{\{\omega\}}(\omega_i) - g_{\beta H}(\omega)\right) - \epsilon^2\right] \\ &= \epsilon n\beta^{-1}\left[\frac{1}{n}\sum_{i=1}^n 1_{\{\omega\}}(\omega_i) - g_{\beta H}(\omega) - \epsilon\right] \end{aligned} \tag{4.2.12}$$

In the second line, we have used $\left(1_{\{\omega\}}(\omega_i) - g_{\beta H}(\omega)\right)^2 \leq 1$. For a given sequence $\xi = \omega_1\omega_2\dots$, if $\sup_n W^{\hat{H}}(\omega_n) < \infty$, there exists a real number C_ξ such that $W^{\hat{H}}(\omega_n) \leq C_\xi$ for all n. We then obtain that

$$\frac{1}{n}\sum_{i=1}^n 1_{\{\omega\}}(\omega_i) \leq g_{\beta H}(\omega) + \epsilon + \frac{C_\xi}{\epsilon n\beta^{-1}} \tag{4.2.13}$$

for all n, which leads to

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{i=1}^n 1_{\{\omega\}}(\omega_i) \leq g_{\beta H}(\omega) + \epsilon. \tag{4.2.14}$$

This means that the strategy (4.2.9) weakly forces (4.2.7). By a similar argument, we find that the strategy

$$\hat{H}(\omega_n|\omega_1\dots\omega_{n-1}) = H^{-,\omega,\epsilon}(\omega_n) \coloneqq H(\omega_n) - \beta^{-1}\ln[1-\epsilon(1_{\{\omega\}}(\omega_n) - g_{\beta H}(\omega))] \tag{4.2.15}$$

weakly forces (4.2.8).

□

Lemma 4.2.3.3. Let $\hat{H}_k$ $(k = 1, 2, \dots,)$ be a strategy that weakly forces $E_k \subseteq \Omega^{\mathbb{N}_+}$. There exists a strategy that weakly forces $\bigcap_{k=1}^{\infty} E_k$.

Proof. Let $\hat{H}_k$ be a strategy that weakly forces $E_k \subseteq \Omega^{\mathbb{N}_+}$. We define a strategy $\hat{H}_*$ as

$$\hat{H}_*(\omega_n|\omega^{n-1}) \coloneqq H(\omega_n) - \beta^{-1}\left[\ln\left(\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^n)}\right) - \ln\left(\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1})}\right)\right]. \tag{4.2.16}$$

We remark that $\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^n)}$ is well-defined. We can prove this fact by induction. If $n = 0$, $\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\square)} = 1$ since $W^{\hat{H}_k}(\square) = 0$. Let us assume that $\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^n)} < \infty$. Since $\sum_{\eta} e^{\beta W^{\hat{H}_k}(\omega^n\eta)} g_{\beta H}(\eta) = e^{\beta W^{\hat{H}_k}(\omega^n)}$, we obtain that $\sum_{\eta}\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^n\eta)} g_{\beta H}(\omega) < \infty$, which implies $\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n+1})} < \infty$ for any ω^{n+1}.

The free energy of the Hamiltonian $\hat{H}_*(\cdot|\omega^{n-1})$ is equal to $F_\beta(H)$ because

$$\begin{aligned}
\sum_{\eta} e^{-\beta H_*(\eta|\omega^{n-1})} &= \sum_{\eta} e^{-\beta H(\eta)} \frac{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1}\eta)}}{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1})}} \\
&= \sum_{\eta} \frac{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1}) - \beta\hat{H}_k(\eta|\omega^{n-1}) + \beta F_\beta(\hat{H}_k(\cdot|\omega^{n-1})) - \beta F_\beta(H)}}{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1})}} \\
&= \frac{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1})} \sum_{\eta} e^{-\beta\hat{H}_k(\eta|\omega^{n-1}) + \beta F_\beta(\hat{H}_k(\cdot|\omega^{n-1}))} e^{-\beta F_\beta(H)}}{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1})}} \\
&= \frac{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1})} e^{-\beta F_\beta(H)}}{\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{n-1})}} = e^{-\beta F_\beta(H)}.
\end{aligned}$$

Under this strategy, the accumulation of the extracted work is given by

$$\begin{aligned}
W^{\hat{H}_*}(\omega^n) &= \sum_{i=1}^{n} [H(\omega_i) - \hat{H}_*(\omega_i|\omega^{i-1})] \\
&= \beta^{-1}\sum_{i=1}^{n}\left[\ln\left(\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^i)}\right) - \ln\left(\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^{i-1})}\right)\right] \\
&= \beta^{-1}\ln\left(\sum_{k=1}^{\infty} 2^{-k} e^{\beta W^{\hat{H}_k}(\omega^n)}\right).
\end{aligned} \tag{4.2.17}$$

We now assume that $\sup_n W^{\hat{H}_*}(\omega^n) < \infty$. From the expression (4.2.17), we conclude that $\sup_n W^{\hat{H}_k}(\omega^n) < \infty$ for all $k = 1, 2, \ldots$. This implies that the strategy $\hat{H}_*$ weakly forces $\bigcap_{k=1}^{\infty} E_k$.

□

We consider the set of strategies

$$\left\{ H^{\sigma,\omega,2^{-k}} : \sigma \in \{-,+\}, \omega \in \Omega, k \geq k_0, 2^{-k_0} < \min_{\omega\in\Omega} R_\omega/2 \right\}. \tag{4.2.18}$$

By Lemma 4.2.3.2, the strategy $H^{+,\omega,2^{-k}}$ weakly forces

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) \leq g_{\beta H}(\omega) + 2^{-k}, \tag{4.2.19}$$

and the strategy $H^{-,\omega,2^{-k}}$ weakly forces

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) \geq g_{\beta H}(\omega) - 2^{-k}. \tag{4.2.20}$$

By Lemma 4.2.3.3, the strategy H_* defined in the proof of Lemma 4.2.3.3 weakly forces

$$g_{\beta H}(\omega) - 2^{-k} \leq \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) \leq \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) \leq g_{\beta H}(\omega) + 2^{-k} \tag{4.2.21}$$

for all $\omega \in \Omega$ and sufficiently large k. This implies that the strategy H_* weakly forces

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{\omega\}}(\omega_i) = g_{\beta H}(\omega) \tag{4.2.22}$$

for all $\omega \in \Omega$. By Lemma 4.3.3.1, it is possible to construct a strategy that forces (4.2.22).

□

4.2.4 Martingale and Ville's theorem

We clarify a general structure behind Theorem 4.2.3.1. Although we construct explicitly a strategy that make the empirical distribution converge the Gibbs distribution in subsection 4.3.3, Theorem 4.2.3.1 follows from a more general theorem proved by Ville [Vil39]. The essence of Ville's theorem is some kind of equivalence between the asymptotic behavior of martingales and almost sure properties.

To see this, we first clarify the martingale property of the exponentiated work. The accumulation of the extracted work (4.2.3) can be written as the logarithmic likelihood ratio function for Gibbs distributions:

$$\beta W^{\hat{H}}(\omega^n) = \ln \frac{\hat{q}(\omega^n)}{g_{\beta H}^n(\omega^n)}, \tag{4.2.23}$$

where $g_{\beta H}^n(\omega^n) \coloneqq \prod_{i=1}^n g_{\beta H}(\omega_i)$ and

$$\hat{q}(\omega^n) \coloneqq \prod_{i=1}^n g_{\beta \hat{H}(\cdot|\omega^{i-1})}(\omega_i). \tag{4.2.24}$$

Since $g_{\beta \hat{H}(\cdot|\omega^{i-1})}$ specifies the conditional probability density conditioned on the past history ω^{i-1}, the function (4.2.24) gives a probability density on Ω^n. Conversely, a stochastic process on $\Omega^{\mathbb{N}_+}$ with strictly positive probability densities $\hat{q}$ specifies a strategy for the agent in our setting through the relation $g_{\beta \hat{H}(\cdot|\omega^{n-1})}(\omega_n) = \hat{q}(\omega_n|\omega^{n-1})$. Therefore, a strategy $\hat{H}$ is identified with a stochastic process having strictly positive densities.

Let us consider a discrete time stochastic process $M : \Omega^* \to \mathbb{R}$. We say that M is a martingale with respect to a probability measure P on $\Omega^{\mathbb{N}_+}$ if $\mathbb{E}_P[M_n|\omega^{n-1}] = M(\omega^{n-1})$ for any $\omega^{n-1} \in \Omega^*$ [Wil91]. Here $M_n(\omega_1\omega_2\dots) \coloneqq M(\omega^n)$ and $\mathbb{E}_P[\,\cdot\,|\omega^n]$ denotes the conditional expectation conditioned on the past history ω^n. For a fixed strategy $\hat{H}$, it is easy to see that the exponential of the accumulation of the extracted work $e^{\beta W^{\hat{H}}}$ is a positive martingale with respect to the infinite product of Gibbs distributions $g_{\beta H}^{\mathbb{N}_+}$,

$$\mathbb{E}_{g_{\beta H}^{\mathbb{N}_+}}[e^{\beta W_n^{\hat{H}}} \mid \omega^{n-1}] = e^{\beta W^{\hat{H}}(\omega^{n-1})},$$

In particular, $\mathbb{E}_{g_{\beta H}^{\mathbb{N}_+}}[e^{\beta W_n^{\hat{H}}}] = e^{\beta W^{\hat{H}}(\square)} = 1$. In fact,

$$\mathbb{E}_{g_{\beta H}}[e^{\beta W_n^{\hat{H}}}|\omega^{n-1}] = \sum_{\omega_n \in \Omega} \frac{\hat{q}(\omega^n)}{g_{\beta H}^n(\omega^n)} g_{\beta H}(\omega_n) = \sum_{\omega_n \in \Omega} \frac{\hat{q}(\omega^{n-1}\omega_n)}{g_{\beta H}^{n-1}(\omega^{n-1})} = \frac{\hat{q}(\omega^{n-1})}{g_{\beta H}^{n-1}(\omega^{n-1})} = e^{\beta W^{\hat{H}}(\omega^{n-1})},$$

where we have used $\sum_{\omega_n \in \Omega} \hat{q}(\omega^{n-1}\omega_n) = \hat{q}(\omega^{n-1})$ in the last line. Conversely, for a given positive martingale M with respect to $g_{\beta H}^{\mathbb{N}_+}$ starting from $M(\square) = 1$, there exists a strategy $\hat{H}$ such that $e^{\beta W^{\hat{H}}(\omega^n)} = M(\omega^n)$ for any $\omega^n \in \Omega^*$. To prove this, we notice that $\hat{q}(\omega_n|\omega^{n-1}) \coloneqq g_{\beta H}(\omega_n) M(\omega^{n-1}\omega_n)/M(\omega^{n-1})$ is a positive probability density on Ω. Indeed, the positivity of M leads to $\hat{q}(\omega_n|\omega^{n-1}) > 0$ and the martingale property of M implies that

$$\sum_{\omega_n \in \Omega} \hat{q}(\omega_n|\omega^{n-1}) = \frac{1}{M(\omega^{n-1})} \sum_{\omega_n \in \Omega} g_{\beta H}(\omega_n) M(\omega^{n-1}\omega_n) = \frac{M(\omega^{n-1})}{M(\omega^{n-1})} = 1.$$

By defining the strategy $\hat{H}$ through the relation $g_{\beta \hat{H}(\cdot|\omega^{n-1})}(\omega_n) = \hat{q}(\omega_n|\omega^{n-1})$, we have that $e^{\beta W^{\hat{H}}(\omega^n)} = M(\omega^n)$ for any $\omega^n \in \Omega^*$ including the case $\omega^n = \square$. Thus, there is a one-to-one correspondence between the processes of the exponentiated extracted work and the positive martingales with respect to the Gibbs distribution.

Ville's theorem [Vil39, SV01, SV19] in our setting claims that *a measurable set $E \subseteq \Omega^{\mathbb{N}_+}$ has probability one with respect to $g_{\beta H}^{\mathbb{N}_+}$ if and only if there exists a positive martingale $M : \Omega^* \to \mathbb{R}_+$ with respect to the same probability measure such that $\lim_{n\to\infty} M_n(\xi) = \infty$ for any $\xi \notin E$*. Using the equivalence between positive martingales and strategies mentioned above, we rephrase Ville's theorem as follows.

Theorem 4.2.4.1 ([Vil39, SV01, SV19])**.** Let $E \subseteq \Omega^{\mathbb{N}_+}$ be a measurable set. The property E happens almost surely with respect to $g_{\beta H}^{\mathbb{N}_+}$ if and only if there exists a strategy $\hat{H} : \Omega^* \to \mathbb{R}^\Omega$ that forces E.

Ville's theorem clarifies the reason why the convergence to the Gibbs distribution occurs under the second law in spite that no probability measure comes into our setting. Let us consider the experiment that we sample initial microscopic states infinitely many times independently and identically according to the Gibbs distribution $g_{\beta H}$. We take a statistical property $E \subseteq \Omega^{\mathbb{N}_+}$ with probability one in the experiment such as the SLLN (4.2.6). According to Ville's theorem, there exists a strategy $\hat{H}$ such that any sequences violate the second law under the strategy if they do not have the property E. In other words, whether a statistical property is consistent with equilibrium statistical mechanics or not is characterized in terms of the violation of the second law of thermodynamics. In addition, our definition of the second law (Definition 4.2.2.1) needs no probability measure. This is an essential structure that makes it possible to argue the emergence of equilibrium statistical mechanics from the second law without referring to any probability measures.

4.3 Analysis of Parametric Hamiltonian

In the setting in Section 4.2, the agent is assumed to have an ability to prepare any complicated Hamiltonian containing non-local and many-body interactions. As a more physical situation, it is natural to restrict the possible operations into changing a small number of parameters in a certain class of Hamiltonians. In this section, we investigate such situation.

4.3.1 Preliminaries: exponential family

We focus on the Hamiltonians on Ω with a certain number of externally controllable parameters. We assume that the Hamiltonians have the form $\beta H_\theta(\omega) \coloneqq -\theta \cdot \phi(\omega) - h(\omega)$, where the parameter θ corresponds to the control parameters taking values in the parameter space $\Theta \subseteq \mathbb{R}^k$, $\phi : \Omega \to \mathbb{R}^k$ is the conjugate variable with respect to θ, the dot " $\cdot$ " denotes the usual inner product in $\mathbb{R}^k$ and $h : \Omega \to \mathbb{R}$ is the remaining part of the Hamiltonian. Θ is assumed to be an open and convex subset of $\mathbb{R}^k$. Moreover, we suppose that the representation of the Hamiltonians is minimal, i.e., both $\{\theta_i : i = 1, \dots, k\}$ and $\{\phi_i : i = 1, \dots, k\}$ are affinely independent.

The Gibbs distribution for these Hamiltonians forms a minimal canonical *exponential family* [BN78] $\mathcal{P}(\Theta, \phi, h)$, which is a family of distributions P_θ on Ω with densities $p_\theta(\omega) \coloneqq g_{\beta H_\theta}(\omega) = e^{\theta \cdot \phi(\omega) + h(\omega) - \psi(\theta)}$. The function $\psi(\theta) \coloneqq \ln \sum_{\omega \in \Omega} e^{\theta \cdot \phi(\omega) + h(\omega)}$ is the Massieu function, which is related to the free energy $F_\beta(H_\theta)$ through $F_\beta(H_\theta) = -\beta^{-1}\psi(\theta)$. The function $\psi(\theta)$ is differentiable infinitely many times and strictly convex on Θ. The expectation value and covariance matrix of ϕ are respectively given by $\mathbb{E}_\theta(\phi) = \nabla_\theta \psi(\theta) =: \mu(\theta)$ and $\mathrm{Cov}_\theta(\phi) = \nabla_\theta \nabla_\theta \psi(\theta) = \mathbb{E}_\theta[-\nabla_\theta \nabla_\theta \ln P_\theta] =: I(\theta)$. $I(\theta)$ is the Fisher information at θ for the family $\mathcal{P}(\Theta, \phi, h)$ and positive definite over Θ. In the language of statistical physics, $I(\theta)$ is

the isothermal susceptibility matrix and the relation $\mathrm{Cov}_\theta(\phi) = I(\theta)$ gives the fluctuation-response relation for a static isothermal response. The strict convexity of ψ implies that the map $\theta \mapsto \mu(\theta)$ is invertible and that the elements of $\mathcal{P}(\Theta, \phi, h)$ can be reparametrized by the mapping $\mu \mapsto P_\theta$. $\theta(\mu)$ denotes the inverse and $\Xi \coloneqq \mu(\Theta)$ the conjugate parameter space. We note that $\nabla_\mu \theta(\mu) = I(\mu)$ gives the Fisher information in μ-parametrization and is equals to $\mathrm{Cov}_\mu^{-1}(\phi)$. The Kullback-Leibler divergence between P_θ and $P_{\theta'}$ is given by

$$D(P_\theta \| P_{\theta'}) = (\theta - \theta') \cdot \mu(\theta) - \psi(\theta) + \psi(\theta'). \tag{4.3.1}$$

The class of Hamiltonians having the above form describes a variety class of models that appears in statistical physics. We give two examples belonging to the above class.

Example 4.3.1.1 (One-dimensional Ising model)**.** The first example is a one-dimensional Ising model. Let N be a positive integer. The state space of N Ising spins on one-dimensional chain is $\Omega = \{-1, 1\}^N$ and a state is described by $\omega = (\sigma_1, \dots, \sigma_N)$, where $\sigma_i \in \{-1, 1\}$ denotes the state of spin at the site $i \in \{1, \dots, N\}$. The Ising model under a homogeneous magnetic field is described by the Hamiltonian $\beta H_\theta = -\beta h_{\mathrm{ex}} \sum_{i=1}^N \sigma_i - \beta J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1}$, where h_{ex} is the external field and J is the coupling constant. If the agent changes only the magnetic field uniformly, the parameter is chosen as $\theta = \beta h_{\mathrm{ex}} \in \Theta \coloneqq \mathbb{R}$, the conjugate variable as the total magnetization $\phi(\omega) = \sum_{i=1}^N \sigma_i$, and the remaining part as $h(\omega) = \beta J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1}$. In this case, the expectation of ϕ is the average magnetization $\mu(\theta) = \sum_{i=1}^n P_\theta(\sigma_i)$ and the Fisher information is $I(\theta) = \beta^{-1}\chi$, where χ is the magnetic susceptibility. If the agent changes the coupling constant J in addition to the magnetic field, the parameter becomes two-dimensional vector $\theta = (\beta h_{\mathrm{ex}}, \beta J)$, the conjugate variable $\phi = (\sum_{i=1}^n \sigma_i, \sum_{i=1}^{n-1} \sigma_i \sigma_{i+1})$ and the remaining part $h = 0$. The parametrization for a given physical model is thus not unique in general.

Example 4.3.1.2 (Positive distribution)**.** The second example is positive distributions on a general finite state space. This case corresponds to that we argued in Section 4.2. Let $\Omega = \{0, 1, 2, \dots, k\}$ be a finite set. Putting $\Theta = \{(\ln(p_i/p_0))_{i=1}^k \in \mathbb{R}^k : 0 < p_i < 1, \sum_{j=1}^k p_j < 1, p_0 = 1 - \sum_{i=1}^k p_i\}$, $\theta_i = \ln(p_i/p_0)$, $\phi_i = 1_{\{i\}}$, $h = 0$, $\psi(\theta) = -\ln(1 - \sum_{i=1}^k p_i) = -\ln p_0$, we have that

$$p_\theta(i) = \exp\left\{ \sum_{j=1}^k \theta_j \phi_j(i) - \psi(\theta) \right\} = p_i \tag{4.3.2}$$

for every $i \in \Omega$. The exponential family is therefore the family of strictly positive distributions on Ω. The empirical mean of ϕ is identified with the empirical density and the expectation $\mu_i(\theta)$ is given by p_i for $i = 1, \dots, k$.

4.3.2 Setup and Result

The protocol we study in this subsection is almost the same as that in Section 4.2. A crucial difference is that the agent cannot prepare an arbitrary Hamiltonian in general and he has to construct a strategy by tuning only a small number of control parameters.

We assume that the initial Hamiltonian has the form $\beta H_\theta = -\theta \cdot \phi - h$ for some $\theta \in \Theta$. A strategy the agent applies is characterized by a function $\hat{\theta} : \Omega^* \to \Theta$. The operation in the n-th cycle is performed according to the following protocol:

(P1) The agent quenches the Hamiltonian adiabatically from the initial Hamiltonian H_θ to another one $H_{\hat{\theta}(\omega^{n-1})}$ when the initial states up to the $(n-1)$-th cycle are $\omega^{n-1} = \omega_1 \dots \omega_{n-1}$.

(P2) The agent equilibrates the system with the inverse temperature β.

(P3) The agent resets the system quasi-statically and isothermally.

The accumulation of extracted work $W^{\hat{\theta}} \coloneqq W^{H_{\hat{\theta}}}$ in this protocol under the strategy $\hat{\theta}$ is given by

$$\beta W^{\hat{\theta}}(\omega^n) = \sum_{i=1}^{n} \left[(\hat{\theta}(\omega^{i-1}) - \theta) \cdot \phi(\omega_i) - \psi(\hat{\theta}(\omega^{i-1})) + \psi(\theta) \right]. \tag{4.3.3}$$

Let $E \subseteq \Omega^{\mathbb{N}_+}$ be an almost sure property under the infinite product of the Gibbs distributions $p_\theta^{\mathbb{N}_+} = g_{\beta H_\theta}^{\mathbb{N}_+}$. Although there exists a strategy $\hat{H}$ that forces E according to Ville's theorem (Theorem 4.2.4.1), the strategy $\hat{H}$ may not be realizable within the above protocol. In general, the decrease in the number of possible strategies the agent can apply leads to the decrease in the variety of almost sure properties forced by the second law. The question we study in this section is what statistical properties are forced when we restrict the ability of the agent.

We expect from the expression (4.3.3) that such properties are restricted to the statistics of the conjugate variable ϕ with respect to the control parameter θ. Indeed, we find that there exists a strategy in the restricted protocol that forces the strong law of large numbers for the conjugate variable.

Theorem 4.3.2.1. Let $\overline{\phi}_n$ be the empirical mean of ϕ defined by

$$\overline{\phi}_n(\omega^n) \coloneqq \frac{1}{n} \sum_{i=1}^{n} \phi(\omega_i). \tag{4.3.4}$$

There exists a strategy $\hat{\theta} : \Omega^* \to \Theta$ that forces

$$\lim_{n \to \infty} \overline{\phi}_n = \mu(\theta). \tag{4.3.5}$$

This is the second main result of the present chapter. Theorem 4.2.3.1 is a special case of Theorem 4.3.2.1 because the empirical mean of ϕ in Example 4.3.1.2 can be identified with the empirical density.

Since $\mu(\theta)$ is the equilibrium value of the conjugate variable, Theorem 4.3.2.1 implies that the infinite sequences satisfying the second law are indistinguishable from random sequences sampled from the equilibrium distribution as long as we observe the conjugate variable. In

contrast to Theorem 4.2.3.1, the relative frequencies of initial microscopic states may not converge to the Gibbs distribution under the strategy. For instance, when the agent controls the homogeneous magnetic field on the free spin system, the average magnetization should converge to the equilibrium value but the relative frequencies of spin configurations having the same magnetization may not be controlled. Thus the statistical properties the agent can force under the second law depend on what kind of physical operations are allowed for the agent.

4.3.3 Proof of Theorem 4.3.2.1

We present a proof of Theorem 4.3.2.1. By the following lemma, it is sufficient to prove the existence of a strategy that weakly forces (4.3.5) and takes values in a compact set containing the initial parameter θ as an interior point. It is proved in the same way as Lemma 3.1 in [SV01].

Lemma 4.3.3.1. Let $E \subseteq \Omega^{\mathbb{N}}$ be an event and $\Theta_0 \subset \Theta$ a compact convex subset with $\theta \in \mathrm{int}\Theta_0$. If there exists a strategy that weakly forces E and takes values in Θ_0, then there is also a strategy that forces E and takes values in the same set Θ_0.

Proof. For a given strategy $\hat{\theta} : \Omega^* \to \Theta_0$ and a positive real number $C > 0$, we define the "stopped strategy" $\hat{\theta}^{(C)}$ as

$$\hat{\theta}^{(C)}(\omega^n) = \begin{cases} \hat{\theta}(\omega^n) & \text{if } \beta W^{\hat{\theta}}(\omega^m) < C \text{ for all } m \leq n \\ \theta & \text{otherwise.} \end{cases} \tag{4.3.6}$$

Let $\hat{\theta}$ be a strategy that weakly forces E and $\hat{\theta}^{(C)}$ its stopped strategy for $C > 0$. Consider the countable number of stopped strategies $\hat{\theta}^{(2^i)}$ for $i = 1, 2, \ldots$. Since $\hat{\theta}^{(2^i)}(\omega^n) \in \Theta_0$ and Θ_0 is compact, $\hat{\theta}^*(\omega^n) := \sum_{i=1}^{\infty} 2^{-i}\hat{\theta}^{(2^i)}(\omega^n)$ exists for any $\omega^n \in \Omega^*$. The closedness and convexity of Θ_0 lead to $\hat{\theta}^*(\omega^n) \in \Theta_0$. Hence, the function $\hat{\theta}^* : \Omega^* \to \Theta_0$ defines a strategy taking values in Θ_0. We obtain from the convexity of ψ and Jensen's inequality that

$$W^{\hat{\theta}^*}(\omega^n) \geq \sum_{i=1}^{\infty} 2^{-i} W^{\hat{\theta}^{(2^i)}}(\omega^n). \tag{4.3.7}$$

For $\omega^\infty \in E^c$, the limit $\lim_n \beta W^{\hat{\theta}^{(C)}}(\omega^n)$ exists and is larger than or equal to C because $\sup_n W^{\hat{\theta}}(\omega^n) = \infty$. Since $\lim_n \beta W^{\hat{\theta}^{(2^i)}}(\omega^n) \geq 2^i$ for $\omega^\infty \in E^c$, $\lim_n W^{\hat{\theta}^*}(\omega^n)$ diverges to infinity for $\omega^\infty \in E^c$. □

Our method of the proof is based on the maximum likelihood strategy. If the agent applies a constant strategy, i.e., $\hat{\theta} = \theta'$, the accumulation of extracted work up to the n-th cycle is given by

$$\beta W^{\hat{\theta}}(\omega^n) = \ln \frac{p^n_{\theta'}(\omega^n)}{p^n_{\theta}(\omega^n)}. \tag{4.3.8}$$

The maximum likelihood strategy is defined as the maximizer of (4.3.8) for each ω^n. Because the maximizer of (4.3.8) over the open set Θ may not exist, we restrict the range of the strategy to a compact set Θ_0 containing the initial parameter θ as an interior point. The maximum likelihood estimator $\hat{\theta}_{\mathrm{ML}} : \Omega^* \to \Theta_0$ with respect to Θ_0 is defined as

$$\hat{\theta}_{\mathrm{ML}}(\omega^n) \coloneqq \arg\max_{\theta' \in \Theta_0} \ln p^n_{\theta'}(\omega^n). \tag{4.3.9}$$

We define the corresponding conjugate parameter space $\Xi_0 \coloneqq \mu(\Theta_0)$, which is also a compact convex subset of Ξ. We set $\hat{\mu}_{\mathrm{ML}}(\omega^n) = \mu(\hat{\theta}_{\mathrm{ML}}(\omega^n))$. Because $\nabla_\theta \ln p_\theta(\omega) = \phi(\omega) - \mu(\theta)$, the maximum likelihood estimator $\hat{\theta}_{\mathrm{ML}}$ and $\hat{\mu}_{\mathrm{ML}}$ are given as the solution of the equation

$$\frac{1}{n}\sum_{i=1}^n \phi(\omega_i) = P_{\hat{\theta}_{\mathrm{ML}}(\omega^n)}(\phi) = \hat{\mu}_{\mathrm{ML}}(\omega^n) \tag{4.3.10}$$

if and only if $\overline{\phi}_n \in \Xi_0$, which does not always hold.

We summarize the basic properties of the maximum likelihood estimator. The convexity of $-\ln p^n_\theta(\omega^n)$ with respect to θ implies that

$$(\theta - \hat{\theta}_{\mathrm{ML}}(\omega^n)) \cdot (\hat{\mu}_{\mathrm{ML}}(\omega^n) - \overline{\phi}_n) \geq 0 \tag{4.3.11}$$

for all $\theta \in \Theta_0$. In fact, assume that it does not hold for some $\theta \in \Theta_0$. Consider the continuous path $s(t) \coloneqq t\theta + (1-t)\hat{\theta}_{\mathrm{ML}}(\omega^n)$ for $t \in [0,1]$. The convexity of Θ_0 implies that $s(t) \in \Theta_0$ for all $t \in [0,1]$. From the assumption, we obtain that $\partial_t[-\ln p^n_{s(t)}(\omega^n)]|_{t=0} = n(\hat{\mu}_{\mathrm{ML}}(\omega^n) - \overline{\phi}(\omega^n)) \cdot (\theta - \hat{\theta}_{\mathrm{ML}}(\omega^n)) < 0$, and thus $-\ln p^n_{s(t)}(\omega^n) < -\ln p^n_{\hat{\theta}_{\mathrm{ML}}(\omega^n)}(\omega^n)$ for a sufficiently small $t \in [0,1]$. This contradicts the minimality of $\hat{\theta}_{\mathrm{ML}}(\omega^n)$ on Θ_0. We have from the condition (4.3.11) that if $\hat{\theta}_{\mathrm{ML}}(\omega^n) \in \mathrm{int}\Theta_0$, then $\overline{\phi}_n \in \Xi_0$ and therefore $\hat{\mu}_{\mathrm{ML}}(\omega^n) = \overline{\phi}_n$. Otherwise, $(\hat{\mu}_{\mathrm{ML}}(\omega^n) - \overline{\phi}_n) \cdot (\theta - \hat{\theta}_{\mathrm{ML}}(\omega^n)) < 0$ for some $\theta \in \Theta_0$ in a neighborhood of $\hat{\theta}_{\mathrm{ML}}(\omega^n)$, which contradicts (4.3.11).

The maximum likelihood strategy was originally studied in the context of universal coding and sequential prediction. Kotłowski and Grünwald obtained the following lower bound in [KG11]:

Lemma 4.3.3.2 ([KG11])**.**

$$\sum_{i=1}^n \left[\ln p_{\hat{\theta}_{\mathrm{ML}}(\omega^{i-1})}(\omega_i) - \ln p_{\hat{\theta}_{\mathrm{ML}}(\omega^n)}(\omega_i)\right] \geq -\frac{I_{\Xi_0}(B + C_{\Xi_0})^2}{2} \ln n + O(1), \tag{4.3.12}$$

where $B \coloneqq \max_{\omega\in\Omega} |\phi(\omega)|$, $C_{\Xi_0} \coloneqq \max_{\mu\in\Xi_0} \|\mu\|$ and $I_{\Xi_0} \coloneqq \max_{\mu\in\Xi_0} \|I(\mu)\|$.

We note that these constants are all finite due to the compactness of Ω and Ξ_0. By using Lemma 4.3.3.2, we obtain Theorem 4.3.2.1 immediately.

Proof of Theorem 4.3.2.1. For any strategy $\hat{\theta}$, $\beta W^{\hat{\theta}}_n$ can be decomposed into two parts:

$$\beta W^{\hat{\theta}}(\omega^n) = \sum_{i=1}^n \left[\ln p_{\hat{\theta}(\omega^n)}(\omega_i) - \ln p_\theta(\omega_i)\right] + \sum_{i=1}^n \left[\ln p_{\hat{\theta}(\omega^{i-1})}(\omega_i) - \ln p_{\hat{\theta}(\omega^n)}(\omega_i)\right]. \tag{4.3.13}$$

First, we investigate the first part for $\hat{\theta}_{\rm ML}$. We have from (4.3.1) and (4.3.11) that

$$\sum_{i=1}^{n}\left[\ln p_{\hat{\theta}_{\rm ML}(\omega^n)}(\omega_i)-\ln p_\theta(\omega_i)\right]=nD(P_{\hat{\theta}_{\rm ML}(\omega^n)}\|P_\theta)+n(\hat{\theta}_{\rm ML}(\omega^n)-\theta)\cdot(\overline{\phi}_n-\hat{\mu}_{\rm ML}(\omega^n))$$
$$\geq nD(P_{\hat{\theta}_{\rm ML}(\omega^n)}\|P_\theta). \tag{4.3.14}$$

Combining Lemma 4.3.3.2 with the above inequality (4.3.14), we have that

$$\beta W^{\hat{\theta}_{\rm ML}}(\omega^n)\geq n\left[D(P_{\hat{\theta}_{\rm ML}(\omega^n)}\|P_\theta)-\frac{I_{\Xi_0}(B+C_{\Xi_0})^2}{2}\frac{\ln n}{n}\right]+O(1). \tag{4.3.15}$$

Suppose that $\sup_n W_n^{\hat{\theta}_{\rm ML}}(\xi)<\infty$ for $\xi=\omega_1\omega_2\ldots$. Then, there exists a real number $C_\xi\in\mathbb{R}$ such that $W^{\hat{\theta}_{\rm ML}}(\omega^n)\leq C_\xi$ for all $n\in\mathbb{N}_+$. From the inequality (4.3.15), we have that

$$D(P_{\hat{\theta}_{\rm ML}(\omega^n)}\|P_\theta)\leq\frac{C_\xi}{n}+\frac{I_{\Xi_0}(B+C_{\Xi_0})^2}{2}\frac{\ln n}{n}+O(n^{-1}). \tag{4.3.16}$$

Since C_ξ is independent of n, we obtain that $D(P_{\hat{\theta}_{\rm ML}(\omega^n)}\|P_\theta)\to 0$ as $n\to\infty$. This implies that $p_{\hat{\theta}_{\rm ML}(\omega^n)}(\omega)$ converges to $p_\theta(\omega)$ for any $\omega\in\Omega$, and therefore $\hat{\mu}_{\rm ML}(\omega^n)\to\mu(\theta)$. For $\xi\in\Omega^{\mathbb{N}_+}$ such that $\hat{\mu}_{\rm ML}\to\mu(\theta)\in{\rm int}\Xi_0$, $\overline{\phi}_n\in\Xi_0$ and $\hat{\mu}_{\rm ML}(\omega^n)=\overline{\phi}_n$ for sufficiently large n, and therefore $\overline{\phi}_n\to\mu(\theta)$. □

We have several remarks before ending the subsection.

(1) The lower bound (4.3.15) has much information on the asymptotic behavior. If the second law is weakly valid, i.e., $\sup_n W^{\hat{\theta}_{\rm ML}}<\infty$, then $P_{\hat{\theta}_{\rm ML}}$ converges to P_θ with respect to the Kullback-Leibler distance with the convergence rate $O(\sqrt{\ln n/n})$ and the convergence factor $\sqrt{I_{\Xi_0}(B+C_{\Xi_0})^2/2}$. We also note that if the convergence (4.3.5) does not hold, the accumulation of the extracted work grows at least linearly. Specifically, the extracted work *per cycle* becomes positive infinitely many times,

$$\limsup_{n\to\infty}\frac{W_n^{\hat{\theta}}}{n}>0, \tag{4.3.17}$$

for sequences violating the strong law of large numbers (4.3.5). For instance, if $\hat{\theta}_{\rm ML}(\omega^n)\to\theta'(\neq\theta)$ as $n\to\infty$, the rate of the extracted work is given by the Kullback-Leibler distance between $P_{\theta'}$ and P_θ. This implies that our results (Theorem 4.2.3.1 and 4.3.2.1) are valid even if we apply (4.3.17) as a definition of the violation of the second law.

(2) The maximum likelihood estimator $\hat{\theta}_{\rm ML}(\omega_n)$ defined by (4.3.9) depends on ω_n through only $\overline{\phi}_n$ because the conjugate variable ϕ is the sufficient statistic for the exponential family $\mathcal{P}(\Theta,\phi,h)$. Therefore, even if only the information on the past history of ϕ is given to the agent, he can perform the maximum likelihood strategy and can forces the strong law of large numbers for the conjugate variable.

4.3.4 Numerical Demonstration

We illustrate the validity of the maximum likelihood strategy numerically for the Ising Hamiltonian under the homogeneous magnetic field for two spins, $\beta H_\theta(\sigma_1, \sigma_2) = -\sigma_1\sigma_2 - \theta(\sigma_1+\sigma_2)$. We consider the situation that the agent changes the magnetic field with fixed coupling constant. We set $\beta J = 1$ for simplicity and write $\theta := \beta h_{\rm ex}$. The initial parameter is set to be $\theta = 0$. The average magnetization $\mu(\theta) = \mathbb{E}_\theta[\sigma_1 + \sigma_2]$ is given by $\mu(\theta) = 2(e^{2\theta} + e^{-2\theta})/(e^{2\theta} + e^{-2\theta} + 2e^{-2})$. The maximum likelihood estimator for θ with respect to a compact set $\Theta_0 = [-\ln 2, \ln 2]$ is given by

$$\hat{\theta}_{\rm ML}(\omega^n) = \begin{cases} -\ln 2 & \text{if } \bar{\phi}_n < \mu(-\ln 2) \\ \theta(\bar{\phi}_n) & \text{if } \bar{\phi}_n \in [\mu(-\ln 2), \mu(\ln 2)] \\ \ln 2 & \text{if } \bar{\phi}_n > \mu(\ln 2), \end{cases} \tag{4.3.18}$$

where

$$\theta(\mu) = \frac{1}{2} \ln \left(\frac{\mu e^{-2} + \sqrt{4 + (e^{-4} - 1)\mu^2}}{2 - \mu} \right) \tag{4.3.19}$$

is the inverse of $\theta \mapsto \mu(\theta)$. If a sequence is generated by the Gibbs distribution with $\theta \neq 0$, the empirical mean for the sequence converges to $\mu(\theta)$ but it is different from $\mu(0)$. In Fig 4.2, we plot the accumulations of extracted work for sequences generated from the Gibbs distribution with $\theta = 0, \pm 0.1, \pm 0.2$. The figure shows that while the accumulations for parameters $\theta = \pm 0.1, \pm 0.2$ diverge to infinity and the second law is violated, the accumulation for $\theta = 0$ remains finite. The linear growth of the accumulation of extracted work is consistent with the lower bound (4.3.15).

4.4 Game-theoretic Interpretation

We discuss a game-theoretic interpretation of the protocol of this chapter. The interpretation is based on the analogy with game-theoretic probability theory [SV01, SV19].

We interpret the work extraction as a game played between two players, Skeptic and Reality. The player Skeptic doubts that the statistical property of the heat engine is described by equilibrium statistical mechanics, particularly the Gibbs distribution. He attempts to test the hypothesis of equilibrium statistical mechanics by actually extracting work from the engine many times and observing whether the second law of thermodynamics holds or not. If the accumulation of the extracted work diverges to infinity, i.e., the second law is violated, then Skeptic rejects the hypothesis. On the other hand, the player Reality decides initial microscopic states of the heat engine. The duty of Reality is to prevent Skeptic from extracting the infinite amount of work and to make the second law of thermodynamics valid in our world.

The algorithm of the game corresponding to the protocol we discussed in Section 4.2 is as follows.

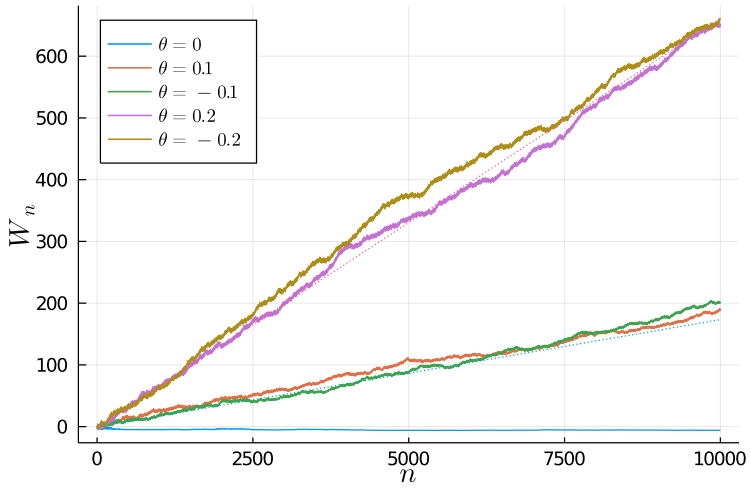

Figure 4.2: The accumulations of the extracted work for sequences generated by the Gibbs distributions with various parameters versus the number of cycles. The top (resp. bottom) dotted line indicates the Kullback-Leibler distance between $P_{\theta=\pm 0.2}$ (resp. $P_{\theta=\pm 0.1}$) and $P_{\theta=0}$.

Protocol 4.4.0.1 (Generic Hamiltonian)**.**

Parameter:

an initial Hamiltonian $H : \Omega \to \mathbb{R}$,

an inverse temperature $\beta > 0$.

Players: Skeptic, Reality.

Protocol:

$W(\square) = 0$.

FOR $n = 1, 2, \ldots$:

Skeptic announces H_n

Reality announces $\omega_n \in \Omega$

$W_n := W_{n-1} + H(\omega_n) - H_n(\omega_n) + F_\beta(H_n) - F_\beta(H)$

END FOR

In Protocol 4.4.0.1, $H_n := \hat{H}(\cdot|\omega^{n-1})$ denotes the quenched Hamiltonian in the n-th cycle and $W_n = W^{\hat{H}}(\omega^n)$ the accumulation of extracted work up to the n-th cycle. We remark that Reality can decides an initial microscopic state ω_n *after* the announcement by Skeptic. Therefore, we allow Reality to move strategically in Protocol 4.4.0.1. Nevertheless, Theorem 4.2.3.1 is still valid even if Reality decides microscopic states strategically because whether a strategy forces a property or not is independent of the strategy Reality applies. Theorem 4.2.3.1 implies that there exists a strategy for Skeptic such that Reality has to converge the empirical distribution to the Gibbs distribution. In other words, Reality is forced to act probabilistically in a manner consistent with equilibrium statistical mechanics due to the second law.

As an application of Theorem 4.2.3.1, we consider a testing procedure of random number generator. Let us suppose that Reality claims that she finds an algorithm to generate random

numbers with respect to the Gibbs distribution. Skeptic doubts her claim and attempts to test it. Although there are several criteria that the random numbers should satisfy, we consider the strong law of large numbers (4.2.6) for the empirical distribution as a criterion here. If the random number obtained from Reality's generator does not satisfy the SLLN, Skeptic rejects her claim and concludes that the random number generator does not work well. According to Theorem 4.3.2.1, Skeptic can confirm the validity of the SLLN by performing cyclic operations for microscopic states prepared by the generator. In other words, the second law of thermodynamics can be used as a test for the random number generators.

Similarly, the algorithm of the protocol in Section 4.3 is as follows.

Protocol 4.4.0.2 (Parametric Hamiltonian)**.**

Parameter:

a control parameter space Θ,
a conjugate variable $\phi : \Omega \to \mathbb{R}^k$,
a remaining Hamiltonian $h : \Omega \to \mathbb{R}$,
an initial parameter $\theta \in \Theta$,
an inverse temperature $\beta > 0$.

Players: Skeptic, Reality.

Protocol:

$W(\square) = 0$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $\theta_n \in \Theta$.
Reality announces $\omega_n \in \Omega$
$W_n \coloneqq W_{n-1} + H_\theta(\omega_n) - H_{\theta_n}(\omega_n) + F_\beta(H_{\theta_n}) - F_\beta(H_\theta)$
END FOR

Theorem 4.3.2.1 can be interpreted game-theoretically in a similar way to the case of Theorem 4.2.3.1.

4.5 Concluding Remarks

In this chapter, we provided a novel formulation of the second law of thermodynamics and showed that there exist strategies for the agent that force statistical properties consistent with equilibrium statistical mechanics, i.e., the strong law of large numbers for the empirical distribution and empirical mean. In the protocol where the agent is able to prepare arbitrary Hamiltonians, the empirical distribution must converge to the Gibbs distribution for the initial Hamiltonian under some strategy. In the protocol where the agent can change a small number of parameters in the Hamiltonian, the maximum likelihood strategy forces the strong law of large numbers for the conjugate variable. Before ending the chapter, we discuss future directions of the study.

Extension to other situations We considered the simple settings where initial microscopic states are prepared independently and identically. It is important to study in the same spirit more complicated settings such as stochastic thermodynamics [Sei12, Sek98], information thermodynamics [SU10], and quantum systems [PW78, Len78]. In particular, it is a future subject to treat the stochastic evolution of microscopic state in our framework and to connect it with thermodynamic quantities. The extension to the quantum settings might be more difficult because the quantum theory has a probabilistic structure different from the classical one. See the chapter 10.6 in [SV19] for game-theoretic formulations for Born's rule and quantum computation.

Assumption on quasi-static isothermal work We assume in our settings that the work extracted in the quasi-static isothermal process is equal to the decrease in free energy defined by $F_\beta(H) = -\beta^{-1} \ln \sum_{\omega \in \Omega} e^{-\beta H(\omega)}$. This assumption allows us to connect the work extracted in the quasi-static isothermal process with the microscopic Hamiltonian, without explicitly referring to the time evolution of the system. However, we should remove it to clarify the emergence of the probabilistic structure from thermodynamics. Although one way to do so is to replace the difference in free energy, $F_\beta(\hat{H}) - F_\beta(H)$, in Protocol 4.4.0.1 by the work in the quasi-static process defined only in terms of mechanics, it is a challenging task because we have to take the dynamical properties of the system into account explicitly.

Macroscopic systems We investigated the asymptotic behavior of the work extracted from a *finite* system as the number of cycles n goes to infinity. While our results are valid for small systems as with stochastic thermodynamics, the second law of thermodynamics is believed to hold almost surely for *macroscopic* systems without repeating the operations. Developing our formulation for macroscopic thermodynamics is a future direction of the study.

Coarse-grained information We assumed in our protocol that the agent can use the information about microscopic states in the past. However, this assumption may be too demanding when the size of the system is large. In a more realistic situation, the agent can measure only the value of extracted work. It is important to study statistical properties forced by strategies under such situation.

Classification of probabilistic laws The approach based on the prequential form allows us to classify the statistical property of the equilibrium state according to the ability of the agent. In the usual formulation of statistical mechanics, the law of large numbers for the empirical distribution and for the empirical mean of the conjugate variable are both formulated as the events with probability one. In this sense, there is no distinction between these probabilistic laws. However, in the prequential approach, what kind of probabilistic laws can be observed depends on our ability. It is interesting to provide a detailed classification of the probabilistic laws of equilibrium states according to the ability of the agent to operate the system.

Algorithmic randomness We mention the probabilistic feature of infinite sequences. Although the empirical distribution for the simple binary sequence $01010101\ldots$ converges to the uniform distribution on $\{0,1\}$, it is not sufficiently random according to our intuition. However, Theorems 4.2.3.1 and 4.3.2.1 refer to only the empirical distribution and mean. To say that an infinite sequence is random with respect to a probability distribution, we have to require stronger conditions on the empirical statistics. The algorithmic theory of randomness provides an idealized notion of randomness on the basis of computability theory and martingale theory [Nie09, DH10]. A class of randomness is defined as the intersection of sets that are weakly forced by strategies with some computability condition. For instance, computable strategies specify the class of randomness called *computable random*. Although there exists no universal computable strategy, i.e., the computable randomness cannot be forced by a single strategy, to investigate such class in the context of thermodynamics may be interesting. We hope that the present study provides an insight into studying thermodynamics from the viewpoints of the martingale structure and computability of strategies. As a study in the same spirit, see Chapter 3 showing the relevance of algorithmic randomness to thermodynamic irreversibility.

Chapter 5

Conclusion

In this thesis, we have studied several problems in statistical physics with a game-theoretic approach. Throughout this thesis, we have seen that the concept of "martingale" is useful to investigate two aspects of stochastic processes. On the one hand, martingale theory was applied to the stopping time problem in stochastic thermodynamics. In Chapter 2, we have studied along this line and derived the kinetic uncertainty relation on first passage time for time-integrated currents via the information inequality at stopping time. On the other hand, martingale processes are used to characterize the randomness in stochastic processes. In Chapter 3 and 4, we have studied along this line. In Chapter 3, by using the theory of algorithmic randomness, we have shown that the random configurations of a macroscopic system exhibit irreversible relaxation behavior. In Chapter 4, we have identified processes of exponentiated extracted work with positive martingale process with respect to the Gibbs distribution and have proved that the second law of thermodynamics implies the Gibbs distribution. Although the former aspect has already been studied in the context of stochastic thermodynamics, the latter aspect in statistical physics was first proposed by the author and should be studied in future works.

Before ending this thesis, we make some speculative remarks.

- In Chapter 2, we have used the information inequality at stopping time, which provides an upper bound on the relative fluctuations of the first passage time in terms of the Fisher information. We recall that the Fisher information has been studied as a geometric object, i.e., Riemannian metric, in the context of information geometry. It is natural to expect that there may be a geometrical understanding of stochastic thermodynamics and thermodynamic inequalities. Recently, the idea along this line has been studied and there is room for further development. Is there a counterpart of this problem in stopping time statistics? For example, consider the following transport problem: for a given initial probability distribution p and final distribution q, find a stopping time τ such that $X_0 \sim p$, $X_\tau \sim q$, and it minimizes some cost function, e.g., the entropy production. Is there a concise characterization of such optimal stopping time? Does the optimal stopping time allow a geometrical understanding? Is it useful to control the thermodynamic system? Is it relevant to the efficiency of biological systems?

- In Chapter 3 and 4, we have discussed the randomness of microscopic states. The same idea can be applied to equilibrium statistical mechanics for large systems. Consider a classical spin system on a lattice with a Hamiltonian H. According to the Birkhoff theorem, for a spatially ergodic Gibbs measure μ, the spatial average of a local quantity f converges to the equilibrium value $\mathbb{E}_\mu(f)$ as the averaged region grows with probability one. It is known that the Birkhoff theorem can be "effectivized". That is to say, for a spatially ergodic *computable* Gibbs measure, the spatial average of any *lower semicomputable* function f converges to the equilibrium value *if and only if* the spin configuration is Martin-Löf random with respect to the Gibbs measure. If we define the "equilibrium state" as the set of all configurations whose spatial average of any "effective" quantity converges to the equilibrium value, this theorem claims that the equilibrium state is completely characterized in terms of Martin-Löf randomness. Moreover, the Kolmogorov complexity density of a Martin-Löf random configuration converges to the thermodynamic entropy density in the thermodynamic limit, despite that the thermodynamic entropy is not usually considered as a function of microscopic states. With this in mind, it would be a challenging task to formulate the second law of thermodynamics for macroscopic systems in terms of (algorithmic) randomness.

Bibliography

[ATM18] Takayuki Ariga, Michio Tomishige, and Daisuke Mizuno. Nonequilibrium energetics of molecular motor kinesin. *Physical Review Letters*, 121:218101, 2018.

[BBDG18] J. P. Boudhaud, J. Bonard, J. Donier, and M. Gould. *Trades, Quotes and Prices*. Cambridge University Press, Cambridge, England, 2018.

[Ben82] Charles H. Bennett. The thermodynamics of computation - a review. *International Journal of Theoretical Physics*, 21:905–940, 1982.

[BN78] O. E. Barndorff-Nielsen. *Information and Exponential Families in Statistical Theory*. Wiley, Chichester, UK, 1978.

[BP12] L. Bienvenu and C. Porter. Strong reductions in effective randomness. *Theoretical Computer Science*, 459:55–68, 2012.

[Bri96] J. Bricmont. Science of chaos or chaos in science? In *The Flight from Science and Reason, Annals of the New York Academy of Sciences*, volume 775, pages 131–175. New York, 1996.

[BS15] Andre C. Barato and Udo Seifert. Thermodynamic uncertainty relation for biomolecular processes. *Physical Review Letters*, 114(114):158101, 2015.

[BS16] Andre C. Barato and Udo Seifert. Cost and precision of brownian clocks. *Physical Review X*, 6(4):041053, 2016.

[BS17] Andre C. Barato and Udo Seifert. Coherence of biochemical oscillations is bounded by driving force and network topology. *Physical Review E*, 95:062409, 2017.

[Cav93] C. M. Caves. Information and entropy. *Physical Review E*, 47:4010, 1993.

[CG11] R. Chetrite and S. Gupta. Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale. *Journal of Statistical Physics*, 143:543, 2011.

[Coo04] S. Cooper. *Computability Theory*. Chapman and Hall/CRC, Boca Raton, 2004.

[CRJ+05] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco Jr., and C. Bustamante. Verification of the crooks fluctuation theorem and recovery of rna folding free energies. *Nature*, 437:231–234, 2005.

[Cro98] G. E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *Journal of Statistical Physics*, 90:1481, 1998.

[CT12] T. M. Cover and J. A. Thomas. *Elements of Information Theory, 2nd Edition*. Wiley, New York, 2012.

[Dan81] H. A. M. Daniëls. Passivity and equilibrium for classical hamiltonian systems. *Letters in Mathematical Physics*, 22:843, 1981.

[Daw84] A. P. Dawid. The prequential approach (with discussion). *Journal of the Royal Statistical Society. Series A*, 147:278–292, 1984.

[DH10] R. G. Downey and D. R. Hirschfeld. *Algorithmic Randomness and Complexity*. Springer-Verlag, New York, 2010.

[Doo53] J. L. Doob. *Stochastic Processes*. Wiley, New York, 1953.

[DS20] Andreas Dechant and Shin-ichi Sasa. Fluctuation-response inequality out of equilibrium. *Proceedings of the National Academy of Sciences of the United States of America*, 117:6430, 2020.

[EMS20] Martin R. Evans, Satya N. Majumdar, and Grégory Schehr. Stochastic resetting and applications. *Journal of Physics A: Mathematical and Theoretical*, 53:193001, 2020.

[EVdB11] M. Esposito and C. Van den Broeck. Second law and landauer principle far from equilibrium. *Europhysics Letters*, 95:40004, 2011.

[Ewe04] W. J. Ewens. *Mathematical Population Genetics*. Springer-Verlag, New York, 2004.

[FE20] Gianmaria Falasco and Massimiliano Esposito. Dissipation-time uncertainty relation. *Physical Review Letters*, 125:120604, 2020.

[Gác] Peter Gács. Lecture notes on descriptional complexity and randomness. arXiv:2105.04704.

[Gác94] Peter Gács. The boltzmann entropy and randomness tests. *Proceedings Workshop on Physics and Computation. PhysComp '94, IEEE*, pages 209–216, 1994.

[Gác01] P. Gács. Quantum algorithmic entropy. *Journal of Physics A: Mathematical and General*, 34:6859–6880, 2001.

[Gác05] Peter Gács. Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science*, 341:91–137, 2005.

[Gar17] Juan P. Garrahan. Simple bounds on fluctuations and uncertainty relations for first-passage times of counting observables. *Physical Review E*, 95:032134, Mar 2017.

[GH17] Todd R. Gingrich and Jordan M. Horowitz. Fundamental bounds on first passage time fluctuations for currents. *Physical Review Letters*, 119:170601, Oct 2017.

[GHR10] S. Galatolo, M. Hoyrup, and C. Rojas. Effective symbolic dynamics, random points, statistical behaviorm complexity and entropy. *Information and Computation*, 208:23–41, 2010.

[GHR11] P. Gács, M. Hoyrup, and C. Rojas. Randomness on computable probability spaces - a dynamical point of view. *Theory of Computing Systems*, 48:465–486, 2011.

[GLTZ06] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N Zanghí. Canonical typicality. *Physical Review Letters*, 96:050403, 2006.

[GMS97] M. Ghosh, N. Mukhopadhyay, and P. K. Sen. *Sequential Estimation*. John Wiley and Sons, New York, 1997.

[GO09] G. A. Gottwald and M. Oliver. Boltzmann's dilemma: An introduction to statistical mechanics via the kac ring model. *SIAM review*, 51:613–635, 2009.

[GP80] J. Gorecki and W. Pusz. Passive states for finite classical systems. *Letters in Mathematical Physics*, 4:433, 1980.

[GRJ20] Alexandre Guillet, Edgar Roldán, and Frank Jülicher. Extreme-value statistics of stochastic transport processes. *New Journal of Physics*, 22:123038, 2020.

[Grü07] P. D. Grünwald. *The Minimum Description Length Principle*. MIT Press, Cambridge, 2007.

[HR09] M. Hoyrup and C. Rojas. Computability of probability measures and martin-löf randomness over metric spaces. *Information and Computation*, 207:830–847, 2009.

[HS05] Takahiro Harada and Shin-ichi Sasa. Equality connecting energy dissipation with a violation of the fluctuation-response relation. *Physical Review Letters*, 95:130602, 2005.

[HS06] Takahiro Harada and Shin-ichi Sasa. Energy dissipation and violation of the fluctuation-response relation in nonequilibrium langevin systems. *Physical Review E*, 73:026131, 2006.

[HTB90] P. Hänggi, P. Talker, and M. Borkoves. Reaction-rate theory: fifty years after kramers. *Reviews of Modern Physics*, 62:251, 1990.

[IBZ16] S. Iyer-Biswas and A. Zilman. First-passage processes in cellular biology. *Advances in Chemical Physics*, 160:261, 2016.

[IKT17] E. Iyoda, K. Kaneko, and Sagawa. T. Fluctuation theorem for many-body pure quantum states. *Physical Review Letters*, 119:100601, 2017.

[Ito16] S. Ito. Backward transfer entropy: Informational measure for detecting hidden markov models and its interpretations in thermodynamics, gambling and causality. *Scientific Reports*, 6:36831, 2016.

[Jar97] C. Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78:2690, 1997.

[Kac59] Mark Kac. *Probability and Related Topics in Physical Science*. Interscience Publishers Inc., New York, 1959.

[Kel56] J. L. Kelly. A new interpretation of information rate. *The Bell System Technical Journal*, 35:917–926, 1956.

[KG11] W. Kotłowski and P. Grünwald. Maximum likelihood vs. sequential normalized maximum likelihood in on-line density estimation. In *Proceedings of the 24th Annual Conference on Learning Theory*. IEEE, 2011.

[KIT19] K. Kaneko, E. Iyoda, and Sagawa. T. Work extraction from a single energy eigenstate. *Physical Review E*, 99:032128, 2019.

[Kol65] A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1:1–7, 1965.

[KT08] M. Kumon and A. Takemura. On a simple strategy weakly forcing the strong law of large numbers in the bounded forecasting game. *Annals of the Institute of Statistical Mathematics*, 60:801, 2008.

[KTT08] M. Kumon, A. Takemura, and K. Takeuchi. Capital process and optimality properties of a bayesian skeptic in coin-tossing games. *Stochastic Analysis and Applications*, 26:1161, 2008.

[LDS+02] Jan Liphardt, Sophie Dumont, Steven B. Smith, Ignacio Tinoco Jr., and Carlos Bustamante. Equilibrium information from nonequilibrium measurements in an experimental test of jarzynski's equality. *Science*, 296:1832–1835, 2002.

[Leb93] Joel L. Lebowitz. Boltzmann's entropy and time's arrow. *Physics Today*, 46:32, 1993.

[Lef13] Raphael Lefevere. Macroscopic diffusion from a hamilton-like dynamics. *Journal of Statistical Physics*, 151:861–869, 2013.

[Lef15] Raphael Lefevere. Fick's law in a random lattice lorentz gas. *Archive for Rational Mechanics and Analysis*, 216:983–1008, 2015.

[Len78] A. Lenard. Thermodynamical proof of the gibbs formula for elementary quantum systems. *Journal of Statistical Physics*, 19:575, 1978.

[Los76] J Loschmidt. Über den zustand des wärmegleichgewichtes eines systems von körpern mit rücksicht auf die schwerkraft. *Sitzungsber. Kais. Akad. Wiss. Wien, Math. Naturwiss. Cl. Abt. II*, 73:128, 1876.

[LPS88] Joel L. Lebowitz, Errico Presutti, and Herbert Spohn. Microscopic models of hydrodynamic behavior. *Journal of Statistical Physics*, 51:841–862, 1988.

[LV08] M. Li and P. M. B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer-Verlag, New York, 2008.

[Mae20] Christian Maes. Frenesy: Time-symmetric dynamical activity in nonequilibria. *Physics Reports*, 850:1, 2020.

[MCH19] Robert Marsland, Wenping Cui, and Jordan M. Horowitz. The thermodynamic uncertainty relation in biochemical oscillations. *Journal of the Royal Society Interface*, 16:20190098, 2019.

[ML66] P. Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[MNS09] C. Maes, K. Netočný, and B. Shergelashvili. A selection of nonequilibrium issues. In *Methods of Contemporary Mathematical Statistical Physics*, volume 1970 of *Lecture Notes in Mathematics*, pages 247–306. Springer, Berlin, Heidelberg, 2009.

[MS20] C. Moslonka and K. Sekimoto. Memory through a hidden martingale process in progressive quenching. *Physical Review E*, 101:062139, 2020.

[MSM+21] Gonzalo Manzano, Diego Subero, Oliver Maillet, Rosario Fazio, Jukka P. Pekola, and Édgar Roldán. Thermodynamics of gambling demons. *Physical Review Letters*, 126:080603, 2021.

[Ner20] Izaak Neri. Second law of thermodynamics at stopping times. *Physical Review Letters*, 124:040601, 2020.

[Nie09] A Nies. *Computability and Randomness*. Oxford University Press, Oxford, 2009.

[NRJ17] Izaak Neri, Édgar Roldán, and Frank Jülicher. Statistics of infima and stopping times of entropy production and applications to active molecular processes. *Physical Review X*, 7:011019, 2017.

[NS19] André Nies and Volkher B. Scholtz. Martin-löf random quantum states. *Journal of Mathematical Physics*, 60:092201, 2019.

[Odi99a] P. Odifreddi. *Classical Recursion Theory, Volume 1.* North-Holland Publishing Company, Amsterdam, 1999.

[Odi99b] P. Odifreddi. *Classical Recursion Theory, Volume 2.* North-Holland Publishing Company, Amsterdam, 1999.

[Pan18] Krzysztof Ptaszyński. First-passage times in renewal and nonrenewal systems. *Physical Review E*, 97:012127, Jan 2018.

[PRR21] Arnab Pal, Shlomi Reuveni, and Saar Rahav. Thermodynamic uncertainty relation for systems with unidirectional transitions. *Physical Review Research*, 3:013273, Mar 2021.

[PSW06] S. Popescu, A. Short, and A. Winter. Entanglement and the foundations of statistical mechanics. *Nature Physics*, 2:754, 2006.

[PW78] W. Pusz and S. L. Woronowicz. Passive states and kms states for general quantum systems. *Communications in Mathematical Physics*, 58:273, 1978.

[RDO08] M. Rigol, V. Dunjko, and M. Olshanii. Thermalization and its mechanics for generic isolated quantum systems. *Nature*, 452:854, 2008.

[Red01] S. Redner. *A Guide to First-Passage Processes.* Cambridge University Press, Cambridge, UK, 2001.

[Rei07] P. Reimann. Typicality for generalized microcanonical ensembles. *Physical Review Letters*, 99:160404, 2007.

[SD16] Keiji Saito and Abhishek Dhar. Waiting for rare entropic fluctuations. *Europhysics Letters*, 114:50004, 2016.

[Sei12] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75:126001, 2012.

[Sek98] Ken Sekimoto. Langevin equation and thermodynamics. *Progress of Theoretical Physics Suppelements*, 130:17–27, 1998.

[Shi21] Naoto Shiraishi. Optimal thermodynamic uncertainty relation in markov jump processes. *Journal of Statistical Physics*, 185:19, 2021.

[SK00] Shin-ichi Sasa and Teruhisa S. Komatsu. Thermodynamic irreversibility from high-dimensional hamiltonian chaos. *Progress of Theoretical Physics*, 103:1–52, 2000.

[SMG+19] Shilpi Singh, Paul Manczel, Dmitry S. Golubev, Ivan M. Khaymovich, Joonas T. Peltonen, Christian Flindt, Keiji Saito, Édgar Roldán, and Jukka P. Pekola. Universal first-passage-time distribution of non-gaussian currents. *Physical Review Letters*, 122:230602, 2019.

[Sol64a] R. J. Solomonoff. A formal theory of inductive inference. part i. *Information and Control*, 7:1–22, 1964.

[Sol64b] R. J. Solomonoff. A formal theory of inductive inference. part ii. *Information and Control*, 7:224–254, 1964.

[SU10] Takahiro Sagawa and Masahito Ueda. Generalized jarzynski equality under nonequilibrium feedback control. *Physical Review Letters*, 104:090602, Mar 2010.

[Sug07] A. Sugita. On the basis of quantum statistical mechanics. *Nonlinear Phenomena in Complex Systems*, 10:192, 2007.

[SV01] G. Shafer and V. Vovk. *Probability and Finance: It's Only a Game!* Wiley, New York, 2001.

[SV19] G. Shafer and V. Vovk. *Game-Theoretic Foundations for Probability and Finance*. Wiley, Hoboken, NJ, 2019.

[Szi64] L. Szilard. Über die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen. *Zeitschrift fr Physik*, 53:840–856, 1964.

[Tas16] Hal Tasaki. Typicality of thermal equilibrium and thermalization in isolated macroscopic quantum systems. *Journal of Statistical Physics*, 163:937–997, 2016.

[TB19] Ivan Di Terlizzi and Marco Baiesi. Kinetic uncertainty relation. *Journal of Physics A: Mathematical and Theoretical*, 52:02LT03, 2019.

[TMMR20] L. Touzo, M. Marsili, N. Merhav, and É. Roldán. Optimal work extraction and the minimum description length principle. *Journal of Statistical Mechanics*, page 093403, 2020.

[Vil39] J. Ville. *Étude Critique de la Notion de Collectif. Monographies des Probabilitités. Calcul des Probabilités et ses Applications.* Gauthier-Villars, Paris, 1939.

[Vit01] P. M. Vitányi. Quantum kolmogorov complexity based on classical descriptions. *IEEE Transactions on Information Theory*, 47:2464–2479, 2001.

[vK07] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, 2007.

[vL87] M van Lambalgen. *Random Sequences*. PhD thesis, University of Amsterdam, Amsterdam, 1987.

[VPM16] D. A. Vinkler, H. H. Permuter, and N. Merhav. Analogy between gambling and measurement-based work extraction. *Journal of Statistical Mechanics*, page 043403, 2016.

[Wil91] D. Williams. *Probability with Martingales*. Cambridge University Press, Cambridge, 1991.

[Zer96] E Zermelo. Über einen satz der dynamik und die mechanische wärmetheorie. *Wied. Ann.*, 57:485–494, 1896.

[ZMR96] D. N. Zubarev, V. Morozov, and G Ropke. *Statistical Mechanics of Nonequilibrium Processes, Volume 1: Basic Concepts, Kinetic Theory*. Wiley, New York, 1996.

[ZMR97] D. N. Zubarev, V. Morozov, and G Ropke. *Statistical Mechanics of Nonequilibrium Processes, Volume 2: Relaxation and Hydrodynamic Processes*. Wiley, New York, 1997.

[Zur89a] W. H. Zurek. Algorithmic randomness and physical entropy. *Physical Review A*, 40:4731, 1989.

[Zur89b] W. H. Zurek. Thermodynamic cost of computation, algorithmic complexity and the information metric. *Nature*, 341:119–124, 1989.

Acknowledgement

I would like to thank my supervisor, Shin-ichi Sasa, for his kind guidance during my doctoral studies. He is not only my supervisor, but also a wonderful role model of a researcher for me. I also thank members of the nonlinear dynamics group, Andreas Dechant, Hiroki Ohta, Michikazu Kobayashi, Masato Itami, Tomohiro Tanogami, Kazuma Nishimura. They have helped me in my research and in my daily life. Finally, I would like to thank my family for supporting me during the graduate course.