INTRODUCTION TO THE THERMODYNAMIC FORMALISM

FINN BRAATEN

ABSTRACT. This paper introduces the thermodynamic formalism for compact metrizable spaces. We introduce Gibbs and equilibrium states for lattice systems and highlight some results for one dimensional lattices. Then, we consider thermodynamics for homeomorphisms of general compact metric spaces. Using this construction, we highlight results for equilibrium states of Smale's axiom A diffeomorphims through their symbolic representation as lattice systems.

CONTENTS

1. INTRODUCTION TO THERMODYNAMIC ENSEMBLES

Consider a finite set of configurations $\Omega = \{\xi_1, \dots, \xi_k\}$ with a probability measure ρ where $\rho(\xi_i) = \rho_i$. This quantifies the probability of being in a particular configuration. For some combination of N elements of Ω, we denote the number of occurrences of each ξ_i by N_i. For some collection of integers $n_1, \dots, n_k$ such that $\sum_{i=1}^k n_i = N$, the probability that $n_i = N_i$ is

$$\frac{N!}{n_1! \cdots n_k!} \rho_1^{n_1} \cdots \rho_k^{n_k}.$$

By Stirling's formula, the logarithm of this value is approximately

$$\sum_{i=1}^k n_i \log\left(\frac{N\rho_i}{n_i}\right) = N \sum_{i=1}^k q_i(\log \rho_i - \log q_i)$$

where $q_i = \frac{n_i}{N}$ is the the *state density* of the system. This probability is maximized for $q_i = \rho_i$ when the second value is zero. So, for large N, the probability that the state density of the system matches the probability measure is approximately one. This relationship identifies the probability measures on a configuration space with the state densities of an arbitrarily large system.

Date: DEADLINES: Draft AUGUST 14 and Final version AUGUST 28, 2021.

The states which match the measure are said to be in "statistical equilibrium," though this is not the same equilibrium as we consider throughout the paper. Further, the logarithm of the number of states that are in statistical equilibrium is $N\sum_{i=1}^{k} -\rho_i \log \rho_i$. A postulate central to statistical mechanics is that these states are equally likely in some sense. So, the system will align with the state density, and thus the probability measure, which corresponds to the highest number of statistical equilibrium states.

Suppose that for each configuration ξ_i, there is an associated energy u_i. Further, suppose the total energy of the combination of configurations is constant U. Then, the average energy may be written as $\sum_i \rho_i u_i = \frac{U}{N} = \bar{U}$. The probability measure ρ that maximizes the number of statistical equilibrium states or the *mean entropy* $s = \sum_{i=1}^{k} -\rho_i \log \rho_i$ subject to this constraint is the *Gibbs ensemble* or *equilibrium state*

$$\rho_i = \frac{e^{-\beta u_i}}{\sum_i e^{-\beta u_i}} \text{ such that } \frac{\sum_i u_i e^{-\beta u_i}}{\sum_i e^{-\beta u_i}} = \bar{U}.$$

This probability measure is that which the collection of configurations will tend to exhibit at equilibrium.

The *partition function* is defined to be $Z(\beta) = \sum_i e^{-\beta u_i}$. From this, the second equation may be expressed as $-\frac{\partial}{\partial \beta} \log Z(\beta) = \bar{U}$. Because this equation is strictly decreasing, the equation has a unique solution for β, which we call the temperature. This quantity is only dependent on the average energy $\bar{U}$ of a combination of configurations, so it serves to quantify the energy of the combination of configurations, independent of their number.

If we assume that the system is at some constant temperature, we may observe that the Gibbs state minimizes the Gibbs free energy

$$\beta\bar{U} - s = \beta \sum_i \rho_i u_i + \sum_i \rho_i \log \rho_i,$$

which is equivalent to $-P(\beta)$, where $P(\beta) = \log Z(\beta)$ is the *pressure*, at the value's minimum. The above discussion is based on the explanation of Boltzmann's entropy in Zinmeister's book [8].

The reason why Gibbs ensembles describe thermal equilibrium for infinite possible configurations is not completely clear; however, the formalism seeks to describe these equilibrium states to reveal more about their properties. The following discussion takes the energy to be multiplied by $-\beta$ in its definition and generalizes the notion of the the Gibbs ensemble at equilibrium for infinite systems. The first part will consider lattice configurations and generalize Gibbs ensembles and pressure. It will also identify Gibbs ensembles with probability measures that maximize the Gibbs free energy $s - \bar{U}$. The discussion continues into a representation of one dimensional lattice systems called subshifts of finite type. Finally, it will generalize the notion of equilibrium states to compact sets, and use subshifts of finite type to deduce properties of these equilibrium states for a particular class of diffeomorphisms.

2. Introduction to Lattice Configuration Space

The following constructions of configurations for lattice systems are due to Ruelle [6]. Let L be a countable set of lattice points. We define a finite set of micro-states Ω_x for all $x \in L$. The product space $\prod_{x\in L} \Omega_x$ then defines the set of all configurations on the lattice.

However, not all of these configurations may be admissible. To define which configurations are possible, a family of allowed configurations is defined on this space via $\mathcal{F}$, a collection of finite subsets of L. This collection is locally finite, so each element of L is contained in finitely many of these sets $\Lambda \in \mathcal{F}$. For each such Λ, there is some $\bar{\Omega}_\Lambda \subset \prod_{x\in\Lambda} \Omega_x$ that restricts the product space to the configuration space

$$\Omega = \left\{ \xi \in \prod_{x\in L} \Omega_x : \xi|_\Lambda \in \bar{\Omega}_\Lambda \text{ (for all } \Lambda \in \mathcal{F}) \right\}.$$

For any subset $M \subset L$, define the M-configuration space

$$\Omega_M = \left\{ \xi \in \prod_{x\in M} \Omega_x : \xi|_\Lambda \in \bar{\Omega}_\Lambda \text{ (for all } \Lambda \in \mathcal{F} : \Lambda \subset M) \right\}.$$

In both settings, the discrete topology on Ω_x extends to the product topology on the configuration spaces. These are compact metric spaces as the at-most-countable product of finite discrete spaces.

Definition 2.1. A *lattice system* is a triple $(L, (\Omega_x)_{x\in L}, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$ described above, as it gives rise to a configuration space for all subsets M of L, which are equipped with the product topology.

Between configuration spaces, there are continuous restriciton maps $\alpha_M : \Omega \to \Omega_M$ such that $\alpha_M\xi = \xi|_M$, and $\alpha_{TM} : \Omega_M \to \Omega_T$ defined similarly for $T \subset M$. It is important to note that this is not a surjection, since configurations of Ω_M may not satisfy the conditions of $\bar{\Omega}_\Lambda$ for $\Lambda \in \mathcal{F}$ if $\Lambda \not\subset M$.

The following results for compact metric spaces may be found in [1]. In particular we consider the configuration space Ω. Let $\mathcal{C}$ be the real continuous functions on Ω. These form a Banach space with respect to the supremum norm. The set of probability measures $\mathcal{M}(\Omega)$ may be identified with continuous linear functionals $\mathcal{C}^*$ via the Riesz-Representation theorem. The weak-$*$ topology on $\mathcal{C}^*$ is associated to the probability measures, so the probability measures form a compact convex metrizable space . In this construction, the functions in $\mathcal{C}$ form observable quantities with the measures $\mathcal{C}^*$ corresponding to statistical states of the system.

Combining considerations of real continuous functions with our discussion of restriction maps, there is an algebra of finite range functions C_Λ for $\Lambda \subset L$ on configuration space defined by $A \circ \alpha_\Lambda$ for $A \in \mathcal{C}(\Omega_\Lambda)$. Corresponding to these restriction maps on the space of functions, there are restrictions on the dual $\beta_{\Lambda M} : C_M^* \to C_\Lambda^*$ such that $\beta_{\Lambda M}\mu = \mu(A \circ \alpha_{\Lambda M})$ for all $A \in \mathcal{C}(\Omega_M)$. This maps the probability measures on Ω_M to those on Ω_Λ, for $M \supset \Lambda$. The set $\bigcup_{\Lambda \text{ finite}} \mathcal{C}_\Lambda$ is dense in $\mathcal{C}$ because for any $A \in \mathcal{C}$, the sequence $(A|_\Lambda \circ \alpha_\Lambda)$ as $\Lambda \to L$ converges to A. The restriction map $\beta_\Lambda : C^* \to C_\Lambda^*$ is defined similarly.

Let $L = \mathbb{Z}^\nu$ for some $\nu \in \mathbb{N}$, and define translation maps $\tau^a : \prod_{x\in M} \Omega_x \mapsto \prod_{x\in M-a} \Omega_x$ such that $(\tau^a\xi)_x = \xi_{x-a}$.

Definition 2.2. The lattice system $(L, (\Omega_x)_{x\in L}, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$ is *translation invariant* if $\Omega_x = \Omega_0$ for all $x \in L$, and

(1) $\Omega_{M-a} = \tau^a \Omega_M$.
(2) $\Omega_T \supset \Omega_M|_T$ if $T \subset M$.
(3) $\Omega \neq \emptyset$, and $\xi \in \Omega \iff \xi|_\Lambda \subset \Omega_\Lambda$ for all Λ.

Though it is not necessary to discuss Gibbs states, we will assume translation invariance in general.

3. Energy and Gibbs States

On these lattice systems, there are physical interactions or energies which determine the behavior of the system. To describe these, we use the following definition.

Definition 3.1. An interaction is a function

$$\Phi : \bigcup_{\Lambda \text{ finite } \subset L} \Omega_\Lambda \to \mathbb{R}$$

such that $\Phi|_{\Omega_\emptyset} = 0$ and for each $x \in L$

$$|\Phi|_x = \sum_{\Lambda \ni x} \frac{1}{|\Lambda|} \sup_{\xi\in\Omega_\Lambda} |\Phi(\xi)| < \infty$$

where $|\Lambda|$ is the cardinality of Λ. Each interaction is the energy associated to some collection of lattice states.

For example, consider a lattice system L that describes the presence of charged particles $\Omega_x = \{e^+, e^-, 0\}$ at each point. A pair of points with their micro-states in space $\Omega_{\{a,b\}}$ for $a, b \in L$ has an associated energy due to the Coulomb force. This energy is described by a real number associated to the configurations for each pair of points on the lattice. Hence, this may be described by an interaction function defined as such for all two-point configuration spaces, with $\Phi_\Lambda(\xi) = 0$ for all other finite sets $|\Lambda| \neq 2$.

Definition 3.2. We say an interaction is *translation invariant* if $\Phi(\tau^a\xi) = \Phi(\xi)$ for all $a \in \mathbb{Z}^\nu$. For a translation invariant $\mathbb{Z}^\nu$ lattice system, these interactions are denoted $\mathscr{A}$. We may also consider the dense subset of *finite range interactions* $\mathscr{A}_0$ the set interactions such that there is some Δ where $\Phi(\xi|_X) = 0$ if $X - a \not\subset \Delta$ for all $a \in X$.

The set of interactions may be identified with a continuous function $A_\Phi : \Omega \to \mathbb{R}$ defined as

$$A_\Phi(\xi) = -\sum_{X\ni 0} \frac{\Phi(\xi|_X)}{|X|}, \tag{3.3}$$

which is bounded by

$$\|A_\Phi\| \leq \sum_{X\ni 0} \frac{\sup_{\xi\in\Omega_X} |\Phi(\xi)|}{|X|} = |\Phi| \tag{3.4}$$

due to our definition of interactions. This value may be interpreted as the energy of the configuration due to a particular point on the lattice.

In the following discussion, the interactions will be further bounded by

$$\|\Phi\|_x = \sum_{\Lambda \ni x} \sup_{\xi \in \Omega_\Lambda} |\Phi(\xi)| < \infty.$$

We denote the banach space of translation invariant interactions with respect to this norm $\mathscr{B}$. For such interactions, the norm is independent of x. As the restriction of finite range interactions forces this norm to be finite, it follows that $\mathscr{A}_0 \subset \mathscr{B} \subset \mathscr{A}$. This is a more physically reasonable space of interactions, as discussed in section 3.18 of Ruelle's book [6], and allows for the association of Gibbs and equilibrium states defined later.

Proposition 3.5. *The function $\varphi : \mathscr{A} \to \mathcal{C}$, defined such that $\Phi \mapsto A_\Phi$, is a surjective continuous linear map. The sets $\varphi\mathscr{A}_0$ and $\varphi\mathscr{B}$ are both dense in $\mathcal{C}$.*

Proof. The function is linear by the definition (3.3). Further, the boundedness condition (3.4) implies continuity. The funciton is surjective as each $A \in \mathcal{C}$ has some interaction $\Phi(\xi|_{\{x\}}) = -A$ for all $x \in \mathbb{Z}^\nu$ and 0 for all other configurations. This provides a $\Phi \in \mathscr{A}$ such that $\varphi\Phi = A$. Finally, the continuity of the map as well as the density of $\mathscr{A}_0$, and thus $\mathscr{B}$, in $\mathscr{A}$ implies their images are dense in $\varphi\mathscr{A} = \mathcal{C}$. □

This allows results for $\mathcal{C}$ to be derived from properties of $\mathscr{A}$. In particular, the restriction to the dense space of finite range interactions allows for results to be derived which, when preserved under a limit, hold for general $A \in \mathcal{C}$.

Definition 3.6. The *total energy* due to interactions on finite subsets of the configuration space is $U_\Lambda^\Phi : \Omega_\Lambda \to \mathbb{R}$ such that

$$U_\Lambda^\Phi(\xi) = \sum_{X \subset \Lambda} \Phi(\xi|_X)$$

Definition 3.7. Let Λ, M be disjoint subsets of L where Λ is finite. The *interaction between subsets* is the energy from interactions that include elements from both subsets.

$$W_{\Lambda M}(\xi) = \sum_{X \subset \Lambda \cup M \ :\ X \cap \Lambda \neq \emptyset} \Phi(\xi|_X).$$

We define the *boundary energy* of Λ to be $W_\Lambda = W_{\Lambda M}$ where $M = L/\Lambda$.

Since

$$\begin{aligned} |W_{\Lambda M}(\xi)| &\leq \sum_{X : X \cap \Lambda \neq \emptyset} \Phi(\xi|_X) \\ &= \sum_{x \in \Lambda} \sum_{X \ni x} \Phi(\xi|_X) \\ &\leq \sum_{x \in \Lambda} \|\Phi\|_x \end{aligned}$$

is bounded as Λ is finite, this quantity is well-defined. Further, for Λ and M finite, this is equivalent to $W_{\Lambda M} = U_{\Lambda \cup M} - U_\Lambda - U_M$.

Definition 3.8. The *Gibbs ensemble* for $\Lambda \subset L$ finite and an interaction Φ is the probability measure

$$\mu_{(\Lambda)}\{\xi\} = \frac{1}{Z_\Lambda^\Phi} \exp\left[-U_\Lambda^\Phi(\xi)\right]$$

where $Z_\Lambda^\Phi = \sum_{\xi\in\Omega_\Lambda} \exp[-U_\Lambda^\Phi]$ is the *partition function* for the interaction.

The thermodynamic limit describes the limiting behavior of these probability distributions as $M_n \to L$. The following proposition makes this notion more precise, and the proceeding theorem describes this limit for Gibbs ensembles.

Proposition 3.9. *Let (M_n) be a sequence of finite subsets of L such that $M_n \to L$. For each n, let $\mu_{(M_n)}$ be a probability measure on Ω_{M_n}. There exists a subsequence (M_n') such that the limit*

$$\lim_{n\to\infty} \alpha_{\Lambda M_n'}\mu_{(M_n')} = \rho_\Lambda \tag{3.10}$$

exists for all finite $\Lambda \subset L$. Further, there exists a unique probability measure ρ on Ω such that $\rho_\Lambda = \alpha_\Lambda\rho$ for all Λ. This is called the thermodynamic limit *of $\mu_{(M_n)}$.*

Proof. As Λ is finite, its configuration space $\Omega_\Lambda = \{\xi_1, \dots, \xi_k\}$ is finite. As measures over discrete spaces are defined by their values over individual points, the limit may be constructed via a diagonalization argument. Since $\alpha_{\Lambda M_n}\mu_{M_n}\{\xi_i\}$ is restricted to the compact set $[0,1]$ for all i, a diagonalization of successive subsequences $(M_{n,i})_{i=1}^k$ such that

$$\begin{aligned}
&\alpha_{\Lambda M_{n,1}}\mu_{M_{n,1}}\{\xi_1\} \to \rho_{\Lambda,1}\\
&\vdots\\
&\alpha_{\Lambda M_{n,k}}\mu_{M_{n,k}}\{\xi_k\} \to \rho_{\Lambda,k}
\end{aligned}$$

yields a sequence of measures $\alpha_{\Lambda M_n'}\mu_{M_n'}$ such that their values on each ξ_i converge to $\rho_{\Lambda,i}$. This defines convergence to a measure where $\rho_\Lambda\{\xi_i\} = \rho_{\Lambda,i}$. By construction, the probability measures have $\sum_{j=1}^k \alpha_{\Lambda M_{n,i}}\mu_{M_{n,i}}\{\xi_j\} = 1$ for all i, and

$$\lim_{n\to\infty}\sum_{i=1}^k \alpha_{\Lambda M_n'}\mu_{M_n'}\{\xi_i\} = \sum_{i=1}^k \lim_{n\to\infty} \alpha_{\Lambda M_n'}\mu_{M_n'}\{\xi_i\}$$

This yields

$$1 = \rho_\Lambda\{\xi_1, \dots, \xi_k\}.$$

So, the established ρ_Λ solving (3.10) is well defined as a probability measure on Ω_Λ.

By (3.10), the restrictions of these measures are consistent as $\alpha_{\Lambda M}\rho_M = \rho_\Lambda$. This allows for $\rho(A \circ \alpha_\Lambda) = \rho_\Lambda(A)$, where $A \in \mathcal{C}(\Omega_\Lambda)$, to be well defined. As there are only finitely many $F \in \mathcal{F}$ which contain points in Λ, there is some finite $M \supset L$ composed of L and any such F. This specifies a finite M where $\{\xi|_\Lambda : \xi \in \Omega_M\} = \{\xi|_\Lambda : \xi \in \Omega\}$. A quick calculation gives

$$|\rho(A\circ\alpha_\Lambda)| = |\rho_\Lambda(A)| = |\rho(A\circ\alpha_{\Lambda M})| \le \|A\circ\alpha_{\Lambda M}\| = \|A\circ\alpha_\Lambda\|.$$

This allows us to establish continuity of ρ as a linear functional, and defines a unique continuous extension from a dense subset $\bigcup_{\Lambda \text{ finite}} \mathcal{C}_\Lambda^*$ onto $\mathcal{C}^*$. Hence, there is a unique probability measure ρ associated to this continuous linear functional such that $\rho_\Lambda = \alpha_\Lambda\rho$ for all finite $\Lambda \in L$. $\square$

Definition 3.11. The probability measure $\sigma \in \mathcal{M}(\Omega)$ is a *Gibbs state* if for all $\Lambda \subset L$ finite, there exists a probability measure $\sigma_{L/\Lambda}$ on $\Omega_{L/\Lambda}$ such that for all $\xi_\Lambda \in \Omega_\Lambda$

$$(\alpha_\Lambda \sigma)\{\xi_\Lambda\} = \int_{\Omega_{L/\Lambda}} \sigma_{L/\Lambda}(d\eta)\mu_{(\Lambda)_\eta}\{\xi_\Lambda\},$$

where

$$\mu_{(\Lambda)_\eta}\xi_\Lambda = \frac{\exp[-U_\Lambda(\xi_\Lambda) - W_{\Lambda,L/\Lambda}(\xi_\Lambda \vee \eta)]}{\sum_{\eta_\Lambda \in \Omega_\Lambda} \exp[-U_\Lambda(\eta_\Lambda) - W_{\Lambda,L/\Lambda}(\eta_\Lambda \vee \eta)]}.$$

The configuration $\xi_\Lambda \vee \eta$ is the element $\zeta \in \Omega$ where $\zeta|_\Lambda = \xi_\Lambda$ and $\zeta|_{L-\Lambda} = \eta$. If such a ζ does not exist, then we let $\mu_{(\Lambda)_\eta}\xi_\Lambda = 0$. This represents the probability of a particular finite configuration as an average of its "Gibbs probabilities" due to the energy it contributes to an arbitrary external configuration.

A result from Ruelle [6] specifies that Gibbs states are defined by their systems of conditional probabilities $\mu_{(\Lambda)_\eta}\xi_\Lambda$ that $\xi|_\Lambda = \xi_\Lambda$ given $\xi|_{L/\Lambda} = \eta$, but proof is omitted here for the sake of brevity.

Theorem 3.12. *The thermodynamic limit of Gibbs ensembles $\mu_{(\Lambda)}$ for Φ is a Gibbs state of Φ. Here, we follow the proof from Ruelle [6].*

Proof. For $\Lambda \subset M$, let $\Omega^M_{\xi_\Lambda} = \{\xi \in \Omega_M : \xi|_\Lambda = \xi_\Lambda\}$. The restriction of a Gibbs ensemble is

$$\begin{aligned}
(\alpha_{\Lambda M}\mu_{(M)})\{\xi_\Lambda\} &= \mu_{(M)}(\chi_{\xi_\Lambda} \circ \alpha_{\Lambda M}) \\
&= \mu_{(M)}\left(\sum_{\xi \in \Omega^M_{\xi_\Lambda}} \chi_\xi\right) \\
&= \sum_{\eta \in \Omega_{M/\Lambda}} \frac{\exp(-U_M(\xi_\Lambda \vee \eta)}{Z_M} \\
&= \sum_{\eta \in \Omega_{M/\Lambda}} \frac{\exp(-U_{M/\Lambda}(\eta)}{Z_M} \exp[-U_\Lambda(\xi_\Lambda) - W_{\Lambda,M/\Lambda}(\xi_\Lambda \vee \eta)] \\
&= \sum_{\eta \in \Omega_{M/\Lambda}} [(\alpha_{M/\Lambda,M}\mu_{(M)})\{\eta\}]\mu_{(\Lambda,M)_\eta}\{\xi_\Lambda\}
\end{aligned}$$

where

$$\mu_{(\Lambda,M)_\nu}\{\xi_\Lambda\} = \frac{\exp[-U_\Lambda(\xi_\Lambda) - W_{\Lambda,M/\Lambda}(\xi_\Lambda \vee \eta)]}{\sum_{\eta_\Lambda \in \Omega_\Lambda} \exp[-U_\Lambda(\xi_\Lambda) - W_{\Lambda,M/\Lambda}(\xi_\Lambda \vee \eta)]}$$

and χ is the indicator function. Note that there is a slight abuse of notation where μ denotes both the probability measure and its associated functional.

By the uniform continuity of exponential functions with domain bounded by

$$|-U_\Lambda(\xi_\Lambda) - W_{\Lambda M}(\xi)| \leq |U_\Lambda(\xi_\Lambda)| + \sum_{x \in \Lambda} \|\Phi\|_x$$

for all $M \subseteq L/\Lambda$ and as the functions

$$\left[\eta \mapsto W_{\Lambda, M_n'/\Lambda}(\xi_\Lambda \vee \eta|_{M_n'/\Lambda})\right] \to \left[\eta \mapsto W_{\Lambda, L/\Lambda}(\xi_\Lambda \vee \eta)\right]$$

converge uniformly by the convergence of $\sum_{X \subset \Lambda} \Phi(\xi|_X)$, the functions $\eta \mapsto \mu_{(\Lambda, M_n')_{(\eta|_{M_n'/\Lambda})}}\{\xi_\Lambda\}$ under the thermodynamic limit $M_n' \to L$ converge uniformly to $\eta \mapsto \mu_{(\Lambda)_\eta}\{\xi_\Lambda\}$. This is a continuous function because $\{\eta \in \Omega_{L/\Lambda} : \xi_\Lambda \vee \eta \in \Omega\}$ is open by the local finiteness of $\mathcal{F}$ and because it is the composition of exponential functions and a uniformly convergent sum of continuous interaction functions.

The thermodynamic limit in 3.9 gives a probability measure $\rho_{L/\Lambda}$ on $\Omega_{L/\Lambda}$ for the sequence of probability measures $\alpha_{M_n'/\Lambda, M_n'}\mu_{(M_n')}$ on $\Omega_{M_n'/\Lambda}$. Thus, the limit of the restricted Gibbs ensembles becomes

$$\lim_{n\to\infty}(\alpha_{\Lambda M_n'}\mu_{(M_n')})\{\xi_\Lambda\} = \int_{\Omega_{L/\Lambda}} \rho_{L/\Lambda}(d\eta)\mu_{(\Lambda)_\eta}\{\xi_\Lambda\},$$

which is a Gibbs state as defined. □

These Gibbs states provide a characterization for the analogue of Gibbs ensembles in the thermodynamic limit. The physical importance of these states on $\mathscr{B}$ will become clearer in the following section.

4. Equilibrium States and the Variational Principle

In this setting, we return to $L = \mathbb{Z}^\nu$ and $\Phi \in \mathscr{A}$. Using the energies or energy density functions on the lattice, notions of pressure defined over finite subsets of the lattice hold under a limit to an infinite part of the lattice.

Definition 4.1. For $S \subset \mathbb{Z}^\nu$, there are configurations $\Omega_S^* = \{\xi : \xi^*|_S = \xi$ for some $\xi \in \Omega\}$. Associated to $A \in \mathcal{C}$ is a *partition function*

$$Z_\Lambda(A) = \sum_{\xi \in \Omega_\Lambda^*} \exp\left[\sum_{x\in\Lambda} A(\tau^x \xi^*)\right]$$

where ξ^* is arbitrarily chosen such that $\xi^*|_S = \xi$. For $A = A_\Phi$, the following results hold analogously for the partition function Z_Λ^Φ as defined above. The results for this partition function may be found in the book by Ruelle [6]. The value $P_\Lambda(A) = \frac{\log Z_\Lambda(A)}{|\Lambda|}$ is called the finite pressure.

There are a few notable properties of the finite pressure which we will note. First, the "directional derivative" of the pressure is

$$\frac{d}{dt}(P_\Lambda(A+Bt)) = \frac{1}{|\Lambda|\ Z_\Lambda(A+Bt)} \sum_{\xi\in\Omega_\Lambda^*}\left[\exp\left(\sum_{x\in\Lambda}(A+Bt)(\tau^x\xi^*)\right)\sum_{x\in\Lambda}B(\tau^x\xi^*)\right].$$

Then, we will denote $\mathbf{A}_\xi = \sum_{x\in\Lambda} A(\tau^x\xi^*)$ and $\mathbf{B}_\xi = \sum_{x\in\Lambda} B(\tau^x\xi^*)$. It follows that

$$
\begin{aligned}
|\Lambda| \frac{d^2}{dt^2}(P_\Lambda(A+Bt))\Big|_{t=0} &= \frac{1}{Z_\Lambda(A)^2} \sum_{\xi\in\Omega^*_\Lambda} \sum_{\eta\in\Omega^*_\Lambda} \left[\exp(\mathbf{A}_\xi + \mathbf{A}_\eta)\ (\mathbf{B}_\xi^2 - \mathbf{B}_\xi \mathbf{B}_\eta)\right] \\
&= \frac{1}{Z_\Lambda(A)^2} \sum_{\xi\in\Omega^*_\Lambda} \sum_{\eta\in\Omega^*_\Lambda} \frac{1}{2} \left[\exp(\mathbf{A}_\xi + \mathbf{A}_\eta)\ (\mathbf{B}_\xi - \mathbf{B}_\eta)^2\right] \\
&\geq 0
\end{aligned}
$$

This implies the pressure map $A \mapsto P_\Lambda(A)$ is convex. Further,

$$
\begin{aligned}
\left|\frac{d}{dt}(P_\Lambda(A+Bt))\right| &\leq \frac{1}{|Z_\Lambda|} \sum_{\xi\in\Omega^*_\Lambda} \left[\exp((\mathbf{A}{+}\mathbf{Bt})_\xi \frac{\sum_{x\in\Lambda} |B(\tau^x \xi^*)|}{|\Lambda|}\right] \\
&\leq \|B\|
\end{aligned}
$$

so

$$
|(P_\Lambda(A) - P_\Lambda(B))| \leq \sup_{0\leq t\leq 1} \left|\frac{d}{dt}(P_\Lambda(A+(B-A)t))\right| \leq \|A-B\|. \tag{4.2}
$$

The following theorem produces the thermodynamic limit of the pressure.

Theorem 4.3. *Given $a_1, \ldots, a_\nu > 0$ let $\Lambda(a) = \{x \in \mathbb{Z} : 0 \leq x_i < a_i\}$ and write $a \to \infty$ for $a_1, \ldots, a_\nu \to \infty$. If $A \in \mathcal{C}$, then $P(A) = \lim_{a\to\infty} P_{\Lambda(a)}(A)$ exists. This function is convex and continuous on $\mathcal{C}$ and has, for all $A, B \in \mathcal{C}$,*

$$
|P(A) - P(B)| \leq \|A - B\|.
$$

Proof. Let $A \in \mathcal{C}_\Lambda$ for some Λ finite. If $\Lambda \subset M$, then

$$
Z_M(A) \leq Z_\Lambda(A) \left(|\Omega_0| e^{\|A\|}\right)^{|M|-|\Lambda|}. \tag{4.4}
$$

If $\Lambda_1 \cap \Lambda_2 = \emptyset$, then as the states determining A are restricted to some finite Λ

$$
Z_{\Lambda_1\cup\Lambda_2} \leq Z_{\Lambda_1}(A) Z_{\Lambda_2}(A) e^{|\Omega_0|^{|\Lambda|}\|A\|}. \tag{4.5}
$$

Define $P(A) = \liminf_{a\to\infty} P_{\Lambda(a)}$. For $\epsilon > 0$ choose b large enough such that

$$
|\Omega_0|^{|\Lambda|} \frac{\|A\|}{|\Lambda(b)|} < \frac{\epsilon}{2} \text{ and } P_{\Lambda(b)}(A) < P(A) + \frac{\epsilon}{2}. \tag{4.6}
$$

This holds by the limit inferior, as we can choose such an arbitrarily large $\Lambda(b)$ that satisfies the second property. If $c_1, \ldots, c_\nu$ are positive integer n multiples of $b_1, \ldots, b_\nu$, then translations of $\Lambda(b)$ and (4.5) gives

$$
P_{\Lambda(c)}(A) \leq P_{\Lambda(b)}(A) |\Omega_0|^{|\Lambda|} \frac{\|A\|}{|\Lambda(b)|}
$$

By (4.6), this yields

$$
P_{\Lambda(c)}(A) < P(A) + \epsilon.
$$

Hence, for any $\Lambda(a)$ large enough, we can choose such an integer multiple $\Lambda(c) \subset \Lambda(a)$ with $\left(|\Omega_0|e^{\|A\|}\right)^{|\Lambda(a)/\Lambda(c)|}$ bounded so that $P_{\Lambda(a)}(A) < P_{\Lambda(c)}(A) + \epsilon$ by (4.6). As this was chosen for arbitrary $\epsilon > 0$, this implies

$$\limsup_{a\to\infty} P_{\Lambda(a)}(A) = P(A). \tag{4.7}$$

This holds for the dense subspace of $\bigcup_{\Lambda \text{ finite}} \mathcal{C}_\Lambda$ as shown. So, by the equicontinuity property (4.2), this holds for all $A \in \mathcal{C}$.

The convexity of P and the inequality follow from the convexity of P_Λ and (4.2) respectively. ☐

Proposition 4.8. *For each interaction Φ, there are alternative partition functions*

$$Z_\Lambda^\Phi = \sum_{\xi\in\Omega_\Lambda} \exp[-U_\Lambda^\Phi(\xi)] \text{ or } Z_\Lambda^{*\Phi} = \sum_{\xi\in\Omega_\Lambda^*} \exp[-U_\Lambda^\Phi(\xi)]$$

such that their pressures are equivalent and correspond to $P(A_\Phi)$, that is

$$P^\Phi = \lim_{a\to\infty} \frac{\log Z^\Phi_{\Lambda(a)}}{|\Lambda(a)|} = \lim_{a\to\infty} \frac{\log Z^{*\Phi}_{\Lambda(a)}}{|\Lambda(a)|} = P(A_\Phi).$$

A full discussion of various definitions of pressure is omitted here for brevity, but may be found in Ruelle [6]. We may note that, while the pressure related to A corresponds more analogously to the thermodynamic formalism on generalized measure spaces discussed in a later section, the pressure related to Φ allows us to simplify results for pressure on the dense subset of finite range interactions $\mathscr{A}_0$.

Definition 4.9. For each $a \in \mathbb{Z}^\nu$ a linear map τ^a in $\mathcal{C}^*$ is defined by

$$(\tau^a\sigma)(A) = \sigma(A \circ \tau^a)$$

for all $\sigma \in \mathcal{C}^*$ and $A \in \mathcal{C}$. This is continuous in the vague topology. Translations also map probability measures E to themselves as the identity is translation invariant. We denote the set of translation invariant probability measures

$$I = \{\sigma \in E : \tau^a\sigma = \sigma \text{ for all } a \in \mathbb{Z}^\nu\}$$

This set I is a convex, compact, and a Choquet simplex whose extremal measures are called ergodic [6].

Definition 4.10. Functionals tangent to P at A, which we denote

$$I_A = \{P(A+B) \geq P(A) + \sigma(B) \text{ for all } B \in \mathcal{C}\}$$

are called *equilibrium measures for A*. The equilibrium measures for Φ are those corresponding to A_Φ.

Theorem 4.11. *For all $A \in \mathcal{C}$, the set I_A is nonempty, convex, compact, a simplex, and a face of I. The set*

$$D = \{A \in \mathcal{C} : I_A \textit{ is a single point}\}$$

is residual in C.

Let $\mathscr{X}$ be a separable Banach space and $\varphi : \mathscr{X} \to \mathcal{C}$ be a continuous linear map such that $\varphi(\mathscr{X})$ is dense in $\mathcal{C}$. Then, for $\Phi \in \mathscr{B}$, the set $I_{\varphi\Phi}$ is the closed convex hull of the set of ρ such that $\rho = \lim_{n\to\infty} \rho_n$ for any sequence $\rho_n \in I_{\varphi\Phi_n}$ defined by interactions $\Phi_n \in \varphi^{-1}(D)$ such that $\lim_{n\to\infty} \|\Phi_n - \Phi\| = 0$.

As will be shown later in (4.21), this defines the measures which minimize the "free energy". The fact that it is a face of I shows that its extremal measures are ergodic, and that every equilibrium state has a unique decomposition into those ergodic phases. The proof to this theorem may be found as Theorem 3.7 of Ruelle's book [6]. We may note that $\mathscr{B}$ is one such space where this theorem applies.

In the following discussion, the notion of entropy is extended to probability measures on the lattice configuration space. This is described via the limit of the average entropy on finite subsets to the whole lattice system.

Definition 4.12. Given a probability measure σ_Λ on Ω_Λ for finite Λ, the *entropy* is

$$S(\sigma_\Lambda) = -\sum_{\xi \in \Omega_\Lambda} \sigma_\Lambda\{\xi\} \log \sigma_\Lambda\{\xi\}$$

One may then observe

$$0 \leq S(\sigma_\Lambda) \leq |\Lambda| \log |\Omega_0|$$

The following boundedness and *strong subadditivity* property also hold. These are shown in detail in Ruelle's book [6]. For $0 < t < 1$ and σ'_Λ, another probability measure on Ω_Λ

$$\begin{aligned} tS(\sigma_\Lambda) + (1-t)S(\sigma'_\Lambda) &\leq S(t\sigma_\Lambda + (1-t)\sigma'_\Lambda) \\ &\leq tS(\sigma_\Lambda) + (1-t)S(\sigma'_\Lambda) + \log 2. \end{aligned} \tag{4.13}$$

For $\sigma \in E$, the function $S(\alpha_\Lambda \sigma)$ is increasing with Λ and satisfies

$$S(\alpha_{\Lambda_1 \cup \Lambda_2}\sigma) + S(\alpha_{\Lambda_1 \cap \Lambda_2}\sigma) \leq S(\alpha_{\Lambda_1}\sigma) + S(\alpha_{\Lambda_2}\sigma) \tag{4.14}$$

Definition 4.15. A sequence of finite sets $\Lambda_n \subset \mathbb{Z}^\nu$ *tends to infinity in the sense of van Hove* if $|\Lambda_n| \to \infty$ and for all $a \in \mathbb{Z}^\nu$

$$\frac{|(\Lambda_n + a)/\Lambda_n|}{|\Lambda_n|} \to 0.$$

This is less restrictive than $\Lambda(a) \to \infty$ as defined prior, as it allows for some points of L to be excluded. However, it preserves the notion of the "boundary" points of Λ_n becoming negligible in the limit. We denote this as $\Lambda_n \nearrow \infty$.

Theorem 4.16. *If $\sigma \in I$, then the limit*

$$s(\sigma) = \lim_{\Lambda_n \nearrow \infty} \frac{S(\alpha_{\Lambda_n}\sigma)}{|\Lambda_n|} = \inf_\Lambda \frac{S(\alpha_\Lambda \sigma)}{|\Lambda|}$$

exists, is well defined over all sequences Λ_n that tend to infinity in the sense of van Hove, and is called the mean entropy*. This function is positive, affine, and upper semicontinuous on I. The following proof follows that of Ruelle [6].*

Proof. For Λ, M finite and disjoint, strong subadditivity implies

$$S(\alpha_{\Lambda \cup M}\sigma) \leq S(\alpha_\Lambda \sigma) + S(\alpha_M \sigma). \tag{4.17}$$

For $\sigma \in I$, translation invariance implies that for all $x \in L$

$$S(\alpha_{(\Lambda + x)}\sigma) = S(\alpha_\Lambda \sigma). \tag{4.18}$$

Let $s = \inf_a |\Lambda(a)|^{-1} S(\alpha_{\Lambda(a)}\sigma)$. Then, for any $\epsilon < 0$ choose some $\Lambda(b)$ such that

$$|\Lambda(b)| S(\alpha_{\Lambda(b)}\sigma) \leq s + \epsilon.$$

Integer translations of $\mathbb{Z}^\nu(b)$ over elements of

$$\mathbb{Z}^\nu(b) = \{x_i \in \mathbb{Z}_\nu : x_i = m_i b_i \text{ where } m_i \in \mathbb{N} \text{ for all } 0 \leq i \leq \nu\}$$

define a partition of L. For each Λ_n, let Λ_n^+ be the union of $\Lambda(b) + x$ for $x \in \mathbb{Z}^\nu(b)$ such that the intersection of $\Lambda(b) + x$ and Λ_n is nonempty. As Λ_n tends to infinity in the sense of van Hove, we have

$$\Lambda_n^+ \supset \Lambda_n \text{ and } |\Lambda_n^+|/|\Lambda_n| = \frac{(|\Lambda_n^+/\Lambda_n| + |\Lambda_n|)}{|\Lambda_n|} \to 1.$$

As $S(\alpha_\Lambda \sigma)$ is increasing with Λ, the equations (4.17) and (4.18) imply

$$S(\alpha_{\Lambda_n}\sigma) \leq S(\alpha_{\Lambda_n^+}\sigma) \leq \frac{\Lambda_n^+}{\Lambda(b)} S(\alpha_{\Lambda(b)}\sigma) \leq |\Lambda_n^+|(s + \epsilon).$$

Combined with the fact that $|\Lambda_n^+|/|\Lambda_n| \to 1$, taking the limit superior yields

$$\limsup_{\Lambda \nearrow \infty} \frac{S(\alpha_{\Lambda_n}\sigma)}{|\Lambda_n|} \leq s + \epsilon. \tag{4.19}$$

As one may observe that $\Lambda(a) \to \infty$ implies $\Lambda(a) \nearrow \infty$, the initial definition $s = \inf_a |\Lambda(a)|^{-1} S(\alpha_{\Lambda(a)}\sigma)$ and equation (4.19) imply

$$\lim_{a \to \infty} \frac{S(\alpha_{\Lambda(a)}\sigma)}{|\Lambda(a)|} = s.$$

For $x \notin \Lambda' \supset \Lambda$, strong subadditivity yields

$$S(\alpha_{\Lambda \cup \{x\}}\sigma) - S(\alpha_\Lambda \sigma) \geq S(\alpha_{\Lambda' \cup \{x\}}\sigma) - S(\alpha_{\Lambda'}\sigma).$$

Then, upon iterating this inequality from any Λ to some arbitrarily large $\Lambda(a)$ via successive steps Λ_n', it follows that

$$S(\alpha_{\Lambda \cup \{x\}}\sigma) - S(\alpha_\Lambda \sigma) \geq \lim_{a \to \infty} \frac{S(\alpha_{\Lambda(a)}\sigma)}{|\Lambda(a)|} = s.$$

Applying this iteratively over each $x \in \Lambda$ obtains $S(\alpha_\Lambda \sigma) \geq |\Lambda| s$ for all Λ finite. From this and (4.19), we obtain the result

$$\lim_{\Lambda \nearrow \infty} \frac{S(\alpha_{\Lambda_n}\sigma)}{|\Lambda_n|} = \inf_\Lambda \frac{S(\alpha_\Lambda \sigma)}{|\Lambda|} = s.$$

Finally, as $S \geq 0$ as defined, the limit also has $s \geq 0$. The limit s scaling over $|\Lambda|^{-1}$ with equation (4.13) implies that the function is affine. Finally, as the functions $\frac{S(\alpha_\Lambda \sigma)}{|\Lambda|}$ are continuous, their infimum is upper semicontinuous. □

Lemma 4.20. *For $\sigma \in E$ and $A \in \varphi(\mathscr{A}_0)$, the set of A_Φ corresponding to finite range interactions, the pressure is bounded by $P(A_\Phi) \geq s(\sigma) + \sigma(A_\Phi)$.*

Proof. First, let E_Λ be the set of probability measures on Ω_Λ. For $\sigma_\Lambda \in E$, the concavity of the logarithm implies

$$S(\sigma_\Lambda) - \sigma(U_\Lambda^\Phi) = \sum_{\xi \in \Omega_\Lambda} \sigma_\Lambda\{\xi\} \log \frac{\exp[-U_\Lambda(\xi)]}{\sigma_\Lambda\{\xi\}} \leq \log \sum_{\xi \in \Omega_\Lambda} \exp[-U_\Lambda(\xi)] = \log Z_\Lambda^\Phi.$$

We may note the equality holds for Gibbs ensembles $\sigma_\Lambda\{\xi\} = \frac{\exp[-U_\Lambda(\xi)]}{Z_\Lambda^\Phi} = \mu_{(\Lambda)}\{\xi\}$.

Using the fact that

$$\sigma(A_\Phi) = -\lim_{a\to\infty} |\Lambda(a)|^{-1}(\alpha_{\Lambda(a)}\sigma)(U^\Phi_{\Lambda(a)}),$$

we may then take the limit over

$$\frac{\log(Z^\Phi_{\Lambda(a)})}{|\Lambda(a)|} \geq \frac{s(\alpha_{\Lambda(a)}\sigma) - \alpha_{\Lambda(a)}\sigma(U^\Phi_{\Lambda(a)})}{|\Lambda(a)|}$$

to obtain

$$P(A_\Phi) \geq s(\sigma) + \sigma(A_\Phi).$$

□

Theorem 4.21. *For all $A \in \mathcal{C}$,*

$$P(A) = \max_{\sigma\in I}[s(\sigma) + \sigma(A)] \tag{4.22}$$

and the maximum holds on I_A. For all $\sigma \in I$, the entropy is

$$s(\sigma) = \inf_{A\in\mathcal{C}}[P(A) - \sigma(A)]. \tag{4.23}$$

Proof. We begin by restricting our consideration to $\Phi \in \mathscr{A}_0$ and their associated A_Φ. For a sequence $\Lambda(a_n) \to \infty$ there is some probability measure

$$\rho_{\Lambda,n}\{\xi\} = |\Lambda(a_n)|^{-1} \sum_{x\in B^\Lambda_n} (\alpha_{\Lambda+x}\mu_{(\Lambda(a_n))})\{\tau^x\xi\}$$

For all finite $\Lambda \in \mathbb{Z}^\nu$ and where the set $B^\Lambda_n = \{x : \Lambda + x \in \Lambda(a_n)\}$ is bounded by $|B^\Lambda_n| \leq |\Lambda(a_n)|$. This holds precisely when Λ is a single point for translations that map Λ to the points of $\Lambda(a)$. For Λ with more points, admissible translations must be a subset of those that map at least one point into $\Lambda(a_n)$. Further, as $(B^\Lambda_n)^c = \{x :\ x + l \notin \Lambda(a_n) \text{ for all } l \in \Lambda\} = \bigcup_{l\in\Lambda} \tau^l((\Lambda + l)/\Lambda)$ and $\Lambda(a_n) \nearrow \infty$, the van Hove limit implies $|B^\Lambda_n|/|\Lambda(a_n)| \to 1$.

This allows the limiting probability measure $\rho_\Lambda = \lim_{n\to\infty} \rho_{\Lambda,n}$ to exist via a diagonalization argument similar to 3.9. By continuous extension, as in 3.9 again, the unique state ρ such that $\rho_\Lambda = \alpha_\Lambda\rho$ for all Λ is a translation invariant probability measure $\rho \in I$. So, this implies

$$\begin{aligned} s(\rho) &= \lim_{b\to\infty} |\Lambda(b)|^{-1} S(\rho_{\Lambda(b)}) \\ &= -\lim_{b\to\infty}\lim_{n\to\infty} |\Lambda(b)|^{-1} \sum_{\xi\in\Omega_{\Lambda(b)}} \rho_{\Lambda(b),n}\{\xi\}\log\rho_{\Lambda(b),n}\{\xi\} \\ &\geq \limsup_{b\to\infty}\limsup_{n\to\infty} |\Lambda(b)|^{-1}|\Lambda(a_n)|^{-1} \sum_{x\in B_n} S(\alpha_{\Lambda(b)+x}\mu_{(\Lambda(a_n))}) \end{aligned}$$

We note that the translations of $\Lambda(b) \in \Lambda(a_n)$ may be separated into $|\Lambda(b)|$ collections of disjoint sets that, along with k_n other points such that $k_n/|\Lambda(a_n)| \to 0$, form disjoint partitions of $\Lambda(a_n)$. So, applying (4.18) to these translates under the limit superior as $n \to \infty$ yields

$$s(\rho) \geq \limsup_{n\to\infty} |\Lambda(a_n)|^{-1} S(\mu_{(\Lambda(a_n))})$$

It follows that

$$\begin{aligned} s(\rho)+\rho(A_\Phi) &\geq \lim_{n\to\infty} |\Lambda(a_n)|^{-1} \sum_{\xi\in\Omega_{\Lambda(a_n)}} \mu_{(\Lambda(a_n))}\{\xi\} \log \frac{\exp\left[-U^{\Phi}_{\Lambda(a_n)}(\xi)\right]}{\mu_{(\Lambda(a_n))}\{\xi\}} \\ &= \lim_{n\to\infty} |\Lambda(a_n)|^{-1} \log Z^{\Phi}_{\Lambda(a_n)} \\ &= P(A_\Phi) \end{aligned}$$

Combined with 4.20, this implies $P(A) = \sup_{\sigma\in I}[s(A) - \sigma(A)]$ for $A \in \varphi(\mathscr{A}_0)$. By continuity of these expressions, the fact that this holds on a dense subset of $\mathcal{C}$ implies the property on all $A \in \mathcal{C}$. Further, by upper semi-continuity of σ, the supremum is obtained, and (4.22) holds.

Now, this implies $s(\sigma) \leq P(A) - \sigma(A)$. To show that these values may be arbitrarily close to each other, consider the space

$$C = \{(\sigma, t) \in C^* \times \mathbb{R} : \sigma \in I \text{ and } 0 \leq t \leq s(\sigma)\}$$

as s is affine and upper semi-continuous from 4.16, C is convex and compact. For some $\rho \in I$ and $u > s(\rho)$ there are $A \in \mathcal{C}$ and $c \in \mathbb{R}$ such that $-\rho(A) + c = u$ and $-\sigma(A) + c \leq s(\sigma)$ for all $\sigma \in I$. This implies $-\sigma(A) + \rho(A) + u > s(\sigma)$.

Choosing σ such that the maximum $P(a) = s(\sigma) + \sigma(A)$ is obtained, this yields

$$\begin{aligned} 0 &\leq P(A) - s(\rho) - \rho(A) \\ &= s(\sigma) + \sigma(A) - s(\rho) - \rho(A) \\ &\leq u - s(\rho). \end{aligned}$$

So, as u is chosen arbitrarily close to $s(\rho)$, the infimum of $s(\rho) - \rho(A)$ over $A \in \mathcal{C}$ is equal to $P(A)$. This shows (4.23).

Now, the measures $\rho \in I_A$ are defined such that

$$P(A+B) \geq P(A) + \rho(B), \text{ for all } B \in \mathcal{C}$$

or equivalently

$$P(A+B) - \rho(A+B) \geq P(A) - \rho(A), \text{ for all } B \in \mathcal{C}.$$

Bounding above by a space is equivalent to bounding above by its infimum, so

$$s(\rho) = \inf_{C\in\mathcal{C}}[P(C) - \rho(C)] \geq P(A) - \rho(A).$$

As 4.20 holds for all $\sigma \in E$, the maximum (4.22) holds precisely on I_A. $\square$

Corollary 4.24. *The limit to infinity in the sense of van Hove for the pressure generalizes* 4.3. *That is*

$$P(A) = \lim_{\Lambda\nearrow\infty} |\Lambda|^{-1} \log Z^*_\Lambda(A)$$

The proof of this fact may be found in [6].

Next, we return our considerations on the space of interactions $\mathscr{B}$ so that both Gibbs states and equilibrium states are well defined. The following theorem is the primary result for thermodynamic formalism on the lattice, as it allows for the precise expression of translation invariant equilibrium states as Gibbs states for a large class of interactions.

The condition on this class of interactions is that there are sequences Λ_n, M_n where $\Lambda_n \nearrow \infty$, $\Lambda_n \subset M_n$, and $|\Lambda_n|/|M_n| \to 1$ such that for every $\xi, \eta \in \Omega$ and $n \in \mathbb{N}$, there are $\zeta_n \in \Omega$ where

$$\zeta_n|_{\Lambda_n} = \xi|_{\Lambda_n} \text{ and } \zeta_n|_{\mathbb{Z}^\nu/M_n} = \eta|_{\mathbb{Z}^\nu/M_n}.$$

Theorem 4.25. *If $\Phi \in \mathscr{B}$, then $I_{\varphi\Phi} \subset K_\Phi \cap I$, so all equilibrium states are translation invariant Gibbs states. Further, if the above condition is satisfied, then $I_{\varphi\Phi} = K_\Phi \cap I$, the translation invariant Gibbs states are precisely the equilibrium states.*

Proof. First, we consider the set $\Phi \in \mathscr{A}_0$. In this case, the measures $\mu_{(\Lambda)_\eta}$ only depend on $\eta|_{(M/\Lambda)}$ for the finite set $M = \Lambda + \Delta$. Using the ρ_Λ as defined in 3.21, we see that for sufficiently large $\Lambda(a_n)$ as to contain translates of M, the probability contribution from interactions which intersect Λ may be separated as

$$\rho_{\Lambda,n} = |\Lambda(a_n)|^{-1} \sum_{x:\Lambda+x\in\Lambda(a_n)} \left(\sum_{\eta\in\Omega_\Lambda} \mu(\Lambda)_\eta\{\tau^x\xi\}\right) (\alpha_{M/\Lambda}+x)\mu_{(\Lambda(a_n))}\{\tau^x\xi\}.$$

So, along with the translation invariance of $\mu_{(\Lambda)_\eta}$, taking the limit as $n \to \infty$ yields

$$\rho_\Lambda\{\xi\} = \sum_{\eta\in\Omega_{M/\Lambda}} \mu(\Lambda)_\eta\{\xi\}\rho_{M/\Lambda}\{\eta\}.$$

This characterizes the state ρ^Φ defined by $\alpha_\Lambda\rho^\Phi = \rho_\Lambda$ as a Gibbs state. From the proof of 3.21, this is a equilibrium state. So, for $\Phi \in \mathscr{A}_0$, there is some $\rho^\Phi \in I_{\varphi(\Phi)} \cap K_\Phi$ where K_Φ is the set of Gibbs states.

As $\mathscr{A}_0$ is dense in $\mathscr{A}$, for any $\Phi \in \varphi^{-1}D \subset \mathscr{A}$ there is a sequence $\Phi_n \in \mathscr{A}_0$ where $\lim_{n\to\infty}\Phi_n = \Phi$. As translation invariance and the inequality $P(A_\Phi + B) \geq P(A_\Phi) + \sigma(B)$ are preserved under the limit, the latter due to the continuity of P and φ, the measure $\rho^\Phi = \lim_{n\to\infty}\rho^{\Phi_n}$ is in $I_{\varphi(\Phi)}$. As the set of Gibbs states is closed, this limit is a Gibbs state. Hence, there exists some $\rho^\Phi \in I_{\varphi\Phi} \cap K_\Phi$ for $\Phi \in D$.

Then, for any $\Phi \in \mathscr{B}$, as $I_{\varphi(\Phi)}$ is the closed convex hull of $\rho \in K_{\Phi_d} \cap I_{\Phi_d}$ for $\Phi_d \in D$ from 4.11, the preservation of equilibrium in the limit implies that $I_{\varphi(\Phi)} \subset K_\Phi \cap I$.

Under the above condition it is true that if $\sigma \in K_\Phi$,

$$\liminf_{n\to\infty} |M_n|^{-1}\left[S(\alpha_{M_n}) - (\alpha_{M_n}\sigma)\left(U^\Phi_{M_n}\right)\right] \geq P^\Phi.$$

The proof is omitted here for the sake of brevity, but may be found in [6]. It then follows that translation invariant Gibbs states have $s(\sigma) + \sigma(A_\Phi) \geq P^\Phi$, so they are Gibbs states. □

5. Lattice Morphisms Introduction to Subshifts of Finite Type

In this section, morphism between lattice systems are described. This allows for restrictions on the system to be reduced to limits on the configruations only between neigboring points on the lattice. This allows for the description of one dimensional systems as subshifts of finite type. The primary result for subshifts of finite type is the reduction of the dynamics for subshifts into transitive and mixing subsystems. As a consequence, this shows that transitive subshifts have only one equilibrium state.

Definition 5.1. Given the lattice system $(L, (\Omega_x)_{x\in L}, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$ and another system $(L', (\Omega'_x)_{x\in L'}, (\bar{\Omega}'_\Lambda)_{\Lambda\in\mathcal{F}'})$, we consider maps $F_x : \Omega'_{M(x)} \to \Omega_x$ where $M(x)$ is a finite subset of L' for each $x \in L$. There are two conditions on these maps. First, sets $(M(x))_{x\in L}$ are locally finite, that is $\{x : x' \in M(x)\}$ is finite for all $x \in L$. Second, if $\xi' \in \Omega_{\cap\{M(x):x\in X\}}$, then $(F_x(\xi'_{M(x)}))_{x\in X}$ is an element $FX_{\xi'}$ of Ω_X whenever $X \in L$.

The map $F : \Omega' \to \Omega$ between configuration spaces where $(F\xi')_x = F_x(\xi'|_{M(x)})$ is then continuous for the discrete topologies on $\Omega_x \mapsto \Omega'_x$, and hence product topologies on $\Omega \mapsto \Omega'$.

We define the set $\Sigma_\xi = \{\eta : \eta_x = \xi_x \text{ for finite } x \in L\} \subset \Omega$ and similarly $\Sigma'_{\xi'} \subset \Omega'$. As defined, F is a map where $\Sigma'_{\xi'} \mapsto \Sigma_{F\xi'}$. We call F a *morphism* if it satisfies the above construction and $F|_{\Sigma'_{\xi'}} : \Sigma'_{\xi'} \to \Sigma_{F\xi'}$ is bijective for all $\xi' \in \Omega'$. We may note that the identity map satisfies these conditions and that they are preserved under composition of two morphisms.

If, for the morphism F, there exists some morphism F' from $(L', (\Omega'_x)_{x\in L'}, (\bar{\Omega}'_\Lambda)_{\Lambda\in\mathcal{F}'})$ to $(L, (\Omega_x)_{x\in L}, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$ such thatFF' and $F'F$ are both the identity map, then F is an *isomorphism*.

If $L = L' = \mathbb{Z}^\nu$ and $F_{x-a}\tau^a = F_x$ for all $a \in \mathbb{Z}^\nu$, then F is a $\mathbb{Z}^\nu$*-morphism*. $\mathbb{Z}^\nu$-isomorphisms are defined similarly as before.

This defines a notion of morphisms between lattice systems. We will not undergo a comprehensive discussion of the important results for lattice morphisms; however, there are some important results, the proofs of which we omit for brevity, which may be found in [6].

Proposition 5.2. *Given a morphism F as above for the interaction Φ, there is an interaciton $F^*\Phi$ such that*

$$(F^*\Phi)(\xi') = \sum_{X:\bigcup\{M(x):x\in X\}=X'} \Phi(F_x\xi') \text{ if } \xi' \in \Omega_{X'}.$$

This is constructed such that if σ' is a Gibbs state on Ω' for $F^*\Phi$, then $F\sigma'$ is a Gibbs state for Φ on Ω.

Proposition 5.3. *If $F : (\mathbb{Z}^\nu, \Omega'_0, (\bar{\Omega}'_\Lambda)_{\Lambda\in\mathcal{F}'}) \to (\mathbb{Z}^\nu, \Omega_0, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$ is a $\mathbb{Z}^\nu$-morphism and σ' is a translation invariant state on Ω', then $s(\sigma') \geq s(F\sigma')$ and $\sigma'(A_{F^*\Phi}) = (F\sigma')(A_\Phi)$. Hence by* 5.2 *and the variational principle* (4.22)*, it follows that $P^{F^*\Phi} = P^\Phi$.*

This demonstrates how these morphisms encode the thermodynamic properties of lattices.

Further, the following alternative characterization of $\mathbb{Z}^\nu$ morphisms from [6] will allow for all one dimensional $\mathbb{Z}$ lattices to be encoded as subshifts of finite type.

Proposition 5.4. *A map $F : \Omega' \to \Omega$ is a $\mathbb{Z}^\nu$ morphism if and only if the following hold:*

(1) F is continuous.
(2) F is equivariant or $\tau^a F = F\tau^a$ for all $a \in \mathbb{Z}^\nu$.
(3) F restricted to $\Sigma'_{\xi'} = \{\eta' \in \Omega' : \lim_{x\to\infty} d'(\tau^x\eta', \tau^x\eta) = 0\}$ is a bijection to the set $\Sigma_{F\xi'} = \{\eta \in \Omega : \lim_{x\to\infty} d'(\tau^x\eta, \tau^x F\eta') = 0\}$, where d and d' are metrics compatible with the topologies on Ω and Ω' respectively.

Corollary 5.5. *For each lattice system* $(\mathbb{Z}^\nu, \Omega_0, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$*, there exists some* $F : (\mathbb{Z}^\nu, \Omega_0', (\bar{\Omega}_\Lambda')_{\Lambda\in\mathcal{F}'}) \to (\mathbb{Z}^\nu, \Omega_0, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$ *where* $\mathcal{F}'$ *only contains sets* $\{x, y\}$ *such that* x *and* y *are neighbors, that is* $\sum_{i=1}^{\nu} |x_i - y_i| = 1$.

Proof. Let $M(x) = \{y \in \mathbb{Z}^\nu : \max_i |x_i - y_i| \le l\}$ where $l \ge 0$ be defined so that if $x \in \Lambda \in \mathcal{F}$, then $\Lambda \subset M(x)$. Further, let $\Omega_x' = \Omega_{M(x)}$.

Now, for the set $\mathcal{F}'$ of all such $\{x, y\}$ as defined prior, the sets

$$\bar{\Omega}'_{\{x,y\}} = \{(\xi, \eta) \in \Omega_{M(x)} \times \Omega_{M(y)} : \xi|_{M(x)\cap M(y)} = \eta|_{M(x)\cap M(y)}\}$$

are then defined such that $F : \Omega' \mapsto \Omega$ such that $(F\xi')_x = (\xi_x')_x$ is equivariant as the definition is identical under translations. It is also a homeomorphism as it maps configurations on finite collections of points to their corresponding finite M configurations, and it maps finite M configurations, with the required $\bar{\Omega}'_{\{x,y\}}$ to their precise configuraions on a finite set of Ω. Thus, it satisfies the properties required in 4.4 and defines a $\mathbb{Z}^\nu$-morphism. As an equivariant homeomorphism carries the same properties required in 5.4 for its inverse, this specifies a $\mathbb{Z}^\nu$-isomorphism. □

Definition 5.6. Let G be a directed graph with finite vertices V and edges $E \subset V \times V$. We assume each vertex has at least one incoming and one outgoing edge.

A (two-sided) subshift of finite type is the set

$$\Sigma(G) = \{x \in V^{\mathbb{Z}} : (x_i, x_{i+1}) \in E \text{ for all } i \in \mathbb{Z}\}.$$

This is also equipped with the product topology, which may be identified with the metric $d(x, y) = \exp(-\min\{|n| : x_n \ne y_n\})$. This topology is generated the cylinder sets $[a_0, \cdots, a_k]_m = \{x \in \Sigma^+(G) : x_{i+m} = a_i \text{ for all } 0 \le i \le k\}$ where $m \in \mathbb{Z}$.

This definition for two-sided sub-shifts is equivalent to the translation invariant one dimensional lattice system with $\mathcal{F}$ as in 4.5 where $V = \Omega_0$ and $E = \mathcal{F}'$. So, the prior results establish that thermodynamic results on subshifts of finite type generalize to those of one-dimensional lattice systems in general.

Now, there are two topological properties for continuous maps on Hausdorff topological spaces which characterize the dynamics on the space and will be important to discussing $\mathbb{Z}$ lattice systems and subshifts of finite type.

Definition 5.7. Let Ω be a nonempty Hausdorff topological space and $f : \Omega \to \Omega$ a continuous map. The dynamical system (Ω, f) is *transitive* if for all nonempty open sets $U, V \subset \Omega$ and $N \ge 0$, there exists some $n > N$ such that $f^n U \cap V \ne \emptyset$.

The dynamical system (Ω, f) is *mixing* if for all nonempty open sets $U, V \subset \Omega$, there exists $N \ge 0$ such that $f^n U \cap V \ne \emptyset$ for all $n > N$.

For subshifts of finite type and one dimensional lattice systems, the continuous map is $f = \tau^1$ on the configuration space Ω. The system may be represented by $V = \Omega_0$ and the transition matrix between $i, j \in V$

$$t_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \bar{\Omega}_{\{x,x+1\}}, \\ 0 & \text{if } (i, j) \in \bar{\Omega}_{\{x,x+1\}}. \end{cases}$$

Definition 5.8. There is a class of $\mathbb{Z}^\nu$-isomorphisms called the *restriction to a subgroup* G. In particular, for the lattice system $(\mathbb{Z}^\nu, \Omega_0', (\bar{\Omega}_\Lambda')_{\Lambda\in\mathcal{F}'})$ choose $M(0) \subset \mathbb{Z}^\nu$ containing one element of each residue class $\mathbb{Z}^\nu \mod G$ and define $M(x) = M(0) + x$ for each $x \in G$. This defines $(M(x))_{x\in G}$ to be a partition of $\mathbb{Z}^\nu$. From this, let $\Omega_x = \Omega'_{M(x)}$ and $F_x : \Omega'_{M(x)} \to \Omega_x$ be the identity map. Let $\Lambda' \in \mathcal{F}'$ if $\Lambda' \cap M(x) \neq \emptyset$ for all $x \in \Lambda$ and define the configuration restrictions to be

$$\bar{\Omega}_\Lambda = \{(\xi'|_{M(x)})_{x\in\Lambda} : \xi' \in \Omega'_{\bigcup\{M(x):x\in\Lambda\}}\}.$$

This defines an isomorphism $F : (\mathbb{Z}^\nu, \Omega'_0, (\bar{\Omega}'_\Lambda)_{\Lambda\in\mathcal{F}'}) \to (G, \Omega'_{M(0)}, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$, where the lattice system $(G, \Omega'_{M(0)}, (\bar{\Omega}_\Lambda)_{\Lambda\in\mathcal{F}})$ is the restriction of the original lattice system to the subgroup G of $\mathbb{Z}^\nu$.

A fact which is shown in proposition 4.14 of [6], is that if $A \in \mathcal{C}$ on Ω', then

$$P\left(\sum_{x\in M(0)} A \circ \tau^x \circ F^{-1}\right) = |M(0)|P(A).$$

The following theorems allow for subshifts of finite type to be reduced to transitive and mixing cases, the proofs may be found in chapter 5 of [6]. These hold for the banach space $\mathscr{B}_1$ of interactions under the norm $\|\Phi\|_1 = |\Phi| + |\Phi|_1$ where

$$|\Phi|_1 = \sum_{X\ni 0} \frac{\operatorname{diam}(X)}{|X|} \sup_{\xi\in\Omega_x} |\Phi(\xi)| \leq \infty.$$

Theorem 5.9. *Given a lattice system $(\mathbb{Z}, \Omega_0, (\Omega_\Lambda)_{\Lambda\in\mathcal{F}})$, there are finitely many transitive $(\Omega_0^{(a)}, t^{(a)})$ subshifts of finite type and injective $\mathbb{Z}$-morphisms*

$$F^{(a)} : (\Omega_0^{(a)}, t^{(a)}) \to (\mathbb{Z}, \Omega_0, (\Omega_\Lambda)_{\Lambda\in\mathcal{F}})$$

where the following hold.

(1) The images of $F^{(\alpha)}\Omega^{(\alpha)}$ are disjoint.

(2) Every Gibbs state for $\Phi \in \mathscr{B}_1$ is a convex combination of Gibbs states $F^{(\alpha)}\sigma^{(\alpha)}$ where $\sigma^{(\alpha)}$ is a Gibbs state for $F^{(\alpha)*}\Phi$ on $(\Omega_0^{(\alpha)}, t^{(\alpha)})$.

(3) If $\xi \in \Omega$ is periodic, that is $\tau^p\xi = \xi$ for $p > 0$, then ξ is in one of the $F^{(\alpha)}\Omega^{(\alpha)}$.

Theorem 5.10. *Given a transitive lattice system $(\mathbb{Z}, \Omega_0, (\Omega_\Lambda)_{\Lambda\in\mathcal{F}})$, there are $N \in \mathbb{Z}^+$ mixing $N\mathbb{Z}$-lattice systems $(\Omega_0^{(\beta)}, t^{(\beta)})$, which are defined as $\mathbb{Z}$ lattice systems identified with a $N\mathbb{Z}$ lattice system via the isomorphism $\mathbb{Z} \mapsto N\mathbb{Z}$. These are defined such that the injective $N\mathbb{Z}$ morphisms*

$$F^{(\beta)} : (\Omega_0^{(\beta)}, t^{(\beta)}) \to (N\mathbb{Z}, \Omega_0, (\Omega'_\Lambda)_{\Lambda\in\mathcal{F}'}),$$

where $(N\mathbb{Z}, \Omega_0, (\Omega'_\Lambda)_{\Lambda\in\mathcal{F}'})$ is the restriction of the original lattice system to the subgroup $N\mathbb{Z}$, obey the following properties.

(1) The images of $F^{(\beta)}\Omega^{(\beta)}$ are disjoint.

(2) Every Gibbs state for $\Phi \in \mathscr{B}_1$ has its corresponding interaction Φ^* defined below for $(N\mathbb{Z}, \Omega_0, (\Omega'_\Lambda)_{\Lambda\in\mathcal{F}'})$. Then, every gibbs state for Φ is a convex combination of Gibbs states $F^{(\beta)}\sigma^{(\beta)}$ for $\sigma^{(\beta)}$ the unique gibbs state for $F^{(\beta)*}\Phi^*$ on $(\Omega_0^{(\beta)}, t^{(\beta)})$.

(3) If $\xi \in \Omega$ is periodic and $\tau^p\xi = \xi$, then p is a multiple of N.

The corresponding interaction $\Phi^* = (F^{-1})^*\Phi'$ is the interaction defined for the inverse morphism of F, the restriction of $\mathbb{Z}$ to the subgroup $N\mathbb{Z}$.

Corollary 5.11. *Let $(\mathbb{Z}, \Omega_0, (\Omega_\Lambda)_{\Lambda\in\mathcal{F}})$ be a $\mathbb{Z}$-lattice system and $\Phi \in \mathscr{B}_1$. If the system is transitive, there is a unique equilibrium state corresponding to the unique* *translation invariant Gibbs state. If the system is mixing, then there is a unique Gibbs state.*

Further, the system described in 4.9 *with* $A \in \mathcal{C}$ *has pressure*

$$P(A) = \max_{\alpha} P(A \circ F^{(\alpha)}).$$

For the system in 5.3 *with* $A \in \mathcal{C}$ *for all* β *has pressure*

$$P(A) = N^{-1} P((A + A \circ \tau + \cdots + A \circ \tau^{N-1}) \circ F^{(\beta)}).$$

and the Gibbs state is of the form

$$N^{-1} \sum_{\beta=0} \tau^{-\beta} \sigma$$

where σ *is a Gibbs state whose support is in a mixing component* $F^{(\beta)}\Omega^{(\beta)}$.

The results of 5.9 reduce the general case to those of finitely many distinct transitive subshifts of finite type. Further, the results of 4.10 allow these cases to be broken down to distinct mixing systems over a finite number of map iterations. These mixing systems, as they are ergodic, must have unique equilibrium states, which extend to transitive systems by cyclic permutations as in 5.10.

The corollary demonstrates how the pressure, and consequently the thermodynamic properties of the system, depend on the transitive and mixing subsystems described above. This presents a method of analyzing thermodynamic properties of subshifts of finite type via study of the much smaller class of mixing systems. Further, the uniqueness of equilibrium states for these transitive and mixing cases poses a strong restriction on the steady state behavior of one-dimensional lattice systems.

6. Generalization to Dynamics on Compact Metric Spaces

The prior sections established a description of equilibrium states for lattice systems. In the following sections, we seek to generalize the notion of equilibrium states on a compact metrizable space X. In this setting, iterations of homeomorphisms $T : X \to X$ serve analogously to iterations of translations τ^1 on $\mathbb{Z}$. The sets $\mathcal{C}$ and the probability measures $E \in \mathcal{C}^*$ are defined as in section 1. Similarly to lattice systems, the aim is to describe equilibrium states which maximize the pressure of the system. In this case, the uncountable space of configurations is identified with finitely many "course grain" locations in a partition, as well as the partition location of a configuration's iterates by T.

As in the prior discussion, we consider the probability measures on the Borel σ-algebra, that is the states of the system $E \in C^*$. In particular, an important class of states is I, the set of T-invariant states, that is the states μ with the property $\mu(T^{-1}(B)) = \mu(B)$ for all B in the Borel σ-algebra $\mathfrak{B}$. In this instance, T is an automorphism of the measure space $(X, \mathfrak{B}, \mu)$.

Definition 6.1. For a measure space $(X, \mathfrak{B}, \mu)$ with a finite measurable partition $\{U_1, \ldots, U_k\} = \mathfrak{U} \subset \mathfrak{B}$, we define

$$H_\mu(C) = \sum_{i=1}^{k} -\mu(U_i) \log(\mu(U_i)).$$

This quantifies a weighted average of the information gained from specifying that a point lies in a particular U_i.

These partitions provide a "coarse-grain" representation of points in X, and the entropy describes the amount of uncertainty that this description provides about the location of points in X. More appropriately for the physical ensemble perspective, this relates to the average number of ways a particular ensemble of micro-states may be arranged, provided their coarse-grain locations are specified. This is a generalization of entropy in 4.12, where entropy in that case is taken over a partition by cylinder sets, which classify points of the configuration space by their micro-states on a finite set $\Lambda \subset L$. From here, we define partitions that specify the elements of $\mathfrak{U}$ to which the points of X will be mapped.

Lemma 6.2. *For two finite partitions $\mathfrak{U}, \mathfrak{V} \subset \mathfrak{B}$, their refinement is*

$$\mathfrak{U} \vee \mathfrak{V} = \{U_i \cap V_j : U_i \in, \mathfrak{U}, V_i \in, \mathfrak{V}\}.$$

So, the value $H_\mu(\mathfrak{U} \vee \mathfrak{V}) \leq H_\mu(\mathfrak{U}) + H_\mu(\mathfrak{V})$.

Proof. This holds as

$$\begin{aligned} H_\mu(\mathfrak{U} \vee \mathfrak{V}) - H_\mu(\mathfrak{U}) &= \sum_{i,j} -\mu(U_i \cap V_j) \log \mu(U_i \cap V_j) + \sum_i \mu(U_i) \log \mu(U_i) \\ &= \sum_{i,j} \mu(U_i) \left(-\frac{\mu(U_i \cap V_j)}{\mu(U_i)} \log \frac{\mu(U_i \cap V_j)}{\mu(U_i)} \right) \\ &\leq \sum_j -\mu(V_j) \log \mu(V_j) = H_\mu(\mathfrak{V}). \end{aligned}$$

□

Lemma 6.3. *Suppose $\{a_m\}_{m=1}^\infty$ such that* $\inf \frac{a_m}{m} \geq -\infty$ *and $a_{m+n} \leq a_m + a_n$ for all $m, n \in \mathbb{N}$. then* $\lim_{m\to\infty} \frac{a_m}{m} = \inf_m \frac{a_m}{m}$.

Proof. Fix $m > 0$. For $j > 0$, let $j = km + n$ with $0 \leq n < m$. It follows that

$$\frac{a_j}{j} \leq \frac{a_{km} + a_n}{km} \leq \frac{ka_m + a_n}{km}.$$

Letting $j \to \infty$ and $k \to \infty$, it follows from the expression that

$$\limsup_j \frac{a_j}{j} \leq \frac{a_m}{m} \leq \inf_m \frac{a_m}{m}.$$

So, as $\liminf_m \frac{a_m}{m} \geq \inf_m \frac{a_m}{m}$, the limit exists and is equal to the infimum. □

Definition 6.4. Given $(X, \mathfrak{B}, \mu)$ and $\mathfrak{U}$ as prior and an automorphism T, let $\mathfrak{U}^\Lambda = \bigvee_{a\in\Lambda} T^{-a}\mathfrak{U}$ where $T^{-a}\mathfrak{U} = \{T^{-a}(U) : U \in \mathfrak{U}\}$. Then, the following quantity exists and is called *entropy*.

$$h_\mu(T, \mathfrak{U}) = \lim_{m\to\infty} \frac{1}{m} H_\mu(\mathfrak{U}^{\{0,\ldots,m-1\}}) = \inf_m \frac{1}{m} H_\mu(\mathfrak{U}^{\{0,\ldots,m-1\}}).$$

This holds by 5.3 as

$$\begin{aligned} H_\mu(\mathfrak{U}^{\{0,\ldots,m+n-1\}}) &\leq H_\mu(\mathfrak{U}^{\{0,\ldots,m-1\}}) + H_\mu(\mathfrak{U}^{\{m,\ldots,n+m-1\}}) \\ &= H_\mu(\mathfrak{U}^{\{0,\ldots,m-1\}}) + H_\mu(T^{-m}\mathfrak{U}^{\{0,\ldots,n-1\}}) \\ &= H_\mu(\mathfrak{U}^{\{0,\ldots,m-1\}}) + H_\mu(\mathfrak{U}^{\{0,\ldots,n-1\}}) \end{aligned}$$

This may be interpreted as the limiting rate at which the partition provides information about a point based on the location of its iterates.

Let the partition $\mathfrak{U}$ be the micro-states Ω_0 for a one-dimensional lattice system. This definition describes the way the map T acts on points of X as translations on the lattice system $\Omega \subset (\Omega_0)^{\mathbb{Z}}$. This encodes the action T on $x \in X$ as it takes iterations $T^{-a}x$ to coarse grain locations $U \in \mathfrak{U}$. In this case, admissible configurations are those where T iterates some x to a particular sequence $(U_i) \in \mathfrak{U}$. More precisely, the lattice system includes configurations with successive elements $\{U_i, U_j\}$ such that $TU_i \cap U_j \neq \emptyset$.

This dynamical entropy is precisely the (mean) entropy as in 3.16 for this lattice system. So, we can alternatively define entropy via limit in the sense of van Hove where

$$h_\mu(T, \mathfrak{U}) = \lim_{\Lambda \nearrow \infty} \frac{1}{|\Lambda|} H_\mu(\mathfrak{U}^\Lambda).$$

We may then define the dynamical *mean entropy* to be

$$h_\mu(T) = \sup_{\mathfrak{U}} h_\mu(T, \mathfrak{U}).$$

The following proposition for expanding maps demonstrates how this notion of mean entropy coincides with that for one dimensional lattice systems. To state the proposition, we call a homeomorphism $T : X \to X$ *expansive* if there is some constant $\epsilon > 0$ such that the property $d(T^k x, T^k y) \leq \epsilon$ for all $k \in \mathbb{Z}$ holds only if $x = y$. In this case, we call ϵ an *expansive constant.* The diameter of a partition is also defined $\mathrm{diam}(\mathfrak{U}) = \sup_{U \in \mathfrak{U}} \mathrm{diam}(U)$.

Proposition 6.5. *If $T : X \to X$ is a homeomorphism and an automorphism of $(X, \mathfrak{B}, \mu)$ where X is a compact metric space and $\mathfrak{U}_n$ is a sequence of partitions such that diam$(\mathfrak{U}_n) \to 0$, then the entropy of the partitions approaches the mean entropy or*

$$\lim_{n\to\infty} h_\mu(T, \mathfrak{U}_\mathfrak{n}) = h_\mu(T).$$

In particular, if $T : X \to X$ has expansive constant $\epsilon > 0$, then $h_\mu(T, \mathfrak{U}) = h_\mu(T)$ whenever $\mathrm{diam}(\mathfrak{U}) \leq \epsilon$.

The proof of this fact may be found in [1]. We may note that the expansive case follows from the first as expansiveness implies $\lim_{n\to\infty} \mathrm{diam}\ \mathfrak{U}^{\{0,\dots,n\}} = 0$.

For the partition defined by cylinder sets $[a]_0$ for $a \in V$ for subshifts of finite type, equivalently the open sets of all configurations with $a \in \Omega_0$ at 0, the translations τ^1 are expansive for any $\epsilon > 0$ with respect to the metric.

As the micro-states $a \in \Omega_0$ correspond exactly to the lattice system defined for the partition in 5.4, with all admissible configurations on the new lattice system being exactly those in the original system, the original mean entropy is defined identically to the dynamical entropy. Hence, the expansiveness of the map implies the equivalence of the two notions of mean entropy. Now that there is a well defined notion of entropy for compact metrizable spaces, we consider the generalization of pressure for the system.

Definition 6.6. For $\mathfrak{U}$ a finite Borel cover of X, given $A \in \mathcal{C}$ and $\Lambda \subset \mathbb{Z}$, the *partition function* is defined to be

$$Z_\Lambda(A, \mathfrak{U}) = \min \left\{ \sum_j \exp\left[\sup_{x \in \mathfrak{B}_j} \sum_{n \in \Lambda} A(T^n x) \right] : (\mathfrak{B}_j) \text{ is a subcover of } \mathfrak{U}^\Lambda \right\}.$$

This definition is similar to that for lattices. In this case, the subcovers specify the partition location of x for its $n \in \Lambda$ iterates. The supremum over the iterative sum of values taken by elements x in a particular subcover is analogous to the arbitrary extension $\xi^* \in \Omega$ for $\xi \in \Omega_\Lambda$, as it chooses a value arbitrarily close to that of some point with iterates specified by U^Λ. Along these lines, the following definition of pressure should be identical when Z_Λ is defined with infimum instead of supremum. It is, as may be seen in [6], though the proof is omitted here for brevity.

Theorem 6.7. *The value defined as*

$$P(A, \mathfrak{U}) = \lim_{n\to\infty} \frac{1}{n} \log Z_{\{0,\dots,n\}}(A, \mathfrak{U}) = \inf_n \frac{1}{n} \log Z_{\{0,\dots,n\}}(A, \mathfrak{U})$$

exists and is finite. Further, the pressure *defined by*

$$P(A) = \lim_{\mathrm{diam}(\mathfrak{U})\to 0} P(A, \mathfrak{U})$$

is well defined, though it may be $+\infty$, and satisfies the variational principle

$$P(A) = \sup_{\mu\in I} \left(h_\mu(T) + \int A \; d\mu \right)$$

The proof to this construction may be found in section 2 of [1]. Another proof for the generalized variational principle may be found in [6]. Alternatively, the pressure may be reasonably defined by the variational principle for these systems; however, we include this construction as it provides a formula for pressure which relates to the partition function. As an aside, the quantity $P(0) = \sup_{\mu\in I} h_\mu(T)$ is often referred to as the *topological entropy* of the system, and describes maximal entropy without reference to a function A.

The pressure serves to define equilibrium analogously to that of lattice systems. It quantifies the exponential growth of the partition function as the coarse-grained approximation becomes arbitrarily precise. Intuitively, this theorem demonstrates how pressure describes equilibrium similarly to before. As the average energy stays roughly constant, this quantity defines the maximum entropy for states of the system, which is attained at equilibrium.

The following facts summarize a few important results regarding equilibrium states on these spaces. These facts may be found in Ruelle's book [6]. Similarly to 4.11, the set

$$I_A = \{\mu \in I : h_\mu(T) + \mu(A)\} = \{\mu \in \mathcal{C}^* : P(A+B) \geq P(A) + \sigma(B) \text{ for all } B \in \mathcal{C}\}$$

is nonempty, convex, compact, a simplex, and a face of I. Similarly to the case of lattices, this property allows for the decomposition of equilibrium states into ergodic "pure thermodynamic" states. These ergodic states are such that the average value of observables, such as A, vary little over time. So, these equilibrium states describe ensembles of "steady" pure states.

7. Equilibrium States of Axiom A Diffeomorphisms

Using this generalized thermodynamic formalism, the prior description allows for the analysis of equilibrium states of particular classes of diffeomorphisms. In particular, we highlight results for diffeomorphisms on the nonwandering set of compact metrizable spaces that obey Smale's axiom A and are equipped with a Holder continuous function.

Definition 7.1. A diffeomorphism $f : M \to M$ of a compact C^∞ manifold M is an *axiom A diffeomorphism* if the set of its nonwandering points

$$\Omega(f) = \left\{ x \in M : U \cap \bigcup_{n>0} T^n U \neq \emptyset \text{ for all open } U \ni x \right\}$$

is hyperbolic and $\Omega(f) = \overline{\{x : x \text{ is periodic}\}}$.

A closed subset $\Lambda \subset M$ is *hyperbolic* if $f(\Lambda) = \Lambda$ and each tangent space $f_x M$ with $x \in \Lambda$ can be written as a direct sum $f_x M = E^u_x \oplus E^s_x$ where the following hold.

(1) The derivative preserves the splitting of the tangent space $Df(E^u_x) = Df(E^u_x)$ and $Df(E^s_x) = Df(E^s_x)$.
(2) There are some constants $c > 0$ and $\lambda \in (0, 1)$ such that, for all $n \geq 0$, the splitting separates the space into expanding, or unstable, and contracting, or stable, parts respectively. That is $\|Df^{-n}(w)\| \leq c\lambda^n \|w\|$ when $w \in E^u_x$ and $\|Df^n(v)\| \leq c\lambda^n \|v\|$ when $v \in E^s_x$.
(3) The component parts of the tangent space, E^u_x and E^s_x vary continuously with x.

The following result on the nonwandering sets of axiom A diffeomorphisms is called Smale's spectral decomposition theorem and may be viewed as an extension of 4.9.

Proposition 7.2. *The nonwandering set of the axiom A diffeomorphism (M, f) is the union of a finite number of disjoint compact sets Ω^α such that $f\Omega^\alpha = \Omega^\alpha$ and $f|_{\Omega^\alpha}$ is transitive. Further, each Ω^α is the union of n_α compact sets $\Omega^{\alpha\beta}$ such that each $f|_{\Omega^{\alpha\beta}}$ is mixing. These $\Omega^{\alpha\beta}$ are disjoint and cyclically permuted by iterations of f. This splitting of the nonwandering set as described is unique. The sets Ω_α are called* basic sets*.*

The proof of this fact is due to Smale in [7]. The properties of these basic sets above allow for the encoding and analysis of dynamics on these spaces in terms of subshifts of finite type.

Definition 7.3. For $x \in \Omega_\alpha$ and $\delta > 0$, the *closed stable and unstable sets of size* δ are defined as

$$W_{\delta,s}(x) = \{y \in \Omega^\alpha : d(f^n(x), f^n(y)) \leq \delta \text{ for all } n \geq 0\}$$
$$W_{\delta,u}(x) = \{y \in \Omega^\alpha : d(f^{-n}(x), f^{-n}(y)) \leq \delta \text{ for all } n \geq 0\}.$$

For small $\epsilon > 0$, there is some $\delta > 0$ such that $W_{\epsilon,s}(x) \cap W_{\epsilon,u}(y)$ is a single point whenever $x, y \in \Omega_\alpha$ and $d(x, y) \leq \delta$, which we denote $[x, y]$. This point is contained in Ω^α and the map

$$[\cdot, \cdot] : \{(x, y) \in \Omega^\alpha \times \Omega^\alpha : d(x, y) \leq \epsilon\} \to \Omega^\alpha$$

is continuous. The proof of this may be found in [2].

Definition 7.4. For $\delta > 0$ small enough and $x \in \Omega(f)$, there are sets $C \subset W_{\delta,s}(x)$ which is the closure of its interior in $W_{2\delta,s}(x)$ and $D \subset W_{\delta,u}(x)$ which is the closure of its interior in $W_{2\delta,u}(x)$. Then $R = [C, D]$ is the closure of its interior in $\Omega(f)$ and is called a *rectangle*.

A *Markov partition* is a finite cover of $\Omega(f)$ by rectangles $\{R_1, \dots, R_k\}$ such that the following hold. First, the partition must satisfy $\text{int}R_i \cap \text{int}R_j = \emptyset$ if $i \neq j$. Second, if $x \in \text{int}R_i \cap \text{int}f^{-1}R_j$, then

$$f[C_i, x] \supset [C_j, f(x)] \text{ and } f[x, D_i] \subset [f(x), D_j],$$

where $R_i = [C_i, D_i]$ for all $1 \leq i \leq k$.

The Markov partition is also defined with the boundaries

$$\partial^s = \bigcup_i \partial^s R_i \text{ and } \partial^u = \bigcup_i \partial^u R_i,$$

where the boundaries of the rectangles are

$$\partial^s R_i = [\partial C_i, D_i] \text{ and } \partial^u R = [C_i, \partial D_i].$$

These have the property that

$$f\partial^s \subset \partial^s \text{ and } f^{-1}\partial^u \subset \partial^u. \tag{7.5}$$

A proof in [2] shows that for any $\epsilon > 0$ there exists a Markov partition with $\text{diam}\{R_1, \dots, R_k\} \leq \epsilon$.

This set constitutes a $V = \{R_1, \dots, R_k\}$. Subsequently, the matrix defined by

$$t_{R_i R_j} = \begin{cases} 1 & \text{if } \text{int}R_i \cap \text{int}f^{-1}R_j \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

specifies a subshift of finite type as described in 4.6 with the shift map $\tau = \tau^1$. This subshift is called the *symbolic dynamics of the system.*

Theorem 7.6. *If $\xi = (\xi_n)_{n\in\mathbb{Z}} \in \Omega$ for the lattice system described by the subshift of finite type above, then $\pi(\xi) = \cap_{n\in\mathbb{Z}} f^{-n}\xi_n$ is a single point in $\Omega(f)$. The map $\pi : \Omega \to X$ obeys the following properties.*

1. *The map π is continuous and surjective*
2. *$\pi \circ \tau = f \circ \pi$*
3. *The map π^{-1} is unique on the residual set $\Omega / \cup_{n\in\mathbb{Z}} f^n(\partial^u \cup \partial^s)$.*
4. *If f is transitive, then the subshift is transitive. If f is mixing, then the subshift is mixing.*

The proof of these properties may be found in section 4 of [1]. This map then encodes each point in the nonwandering set by the rectangles it is mapped to in the Markov partition. The uniqueness or bijectivity of this representation only fails on the boundaries as in (3). From these properties of the map, this allows for the expression of thermodynamic quantities and equilibrium states on the diffeomorphism in terms of its corresponding subshift.

Theorem 7.7. *If f is transitive on $\Omega(f)$ and $A \in \mathcal{C}(\Omega(f))$, then the pressure of A on the nonwandering set is*

$$P_f(A) = P_\tau(A \circ \pi).$$

Further, if A is holder continuous, then there is a unique equilibrium state $\rho_A = \pi\rho$ on $\Omega(f)$, where ρ is the unique equilibrium state of $A \circ \pi$ on the subshift Ω as discussed in 4.11. The map π from (Ω, ρ) to $(\Omega(f), \rho_A)$ is an isomorphism of dynamical systems.

The details of this proof may be found in [6]. Due to spectral decomposition, the case for the nonwandering set may be reduced to the transitive case. The holder continuous case allows for the function's representation on the subshift to be $A \circ \pi \in \mathscr{B}_1$.

The application of results on subshifts allows for the uniqueness of equilibrium measures, and the results for the general $\mathcal{C}$ case hold by the density of holder continuous functions in the space of continuous functions. As π is surjective, this implies that $P_f(A) \leq P_\tau(A \circ \pi)$. Then, for the unique equilibrium measure ρ for $A \circ \pi$ on the subshift Ω, by the ergodicity of the unique equilibrium measure and (7.5), the boundaries where π is not bijective are a set of measure zero. Hence, $\pi : (\Omega, \rho) \to (\Omega(f), \pi\rho)$ where $\pi\rho(A) = \rho(A \circ \pi)$ is an isomorphism of dynamical systems. This implies $h_f(\pi\rho) = h_\tau(\rho)$, so equivalence of pressure is established by

$$P_f(A) \geq h_f(\pi\rho) + (\pi\rho)(A) = h_\tau(\rho) + \rho(A \circ \pi) = P_\tau(A \circ \pi).$$

Uniqueness of this equilibrium state follows from the fact that any equilibrium state σ for A on $\Omega(f)$ has a corresponding invariant state μ on Ω where $P_\tau(A \circ \pi) \geq P_f(A)$, which makes μ an an equilibrium state on Ω so $\mu = \rho$ and $\sigma = \pi\rho$.

Now, this completes an introduction to the concepts of the thermodynamic formalism, as well as an example of the essential relationship between the thermodynamics of lattice systems and those on compact metric spaces in general. From this, there are a few more notable results and extensions of the formalism. First are results for the analyiticity of pressure for one dimensional shifts and their corresponding axiom A diffeomorphisms, which may be found in [6]. This implies a smooth transition of behavior between equilibrium states with different energetic properties, which describes an absence of phase transitions for these systems.

Beyond results for diffeomorphisms, an extension of the thermodynamic formalism to axiom A flows due to Policott [5] produces further applications. This leads to results such as estimates on the number and length of periodic orbits for axiom A flows, such as geodesic flows on surfaces of negative curvature, which may be found due to Lalley in [3].

Acknowledgments

I would like to thank my mentor, DeVon Ingram for introducing me to a wide variety of mathematical topics, including ergodic theory and thermodynamics. His guidance and enthusiasm in suggesting reading material and explaining various concepts has been immensely helpful and informative. I would like to thank Professor Peter May for running this amazing REU program, which provides these great opportunities for undergraduate students.

I would also like to thank Dr. Tracy Bibelnieks, who, as the director of the UMTYMP mathematics program in Duluth, Minnesota, while I attended from 2014 to 2020, inspired me to pursue my interest in mathematics. Without her instruction and passion for teaching, I would not be in this program today.

References

[1] Bowen, Rufus. "Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms." Lecture Notes in Mathematics, 470, Springer, 1975. https://doi.org/10.1007/bfb0081279.

[2] Bowen, Rufus. "Markov Partitions for Axiom a Diffeomorphisms." American Journal of Mathematics, vol. 92, no. 3, 1970, p. 725., https://doi.org/10.2307/2373370.
[3] Lalley, S.P. "Distribution of Periodic Orbits of Symbolic and Axiom a Flows." Advances in Applied Mathematics, vol. 8, no. 2, 1987, pp. 154–193., https://doi.org/10.1016/0196-8858(87)90012-1.
[4] Mane Ricardo, and Silvio Levy. Ergodic Theory and Differentiable Dynamics. Springer, 1987.
[5] Pollicott, Mark. "Symbolic Dynamics for Smale Flows." American Journal of Mathematics, vol. 109, no. 1, 1987, p. 183., https://doi.org/10.2307/2374558.
[6] Ruelle, David. Thermodynamic Formalism: The Mathematical Structures of Classical Equilibrium Statistical Mechanics. Cambridge University Press, 2004.
[7] Smale, S. "Differentiable Dynamical Systems." Bulletin of the American Mathematical Society, vol. 73, no. 6, Nov. 1967, pp. 747–817., https://doi.org/10.1090/S0002-9904-1967-11798-1.
[8] Zinsmeister, Michel. Thermodynamic Formalism and Holomorphic Dynamical Systems. American Mathematical Society, 2000.