NONEQUILIBRIUM THERMODYNAMICS AS A SYMPLECTO-CONTACT REDUCTION AND RELATIVE INFORMATION ENTROPY

JIN-WOOK LIM AND YONG-GEUN OH

Abstract. Both statistical phase space (SPS), which is $\Gamma = T^* \mathbb{R}^{3N}$ of $N$-body particle system, and kinetic theory phase space (KTPS), which is the cotangent bundle $T^* \mathcal{P}(\Gamma)$ of the probability space $\mathcal{P}(\Gamma)$ thereon, carry canonical symplectic structures. Starting from this first principle, we provide a canonical derivation of thermodynamic phase space (TPS) of nonequilibrium thermodynamics as a contact manifold. Regarding the collective observation of observables in SPS, as a moment map defined on KTPS, we apply the Marsden-Weinstein reduction and obtain a mesoscopic phase space in between KTPS and TPS as a (infinite dimensional) symplectic fibration. We then show that the $J$-reduction of the relative information entropy (aka Kullback-Leibler divergence) defines a generating function that provides a covariant construction of thermodynamic equilibrium as a Legendrian submanifold. This Legendrian submanifold is not necessarily 0-holonomic (or graph-like). We interpret the Maxwell construction as the procedure of finding a continuous, not necessarily differentiable, thermodynamic potential and explain the associated phase transition by identifying the procedure with that of finding a graph selector in symplecto-contact geometry and in the Aubry-Mather theory of dynamical system.

Contents

1. Introduction  2
1.1. Relative information entropy and thermodynamic entropy  3
1.2. Action of diffeomorphism group  4
1.3. Covariant construction of thermodynamic equilibrium  5

Date: September 20, 2022; Revised on May 13, 2023.

2010 Mathematics Subject Classification. Primary 80A10; Secondary 53D42, 53D20.

Key words and phrases. kinetic theory phase space (KTPS), statistical phase space (SPS), thermodynamic phase space (TPS), relative information entropy, (local) observables, observations, thermodynamic equilibrium, Maxwell construction; cotangent bundle, moment map, Marsden-Weinstein reduction, reduced space, Legendrian submanifold, generating function, Legendrian graph selector.

This work is supported by the IBS project # IBS-R003-D1.

arXiv:2209.10660v4 [math-ph]  16 Jun 2023
1. Introduction

Starting from Caratheodory [Car09] and Hermann [Her73], contact geometry is utilized in the geometric formulation of the thermodynamics, especially of equilibrium thermodynamical process. It has been observed that the state of thermodynamic equilibrium can be interpreted as a Legendrian submanifold in the *thermodynamic phase space* (TPS) equipped with a contact form. The expression of the contact form is not unique in that it depends on the choice of coordinates. In this regard, more fundamental geometric structure is not the contact form but the contact structure. We call this contact structure the *thermodynamic contact structure*.

In fact, TPS exists only abstractly which has been described through coordinates in the literature. One of the main goals of the present paper is to derive TPS as a contact manifold by a canonical and covariant reduction process starting from the *kinetic theory phase space* (KTPS), which is the cotangent bundle

\[ T^*P(\Gamma) \]
of the space $\mathcal{P}(\Gamma)$ of probability distributions of the statistical phase space (SPS) $\Gamma$ which is the cotangent bundle $T^*\mathbb{R}^{3N}$ of $\mathbb{R}^{3N}$ (or its open subset). The latter is nothing but the $N$ body particle system’s phase space.

Such kind of an effort has been made by Grmela [Grm14]. This is manifest in the following phrase he made in [Grm14, p.1666]:

“Again, the main difficulty on this route is in the reduction process. The question that arises on both routes is the following: how shall we make the reductions in order that the essential features of the mechanics put into the original microscopic formulation is kept intact and the compatibility with thermodynamics of both microscopic and mesoscopic formulations is guaranteed.”

One of the main results of the present paper is to derive this reduction process described in this quoted phrase as a symplecto-contact reduction in the symplecto-contact Hamiltonian geometry focusing on the conceptual framework without delving into the functional analytical issues arising from dealing with the infinite dimensional spaces such as $\mathcal{P}(\Gamma)$, $T^*\mathcal{P}(\Gamma)$ and others. For example, we will ignore the fact that the canonical symplectic form of $T^*\mathcal{P}(\Gamma)$ is only a weak symplectic form which is weakly nondegenerate. (See [AM78] for its definition.)

The canonical reduction process in symplectic geometry is introduced by Marsden and Weinstein [MWr74] when there is a Hamiltonian $G$ action by a Lie group $G$ on a symplectic manifold $(M, \omega)$. The notion of moment map is crucial in this reduction process that is also introduced therein as a $G$-equivariant map

$$J : M \to g^*$$

where $g^*$ is the dual of the Lie algebra $g = \text{Lie}(G)$ equipped with the coadjoint action.

We apply the reduction process to the infinite dimensional symplectic manifold $T^*\mathcal{P}(\Gamma)$ under the canonical symplectic action reduced by certain subgroup of symplectomorphism group $\text{Symp}(T^*\mathcal{P}(\Gamma))$ that canonically arise from the given system of observables on the particle phase space, SPS.

1.1. Relative information entropy and thermodynamic entropy. We remark that the particle phase space $\Gamma := T^*\mathbb{R}^K$, which we call SPS, of the gas (i.e., many particle systems) carries the canonical Lebesgue measure $\nu_0 := d\Gamma$ as the reference measure. Here $K = 3N$ where $N$ is the number of particles. A positive density $D$ on the space $\Gamma$ can be expressed as

$$D = f \, d\Gamma, \quad \text{for} \quad f \geq 0 \quad \& \quad \int_\Gamma f \, d\Gamma > 0$$

where the function $f$ is the Radon-Nikodym derivative

$$f = \frac{\partial D}{\partial \nu_0}.$$  

We denote by $\mathcal{D}^+(\Gamma)$ the set of positive densities.

Our reduction process can be viewed as a statistical mechanics derivation of thermodynamics starting from the kinetic approach of the study of many particle systems such as gas which is going back to Bernoulli and Clausius. It was Maxwell [Max07] who first laid down the true base thereof. (See [Vil02, Section 1.1] for a nice historical introduction of the kinetic approach.) This kinetic approach makes appearance of cotangent bundle of $\mathcal{P}(\Gamma)$ in the story natural.
Another natural source of motivation regarding $\mathcal{P}(\Gamma)$ as the starting point arises from the relationship between Shannon’s information entropy and the thermodynamic entropy. The Shannon entropy is defined for a continuous observable (or a random variable) with a probability distribution $\rho$ on $\Gamma$, i.e., as a function on $\mathcal{P}(\Gamma)$, and the thermodynamic entropy can be canonically derived through some ‘reduction’ process. (See [Jay57], [MNSS90] or Section 8 for such a derivation.) Information entropy is defined for probability measures, i.e., positive densities of mass 1. (See Appendix A for a brief recollection of basic definitions from Information theory.)

**Definition 1.1 (Information measure).** Let $\rho \in \mathcal{P}(\Gamma)$ be a probability density. We define the relative information measure, denoted by $I = I(\rho)$, of $\rho$ to be

$$I(\rho) = -\log \left( \frac{\partial \rho}{\partial \nu_0} \right)$$

for the Radon-Nikodym derivative $\frac{\partial \rho}{\partial \nu_0}$. The (relative) information entropy is the function $S : \mathcal{P}(\Gamma) \to \mathbb{R}$

$$S(\rho) := \int_{\Gamma} I(\rho) \rho = D_{KL}(\rho \| \nu_0)$$

**Remark 1.2.** We refer to [KL51], [Mac03] for the general properties of Kullback-Leibler divergence (aka relative information entropy) $D_{KL}(\rho \| \nu_0)$. We will not use this term beyond the above definition except when we want to emphasize ‘being not Shannon’s original information entropy’ and that $S$ is preserved under the action by $\nu_0$-measure preserving diffeomorphisms. This invariance will be of fundamental importance in our derivation of nonequilibrium thermodynamics as a symplecto-contact reduction of statistical mechanics later.

### 1.2. Action of diffeomorphism group.

Let $\text{Diff}(\Gamma)$ be the diffeomorphism group of SPS $\Gamma = T^*\mathbb{R}^{3N}$. We equip $\Gamma$ with the canonical symplectic form $\omega_\Gamma = \Omega_0$. Then the Lebesque measure $\nu_0$ coincides with the Liouville measure of the symplectic manifold $(\Gamma, \Omega_0)$ in that $\nu_0$ is the integration measure of the Liouville volume form $\frac{1}{(3N)!} \omega_\Gamma^{3N}$. The (extended) entropy function canonically arise from the probability density, independently of the choice of other observables, and is invariant under the action of the subgroup

$$\tilde{\text{Diff}}(\Gamma) \subset \text{Symp}(T^*\mathcal{P}(\Gamma))$$

on KTPS, which consists of the symplectomorphisms on $T^*\mathcal{P}(\Gamma)$ naturally lifted from the action of $\text{Diff}(\Gamma)$ on the base $\mathcal{P}(\Gamma)$. It is given by the pushforward operation of the microscopic group action thereof on SPS. We will also consider the microscopic action of the subgroup $\text{Symp}(\Gamma, \Omega_0) \subset \text{Diff}(\Gamma)$, which is the subgroup consisting of symplectic diffeomorphisms of the symplectic manifold $(\Gamma, \Omega_0)$, and has its lifting of

$$\tilde{\text{Symp}}(\Gamma, \Omega_0) \subset \tilde{\text{Diff}}(\Gamma)$$

acted upon $T^*\mathcal{P}(\Gamma)$. We will derive some consequences arising from the restriction of the canonical action of $\text{Symp}(T^*\mathcal{P}(\Gamma))$ on $T^*\mathcal{P}(\Gamma)$ to that of $\text{Symp}(\Gamma, \Omega_0)$, especially from the measure-preserving property of the action of $\text{Symp}(\Gamma, \Omega_0)$ on $\Gamma$. 
CONTACT REDUCTION

In the statistical mechanics derivation of thermodynamics, points in thermodynamic phase space represent macroscopic observations of the relevant microscopic system, which we will denote by $\mathcal{T}$, through the averaging process. In the kinetic theory a state is represented by a distribution density $D$ (say, in Boltzmann’s study of the kinetic theory). See Definition 4.1 below for the precise definition of what we call the observation.

Therefore we may regard the macroscopic observation or the measurement in the experiment of the given statistical system as a reduction process that reduces the infinite dimensional symmetry of measure-preserving diffeomorphisms via Nöther’s principle. Our first motto is

“Observation is a moment map.”

The usual TPS is then the outcome of some kind of reduction of KTPS, where the relative information entropy function universally arises as a reduced function on any thermodynamic phase space, i.e., in any thermodynamic model: It does not carry its conjugate partner in the nonequilibrium thermodynamic space TPS, unlike other observables. This will be the source of odd dimensionality of TPS. Furthermore the temperature $T$, which is usually regarded as the conjugate of entropy, does not appear in SPS until the state reaches a thermodynamic equilibrium.

A thermodynamic equilibrium is defined to be the state that realizes the maximum of the entropy in a given system. Traditionally in the thermodynamics literature, the equilibrium is identified by first solving a constrained extremal problem arising from the given observation data and the obvious constraint

$$\int \rho = 1$$

inside the space $\mathcal{D}^+(D)$ of positive densities, and then carrying out the macroscopic thermodynamic analysis. To solve the first problem, the method of Lagrange multipliers has been utilized in the thermodynamic literature (e.g. [Mru78]) in terms of preferred variables. We will geometrize and globalize this constrained extremal problem as a part of our reduction process and solve the extremal problem globally and covariantly without using coordinates. In this procedure, we perform two constructions in order:

1. **Step 1:** We apply Marsden-Weinstein reduction for each collective observation of the given statistical system, denoted by $\mathcal{T}$, and obtain an intermediate reduced space $\mathcal{M}^\mathcal{T}$ as a symplectic fibration over the observation data.

2. **Step 2:** We employ the method of generating functions in symplecto-contact geometry by taking the reduced function, $S^\mathcal{T}_{\text{red}}$, as the global canonical generating function of the thermodynamic equilibrium state. (See Appendix B for the definition.)

More detailed explanation of these two operations are now in order.

1.3. **Covariant construction of thermodynamic equilibrium.** We start with what we mean by a (microscopic) statistical system $\mathcal{T}$. We call a collection of observables, i.e., of functions

$$\mathcal{T} = \{F_1, \ldots, F_n\}; \quad F_i : \Gamma = T^*\mathcal{P}(\Gamma) \to \mathbb{R}$$

a statistical system on SPS. We consider the symmetry group consisting of the symplectic diffeomorphisms of $T^*\mathcal{P}(\Gamma)$ induced from the microscopic action of the
subgroup $\text{Diff}(\Gamma, \nu_0) \subset \text{Diff}(\Gamma)$ on the base $\Gamma$: The subgroup $\text{Diff}(\Gamma, \nu_0)$ consists of $\nu_0$-measure preserving diffeomorphisms, e.g., symplectic diffeomorphisms of $(\Gamma, \Omega_0)$.

**Remark 1.3.** Kapranov [Kap] looked at the aforementioned statistical system in the point of view of thermodynamics *with several Hamiltonians*. He also related them to the toric moment map for the case of commuting Hamiltonians. Here we single out the Hamiltonian as the observable that drives the dynamics and other ones as the collection that define the given statistical system such as volume, mole number and other observables. Our consideration of statistical system is a noncommutative extension of Kapranov's treatment of commuting Hamiltonians. However our first motto “Observation is a moment map.” is in the similar spirit as that of [Kap].

We find it convenient for the further discussion to introduce the following intuitive terminology.

**Definition 1.4 (Observation (aka Measurement)).** For given local observable $F$ on SPS, we define its *observation*, denote by $\mathcal{O}_F$, to be the function of taking the macroscopic average

$$\mathcal{O}_F : \mathcal{P}(\Gamma) \to \mathbb{R}$$

given by $\mathcal{O}_F(\rho) = \int_{\Gamma} F \rho$. We will also use the term *measurement* equally with the same meaning as the term *observation*, depending on the given circumstances.

(Compare these macroscopic average functions with the entropy function (1.1) for which $S$ would be the observable given by

$$-\log \frac{\partial \rho}{\partial \nu_0}$$

*which depends solely on the distribution* $\rho$ *not involving any other observable.*)

The aforementioned action of $\text{Diff}(\Gamma, \nu_0)$ preserves both the relative information entropy (1.2) and the observations of observables $\mathcal{O}_F = \{F_1, \ldots, F_n\}$ of the given statistical system which enable us to apply the standard Marsden-Weinstein reduction: Their collective observations denoted by

$$\mathcal{O}_\mathcal{F} = \{\mathcal{O}_{F_1}, \ldots, \mathcal{O}_{F_n}\} : \mathcal{P}(\Gamma) \to \mathbb{R}^n$$

canonicaly lift to the functions

$$\widetilde{\mathcal{O}}_\mathcal{F} = \{\widetilde{\mathcal{O}}_{F_1}, \ldots, \widetilde{\mathcal{O}}_{F_n}\} : T^*\mathcal{P}(\Gamma) \to \mathbb{R}^n.$$ 

We then formally extend the symplectic reduction construction to the infinite dimensional *cotangent bundle* $T^*\mathcal{P}(\Gamma)$ of the set $\mathcal{P}(\Gamma)$ of probability densities on $\Gamma$, as usually done in the Hamiltonian formalism of the field theory. See Remark [2.7] for the expression of the canonical symplectic structure on $T^*\mathcal{P}(\Gamma)$. (We highlight the fact that we consider a field theory on the *affine space* of $\mathcal{P}(\Gamma)$, not on a vector space on which a usual Hamiltonian formalism of a field theory takes place.) Ultimately this will be responsible for the odd dimensionality of TPS that we construct.

We will first construct a (infinite dimensional) *symplectic fiber bundle*

$$\mathcal{M}^{\mathcal{F}} \to B_{\mathcal{F}}^2 \subset \mathbb{R}^n$$

(over some open subset $B_{\mathcal{F}}^2$) after the Marsden-Weinstein reduction applied to each fiber of which the original system $\mathcal{F} = \{F_1, \ldots, F_n\}$ induce a collective kinematics.
We will regard this reduced space as the *mesoscopic phase space* in between KTPS and TPS on which the canonical reduction of the function $S$ is defined.

In the literature of thermodynamics, the (equilibrium) thermodynamics is attempted to be derived from statistical mechanics in the point of view of Shannon’s information theory, from the first principle of statistics and the information theory. For example in [Jay57], [Mru78], [MNSS90], the authors provide a derivation of the thermodynamics equilibrium in terms of information entropy and contact geometry. However the description is not satisfactory enough in that in their derivation it is assumed that the associated Legendrian submanifold is *graph-like* relative to the given variables, i.e., it is assumed that the Legendrian submanifold $R_{\mathcal{F};S}$ admits a single-valued potential function $f : B_{\mathcal{F};S} \to \mathbb{R}$ such that

$$R_{\mathcal{F};S} = \{ (\mu, df(\mu), f(\mu)) \mid \mu \in B_{\mathcal{F};S} \}.$$

**Definition 1.5** (*0-holonomic*). Let $B$ be a smooth (possibly infinite dimensional) manifold. We say a Legendrian submanifold $R \subset J^1B$ is *0-holonomic* if it is graph-like, i.e., if it is the image of the 1-jet map $j^1f : B \to J^1B$ of a differentiable function $f : B \to \mathbb{R}$.

**Remark 1.6.** Our usage of the term 0-holonomic in this definition can be regarded as the $C^0$-version of the standard notion of holonomic section $\phi : V \to J^rX$ which is the $k$-jet of some $C^k$-section $f : V \to X$ for the given $C^\infty$-fibration. (See [Gro86, Section 1.1.1] for the definition.) In our current situation, we consider the 1-jet bundle $J^1B \to B$ and regard $R \subset J^1B$ as a (possible partially defined) multi-valued section $B \to J^1B$. We interpret the condition of a Legendrian submanifold $R$ being graph-like as the $C^0$-homonmicity in that $R$ becomes the image of global section $\alpha$ of the projection $J^1B \to B$ given by $j^1f := (df, f) : B \to J^1B$ for a single-valued $C^1$-function $f : B \to \mathbb{R}$ so that the map $j^1f$ is $C^0$.

This terminology 0-holonomic being introduced, we will just simply call holonomic instead throughout the paper when $R$ is graph-like.

The following globalization of the description of the thermodynamic equilibrium complements that of described in [Jay57], [Mru78], [MNSS90] by allowing non-holonomic equilibrium states in the description thereof, which is our second motto:

“Relative information entropy is the generating function of thermodynamic equilibria.”

Now we can turn the formal restatement briefly mentioned at the end of the previous subsection into the following precise mathematical statement. (See Section B for the proofs.)

**Theorem 1.7** (Universality of information entropy). Let $\mathcal{F} = \{ F_1, \cdots, F_n \}$ be any observable system on SPS.

1. The collective observation

$$\mathcal{O}_\mathcal{F} = (\mathcal{O}_{F_1}, \cdots, \mathcal{O}_{F_n})$$

is the moment map whose reduced spaces define a symplectic fiber bundle $M^\mathcal{F}$ over an open subset of $\mathbb{R}^n$, which consists of the set of regular values of the collective observation map $\mathcal{O}_\mathcal{F}$.

2. The relative entropy function $S : T^*P(\Gamma) \to \mathbb{R}$ induces a generating function $S^\text{red} : M^\mathcal{F} \to \mathbb{R}$ that generates the thermodynamic equilibrium $R_{\mathcal{F};S}$ in $J^1g^\mathcal{F}_\mathcal{O} \cong J^1\mathbb{R}^n$ associated to the system $\mathcal{F}$.
We call the function \( S^{\text{red}}_{\mathcal{F}} : \mathcal{M}^\mathcal{F} \rightarrow \mathbb{R} \) the \( \mathcal{F} \)-reduced information entropy.

Recall that as a statistical observable, our relative information entropy \( S \) is a universal observable in SPS (as a measure space equipped with the Liouville measure), corresponds to the Radon-Nikodym derivative \( F = -\log \frac{\partial \rho}{\partial \nu} \). Because of the constraint \( \int_F \rho = 1 \), it will not carry any thermodynamic conjugate partner in our reduction process. This is responsible for the odd dimensionality of TPS: the \( \mathcal{F} \)-reduced entropy is the generating function of a thermodynamic equilibrium, a Legendrian submanifold in the one-jet bundle.

1.4. Volume, a non-local observable. After these we revisit the previously well-studied thermodynamical models through the eyes of our reduction. We first provide how to derive the volume variable \( V \) as one of the variables of TPS from the first principle of SPS and KTPS.

In this regard, we would like to highlight that the volume is not a local observable in statistical mechanics. It is not a dynamical variable in the level of SPS but a system variable. It appears to describe the state of particles confined is an expandable container. In the information theoretic point of view, the volume is the parameter describing the state of the container, which is a state determined by another event independent of the state of particles contained in the container.

Because of this, we need to modify the definitions of SPS and KTPS by incorporating this difference in our kinetic theory framework, especially in our consideration of the probability distributions and relevant information entropy: The relevant KTPS will be now

\[
T^*\mathcal{P}(\Gamma \times \mathbb{R}_+).
\]

See Section 9 for the details. The volume variable will not appear in SPS but it will appear as a phase space variable in the macroscopic TPS level. It then carries a conjugate partner, which is nothing but the pressure.

1.5. Maxwell construction and phase transition. The resulting equilibrium state arising in Theorem 1.7 is not necessarily 0-holonomic in terms of given preferred variables. We will illustrate by van der Waals model that non-holonomicity indeed occurs in a physical model which is responsible for the phase transition.

Then in a thermodynamic equilibrium or nearby, say in the system \( \mathcal{F} \) of ideal gas or van der Waals gas equation, the relevant state variables are given by

\[
(U, S, T, P, V)
\]

where \( U \) is energy, \( S \) the thermodynamic entropy, \( T \) temperature, \( P \) pressure and \( V \) is the volume. The first law of equilibrium thermodynamics is expressed as

\[
dS = \frac{1}{T}dU + \frac{P}{T}dV
\]

in the entropy representation. In our construction, the relevant thermodynamic equilibrium is given by the Legendrian submanifold \( R \) generated by the aforementioned (reduced) relative entropy function \( S^{\text{red}}_{\mathcal{F}} : \mathcal{M}^\mathcal{F} \rightarrow \mathbb{R} \) where

\[
\begin{array}{ccc}
\mathcal{M}^\mathcal{F} & \xrightarrow{S^{\text{red}}_{\mathcal{F}}} & \mathbb{R} \\
\downarrow & & \downarrow \\
\mathbb{R}^2 & & \\
\end{array}
\]
is a (infinite dimensional) fibration.

Since the aforementioned thermodynamic equilibrium may not be globally holonomic in the preferred variables, which are $P, T$ for the van der Waals model, the equilibrium equation cannot be globally solved in terms of the preferred variables in a direct way by simple quadratures. An effort of correcting this deficiency is precisely the celebrated Maxwell construction [Max67] in the equilibrium thermodynamics.

It turns out that the Maxwell construction is the same process as that of constructing the so called graph selector out of non-holonomic Lagrangian (resp. Legendrian) submanifold in symplectic (resp. in contact) geometry and in dynamical systems. (See [Oh97], [Arn10], [BOdS12] and [AOOdS18] for example.) In Section [11] we relate the construction of a graph selector with Gibbs free energy $G = U + PV - TS$ as its thermodynamic potential to the relevant Maxwell construction and the phase transition of the isotherms with respect to the pressure-volume diagram (abbreviated as the $PV$-diagram). The outcome will be a construction of thermodynamic potential that is Lipschitz continuous but not differentiable across the Maxwell pressure at which Maxwell equal area law holds. (See Theorem [11.4].)

A brief organization of the paper is in order. Section 2 introduces the spaces SPS and KTPS and Section 3 explains the underlying symmetry group of KTPS. Section 4–7 explains the Marsden-Weinstein reduction of KTPS for the system of local observables. Section 4 explains what we mean by local observables and observations and their invariance property on SPS. Section 5 carries out the Marsden-Weinstein connection following our first motto and construct the $F$-reduced KTPS $M^F$. Section 6 constructs the $F$-reduced information entropy $S_{red}$ and shows that it generates a thermodynamic equilibrium in the sense of the theory of generating functions (or of the Morse family) See Appendix [B]. Then Section 7 derives the simplest Boltzmann equation without collision terms associated to the total energy Hamiltonian $H$ on SPS as the internal mesoscopic dynamics of the obtained thermodynamic equilibrium.

Then in Section 9, we explain a modification of the reduction process for the system involving non-local observables of the volume variable of the gas models. In Section 10 we derive the gas equations for the ideal gas. In Section 11 we identify the Maxwell construction as the process of selecting a graph-selector in the sense of [Oh97], [Arn10], [Oh16] and interpret the relevant phase transition for the van der Waals model in our framework. Finally in Section 12 we have some discussion on the dynamical perspective of our geometric study mostly of the kinematics of nonequilibrium thermodynamics, and propose some future direction of research. In this regard, we mention the relationship between our study and those given in the recent interesting articles [EGP22a], [EGP22b] by Esen-Grmela-Pavelka in particular.

Acknowledgements: The second named author thanks Sang-Jin Sin for his interest in this work and illuminating discussions on the nature of volume variable $V$ which greatly helps in finding our formulation of the volume, a nonlocal thermodynamic observable, from a similar kind of viewpoint applied to the local observables. We thank the referees for providing many useful comments and suggestions which help us to much improve the presentation of the present article.

Competing interest: There is no competing interest.
Notations and Vocabulary

- SPS: statistical phase space $(\Gamma, \omega_\Gamma) = (T^*\mathbb{R}^{3N}, \Omega_0)$,
- $\mathcal{P}(\Gamma)$: the space of probability distributions,
- $I = I(\rho) = D_{KL}(\rho\|\nu_0)$: Relative information measure (aka Kullback-Leibler divergence),
- $S(P(\Gamma)) \rightarrow \mathbb{R}$: relative information entropy,
- $KTPS$: kinetic theory phase space $T^*(P(\Gamma))$,
- $\mathcal{F}_F$: observable system or statistical system,
- $S_{\text{red}}: \mathcal{F}_F \rightarrow \mathbb{R}$: $\mathcal{F}$-reduced entropy function,
- TPS: thermodynamic phase space $J^1B$ for some open subset $B \subset \mathbb{R}^n$.

2. Symplectic structure of the kinetic theory phase space

Let $(\mathbb{R}^{2K} = T^*\mathbb{R}^K, \Omega_0)$ be the Hamiltonian phase space equipped with the standard symplectic form

$$\omega_\Gamma = \sum_{i=1}^{K} dx_i \wedge dy_i.$$  

Our main interest will be a particle phase space in gas dynamics, i.e., the classical statistical phase space (SPS) of many body systems where $K = 3N$ with $N$ being the number of particles in the system.

Remark 2.1. In our derivation, we will ignore the effect of collisions of the particles in the story. To incorporate the effect of collisions as in the collisional kinetic theory [Vil02], we should consider the statistical phase space such as

$$T^*(\mathbb{R}^{3N} \setminus \Delta_n) =: \Gamma$$

where $\Delta_n$ is the big diagonal

$$\Delta_N = \{ (q_1, \ldots, q_N) \in \mathbb{R}^{3N} \mid q_i \in \mathbb{R}^3, q_i = q_j, i \neq j \}.$$  

Remark 2.2. The whole discussion can be applied to any smooth manifold with a reference smooth measure, e.g., to any symplectic manifold equipped with the Liouville measure on it. However in the present paper, we will restrict ourselves to SPS, except when we consider the nonlocal observable, the volume where we employ a more general configuration space in Section 10.

We regard SPS as a symplectic manifold and denote it by

$$(\Gamma, \omega_\Gamma), \quad \omega_\Gamma = \Omega_0.$$  

We may use $\Omega_0$ and $\omega_\Gamma$ as we feel more appropriately without distinction depending on circumstances. Following the notation of [MNSS90], we denote by $d\Gamma$ the density form of the Lebesgue measure of $\Gamma = \mathbb{R}^{2K}$ which coincides with the Liouville volume form

$$\frac{1}{K!} \Omega_0^K = d\Gamma.$$  

Again we will also denote by $\nu_0$ the associated Liouville measure (or the Lebesgue measure). We use $\nu_0$ and $d\Gamma$ interchangeably as we feel more appropriately without distinction depending on circumstances.
In this section, we apply the theory of symplectic reduction to the infinite dimensional symplectic manifold

$$T^*\mathcal{P}(\Gamma)$$

which is the cotangent bundle of the probability density space $\mathcal{P}(\Gamma)$ as the Hamiltonian phase space.

**Definition 2.3.** We denote by $\mathcal{D}^+(\Gamma)$ the space of nonnegative densities, denoted by $D$, with finite total mass

$$M(D) := \int_\Gamma D < \infty.$$  

We denote by $\mathcal{P}(\Gamma)$ its subset consisting of probability densities, denoted by $\rho$, whose total mass is 1, i.e., $M(\rho) = 1$.

We note that $\mathcal{D}^+(\Gamma)$ is an open subset of a linear space, and hence its tangent space of $\mathcal{D}^+(\Gamma)$ at $D$ can be identified with the set of signed densities which we denote by

$$\mathcal{D}(\Gamma)$$

which is naturally a vector space. On the other hand $\mathcal{P}(\Gamma)$ is a convex hypersurface given by the integral constraint

$$\int \rho = 1.$$  

Because of this, we start our discussion with the cotangent bundle of $\mathcal{D}^+(\Gamma)$.

Thanks to the presence of the Liouville volume form, any density $D$ has the form

$$D = f \, d\Gamma$$

for a nonnegative function $f : \Gamma \to \mathbb{R}^+$ with its $L^1$-norm finite, i.e.,

$$0 \leq \int_\Gamma f \, d\Gamma < \infty$$

provided $D$ has finite mass. We denote by $\mathcal{D}^+(\Gamma)$ the set of nonnegative densities. We denote by $\mathcal{D}(\Gamma)$ the set of densities $D$ with $|D|$ having finite mass where $|D| = D^+ + D^-$ where $D^\pm \geq 0$ are the positive and negative parts of $D$ so that $D = D^+ - D^-$.  

The space $\mathcal{D}^+(\Gamma)$ is a principal $\mathbb{R}_+$ homogeneous space that gives rise to the following nonlinear exact sequence

$$0 \to \mathcal{P}(\Gamma) \to \mathcal{D}^+(\Gamma) \xrightarrow{M} \mathbb{R}_+ \to 0 \quad (2.1)$$

where the last map is given by the function $D \to M(D)$. By taking the logarithm $\log M(D)$, we get

$$0 \to \mathcal{P}(\Gamma) \to \mathcal{D}^+(\Gamma) \xrightarrow{\log M} \mathbb{R} \to 0 \quad (2.2)$$

another exact sequence.

Note that because $\mathcal{D}^+(\Gamma)$ is an open subset of the vector space of densities, the set $\mathcal{D}^+(\Gamma)$ of infinitesimal variation thereof can be canonically identified with the space $\mathcal{D}(\Gamma)$ of $L^1$ densities on $\Gamma$ mentioned above. We adopt physicists' notation

$$\dot{D}$$

to represent a tangent vector or the first variation of $D$ in $\mathcal{D}^+(\Gamma)$. 

Notation 2.4. Likewise we may denote by $\hat{D}(\Gamma)$ the representative vector space $\mathcal{D}(\Gamma)$ of the fiber of the tangent bundle $T\mathcal{D}^+(\Gamma)$. Then we have

$$\hat{D}(\Gamma) \cong T_D \mathcal{D}^+(\Gamma) \cong \mathcal{D}(\Gamma)$$

for all $D \in \mathcal{D}^+(\Gamma)$.

For our purpose of reduction, it is natural to start with the cotangent bundle

$$T^* \mathcal{D}^+(\Gamma) = \bigcup_{D \in \mathcal{D}^+(\Gamma)} \{D\} \times T_D \mathcal{D}^+(\Gamma)$$

equipped with the canonical (weak) symplectic form on it. (See (2.5) below.) We note that the $\mathbb{R}^+$-action on $\mathcal{D}^+(\Gamma)$ is given by $(c, D) \mapsto c D$ which induces the action on $T\mathcal{D}^+(\Gamma)$ given by

$$(c, (D, \dot{D})) \mapsto (c D, dR_c(\dot{D})) :$$

Here we denote by $R_c : \mathcal{D}^+(\Gamma) \to \mathcal{D}^+(\Gamma)$
this multiplication map by a constant $c > 0$. We denote its dual action on $T^* \mathcal{D}^+(\Gamma)$ by

$$(c, (D, \alpha)) \mapsto (c D, (dR_c^{-1})^\ast \alpha)).$$

In the level of tangent spaces, (2.2) also induces the exact sequence

$$0 \to T_{\rho} \mathcal{P}(\Gamma) \to T_{\rho} \mathcal{D}^+(\Gamma) \to T_{\rho} \mathcal{D}^+(\Gamma)/T_{\rho} \mathcal{P}(\Gamma) \to 0. \quad (2.3)$$

By taking the dual sequence, we obtain

$$0 \to \nu^*_{\rho} \mathcal{P}(\Gamma) \to T^*_{\rho} \mathcal{D}^+(\Gamma) \to T^*_{\rho} \mathcal{P}(\Gamma) \to 0$$

where

$$\nu^*_{\rho} \mathcal{P}(\Gamma) := (T_{\rho} \mathcal{P}(\Gamma))^\perp$$

is the conormal space of $T_{\rho} \mathcal{P}(\Gamma)$ which is defined to be the subspace

$$(T_{\rho} \mathcal{P}(\Gamma))^\perp = \{ \eta \in T_{\rho} \mathcal{P}(\Gamma) \mid \eta(\xi) = 0 \forall \xi \in T_{\rho} \mathcal{P}(\Gamma) \} \subset T_{\rho} \mathcal{D}^+(\Gamma).$$

We have a canonical identification $T_{\rho} \mathcal{D}^+(\Gamma) \cong \mathcal{D}(\Gamma)$.

Definition 2.5 (Extended information measure). Let $D \in \mathcal{D}^+(\Gamma)$. We call the function

$$\hat{I}(D) = I(D/M(D)) = -\log \left( \frac{\partial D/M(D)}{\partial v_0} \right) = -\log \left( \frac{\partial D}{\partial v_0} \right) + \log M(D) \quad (2.4)$$

the extended information measure of the distribution $D$.

Definition 2.6 (Kinetic theory phase space (KTPS)). We call the cotangent bundle

$$T^* \mathcal{P}(\Gamma)$$

with its canonical symplectic structure the kinetic theory phase space (KTPS).

Remark 2.7. A precise form of the canonical symplectic structure does not play much role except its general properties. The most convenient way of writing down the (weak) symplectic structure on $T^* \mathcal{P}(\Gamma)$ as in the Hamiltonian formalism of the general field theory may be in terms of the associated Poisson bracket: For given
real-valued functions $F = F(\rho, \beta)$, $G = G(\rho, \beta)$ on $T^*\mathcal{P}(\Gamma)$, the canonical Poisson bracket between $F$ and $G$ are given by

$$\{F, G\}(\rho, \beta) = \int_{\Gamma} \left( \frac{\delta F}{\delta \rho} \frac{\delta G}{\delta \beta} - \frac{\delta G}{\delta \rho} \frac{\delta F}{\delta \beta} \right) \rho \quad (2.5)$$

where $\frac{\delta F}{\delta \rho}$ stands for the variational derivative. (See e.g., [MWS83] for such a formula and the precise meaning of the variational derivative.)

3. Diff$(\Gamma, \nu_0)$-symmetry and its moment map

We will use the following notation systematically.

**Notation 3.1.** For any subgroup $\mathcal{S} \subset \text{Diff}(\mathcal{P}(\Gamma))$, we denote by

$$\widehat{\mathcal{S}} \subset \text{Symp}(T^*\mathcal{P}(\Gamma))$$

the lifted group to $T^*\mathcal{P}(\Gamma)$.

The following is obvious.

**Lemma 3.2.** Any diffeomorphism on $\mathcal{P}(\Gamma)$ canonically lifts to a symplectic diffeomorphism on $T^*\mathcal{P}(\Gamma)$. In particular the induced action of microscopic action of $\text{Diff}(\Gamma)$ on $\mathcal{P}(\Gamma)$ by push-forward is a symplectic action on $T^*\mathcal{P}(\Gamma)$.

**Proof.** Recall that for any diffeomorphism $\psi : N \to N$ of a manifold $N$ (whether it is of finite dimension or not), the formula

$$\bar{\psi}(x, \beta) := (\psi(x), (d_x \psi^{-1})^*(\beta))$$

provides the lifting to $T^*\mathcal{P}(\Gamma)$ which is symplectic.

Furthermore this lifting is linear in fiber and in particular preserves the zero section. This shows that it canonically induces a symplectomorphism on $T^*\mathcal{P}(\Gamma)$. \qed

We denote by

$$\widehat{\text{Diff}}(\Gamma) \subset \text{Symp}(T^*\mathcal{P}(\Gamma))$$

the subgroup consisting of the canonical lifting of the aforementioned microscopic action of $\text{Diff}(\Gamma)$.

Next we describe the moment map of this symplectic action of $\text{Diff}(\Gamma)$ on $T^*\mathcal{P}(\Gamma)$ more explicitly. For this purpose, we note that $\text{Diff}(\Gamma)$ action on $\mathcal{P}(\Gamma)$ is nothing but the restriction of the natural action on the bigger space $\mathcal{D}^+(\Gamma) \supset \mathcal{P}(\Gamma)$: For $\phi \in \text{Diff}(\Gamma)$, it induces an action by the push-forward $(\phi, D) \mapsto \phi_\ast D$ on $\mathcal{D}^+(\Gamma)$ and in turn by the action

$$(\phi, (D, \alpha)) \mapsto (\phi_\ast D, (d\phi^{-1})^\ast \alpha) \quad (3.1)$$

on $T^*\mathcal{D}^+(\Gamma)$. Considering the identity component of $\text{Diff}(\Gamma)$, we can identify its tangent vector with a vector field $X \in \mathcal{X}(\Gamma)$ after a right translation. Denote by

$$\mathcal{J}_{\text{Diff}(\Gamma)} : T^*\mathcal{D}^+(\Gamma) \to \mathcal{X}(\Gamma)^\ast$$

the universal moment map of the action of $\text{Diff}(\Gamma)$ on $T^*\mathcal{D}^+(\Gamma)$. (See [AM78] Theorem 4.2.10.) Then it is determined by its defining equation

$$\langle \beta, X \rangle = \alpha(\hat{X}) \quad (3.2)$$

where $\hat{X}$ is the vector field of $\mathcal{D}^+(\Gamma)$ associated to $X$ under the lifting of the linearized action of $\widehat{\text{Diff}}(\Gamma)$.
Example 3.3. The tangent vector $\hat{X} \in T_D \mathcal{D}^+(\Gamma) \cong \mathcal{D}(\Gamma)$ associated to $X \in \mathfrak{X}(\Gamma)$ has the canonical pairing
\[ \alpha(\hat{X}) = \int_{\Gamma} \langle \alpha, X \rangle \, d\Gamma \] (3.3)
with each element $\alpha \in T_D^* \mathcal{D}^+(\Gamma) \cong \mathcal{D}^*(\Gamma)$, by the canonical pairing between (vector-valued) functions and (vector-valued) distributions on $\Gamma$.

We note that as a subset of $T_{\rho} \mathcal{D}^+(\Gamma)$ each tangent vector at $\rho$ of the Diff(\Gamma)-orbit has the form $L_X \rho$ and so the tangent space of a Diff(\Gamma)-orbit thereat is given by the subspace
\[ \tilde{\mathcal{X}}_0(\Gamma; \rho) = \left\{ \hat{X} \in \mathcal{D}(\Gamma) \mid \int_{\Gamma} \mathcal{L}_X \rho = 0 \right\}. \] (3.4)

We recall that we have the natural isomorphism for the cotangent space,
\[ T^* \mathcal{P}(\Gamma) \cong \mathcal{D}(\Gamma) \]
under the identifications (3.4) and (3.5). Here we regard the mass map $M$ as a linear functional on $\mathcal{D}(\Gamma)$.

Theorem 3.4 (Universal moment map). The moment map of the action of Diff(\Gamma, \nu_0)
\[ \mathcal{J}_{\text{Diff}(\Gamma, \nu_0)} : T^* \mathcal{P}(\Gamma) \to \tilde{\mathcal{X}}_0(\Gamma; \rho)^* \]
is expressed as
\[ \mathcal{J}_{\text{Diff}(\Gamma, \nu_0)}(\rho, \beta) = \left( \beta - \frac{\beta(D)}{M(D)} M \right) |_{\mathcal{X}_0(\Gamma; \rho)} \] (3.6)
under the identifications (3.4) and (3.5). Here we regard the mass map $M$ as a linear functional on $\mathcal{D}(\Gamma)$.

4. Statistical phase space and systems

Denote by $\Gamma$ an open subset of $T^* \mathbb{R}^{3N} \cong \mathbb{R}^{6N}$ the phase space of $N$ particles in $\mathbb{R}^3$, which we will call a statistical phase space (SPS). We equip $T^* \mathbb{R}^K$ with the canonical symplectic form
\[ \omega_{\Gamma} = \sum_{i=1}^{K} dq_i \wedge dp_i, \quad K = 3N. \]
We denote by
\[ \mathfrak{X}(\mathcal{P}(\Gamma)) \]
the set of vector fields on $\mathcal{P}(\Gamma)$. Since any symplectic diffeomorphism satisfies $\phi^* \omega_{\Gamma} = \omega_{\Gamma}$, it also preserves the Liouville measure $\phi^* d\Gamma = d\Gamma$.

We adopt the following terminologies from statistical mechanics.

Definition 4.1 (Observables and observations). (1) Let $\mathcal{O} : \mathcal{P}(\Gamma) \to \mathbb{R}$ be a function on the probability space, which we call an observable. A local observable is an observable $\mathcal{O}_F$ arising from a function $F : \Gamma \to \mathbb{R}$. We call a collection of functions
\[ \mathcal{F} = \{ F_1, \ldots, F_n \} \]
a local observable system.
(2) Each probability distribution $\rho \in \mathcal{P}(\Gamma)$ with respect to which $F$ is an $L^1(\rho)$-function defines the average

$$O_F(\rho) = \langle F \rangle_\rho := \int F \rho. \quad (4.1)$$

(3) We call this macroscopic average of $F$ an *observation* of the local observable $F$ relative to $\rho$.

(4) We define a *collective observation* to be the map $O_F : \mathcal{P}(\Gamma) \to \mathbb{R}^n$ given by

$$O_F = (O_{F_1}, \ldots, O_{F_n}). \quad (4.2)$$

We may also denote the value of $O_{F_i}$ at $\rho$ by the standard notation $O_{F_i}(\rho) = \langle F_i \rangle_\rho$ in statistical mechanics.

We regard the relative information entropy $S$ as an observable which is intrinsic and universal in the sense that $S$ depends only on the probability density $\rho$ independent of other local observables.

**Definition 4.2** (Relative information entropy). We recall the relative information entropy function $D_{KL}(\rho\|\nu_0)$ (aka Kullback-Leibler divergence) is defined to be

$$S(\rho) = -\int \rho \log \frac{\partial \rho}{\partial \nu_0}, \quad \nu_0 = d\Gamma.$$

The following is obvious. However we would like to remark that consideration of relative information entropy $D_{KL}(\rho\|\nu_0)$ in the presence of the background Liouville measure $\nu_0 = d\Gamma$, not Shannon’s original continuous entropy, is crucial to have this invariance property.

**Proposition 4.3.** The (relative) information entropy is invariant under the action of $\nu_0$-measure preserving diffeomorphisms. In particular it is preserved under the action by symplectic diffeomorphisms on SPS.

**Proof.** Consider (relative continuous) entropy,

$$D_{KL}(\rho\|\nu_0) = -\int_\Gamma \rho \log \frac{\partial \rho}{\partial \nu_0}.$$

Recall the action of a diffeomorphism $\phi : \Gamma \to \Gamma$ is given by

$$(\phi, \rho) \mapsto \phi_* \rho$$

and by the definition of the Radon-Nikodym derivative that

$$\frac{\partial (\phi_* \rho)}{\partial \nu_0} = \frac{\partial \rho}{\partial \nu_0} \circ \phi^{-1}.$$

Therefore we compute

$$D_{KL}(\phi_0 \rho\|\nu_0) = -\int_\Gamma \phi_* \rho \log \frac{\partial \phi_* \rho}{\partial \rho} = \int_\Gamma \left( \frac{\partial \rho}{\partial \nu_0} \circ \phi^{-1} \phi_* \rho \right) \log \frac{\partial \rho}{\partial \nu_0} \circ \phi^{-1} = \int_\Gamma \left( \frac{\partial \rho}{\partial \nu_0} \nu_0 \right) \log \frac{\partial \rho}{\partial \nu_0} = D_{KL}(\rho\|\nu_0)$$

where the last equality follows since $\phi_* \nu_0 = \nu_0$. This finishes the proof. \qed
5. Mesoscopic reduction of KTPS: the first motto

In this section, we perform the first stage of thermodynamic reduction by decomposing KTPS over the collective observation data by applying Marsden-Weinstein reduction.

Remark 5.1. It appears that this reduction process is related to the theory of coarse graining in statistical mechanics. Coarse-graining is the process of grouping together of a large number of small entities into some larger size ensemble with important characteristics of small entities intact and then analysing this new system. Any description in between the microscopic level of statistical mechanics and the macroscopic of thermodynamics is called a mesoscopic level. (See [Esp04] for example.)

We start with the following obvious lemma.

Lemma 5.2. Consider any system $\mathcal{F} = \{F_1, \cdots, F_n\}$ of local observables. The collective observation $\{O_{F_i}\}$ is invariant under the action of the group $\text{Diff}(\Gamma, \nu_0)$.

Proof. Again this follows by the chain rule this time more easily using the measure-preserving hypothesis of the action as in the proof of Proposition 4.3. □

Suppose we are given a local observable system $\mathcal{F}$. We consider its collective observation function $\{O_{F_i}\}_{i=1}^n$. We then lift each function $O_{F_i}$ to $\tilde{O}_{F_i}$ to $T^*\mathcal{P}(\Gamma)$ by pull-back, i.e.,

$$\tilde{O}_{F_i}(\rho, \beta) := O_{F_i}(\rho).$$

For the simplicity of notation, we abuse our notation by writing $\tilde{O}_F = O_F$ as long as there is no danger of confusion from now on.

Then the functions $\{O_{F_i}\}_{i=1}^n$ Poisson commute with respect to the canonical symplectic structure on $T^*\mathcal{P}(\Gamma)$ given by (2.5) and so their collective Hamiltonian flows on $T^*\mathcal{P}(\Gamma)$ defines a Hamiltonian action of the abelian group $\mathbb{R}^n$. Each element of these Hamiltonian flows is derived from a $\nu_0$-measure preserving (local) diffeomorphism group symmetries of $\Gamma$.

Definition 5.3. For the given set of observables $\mathcal{F} = \{F_1, \cdots, F_n\}$ on SPS $\Gamma$, we consider

$$\mathfrak{g}_{\mathcal{F}} \subset \mathfrak{X}_{\text{sym}}(T^*\mathcal{P}(\Gamma))$$

the Lie algebra generated by the commuting Hamiltonian vector fields

$$\{X_{O_{F_1}}, \cdots, X_{O_{F_n}}\}$$

on $T^*\mathcal{P}(\Gamma)$.

We recall the inclusions

$$\mathcal{P}(\Gamma) \subset \mathcal{D}^+(\Gamma) \subset \mathcal{D}(\Gamma)$$

where the last space is a vector space, the second is an open subset thereof and the first is a hypersurface of the second. Furthermore the observation function can be linearly extended to $\mathcal{D}(\Gamma)$ by the same formula

$$\int_\Gamma F \, D =: O_F(D), \quad D \in \mathcal{D}(\Gamma).$$
We also have the decomposition
\[ T_\rho D(\Gamma) = T_\rho \mathcal{P}(\Gamma) \oplus \mathbb{R}(\rho) \]
given by the explicit formula
\[ \dot{D} \mapsto \left( \dot{D} - \left( \int_\Gamma \dot{D} \right) \rho, \left( \int_\Gamma \dot{D} \right) \rho \right). \] (5.2)

Now we derive the formula for the associated moment map \( J_{\mathcal{F}} : T^* \mathcal{P}(\Gamma) \to g^*_\mathcal{O}_\mathcal{F} \).

Recall that the assignment \( F \mapsto X_{\mathcal{O}_F} \) defines a map
\[ C^\infty(\Gamma) \to \mathfrak{x}^{\text{symp}}(T^* \mathcal{P}(\Gamma)). \]
This in turn induces the dual map
\[ \mathfrak{x}^{\text{symp}}(T^* \mathcal{P}(\Gamma))^* \to (C^\infty(\Gamma))^* \cong \mathfrak{d}(\Gamma). \]

The following proposition is a crucial link between the statistical system and the thermodynamic system. We phrase this proposition as our first motto:

"Observation is a moment map."

**Proposition 5.4.** The moment map
\[ \beta_{\mathcal{F}} : T^* \mathcal{P}(\Gamma) \to g^*_\mathcal{O}_\mathcal{F} \]
of the action of \( \mathcal{G}_\mathcal{F} \) on \( T^* \mathcal{P}(\Gamma) \) is characterized by the formula
\[ \left\langle \beta_{\mathcal{F}}(\rho, \beta), \frac{\partial}{\partial F_i} \right\rangle = \mathcal{O}_{F_i}(\rho) - \beta(\rho) \] (5.3)
for all \( i \).

**Proof.** We first note that the function \( \mathcal{O}_{\mathcal{F}} : \mathcal{P}(\Gamma) \to \mathbb{R} \) is the natural restriction of the linear map defined on the bigger space \( \mathfrak{d}(\Gamma) \). We denote this latter map by
\[ \hat{\mathcal{O}}_{\mathcal{F}} : \mathfrak{d}(\Gamma) \to \mathbb{R}. \]
For the latter map which is a linear function on the vector space \( \mathfrak{d}(\Gamma) \), we apply the following standard lemma.

**Lemma 5.5.** Let \( V \) be a vector space and consider its cotangent bundle \( \pi : T^* V \to V \). Let \( f : V \to \mathbb{R} \) be a linear function and consider the function \( f \circ \pi : V \oplus V^* \cong T^* V \to \mathbb{R} \) where \( \pi : T^* V \to V \) is the cotangent projection. Denote by \( \phi^{f \circ \pi}_t \) be the one-parameter subgroup of the Hamiltonian flow of \( f \circ \pi \) regarded as the action of the Lie group \( \mathbb{R} \) acting on \( T^* V \). Then its moment map \( J : T^* V \to \mathbb{R}^* \) is characterized by
\[ \left\langle J(v, \beta), \frac{\partial}{\partial t} \right\rangle = f \circ \pi(v, \beta)(= f(v)). \]

**Proof.** By the defining condition of the moment map,
\[ \left\langle J(v, \beta), \frac{\partial}{\partial t} \right\rangle \]
is the Hamiltonian generating the flow \( \phi^{f \circ \pi}_t \) which is obviously given by \( f \circ \pi \). \( \square \)
Therefore the Hamiltonian generating this flow associated the observable $F_i$ is precisely the function $\mathcal{O}_{F_i} \circ \pi_{T^* \mathcal{D}(\Gamma)}$. In the current case, we are considering the $n$-dimensional Lie algebra whose generating vector fields are those whose flows are the linear translations

$$(\rho, \beta) \mapsto (\rho, \beta + d\mathcal{O}_{F_i})$$

respectively for each $i$ by definition of the Lie algebra action of $\mathfrak{G}_{\mathcal{O}_y}$. (See Definition 5.3.)

Since the Hamiltonian vector field $X_{\mathcal{O}_{F_i}}$ on $T^* \mathcal{D}(\Gamma)$ is vertical, the flow preserves the subset $T^* \mathcal{D}(\Gamma)|_{\mathcal{P}(\Gamma)} \subset T^* \mathcal{D}(\Gamma)$.

We also have the natural projection map $T^* \mathcal{D}(\Gamma)|_{\mathcal{P}(\Gamma)} \to T^* \mathcal{P}(\Gamma)$.

Applying (3.6), we derive

$$J_{F_i}(\rho, \beta), \frac{\partial}{\partial F_i} \mapsto \mathcal{O}_{F_i}(\rho) - \beta(\rho)\frac{M(\rho)}{M(\rho)}X_{\mathcal{O}_{\mathcal{P}(\Gamma)}}(\Gamma; \rho) = \mathcal{O}_{F_i}(\rho) - \beta(\rho)$$

using $M(\rho) = 1$ under the identification (3.5). This finishes the proof. □

We have the commutative diagram

$$T^* \mathcal{P}(\Gamma) \xrightarrow{\mathcal{J}_{\mathcal{O}_{\mathcal{P}(\Gamma)}}} \mathbb{R}$$

Definition 5.6 (Observation data set). We denote by $B_{\mathcal{O}_y}$ the image of the moment map $\mathcal{J}_{\mathcal{O}_y}$ and by $B_{\mathcal{O}_y}^\circ$ the set of its regular values. We call $B_{\mathcal{O}_y}$ the observation data set.

Then we have decomposition

$$T^* \mathcal{P}(\Gamma) = \bigcup_{\mu \in B_{\mathcal{O}_y}} \{\mu\} \times \mathcal{J}_{\mathcal{O}_y}^{-1}(\mu)$$

and

$$T^* \mathcal{P}(\Gamma)/\mathfrak{G}_{\mathcal{O}_y} = \bigcup_{\mu \in B_{\mathcal{O}_y}} \{\mu\} \times \mathcal{J}_{\mathcal{O}_y}^{-1}(\mu)/\mathfrak{G}_{\mathcal{O}_y}.$$ 

We attract readers’ attention that the coadjoint isotropy group of $\mathfrak{G}_{\mathcal{O}_y}$ is the full group

$$\mathfrak{G}_{\mathcal{O}_y, \mu} = \mathfrak{G}_{\mathcal{O}_y}$$

for all $\mu$ and hence the reduced space at $\mu$ [MW74] becomes

$$\mathcal{J}_{\mathcal{O}_y}^{-1}(\mu)/\mathfrak{G}_{\mathcal{O}_y} =: \mathcal{M}_\mu$$

for all regular values of $\mu$.

We summarize the above discussion into

Corollary 5.7. Let $\mu = (\mu_1, \cdots, \mu_n)$ be a collective observation of $\mathcal{F} = \{F_1, \cdots, F_n\}$ and assume $\mu \in \mathfrak{g}_{\mathcal{O}_y}$ is a regular value of the moment map $\mathcal{J}_{\mathcal{O}_y}$. Then the reduced space denoted by

$$\mathcal{M}_\mu^\circ := \mathcal{J}_{\mathcal{O}_y}^{-1}(\mu)/\mathfrak{G}_{\mathcal{O}_y}$$

is a (infinite dimensional) symplectic manifold.
Therefore we have constructed the following fibration over $B^\mathcal{F}_\mathcal{F} \subset g_{\mathcal{O}_\mathcal{F}}$, which we call the collective observation data of $\mathcal{F}$.

**Definition 5.8.** Let $\mathcal{F} = \{F_1, \cdots, F_n\}$ on SPS be an observable system. We call the union

$$\mathcal{M}^\mathcal{F} := \bigsqcup_{\mu \in B^\mathcal{F}_\mathcal{F}} \{\mu\} \times \mathcal{M}_\mu^\mathcal{F}; \quad \mu := (\mu_1, \cdots, \mu_n)$$

the $\mathcal{F}$-reduced kinetic theory phase space ($\mathcal{F}$-reduced KTPS) associated to $\mathcal{F}$, where $\mathcal{J} = \mathcal{J}_\mathcal{F}$ is the moment map associated to the symmetry group generated by the induced Hamiltonian flows on $T^*\mathcal{P}(\Gamma)$ (KTPS).

The projection $\mathcal{M}^\mathcal{F} \to B^\mathcal{F}_\mathcal{F} \subset g_{\mathcal{O}_\mathcal{F}}$ forms a symplectic fiber bundle.

The following is obvious.

**Corollary 5.9.** Suppose that the set $\{X_{\mathcal{O}_{F_i}}\}_{i=1}^n$ is linearly independent. Then the map $\mathbb{R}^n \to g^*_{\mathcal{O}_\mathcal{F}}$ defined by

$$(c_1, \cdots, c_n) \mapsto \sum_{i=1}^n c_i \mathcal{O}_{F_i}$$

is an isomorphism.

We would like to mention that the hypothesis of this corollary is a very weak one. For example, the hypothesis holds for any local observable system that has a point $x \in \Gamma$ at which the differentials

$$\{dF_1(x), \cdots, dF_n(x)\}$$

are linearly independent.

### 6. Reduced Entropy as a Generating Function of Thermodynamic Equilibrium: the Second Motto

We have shown before that the relative information entropy $S : \mathcal{P}(\Gamma) \to \mathbb{R}$ is invariant under the action of $\text{Diff}(\Gamma, \nu_0)$ in Proposition 4.3.

The following will enable us to derive the folklore that the thermodynamic entropy is derived as the reduction of the information entropy.

**Corollary 6.1 ($\mathcal{F}$-reduced entropy).** The pull-back function $S \circ \pi : T^*\mathcal{P}(\Gamma) \to \mathbb{R}$ of the information entropy $S : \mathcal{P}(\Gamma) \to \mathbb{R}$ is universally reduced to a well-defined function

$$S_{\mathcal{F}}^\text{red} : \mathcal{M}^\mathcal{F} \to \mathbb{R}.$$  

We call $S_{\mathcal{F}}^\text{red}$ the $\mathcal{F}$-reduced entropy function.

In fact the above definition can be put into the general construction of Legendrian generating function as follows. (See Appendix B for the definition.)

The procedure of finding a critical point of $S_{\mathcal{F}}^\text{red}$ can be decomposed into the two steps. We choose an Ehresmann connection

$$TM^\mathcal{F} = T^\mathcal{V}M^\mathcal{F} \oplus T^hM^\mathcal{F}$$

of the fibration $\pi : \mathcal{M}^\mathcal{F} \to g^*_{\mathcal{O}_\mathcal{F}}$, and decompose the differential $dS_{\mathcal{F}}^\text{red}(\rho)$ into

$$dS_{\mathcal{F}}^\text{red}(\rho) = d^\mathcal{V}S_{\mathcal{F}}^\text{red}(\rho) + D^hS_{\mathcal{F}}^\text{red}(\rho)$$

into the vertical and the horizontal components thereof. Recall that the vertical differential is canonically defined but the horizontal differential of $S_{\mathcal{F}}^\text{red}$ needs the
use of connection. In general the horizontal component $D^h S(\rho)$ depends on the connection but it will be independent thereof at the point $\rho$ where $d^v S(\rho) = 0$.

Consider the following diagram induced by the diagram \(5.4\)

\[
\begin{array}{ccc}
\mathcal{M}^\tau & \xrightarrow{g^\text{red}_\tau} & \mathbb{R} \\
\downarrow \pi^\tau & & \\
\mathcal{O}_{\mathcal{F}}^\tau
\end{array}
\]

**Definition 6.2 (Covariant thermodynamic equilibrium).** We call the Legendrian submanifold defined by

\[
R_{\mathcal{F}, S} := \{ (\mu, D^h S^\text{red}_\tau(\rho, [\beta]), S^\text{red}_\tau(\rho, [\beta])) \in J^1(\mathcal{O}_{\mathcal{F}}^\tau) \mid d^v S^\text{red}_\tau(\rho, [\beta]) = 0, \\
\mu = \beta_{\mathcal{F}}(\rho, [\beta]) \}
\]

the **covariant thermodynamic equilibrium** of observable system $\mathcal{F}$. We regard the diagram \(6.1\) as a generating function of the Legendrian submanifold $R_{\mathcal{F}, S}$.

**Proposition 6.3.** $R_{\mathcal{F}, S}$ is a Legendrian submanifold of the 1-jet bundle $J^1(\mathcal{O}_{\mathcal{F}}^\tau)$ for the contact form

\[
\lambda = dw - \sum_{i=1}^n q_idp_i
\]

whose potential function is given by $S^\text{red}_\tau$.

**Proof.** We will show that the subset

\[
\mathcal{L}_{\mathcal{F}, S} := \{ (\mu, D^h S^\text{red}_\tau(\rho, [\beta]))(\rho, [\beta])) \in T^* \mathcal{O}_{\mathcal{F}}^\tau \mid d^v S^\text{red}_\tau(\rho, [\beta]) = 0 \}
\]

is a Lagrangian submanifold of $T^* \mathcal{O}_{\mathcal{F}}^\tau$. We denote by $\iota : \mathcal{L}_{\mathcal{F}, S} \to T^* \mathcal{O}_{\mathcal{F}}^\tau$ the canonical inclusion. Then it satisfies

\[
dg = \iota^* \theta, \quad \theta = \sum_{i=1}^n q_idp_i
\]

with $g = \iota^* S^\text{red}_\tau$. The latter is because by definition we have

\[
dS^\text{red}_\tau = D^h S^\text{red}_\tau
\]

on $\mathcal{L}_{\mathcal{F}, S}$. In particular, we have $0 = \iota^* d\theta = -\iota^* \omega_0$ which shows that $\mathcal{L}_{\mathcal{F}, S}$ is a Lagrangian submanifold whose Liouville primitive is given by the function $g$. □

This construction is the coordinate free construction of the local coordinate description, given in the literature (e.g., in \[MNSS90\]), of a thermodynamic equilibrium: The latter construction has been carried out by the Lagrangian multiplier method for the entropy maximum principle. For the readers’ convenience, we summarize this local coordinate description of thermodynamic equilibrium in Section 8.

In the mean time, we now describe our construction of $R_{\mathcal{F}, S}$ more closely in step by step. We first recall the definition of the conormal bundle.

**Definition 6.4.** Let $N \subset M$ be a submanifold of $M$. Then, its **conormal bundle** $\nu^* N$ is the bundle over $N$ defined by

\[
\nu^* N := \{ (x, \alpha) \in T^* M \mid x \in N, \alpha \in T^*_x N \text{ satisfying } T_x N \subset \ker \alpha \}.
\]
We also consider its contact analog, the conormal 1-jet bundle
\[
\tilde{\nu}^* N = \{(x, df(x), f(x)) \in J^1 M \mid x \in N, f \text{ is a germ of function at } x\} \\
\cong \nu^* N \times \mathbb{R} \subset T^* N \times \mathbb{R}.
\]
This is an example of co-Legendrian submanifolds introduced in [Oh21].

6.1. Step 1: Observation and solving of constrained maximum entropy problem. In our framework, the thermodynamic equilibrium state corresponding to \( \mu \in \mathbb{R}^n \) is the subset of \( \mathcal{P}(\Gamma) \),
\[
\{\rho \in \mathcal{P}(\Gamma) \mid d(S \circ \pi)(\rho, \beta) \in \nu^* N(J^{-1}_F(\mu)) \text{ for some } \beta \in T^* \mathcal{P}(\Gamma)\}.
\]
We now introduce the conormal 1-jet bundle \( \tilde{\nu}^* N \) as the following intermediate thermodynamic equilibrium state.

**Definition 6.5.** Let \( \mathcal{F} = \{F_1, \ldots, F_n\} \) be a set of observables in \( \Gamma \).

1. The conormal 1-jet at \( \rho \) is given by
\[
\tilde{\nu}^*(J^{-1}_F(\mu)) = \left\{ (\rho, df|\rho, f(\rho)) \in J^1 \mathcal{P}(\Gamma) \mid \rho \in J^{-1}_F(\mu), df|\rho \in (T^* \mathcal{P}(\Gamma))_\rho \right\}.
\]
2. We call the union \( \tilde{\nu}^*(J^{-1}_F) = \bigcup_{\mu \in B^*_\mathcal{F}} \{\mu\} \times \tilde{\nu}^*(J^{-1}_F(\mu)) \) the \( \mathcal{F} \) iso-data KTPS with \( B^*_\mathcal{F} \subset \mathfrak{g}^*_{\mathcal{O}_\mathcal{F}} \).

6.2. Step 2: Construction of mesoscopic reduced KTPS. Now we consider the information entropy function \( S : \mathcal{P}(\Gamma) \to \mathbb{R} \) and its full 1-jet graph
\[
\mathcal{C}_S := \{(\rho, dS(\rho), S(\rho)) \in J^1 \mathcal{P}(\Gamma) \mid \rho \in \mathcal{P}(\Gamma)\}.
\]

We also consider the \( \mathcal{F} \)-reduced entropy function
\[
S^{\text{red}}_{\mathcal{F}} : M^{\mathcal{F}}_\mu \to \mathbb{R}.
\]
We recall \( M^\mathcal{F} \subset T^* \mathcal{P}(\Gamma)/\mathcal{G}_{\mathcal{O}_\mathcal{F}} \) and \( \mathcal{S} \circ \pi \) descends to the quotient in the way that the value of \( S^{\text{red}}_{\mathcal{F}}(\rho, \beta) \) depends only on \( \rho \). In this regard, we can regard the function \( S^{\text{red}}_{\mathcal{F}} \) restricted to \( M^{\mathcal{F}}_\mu \) as a function on \( \mathcal{P}(\Gamma) \) with the constraints \( \int_{\Gamma} \rho = 1 \) and
\[
\int_{\Gamma} F_i \rho = \mu_i, \quad i = 1, \cdots, n.
\]
We consider the intersection
\[
\mathcal{E}_\mu := \tilde{\nu}^*(J^{-1}_F(\mu)) \cap \mathcal{C}_S \subset J^1 \mathcal{P}(\Gamma).
\]
and write the union
\[
\mathcal{E}_{\mathcal{F},S} := \bigcup_{\mu \in B^*_\mathcal{F}} \{\mu\} \times \mathcal{E}_\mu.
\]
Following the standard procedure of the theory of generating function, we consider the vertical critical point set
\[
\Sigma_{\mathcal{F},S} := \{\rho \in \mathcal{P}(\Gamma) \mid d^\mathcal{S}(\rho) = 0\},
\]
and get the canonical embedding
\[
\Sigma_{\mathcal{F},S} \hookrightarrow \mathcal{E}_{\mathcal{F},S}: \rho \mapsto (\rho, d\mathcal{S}(\rho), \mathcal{S}(\rho))
\]
into \( \mathcal{E}_{\mathcal{F},S} \). (See Appendix [B] for a summary of generating function.)
By the $\mathcal{G}_{\mathcal{O}_S}$-invariance, the embedding descends to the quotient

$$\Sigma_{\mathcal{F};S}/\mathcal{G}_{\mathcal{O}_S} \rightarrow \mathcal{E}_{\mathcal{F};S}/\mathcal{G}_{\mathcal{O}_S}. $$

The latter quotient is the *mesoscopic KTPS parent* of the thermodynamic equilibrium which deserves a name.

**Definition 6.6 (\mathcal{F} iso-data mesoscopic KTPS).** We call the subset $\mathcal{E}_{\mathcal{F};S}/\mathcal{G}_{\mathcal{O}_S}$ the $\mathcal{F}$ *iso-data mesoscopic KTPS*. 

### 6.3. Step 3: Construction of thermodynamic equilibrium state $R_{\mathcal{F};S}$. We solve the $\mathcal{F}$ iso-data critical point problem of $S^{\text{red}}_{\mathcal{F}}$ as follows.

1. We have already obtained the following commutative diagram

2. We actually have the natural projection

$$ \text{pr}_h : \mathcal{E}_{\mathcal{F};S} \rightarrow J^1\mathfrak{g}_{\mathcal{O}_S} $$

given by the horizontal projection

$$(\rho, dS(\rho), S(\rho)) \mapsto (D^h S(\rho), S(\rho))$$

on $\Sigma_{\mathcal{F};S}$ where $d^h S(\rho) = 0$. This descends to the map

$$ \text{pr}_h : \mathcal{E}_{\mathcal{F};S}/\mathcal{G}_{\mathcal{O}_S} \rightarrow J^1\mathfrak{g}_{\mathcal{O}_S} $$

in the above diagram. We write

$$ R_{\mathcal{F};S} := \text{Image}[\text{pr}] \circ [j^1S] = j^1\mathfrak{g}^{\text{red}}_{\mathcal{O}_S}(\Sigma_{\mathcal{F};S}). $$

Then we obtain the following description of $R_{\mathcal{F};S}$ which is nothing but an unravelling of the definition. This is our second motto:
“Relative information entropy is the generating function of thermodynamic equilibria.”

The following is the definition of thermodynamic equilibrium in our framework.

**Definition 6.7 (Thermodynamic equilibrium)**. Consider the horizontal 1-jet of $S j_h^1 S^\text{red}_\mathcal{F} : \Sigma_{\mathcal{F}; S} \to J^1 g^*_\mathcal{O}_\mathcal{F}$ given by the composition $j_h^1 S^\text{red}_\mathcal{F} = pr_h \circ j^1 S^\text{red}_\mathcal{F}$. More explicitly the map is given by

$$j_h^1 S^\text{red}_\mathcal{F}([\rho]) = (\mu, D h^1 S^\text{red}_\mathcal{F}([\rho])).$$

Here $D h^1 S^\text{red}_\mathcal{F} : \Sigma_{\mathcal{F}; S} \to T^* g^*_\mathcal{O}_\mathcal{F}$ is the horizontal derivative of $S^\text{red}_\mathcal{F}$ with respect to the fibration $M^\mathcal{F} \to g^*_\mathcal{O}_\mathcal{F}$. Then we call the Legendrian submanifold $R_{\mathcal{F}; S} := \text{Image} j_h^1 S^\text{red}_\mathcal{F}$ the thermodynamic equilibrium of the statistical system $\mathcal{F}$.

Non-holonomicity of $R_{\mathcal{F}; S}$ arises from the last operation of ‘taking the horizontal projection’.

**7. Mesoscopic dynamics of thermodynamic equilibrium**

In the previous section, we have constructed thermodynamic equilibrium state as the Legendrian submanifold generated by the $\mathcal{F}$-reduced entropy function. In this section, we describe the mesoscopic dynamics of this thermodynamic equilibrium state under the reduction of the Hamiltonian flow on KTPS $T^* P(\Gamma)$ generated by the Hamiltonian flow on SPS of the given Hamiltonian $H$.

Let $H = H(t, q_1, \ldots, q_{3N}, p_1, \ldots, p_{3N})$ be the Hamiltonian driving the dynamics on SPS and consider the induced dynamics on KTPS.

**Remark 7.1.** In the physical thermodynamical system, the Hamiltonian is nothing but the total energy

$$H = \frac{m}{2} \sum_{i=1}^N |\vec{P}_i|^2 + \sum_{i \neq j} U_{ij}(Q_i - Q_j) + \sum_{\ell=1}^N U(Q_\ell).$$

**Lemma 7.2.** The flow $\Phi^t_H := \phi^t_H : T^* P(\Gamma) \to T^* P(\Gamma)$ induced by the Hamiltonian flow $\phi^t_H : \Gamma \to \Gamma$ preserves the observations $O_{F_i}$ and $S$.

**Proof.** We note that $\phi^t_H$ preserves the Liouville measure $\nu_0$. Therefore by the change of variable formula, we have

$$O_{F_i} \circ \Phi^t_H = O_{F_i}, \quad S \circ \Phi^t_H = S.$$

Then lemma follows from Proposition [4.3] and Lemma [5.2].

Therefore the flow $\Phi^t_H$ descends to $M^\mathcal{F}$. We denote this flow by

$$\Phi^t_H : M^\mathcal{F} \to M^\mathcal{F}.$$

Recall that the induced flow on $T^* P(\Gamma)$ of $\Phi^t_H$ preserve the level sets of $O_{F_i}$ and commute with $X_{O_{F_i}}$ for all $i = 1, \ldots, k$. We also recall that

$$\mathcal{F}^{-1}(\mu) \subset T^* P(\Gamma)$$

and the flow on $M^\mathcal{F}$ is the restriction of the quotient flow of $T^* P(\Gamma)/g^*_O$.
Lemma 7.3. Let $\mathcal{L}_S \subset T^*\mathcal{P}(\Gamma)$ (resp. $\mathcal{R}_S \subset J^1\mathcal{P}(\Gamma)$) be the Lagrangian submanifold (resp. the Legendrian submanifold) given by

$$\mathcal{L}_S := \text{Image } dS \quad \text{(resp. } \mathcal{R}_S = \text{Image } j_{1S})$$

(resp. $\mathcal{R}_S$ = Image $j_{1S}$). Then the Hamiltonian flow $\Phi^t_H$ (resp. its lifting to $J^1\mathcal{P}(\Gamma)$) preserves the Lagrangian submanifold $\mathcal{L}_S$ (resp. $\mathcal{R}_S$) for all $t$. In particular the reduced flow also preserves the value of the $\mathcal{F}$-reduced entropy $S_{\mathcal{F}}^{\text{red}}$.

Proof. First we recall the fact that all $\mathcal{O}_F, S$ are defined on $\mathcal{P}(\Gamma)$ which are lifted to $T^*\mathcal{P}(\Gamma)$ and so they Poisson-commute. In particular, their level sets are invariant under the Hamiltonian flows. In particular the intersection

$$\mathcal{L}_{\mathcal{F}, S} := \text{Image } dS \cap \mathcal{O}_F^{-1}(\mu) / \mathcal{G}_F$$

is naturally a subset of $\mathcal{M}_{\mathcal{F}} \subset T^*\mathcal{P}(\Gamma) / \mathcal{O}_F$. This proves $E_{\mathcal{F}, S} = \text{Image } \Phi^t_H$. Since the moment map $J_F$ is invariant under the flow $\Phi^t_H$, the induced flow on $E_{\mathcal{F}, S}$ is a fiberwise flow of $(J_F)^{-1}(\mu)$.

Furthermore we have

$$d(S \circ \Phi^t_H) = dS \circ d\Phi^t_H = (\Phi^t_H)^*(dS).$$

We also have

$$(\Phi^t_H)^*(dS)(\rho, dS(\rho)) = (\Phi^t_H(\rho), d(S \circ \Phi^t_H(\rho)))$$

(7.1)

and

$$S(\Phi^t_H(\rho)) = S(\rho).$$

(7.2)

Combining (7.3) - (7.2), we conclude that their horizontal projections is the same as

$$\text{pr}_h = \text{pr}_h \circ (\Phi^t_H)^{-1}. \quad (7.3)$$

(See Diagram [6.5]) This proves the invariance of $R_{\mathcal{F}, S}$ under the reduced flow. □

Even in the thermodynamic equilibrium, the original kinetic system still carry the internal flow induced by the time-evolution

$$t \mapsto \Phi^t_H(\rho_0) = (\phi^t_H), \rho_0$$

of the probability distribution in SPS (or in KTPS) may also undergo nontrivial time-evolution.

More specifically the flow of the probability distribution satisfies the reduced equation of the kinetic equation

$$\frac{\partial \rho}{\partial t} = -\mathcal{L}_{X_H}, \rho(0) = \rho_0. \quad (7.4)$$

By writing $\rho = f \nu_0$ with $f = \frac{\partial \rho}{\partial \nu_0}$, the latter equation is equivalent to

$$\frac{\partial f}{\partial t} = -\{f, H\}, \quad f(0) = f_{\rho_0}. \quad (7.5)$$

By differentiating the macroscopic entropy $S(\rho_t)$ in time $t$, we get

Proposition 7.4. Let $H = H(t, x)$ be the Hamiltonian of SPS and let $\rho_t = (\phi^t_H), \rho_0$ for any $\rho_0 \in \mathcal{P}(\Gamma)$. Then

$$\frac{\partial}{\partial t}(S(\rho_t)) = \int_{\Gamma} (\log f_t + 1)\{f_t, H\} d\Gamma. \quad (7.6)$$
Proof. We compute

\[
\frac{\partial}{\partial t} (S(\rho_t)) = \int_{\Gamma} \frac{\partial}{\partial t} (f_t \log f_t) \, d\Gamma \\
= \int_{\Gamma} (\log f_t + 1) \frac{\partial f_t}{\partial t} \, d\Gamma \\
= \int_{\Gamma} (\log f_t + 1) \{f_t, H\} \, d\Gamma.
\]

This finishes the proof. \(\Box\)

The equation (7.5) is the simplest form of the Boltzmann equation when \(H\) is just the kinetic energy (or in vacuum). We refer readers to Section 12 for further discussion on the more general form of kinetic equations.

8. Non-holonomicity of thermodynamic equilibrium and phase transition

In the literature of thermodynamics, derivation of thermodynamic phase space is attempted from the first principle of statistics and the information theory. For example in [Jay57], [Mru78], [MNSS90], the authors provide a derivation of the thermodynamics equilibrium in terms of information entropy and contact geometry motivated by the second law of thermodynamics: "In the thermodynamic equilibria, the entropy is maximized."

However the description is not satisfactory in that the description is given under the premise that the Legendrian submanifold \(R^F;S\) admits a single-valued potential function \(f: B^F;S \rightarrow \mathbb{R}\) such that

\[R^F;S = \{ (\mu, df(\mu), f(\mu)) \mid \mu \in B^F;S \} .\]

Our global derivation of thermodynamic equilibrium given in the previous section indicates that the associated Legendrian submanifold is not necessarily holonomic. We will illustrate by van der Waals model that non-holonomicity indeed occurs in a physical model which is responsible for the phase transition. Then the well-known Maxwell’s construction is an attempt to overcome non-holonomicity in computing the relevant thermodynamic potential function.

We first review the exposition of [Jay57], [MNSS90] on the thermodynamic equilibrium in terms of the variables \((q^1, \cdots, q^n)\) which is a coordinate system of \(B^F;S \subset \mathbb{R}^n\). We will closely follow the exposition of [MNSS90], especially try to be consistent therewith in the usage of letters for the relevant variables except the change of \(x^i\)’s by \(q^i\)’s.

We first give the definition of TPS associated to the observable \(F = \{F_1, \ldots, F_n\}\).

**Definition 8.1.** A thermodynamic phase space \(\mathcal{T}\) is an open subset of \(\mathbb{R}^{2n+1}\) equipped with the contact structure of the 1-jet bundle associated to a statistical mechanical system of observables \(F_1, \cdots, F_n\) on SPS.

In the contact geometric formulation of TPS, it is described by a coordinate system \((q^1, \ldots, q^n; p_1, \ldots, p_n; z)\) where \(q^1, \ldots, q^n\) are called configuration variables, and the relevant contact structure is defined by the kernel of the contact form of the type

\[dz - \sum_i q^i dp_i.\]
In this formulation, the contact structure on TPS is related to the first law of thermodynamics, and Legendrian submanifold is related to an equilibrium state. Note that $\nu_0$ is the Lebesgue measure and that we can express any nonnegative density $D$ as $D = f \nu_0$ for the Radon-Nikodym derivative

$$\frac{\partial D}{\partial \nu_0} = f$$

which is a nonnegative $L^1$-function on SPS.

We now consider the system with $n$-observables, $F_1, \ldots, F_n$. Then, we want to find a probability density $\rho \in \mathcal{P}(\Gamma)$ that maximizes the relative entropy functional

$$S(\rho) = -\int_{\Gamma} \rho \log \left( \frac{\partial \rho}{\partial \nu_0} \right),$$

under the constraints

$$q_i = \int_{\Gamma} F_i \frac{D}{D} = \frac{\langle F_i \rangle_D}{M(D)}$$

for given explicit observations $(q^1, \ldots, q^n)$ [Jay57].

We note that while $\mathcal{D}^+(\Gamma)$ is an open subset of a linear space, $\mathcal{P}(\Gamma)$ is not which is an affine hypersurface given by the integral constraint

$$\int \rho = 1.$$ (8.2)

Instead of solving this maximization problem on $\mathcal{P}(\Gamma)$, we solve it on the (open subset of) linear space $\mathcal{D}^+(\Gamma) \subset \mathcal{D}(\Gamma)$ by considering (8.2) itself as a constraint like (8.1).

In particular at any extremum point $D \in \mathcal{P}(\Gamma)$, there exist some constants $(w, \lambda_1, \ldots, \lambda_n)$ such that we can express the one-form $dS|_D$ as a linear combination

$$dS|_D = w \cdot dM|_D - \sum_i \lambda_i d\mathcal{O}_{F_i}|_D$$ (8.3)

(“The method of the Lagrange multiplier”). Presence of the negative sign is nothing significant but just a traditional convention.

More explicitly, by substituting $\dot{D}$ into (8.3), we derive the identity

$$\dot{D}[S] = w \dot{D}[M] - \sum_i \lambda_i \dot{D}[(F_i)]$$

which is equivalent to the vanishing of the directional derivative

$$\dot{D} \left[ S - w M + \sum_i \lambda_i \mathcal{O}_{F_i} \right] = 0$$

for all $\dot{D} \in \mathcal{D}^+(\Gamma) \equiv T_D \mathcal{D}^+(\Gamma) \equiv \mathcal{D}(\Gamma)$. This can be explicitly written into

$$-\int_{\Gamma} \dot{D} \left( \log \left( \frac{\partial D}{\partial \nu_0} \right) + w - \sum_i \lambda_i F_i \right) = 0$$

by the linearity of $\mathcal{O}_{F_i}$ and by definition of the mass function $M$. By writing

$$f_D := \frac{\partial D}{\partial \nu_0},$$
we obtain
\[ \log f_D + w - \sum_{i=1}^{n} \lambda_i F_i = 0 \]
and hence
\[ f_D = e^{-w + \sum \lambda_i F_i} \]
(8.4)

Therefore we have proved

**Proposition 8.2.** At any extreme point \( D \) of the entropy under the constraint (8.1), there exists a Lagrange multiplier \((w, \lambda_1, \ldots, \lambda_n)\) such that \( D = e^{-w + \sum \lambda_i F_i} \nu_0 \)

Then \([\text{MNSS90}]\) proceeds further using the normalization condition
\[ 1 = \int_{\Gamma} D = \int_{\Gamma} e^{-w + \sum \lambda_i F_i} \nu_0 \]
which gives rise to the expression of \( w \) in terms of other Lagrange multipliers
\[ e^w = \int_{\Gamma} e^{\sum \lambda_i F_i} \nu_0. \]
(8.5)

By taking the logarithm thereof, we get
\[ w = \log \left( \int_{\Gamma} e^{\sum \lambda_i F_i} \nu_0 \right). \]
(8.6)

**Corollary 8.3.** For the probability density, \( \rho_\lambda = e^{-w(\lambda) + \sum \lambda_i F_i} \nu_0 \), any equilibrium state of TPS associated to the observables \( \mathcal{F} = \{F_1, \ldots, F_n\} \) and the Lagrange multipliers \( \lambda = (\lambda^1, \ldots, \lambda^n) \) in the previous proposition, the followings also hold:

1. At any extreme point \( \rho \), \( \rho \) has the form
\[ \rho_\lambda = \left( \frac{e^{\sum \lambda_i F_i}}{\int_{\Gamma} e^{\sum \lambda_i F_i}} \right) \nu_0. \]
(8.7)

2. The observations associated to the variables \( \lambda = (\lambda_1, \ldots, \lambda_n) \) have their values
\[ q^i(\lambda) := \int_{\Gamma} F_i \rho_\lambda = \frac{\partial w}{\partial \lambda_i} \]

3. We denote by \( \rho_{(w,q,p)} := \rho_\lambda \) the probability density \( \rho_\lambda \) on \( \Gamma \) determined above. Then in terms of the variables \( \lambda_1, \ldots, \lambda_n =: (p_1, \ldots, p_n) \)
the (relative) entropy is expressed as
\[ S(\rho_{(w,q,p)}) = -\int_{\Gamma} \rho_{(w,q,p)} \log \frac{\partial \rho_{(w,q,p)}}{\partial \nu_0}. \]

**Remark 8.4.**

1. The relationship between the relative information entropy \( S(\rho_{(q,p,w)}) \) and \( w \) on the Legendrian submanifold \( R_{\mathcal{F}, \lambda} \) is the equation
\[ d(S(\rho_{(\cdot)}))_{\mid R_{\mathcal{F}, \lambda}} = \alpha_{R_{\mathcal{F}, \lambda}} \]
for the contact form \( \alpha = dw - \sum_{i=1}^{n} q_i dp_i \) by the general property of Legendrian generating function and by the way how the variables \( (p_1, \ldots, p_n) \) above are chosen.
28 JIN-WOOK LIM AND YONG-GEUN OH

(2) [MNSS90] goes further and consider the expressions
\[ s = -\log \rho = w - \sum_i p_i F_i \]
\[ ds = dw - \sum_i F_i dp_i \]
by saying “These are both functions of the microscopic variables \( \Gamma \) (by way of \( F_i(\Gamma) \)), and of the parameters \( w, p_1, \ldots, p_n \). Nevertheless, differentiation is understood to be only with respect to the variables \( w, p_1, \ldots, p_n \).”

(3) The upshot of Section 6 is that it provides precise mathematical statement in the quotation mark in a mathematically consistent way.

9. Statistical description of gas in a piston and volume variable

So far we have considered the configuration space \( \mathbb{R}^{3N} \), i.e., the set of \( N \) identical particles on the full space \( \mathbb{R}^3 \). In this section, we explain in our framework how we can involve the thermodynamic volume variable, which is \textit{not} a local observable in the gas model of statistical mechanics.

Consider a thermodynamic system of gas, i.e., a system of a finite fixed number of particles within the cylindrical container
\[ D^2(r) \times (0, h) =: C_h \] (9.1)
that is (vertically) expandable from the base. In other words, we assume that its volume varies by some outside work. Following the common words, we will call a piston any expandable container.

We denote by \( h > 0 \) the parameter with respect to which the volume grows linearly in terms of the variable \( h \). We may regard \( h \) the height of the cylinder at a given moment and so the volume of the container in \( \mathbb{R}^3 \) associated to the parameter \( h \) will be given by
\[ F_2(h) = Ch \]
for some fixed constant \( C > 0 \).

Then we consider the configuration space \( M \subset \mathbb{R}^K \times \mathbb{R}_+ \) with \( K = 3N \) defined by
\[ M = \bigcup_{h \in \mathbb{R}^+} (D^2 \times (0, h))^N \times \{ h \} \] (9.2)
which is an open subset of \( \mathbb{R}^{3N} \times \mathbb{R}_+ \subset \mathbb{R}^{3N} \times \mathbb{R} = \mathbb{R}^{3N+1} \). We write the natural coordinates of \( M \) coming from \( \mathbb{R}^{3N+1} \) as
\[ (q_1, \ldots, q_{3N}, h), \quad h \in \mathbb{R}_+ . \]
We then consider
\[ \Gamma := \bigcup_{\Lambda \in \mathbb{R}_+} T^*(D^2 \times (0, h))^N \times \{ \Lambda \} \]
and the natural projection \( \pi_h : \Gamma \to \mathbb{R}_+ \), where we denote
\[ \Lambda := \text{vol}((D^2 \times (0, h))^N) = (\pi r^2 h)^N . \] (9.3)
(Here we use \( \Lambda \) instead of \( h \) to emphasize that our SPS is \( T^*\mathbb{R}^{3N} \), not \( T^*\mathbb{R}^3 \).)

We write the natural coordinates of
\[ \Gamma \cong \bigcup_{\Lambda \in \mathbb{R}_+} (D^2 \times (0, h))^N \times \mathbb{R}^{3N} \times \{ \Lambda \} \subset \mathbb{R}^{3N} \times \mathbb{R}^{3N} \times \mathbb{R}_+ \]
by

\[ (q_1, \cdots, q_{3N}, p_1, \cdots, p_{3N}, h). \]

The canonical symplectic form on \( T^*(D^2 \times (0, h))^{3N} \) is given by the restriction of

\[ \Omega_N = \sum_{i=1}^{3N} dq_i \wedge dp_i. \]

**Remark 9.1.** The SPS in the present case is a degenerate symplectic manifold, or more precisely a pre-symplectic manifold with its null distribution is spanned by \( \frac{\partial}{\partial h} \). For our purpose of deriving the equations of the ideal gas or the van der Waals models, \( h \) does not appear as a dynamical variable and so we do not need to consider dynamics on the variable \( h \). An upshot of our framework of the gas in a piston is that in the statistics or in the information theoretic point of view, we regard the motion of the gas in the piston as the composition of two independent ‘events’, one followed by the other:

\[ \text{(E1) The configuration space motion of the piston encoded by the parameter } h \in \mathbb{R}_+, \text{ \( (E2) \) The phase space motion of the gas inside the container.} \]

We hope to come back to the study of a more general thermodynamical system elsewhere, and restrict ourselves to the cases of ideal gas and van der Waals models in the present paper.

This being said, let \( \mathcal{F} = \{ F_1, \cdots, F_n \} \) on \( T^*\mathbb{R}^{3N} \) be a system of local observables. We consider the observable system on \( \Gamma \) given by

\[ \{ F_1, \cdots, F_n, \Lambda \} =: \mathcal{F} * \Lambda \]

where \( F_i \) are local functions on \( T^*C^N_h \) for the cylinder [9.1], which are independent of \( \Lambda \), i.e., have the form

\[ F_i = F_i(q, p), \quad i = 1, \cdots, n. \]

We apply all of our construction preformed in the previous sections to this observable system with the following modifications.

We define the collective observation to be

\[ J_{\mathcal{F} \Lambda} = (\mathcal{O}_{F_1}, \cdots, \mathcal{O}_{F_n}, \Lambda) \]  

(9.4)

and write the observation data set

\[ B_{\mathcal{O}_{J_{\mathcal{F} \Lambda}}} = \text{Image } J_{\mathcal{F} \Lambda} \subset g_{J_{\mathcal{F} \Lambda}} \]

Then we have natural fibration

\[ B_{\mathcal{O}_{J_{\mathcal{F} \Lambda}}} \to (0, \Lambda) \]

and hence the \( \mathcal{F} \Lambda \)-reduced KTPS carries the variable \( \Lambda \) only as a configuration variable without conjugate partner in the SPS level. (Recall the direction of \( h \) is the null direction of the symplectic manifold \( \Gamma \).) By taking the relevant SPS and KTPS to be

\[ \Gamma = T^*C^N_h \times \mathbb{R}_+ \]

and

\[ T^*\mathcal{F}(\Gamma) \]

respectively, we still take the associated TPS to be

\[ \mathcal{F} = J^1(g_{\mathcal{F} \Lambda}). \]
Then, similarly to (5.4), the $F^*\Lambda$-reduced entropy function $S^\text{red}_{F^*\Lambda}$ has the diagram

$$
M_{F^*\Lambda} \xrightarrow{S^\text{red}_{F^*\Lambda}} \mathbb{R}
$$

and generates the thermodynamic equilibrium of $F^*\Lambda$. In addition to the variables $(p_1, \cdots, p_n)$ with $p_i = \langle F_i \rangle$, we have additional variable coming from $\Lambda$. We call this the \textit{volume variable} of the system $F^*\Lambda$ and its conjugate variable in $J^1(g_{F^*\Lambda}) \cong T^*(g_{F^*\Lambda}) \times \mathbb{R}$ the \textit{pressure variable} and denote it by $P$.

10. THERMODYNAMICS OF IDEAL GAS

In this section and the next, we recall the well-known study on the equilibrium gas equations in the point of view of statistical mechanics and reconstruct them by our geometric reduction.

In thermodynamic equilibrium, say in the system $F$ of ideal gas or van der Waals gas equation, the state variables are given by

$$(U, T, V, P, S) :$$

- $U = \langle H \rangle$ is the energy,
- $S := S(\rho)$ is the thermodynamic entropy,
- $T$ is the temperature,
- $V$ is the volume, and $P$ the pressure.

The first law of equilibrium thermodynamic is written as

$$dS = \frac{1}{T} dU + \frac{P}{T} dV$$

in the entropy representation.

Deriving the gas equation in an equilibrium state in thermodynamics means to express an element in our Legendrian submanifold $R_{F,S}$ in the form

$$(df(p), p, f(p))$$

for some function $f : \mathbb{R}^n \to \mathbb{R}$ called a \textit{thermodynamic potential}. However such a global function may not exist in general as illustrated by the van der Waals gas model.

**Remark 10.1.** We would like to attract readers’ attention to the way how we write the expression $(df(p), p, f(p))$ differently from $(q, df(q), f(q))$ which is the standard way of writing the 1-jet graph of $f$ in contact geometry. We will see that this is reason why the \textit{horizontal} line is drawn in the Maxwell construction, not the \textit{vertical} line drawn in the construction of graph selectors in the cotangent or in the 1-jet bundle in symplectic and contact geometry. (See [Oh97], [Oh10], [AOOdS18].)

10.1. Nonequilibrium thermodynamics of ideal gas. Consider the ideal gas of $N$ identical particles contained in the piston $C^N$ described in Section 10.

Recall that in the ideal gas model we start with the kinetic energy Hamiltonian

$$K = \frac{1}{2m} \sum_{i=1}^N |P_i|^2$$
on SPS which drives the dynamics of many-body system not only on SPS but also
induces the kinetic equation (7.4) which is equivalent to the equation for the density
\[ \frac{\partial f}{\partial t} = -\{f, K\}, \quad f = f_{\rho}(t, Q, P) \]
The Poisson bracket of the right hand side is computed to be
\[ \{f, K\} = P \cdot \nabla_{Q} f \]
and hence the equation is reduced to
\[ \frac{\partial f}{\partial t} + P \cdot \nabla_{Q} f = 0. \] (10.1)

In this subsection, we would like to describe the reduced flow of this equation on TPS, i.e., on \( J_{1} \mathcal{R}_{2}(U,V) \) in terms of the coordinates \((U, V, T, P, S)\) for
\[ U := \langle K \rangle_{\rho}, \quad S := S(\rho). \]
For this purpose, let us first give concrete description of the quotient space
\[ M_{\text{int}}^{T} := \bigcup_{\mu \in B_{2}} M_{\mu}^{T} = \bigcup_{\mu \in B_{2}} \mathcal{F}^{-1}(\mu)/\mathcal{G}_{o} \]
and the \( \{K\} \)-reduced entropy function
\[ S_{\text{red}}^{K} : \mathcal{F}^{-1}(\mu)/\mathcal{G}_{o} \rightarrow \mathbb{R}. \] (10.2)

We first note that each element in (10.2) represents a trajectory of a point \((\rho, \beta) \in \mathcal{F}^{-1}(\mu) \subset T^{*}\mathbb{P}(\Gamma)\) induced by the Hamiltonian flow of \( K \) on SPS \( \Gamma \) with
energy expectation value (or energy observation) \( \mu \). In other words it is the moduli space of solutions of the system of PDE
\[ \frac{\partial \rho}{\partial t} = -\mathcal{L}_{X_{K}} \rho, \quad \frac{\partial \beta}{\partial t} = \mathcal{L}_{X_{K}} \beta \] (10.3)
for \((\rho, \beta)\) in
\[ T^{*}\mathbb{P}(\Gamma) \cong \mathbb{P}(\Gamma) \times D_{0}(\Gamma) : \]
We recall the action of \( \mathcal{G}_{\{K\}} \) on \( T^{*}\mathbb{P}(\Gamma) \) is given by
\[ (\phi, (\rho, \beta)) \mapsto (\phi*\rho, \phi*\beta). \]
We now consider the time-evolution of the Lagrangian submanifold
\[ \text{Image } dS \subset T^{*}\mathbb{P}(\Gamma) \]
under the flow (10.3). By the relation, \( \beta = dS(\rho) \) on \( \text{Image } dS \subset T^{*}\mathbb{P}(\Gamma) \), the evolution thereof is determined by the evolution of \( \rho \)-component which is
\[ \frac{\partial \rho}{\partial t} + \mathcal{L}_{X_{K}} \rho = 0. \] (10.4)
This is equivalent to the Cauchy problem
\[ \frac{\partial f}{\partial t} + P \cdot \nabla_Q f = 0 \] (10.5)
of the probability density function \( f \) of \( \rho \). One may interpret this kinetic equation as the mesoscopic nonequilibrium thermodynamics of the ideal gas. More specifically each solution \( f = f(t, x) \) of (10.4) defines an isotopy of Lagrangian submanifolds
\[ (t, \rho) \mapsto \text{Image } dS(\Phi(t, \rho)) \]
where \( \Phi(t, \rho) := \rho_t \) is the solution of (10.5) at time \( t \) with initial condition \( \rho = f \nu_0 \).\footnote{Remark 10.2. (1) For the ideal gas in the piston \( D^2(1) \times \mathbb{R} \), one should put a suitable boundary condition for \( f \) on \( \partial D^2(1) \times \mathbb{R} \) depending on the given circumstances.
(2) The equation (10.4) is precisely the equation of free transport in the kinetic theory. It has been known that complemented with suitable boundary conditions, it is the right equation for describing the ideal gas. (See [Vil02, Section 1.2] for example.)
(3) Indeed recalling that the operator \( P \cdot \nabla_Q \) generates the \( Q \)-translation with velocity \( P \), we can easily check that the function
\[ f(t, Q, P) = f_0(Q - tP, P) \]
is the unique solution of the initial valued problem of (10.5) for any given (smooth) initial condition \( f(0, Q, P) = f_0(Q, P) \).

10.2. Ideal gas equation at equilibrium. The energy and the volume of \( (D^2(r) \times (0, h))^3N \) are given as
\[ F_1 = \frac{1}{2m} \sum_{i=1}^{3N} |P_i|^2 \]
\[ F_2 = Ch, \quad C = \pi v^2 \]
for some positive constant \( m \). Set \( P = (P_1, \cdots, P_{3N}) \in \mathbb{R}_{3N}^3 \) and
\[ Q = (Q_1, \cdots, Q_{3N}, h) \in M \]
for \( M \) given in (9.2).
We have the formulae for the variable \( U \) and \( V \)
\[ U = \langle F_1 \rangle_\rho, \quad V = \langle F_2 \rangle_\rho \] (10.6)
by definition. The first thermodynamic law is written as
\[ dS = \frac{1}{T} dU + \frac{P}{T} dV. \]
Now we express the thermodynamic entropy \( S \) as a function of \( (U, V) \) by computing the formulae for the expectation value of \( F_1 \) and \( F_2 \) on
\[ \Gamma = T^* \mathbb{R}^3N \times \mathbb{R}_+ \]
with respect to the probability distribution \( \rho \) thereon.
We compute the expectation value

\[ f = \frac{1}{\mathcal{Q}_N^N} e^{\frac{N}{2m} \sum |P_i|^2 + \lambda_2 Ch} \int d\Gamma. \tag{10.7} \]

Therefore

\[ \langle F_1 \rangle_{\rho_{\lambda}} = \frac{\int F_1 e^{\frac{N}{2m} \sum |P_i|^2 + \lambda_2 Ch} d\Gamma}{\int e^{\frac{N}{2m} \sum |P_i|^2 + \lambda_2 Ch} d\Gamma}. \tag{10.8} \]

We compute the expectation value

\[ \langle F_1 \rangle_{\rho_{\lambda}} = \frac{\int_{\Gamma} \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right) e^{\lambda_1 \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right) + \lambda_2 Ch} dP dQ dh}{\int_{\Gamma} e^{\lambda_1 \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right) + \lambda_2 Ch} dP dQ dh} = \frac{\int_{\mathcal{R}_N^N} \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right) e^{\lambda_1 \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right) + \lambda_2 Ch} dP}{\int_{\mathcal{R}_N^N} e^{\lambda_1 \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right) + \lambda_2 Ch} dP} \]

where \( dP = \wedge_{i=1}^N (dp_1^i \wedge dp_2^i \wedge dp_3^i) \) with \( P_i = (p_1^i, p_2^i, p_3^i) \). We evaluate the integral using the following recurrence relation. We set

\[ E_N := \frac{\int_{\mathcal{R}_N^N} \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right) e^{\lambda_1 \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right)} dP}{\int_{\mathcal{R}_N^N} e^{\lambda_1 \left( \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \right)} dP} \]

and \( U_N = \frac{1}{2m} \sum_{i=1}^N |P_i|^2 \).

**Lemma 10.3.** For \( N \geq 2 \), we have

\[ E_N = E_1 + E_{N-1}, \]

and \( E_1 = - \frac{3}{2m} \). In particular, we have \( E_N = - \frac{3N}{2m} \).

**Proof.** We can express the last integral as

\[ E_N = \frac{\int_{\mathcal{R}_3^N} \int_{\mathcal{R}_3^{N-1}} (U_1 + U_{N-1}) e^{\lambda_1 (U_1 + U_{N-1})} dP_1 \wedge dP_2 \wedge \cdots \wedge dP_N}{\int_{\mathcal{R}_3} e^{\lambda_1 U_1} dP_1 \int_{\mathcal{R}_3^{N-1}} e^{\lambda_1 U_{N-1}} dP_2 \wedge \cdots \wedge dP_N} \]

for \( U_1 = \frac{1}{2m} |U_1|^2 \) and \( U_{N-1} = \frac{1}{2m} \sum_{i=2}^2 |U_i|^2 \). It becomes

\[ E_N = \frac{\int_{\mathcal{R}_3} U_1 e^{\lambda_1 U_1} dP_1}{\int_{\mathcal{R}_3} e^{\lambda_1 U_1} dP_1} + \frac{\int_{\mathcal{R}_3^{N-1}} U_{N-1} e^{\lambda_1 U_{N-1}} dP_2 \wedge \cdots \wedge dP_N}{\int_{\mathcal{R}_3^{N-1}} e^{\lambda_1 U_{N-1}} dP_2 \wedge \cdots \wedge dP_N} = E_1 + E_{N-1}. \]

We now express

\[ \frac{\int_{\mathcal{R}_3} U_1 e^{\lambda_1 U_1} dP_1}{\int_{\mathcal{R}_3} e^{\lambda_1 U_1} dP_1} = \frac{\int_0^{2\pi} \int_0^\pi \int_0^\infty \frac{\lambda_1}{2m} r^2 \sin \theta dr d\phi d\theta}{\int_0^{2\pi} \int_0^\pi \int_0^\infty e^{\frac{\lambda_1}{2m} r^2} \sin \theta dr d\phi d\theta} = \frac{\int_0^\infty r^2 e^{\frac{\lambda_1}{2m} r^2} dr}{2m \int_0^\infty r^2 e^{\frac{\lambda_1}{2m} r^2} dr} = \frac{3}{2\lambda_1}. \]

\[ \square \]
We note that for the integral to converge, we need to choose $\lambda_1 < 0$. In that case a direct calculation gives rise to

**Lemma 10.4.**

\[
U = \langle F_1 \rangle_{\rho_\lambda} = -\frac{3N}{2\lambda_1}. \tag{10.9}
\]

10.2.2. *Calculation of* $V$. Similar calculation, which is easier, leads to

\[
\langle F_2 \rangle_{\rho_\lambda} = \int_{\Gamma} (Ch)e^{\lambda_1 F_1 + \lambda_2 Ch} dQ \wedge dh = \int_{C_0^\infty \times R^+} (Ch)e^{\lambda_2 Ch} dQ \wedge dh = \int_{C_0^\infty} C h^N e^{\lambda_2 Ch} dh
\]

We now evaluate integral

\[
\int_0^\infty (Ch)(\pi h)^N e^{\lambda_2 Ch} dh = \left(-\frac{1}{\lambda_2 C}\right)^{N+1} (N+1)! C N^N
\]

\[
\int_0^\infty (\pi h)^N e^{\lambda_2 Ch} dh = \left(-\frac{1}{\lambda_2 C}\right)^{N+1} (N)! N^N. \tag{10.10}
\]

This proves

**Lemma 10.5.**

\[
V := \langle F_2 \rangle_{\rho_{(w,q,p)}} = -\frac{N+1}{\lambda_2}. \tag{10.11}
\]

10.2.3. *Evaluation of* $S$. Finally we compute

\[
S(\rho_\lambda) = \int_{\Gamma} -\rho_\lambda \frac{\partial \rho_\lambda}{\partial v_0} = \int_{\Gamma} (w - \lambda_1 F_1 - \lambda_2 F_2) \rho_\lambda
\]

\[
= w - \lambda_1 \langle F_1 \rangle_{\rho_\lambda} - \lambda_2 \langle F_2 \rangle_{\rho_\lambda}.
\]

From (10.7), we have

\[
\rho_\lambda = f \Gamma = \frac{e^{\lambda_1 F_1 + \lambda_2 F_2} d\Gamma}{\int_{\Gamma} e^{\lambda_1 F_1 + \lambda_2 F_2} d\Gamma}.
\]

We also have

\[
w = \log \left( \int_{\Gamma} e^{\lambda_1 F_1 + \lambda_2 F_2} d\Gamma \right)
\]

from (8.6) for which we write

\[
\int_{\Gamma} e^{\lambda_1 F_1 + \lambda_2 F_2} d\Gamma = \int_{R^3N} e^{\frac{\lambda_1}{2} |P|^2} dP \cdot \int_0^\infty (\pi h)^N e^{\lambda_2 Ch} dh.
\]

We have already computed

\[
\int_0^\infty (\pi h)^N e^{\lambda_2 Ch} dh = \left(-\frac{1}{\lambda_2 C}\right)^{N+1} N! N^N
\]

and compute

\[
\int_{R^3N} e^{\frac{\lambda_1}{2} |P|^2} dP = \left(2 \sqrt{-\frac{2\pi m}{\lambda_1}} \right)^{3N}.
\]

This proves

\[
w = \log \left( \left(-\frac{1}{\lambda_2 C}\right)^{N+1} N! N^N \right) + \log \left(2 \sqrt{-\frac{2\pi m}{\lambda_1}} \right)^{3N}.
\]
Combining the above, we have computed

Lemma 10.6.

\[ w = \frac{3N}{2} \log \left( -\frac{2\pi m}{\lambda_1} \right) + (N + 1) \log \left( -\frac{1}{\lambda_2 C} \right) + \log(N!) + N \log(2\pi). \]  

10.2.4. Derivation of ideal gas law. At the equilibrium state \( R_{\mathcal{F}, \mathcal{S}} \), a thermodynamic potential is nothing but \( S_{\mathcal{F}}^\text{red} \) restricted to the associated Legendrian submanifold regarded as a global function in that a point \((q, p, w) \in R_{\mathcal{F}, \mathcal{S}}\) can be expressed as

\[ q = \mu, \quad p = df(\mu), \quad w = f(\mu). \]

In general such a function may or may not exist as a single-valued function. The following definition deserves attention which is a standard terminology.

Definition 10.7. Let \((q_1, \cdots, q_n, p_1, \cdots, p_n, w)\) be the canonical coordinates of the 1-jet bundle \( J^1M \) and consider the canonical contact form \( \lambda = dw - \sum_{i=1}^n p_idq_i \). We call a Legendrian submanifold \( R \subset J^1M \) holonomic with respect to the variables \((q_1, \cdots, q_n)\) if it admits a function \( g : M \to \mathbb{R} \) such that \( R = \text{Image} j^1g \). We say \( R \) is non-holonomic with respect to the variables \((q_1, \cdots, q_n)\) otherwise.

We will see that for the ideal gas, such a function indeed exists which we denote by \( S = S(U, V) \) of the variables \((U, V)\). By definition, we have the formula

\[ S = \int_{\Gamma} \rho \lambda \frac{\partial \rho}{\partial \nu_0} \]

Proposition 10.8. Let \( R_{\mathcal{F}, \mathcal{S}} \subset \mathbb{R}^5 \) be as above. Then \( R_{\mathcal{F}, \mathcal{S}} \subset \mathbb{R}^5 \) is holonomic with respect to \((U, V)\) and have the formula for the thermodynamic potential

\[ S(U, V) = \frac{3N}{2} \log \left( \frac{4\pi mU}{3N} \right) + (N + 1) \log \left( \frac{V}{C(N + 1)} \right) + \log(N!) + N \log 2\pi. \]

Proof. From the explicit formulae for \( U \) and \( V \) in (10.9), (10.11), we can invert the map \((\lambda_1, \lambda_2) \to (U, V)\) and get

\[ \lambda_1 = -\frac{3N}{2U}, \quad \lambda_2 = -\frac{N + 1}{V} \]

into (10.12). The formula for \( S \) as a single-valued function follows. \( \square \)

Since the conjugate variables of \( U, V \) are \( \frac{1}{T}, \frac{P}{T} \) by the first thermodynamic law (??), we have derived the equation

\[ \frac{1}{T} = \frac{\partial S}{\partial U} = \frac{3N}{2U}, \quad P = \frac{\partial S}{\partial V} = \frac{N + 1}{V} \]

which is equivalent to the well-known ideal gas equation.

Corollary 10.9 (Ideal gas equation). Let \( N \) be the number of particles in a piston of ideal gas. Then we have the relations

\[ U = \frac{3}{2} NT, \quad V = (N + 1) \frac{T}{P} \]  

(modulo a multiplication by constant).
Here we would like to highlight the fact that the equilibrium state $R_{F,S}$ has constant energy in the isotherm, and hence it has the form of a cylinder over the one dimensional curve in the $PV$-plane. In particular its $PV$-diagram undergoes no phase transition.

11. Phase transition, Maxwell construction and graph selector

In our construction, a relevant thermodynamic equilibrium is given by the Legendrian submanifold $R$ generated by the $F$-reduced entropy function $S_{\text{red}}: M^F \to \mathbb{R}$ where $M^F \to \mathbb{R}^2 = \mathbb{R}^2_{U,V}$ is a (infinite dimensional) symplectic fibration. The resulting Legendrian submanifold is not necessarily holonomic (or graph-like) in the given preferred variables. This indeed occurs for the case of the van der Waals gas for the preferred variables $(P,T)$. As a result, the aforementioned thermodynamic variable cannot be globally holonomic (at least so over the $PT$-plane or over $PU$-plane). An effort of correcting this deficiency is precisely the celebrated Maxwell construction in equilibrium thermodynamics the explanation of which is now in order.

11.1. Van der Waals model and its phase transition. The van der Waals gas is the model describing the gas whose assumptions are more rigid than the ideal gas. It needs two more restrictions dictated by 2 additional parameters $a, b$:

1. **Potential hypothesis**: There exists attractive potential between the particles with the constant $a$.
2. **Volume hypothesis**: Each particle has non-zero size with the internal volume of $b$.

Then the observables, volume and energy, should have the form of

\[
U(\vec{x}, \vec{p}, h) = \frac{1}{2m} \sum_i |p_i|^2 - \frac{aN^2}{C} \cdot \frac{1}{h}
\]

\[
V(\vec{x}, \vec{p}, h) = Ch - Nb
\]

for some constant $m$ and $C$.

Consider the new observables $X, Y$ defined as

\[
X = V + Nb = Ch \quad \text{(volume of container)}
\]

\[
Y = U + aN^2 \cdot \frac{1}{V + Nb} = \frac{1}{2m} \sum_i |p_i|^2. \quad \text{(kinetic energy)}
\]

Note that the conjugate variables $(\lambda_X, \lambda_Y)$ of $(X, Y)$ are related to those $(\lambda_U, \lambda_V)$ conjugate to $(U, V)$ by the equation

\[
dS - \lambda_X dX - \lambda_Y dY = dS - \lambda_U dU - \lambda_V dV
\]

where the integration of $(Ch - bN)^N$ is the reason why the volume of the space each particles available is the only $(Ch - bN)$ for fixed $h$. Therefore, we derive

\[
\lambda_X = \lambda_V + aN^2 \cdot \frac{1}{(V + Nb)^2}
\]

\[
\lambda_Y = \lambda_U.
\]
By the same kind of calculation as in the ideal gas case, we compute
\[
X = \frac{\partial w}{\partial \lambda_X} = -(N + 1) \frac{1}{\lambda_X} + bN \simeq NT \frac{1}{P + a \frac{N^2}{(V+N)^2}} + bN
\]
\[
Y = \frac{\partial w}{\partial \lambda_Y} = -\frac{3}{2} (N + 1) \frac{1}{\lambda_Y} \simeq \frac{3}{2} NT.
\]
Therefore, we derive the Van der Waals gas equation
\[
\begin{aligned}
P + a \frac{N^2}{(V+N)^2} V &= NT \\
U &= \frac{3}{2} NT - a \frac{N^2}{V^2}.
\end{aligned}
\tag{11.1}
\]
In terms of the effective volume of container as
\[
V_{\text{eff}} = Ch = V + Nb,
\]
the equation is equivalent to
\[
\begin{aligned}
(P + a \frac{N^2}{(V+N)^2}) V &= NT \\
U &= \frac{3}{2} NT - a \frac{N^2}{V^2}
\end{aligned}
\tag{11.2}
\]
where \( \overline{V} := V_{\text{eff}} \).

We now identify the associated thermodynamic potential relative to the preferred variable \((T, P)\) by rewriting the first thermodynamic law \(dU = TdS - PdV\) into the following form
\[
d(U + PV - TS) = -SdT + VdP.
\tag{11.3}
\]
The function \( \Upsilon : \mathbb{R}^5 \to \mathbb{R} \) defined by
\[
\Upsilon = U + PV - TS (= H - TS)
\tag{11.4}
\]
is nothing but the well-known Gibbs free energy (or available energy in Gibbs’ own term) \cite{Gib73}, where \( H := U + PV \) is called the enthalpy in thermodynamics. This \( \Upsilon \) will play the role of the thermodynamic potential associated to the preferred thermodynamic variables \((P, T)\). (We avoid using the common alphabet ‘G’ to write the Gibbs free energy not to confuse readers here since the same alphabet \( G \) appears in the graph of Figure 2 where we use the same alphabets as Maxwell did in \cite{Max75}.)

**Question 11.1.** Can we express the equilibrium Legendrian submanifold \( R_{T,S} \) as the image of one-jet graph of differentiable function \( f : \mathbb{R}^2(T,P) \to \mathbb{R} \), i.e., such that
\[
R_{T,S} = \{ (df(p), p, f(p)) \mid p \in \mathbb{R}^2(T,P) \}?
\]
I.e., is \( R_{T,S} \) is holonomic with respect to the variables \((T, P)\)?

We will now illustrate by the van der Waals model that such a task is not possible in general, and that the Maxwell construction is precisely the outcome of Maxwell’s effort to overcome this deficiency. It is easy to show that the change of coordinates
\[
(U, V, T, P, S) \mapsto (T, -S, V, P, \Upsilon)
\]
preserves the thermodynamic contact structure
\[
\xi := \ker \left( dS - \frac{1}{T} dU + \frac{P}{T} dV \right) (= \ker (dU - TdS + PdV)).
\]
With this preparation, we are now ready to do the analysis of van der Waals model. When \( T = T_0 \) is given, the first equation of (11.1) becomes a cubic equation of \( V \),

\[
(P + a \frac{N^2}{V^2}) (\bar{V} - bN) = NT_0.
\]

When \( P = P_0 \) is given in addition to \( T \), we have the cubic equation of \( \bar{V} \)

\[
\bar{V}^3 + a\bar{V}^2 + \beta \bar{V} + \gamma = 0 \tag{11.5}
\]

with

\[
\alpha = -\left(bN + \frac{T_0}{P_0}\right), \quad \beta = \frac{aN^2}{P_0}, \quad \gamma = -\frac{abN^3}{P_0}.
\]

The second equation of (11.2) is of the form

\[
U = -\frac{aN^2}{V^2} + \frac{3}{2} NT_0. \tag{11.6}
\]

Remark 11.2. A fundamental difference of the van der Waals equation (11.6) from the ideal gas equation (10.13) is that the energy \( U \) is no longer constant in the isotherm of the \( PV \)-diagram.

The cubic equation (11.5) shows that there is a phase transition in the \( PV \)-diagram at the critical temperature \( T = T_c \) at which the discriminant

\[
D = 18\alpha \beta \gamma - 4\alpha^3 \gamma + \alpha^2 \beta^2 - 4\beta^3 - 27\gamma^2 \tag{11.7}
\]

of the cubic polynomial vanishes: When \( T < T_c \), (11.5) has 3 distinct roots, while when \( T > T_c \), it has a unique root.

11.2. Metamorphosis of the isotherms in the \( PV \)-diagram. In thermodynamic equilibrium, a necessary condition for stability is that pressure \( P \) does not increase with volume \( V \), i.e., that volume should decrease as pressure increases. As we see in Figure 1, this fails when temperature \( T \) is below the critical temperature \( T_c \). The Maxwell construction is a way of correcting this deficiency by considering a continuous thermodynamic potential function, if we can.

The above two cases can be also differentiated by the geometry of associated equilibria whose explanation is now in order.

For this purpose, we first introduce the notion of graph selector of a Legendrian submanifold \( R \subset J^1 M \). The following definition is the Legendrian analog in the 1-jet bundle \( J^1 M \) to the notion of graph selectors for the Lagrangian submanifolds in the cotangent bundle \( T^* M \). (See [Arn10], [AOOdS18] for the definition of symplectic graph selectors.)

**Definition 11.3** (Legendrian graph selector). Let \( N \) be a smooth manifold and \( \iota : N \rightarrow J^1 M \) be a compact, Legendrian embedding into \( J^1 M \). Denote by \( R \) its image Legendrian submanifold. A graph selector of \( R \subset J^1 M \) is a Lipschitz function \( f : M \rightarrow \mathbb{R} \) such that \( f \) is differentiable on a dense open set \( U \subset M \) of full measure and for all points \( q \in U \) we have

\[
(q, df(q), f(q)) \in R.
\]

With this definition in our disposal, we can differentiate the two cases as follows.

**Theorem 11.4.** Let \( T_c \) be the aforementioned critical temperature. Then

1. When \( T > T_c \), \( R_{\bar{\tau},S} \) is a 1-jet graph of a function \( f : \mathbb{R}^2_{P,T} \rightarrow \mathbb{R} \).
(2) When $T < T_c$,

(a) such a globally single-valued function does not exist.

(b) Maxwell construction provides a graph selector in the sense of Definition 11.3 over the PT-plane, i.e., regarding $(P, T)$ as the variables.

\textbf{Proof.} Statement (1) is obvious and so we focus on Statement (2).

In the latter case, let $T_0$ be any temperature $0 < T_0 < T_c$. From the first equation of (11.2)

\[ \left( P + a \frac{N^2}{V^2} \right) (\bar{V} - bN) = k_B N T_0, \]

we obtain the pressure formula

\[ P = -a \frac{N^2}{V^2} + \frac{k_B N T_0}{V - bN}. \] (11.8)

There exists a unique critical pressure $P_c = P_c(T_c)$ such that

- when $P > P_c$, (11.5) has a unique solution $\bar{V}$,
- when $P < P_c$, it has 3 distinct values of $\bar{V}$.

(See Figure 1)

For each $0 < T < T_c$, there is a unique pressure denoted by $P_{mx} = P_{mx}(T)$ which we call Maxwell pressure at which Maxwell equal area law holds. (See Figure 2)
Now we will attempt to find a function of the type \( f = f(P, T) \) so that
\[
\left( T, P, \frac{\partial f}{\partial T}(T, P), \frac{\partial f}{\partial P}(T, P), f(T, P) \right) \in R_{\mathcal{F}, S} \tag{11.9}
\]
explicitly by integration regarding the \( PV \)-diagram as the differential of a multi-valued function whose differential is determined by the Legendrian submanifold \( R_{\mathcal{F}, S} \) (See [Arn10, AOOdS18] for such a procedure.): Regard \( \mathbb{R}^2_{(P, V)} \) as the cotangent bundle \( T^*\mathbb{R}_P \) and then the isotherm in the \( PV \)-diagram can be regarded as the Lagrangian submanifold
\[ (P, V) \in \mathbb{R}^2_{(P, V)} \cong T^*\mathbb{R}_P \]
where \( \alpha = V dP \in T^*_P \mathbb{R}_P \).

The occurrence of this phase transition is ultimately tied to the thermodynamic stability that pressure \( P \) should not increase with volume \( V \), i.e., that volume should decrease as pressure increases. To see how the thermodynamic stability fails, we consider the commutative diagram of the projections
\[
\mathbb{R}^2_{(V, P)} \leftarrow R_{\mathcal{F}, S} \cap \{ T = T_0 \} \rightarrow R_{\mathcal{F}, S}
\]
\[
\mathbb{R}_P \leftarrow \mathbb{R}_{(P, T_0)} \rightarrow \mathbb{R}^2_{(P, T)}
\]
and try to find a single-valued section \( \Upsilon = \Upsilon(P, T) \) of the RHS projection. The \( PV \)-diagram is the union of the projections of the top middle term to the top left plane \( \mathbb{R}^2_{(V, P)} \) as \( T_0 \) varies.

11.3. Maxwell construction and non-differentiable thermodynamic potential. For the simplicity of notation for the following discussion, we will just write \( V \) for \( V^e \) with \( V \) representing the effective volume of container.

We then consider the front projection
\[
R_{\mathcal{F}, S} \rightarrow \mathbb{R}^3_{(T, P, S)}: (U, T, V, P, S) \rightarrow (T, P, S)
\]
and followed by the projection \( \mathbb{R}^3_{(T, P, S)} \rightarrow \mathbb{R}^2_{(T, P)} \). We call the latter composition \( R_{\mathcal{F}, S} \rightarrow \mathbb{R}^2_{(T, P)} \) the base projection. At each fixed time \( T \), we would like to express the base projection of the isotherm of the equilibrium \( R_{\mathcal{F}, S} \) as the graph \( (P, g_T(P)) \) of a single-valued function \( V = g_T(P) \), if possible.

Recalling that \( P \) and \( V \) are conjugate variable, we make the identification
\[
\mathbb{R}^2_{(V, P)} \cong T^*\mathbb{R}_P
\]
so that the \( V \)-coordinate is the one for the coordinate representation of the \( \alpha \in T^*_P \mathbb{R}_P \) as
\[
\alpha = V dP, \quad V = \frac{\partial f_T}{\partial P}(P) \tag{11.10}
\]
for a germ \( g \) of functions at \( P \). As seen in the isotherm pictured around the shaded region in Figure 2, it has 3 different branches of the graph seen horizontally.

Here we attempt to construct the function \( f_T \) by assigning the value \( f_T(P) \) of \( f_T \) as follows: Denote by \( C^{\text{min}} \) and \( C^{\text{max}} \) be the minimal and the maximal branches
Figure 2. Maxwell equal area law

of the isotherm in the $PV$ diagram for the region $P > P_{mx}(T)$ and $P < P_{mx}(T)$ respectively. We express

$$C_{\text{min}} = \{(P, \alpha^-(P)) \mid P > P_{mx}\}$$
$$C_{\text{max}} = \{(P, \alpha^+(P)) \mid P < P_{mx}\}.$$

Then we define

$$f_T(P) = \begin{cases} 
- \int_P^{\infty} \alpha^-(p) \, dp & \text{for } P > P_{mx} \\
- \int_{P_{mx}}^{\infty} \alpha^-(p) \, dp - \int_{P_{mx}}^{P} \alpha^+(p) \, dp & \text{for } P < P_{mx}
\end{cases}$$

where the subscript ‘mx’ stands for ‘Maxwell’. It follows from Maxwell’s equal area condition that this function $g_T$ continuously extends across the Maxwell pressure point $P = P_{mx}(T)$. (See Figure 3.)

By varying $T$, this determines $V$ as a continuous function of $(T, P) \mapsto f_T(P)$. Then substituting we determine $U = U(P, T)$ from (11.1).

Next we determine the thermodynamic potential $S = S(P, T)$ by integrating the equation of the thermodynamic first law

$$dS = \frac{1}{T} \, dU + \frac{P}{T} \, dV$$

with $T$ fixed: the equation is integrable since the function $V = V(P, T)$ and $U = U(P, T)$ are Lipschitz functions and hence their derivatives are bounded. The resulting integral defines a function that is continuous even across the Maxwell
pressure hypersurface

\[
\{(U, V, T, P, S) \mid P = P_{mx}(T)\}
\]

by the equal area law.

This enables us to express Gibbs free energy \( \Upsilon \) on the equilibrium \( R_{F;S} \) as a continuous function \( \Upsilon = f(T, P) := f_T(P) \) where \( f \) is continuous but not differentiable across the curve \( \{(T, P) \mid P = P_{mx}(T)\} \). The derivative \( \frac{\partial f_T}{\partial P} \), which we know from (11.10) is the same as the volume \( V \), jumps on \( P = P_{mx}(T) \) for \( 0 < T < T_c \).

(See the highlighted part of Figure 4.) This finishes construction of a graph selector of the equilibrium Legendrian submanifold \( R_{F;S} \) over the \( PT \)-plane, i.e., regarding \( (P, T) \) as the preferred variables.

\[\square\]

The image of the one-jet map \( j^1 f_T : \mathbb{R}^2_{(T,P)} \to \mathbb{R}^5 \) has its image contained in \( R_{F;S} \) by construction where it is defined and satisfies \( U \, dT + V \, dP = df_T(T, P) \). It has the following properties (See Figure 4):

1. its cotangent projection contains jump-discontinuity across the set \( \{P = P_{mx}\} \).
2. the function \( f \) is defined on the open dense subset

\[\pi_{(T,P)}(R_{F;S} \setminus \{P \neq P_{mx}(T), 0 < T < T_c\})\].

3. the function, however, continuously extends across the discontinuity locus

\[\{(T, P) \mid P = P_{mx,T}, T \in \mathbb{R}_+\}\].
(4) The aforementioned jump discontinuity can be canonically filled in by the Maxwell construction on \( \{ P \neq P_{mx}(T), 0 < T < T_c \} \) using the equal area law.

**Definition 11.5** (Maxwell adjustment). Denote by \( R^{mx}_{T;S} \) the above adjustment performed on the Legendrian submanifold \( R_{T;S} \) the Maxwell adjustment of the thermodynamic equilibrium. We denote by \( R^{mx}_{T;S} \) the Maxwell adjustment.

(See [Oh12, Section 3 & 4] for detailed construction of the general Maxwell adjustment and its explanation in general, and Appendix [C] for a summary.)

**Remark 11.6.** (1) The certain jump of the volume \( V = \frac{\partial f_T}{\partial T} \) at the Maxwell pressure \( P = P_{mx} \) maintaining continuity of the potential \( \Upsilon = H - TS = U + PV - TS \) across the pressure has the following thermodynamic interpretation. The state whose pressure is above that of the equilibrium \( R_{T;S} \) describes a liquid while that below the Maxwell pressure describes vapour. Maxwell [Max75] himself described his construction of replacing the middle sinusoidal part by the horizontal straight line in Figure 4 by saying “...... Since the temperature has been constant throughout, no heat can have been transformed into work. Now the heat transformed into work is represented by the excess of the area FDE over BCD. Hence the condition which determines the maximum pressure of the vapour at given temperature is that the line BF cuts off equal areas from the curve above and below.”
(2) It is worthwhile to point out that the Maxwell construction is the simplest scenario for which existence of such a continuous single-valued selector can be directly established by simple integration. In the more complex system, establishing existence of such a single-valued selector is a nontrivial problem in general, as illustrated by [Oh97]: The proof of existence of such a single-valued selector involves either the Floer theory or the stable Morse theory in symplectic geometry.

(3) We would like to emphasize that the $C^0$-holonomicity depends on the choice of preferred variables. For example, van der Waals model equilibrium $R_{\mathcal{F}}; S$ considered above can be checked to be holonomic with respect to the original variable $(U, V)$ used in the entropy representation for which the entropy $S$ is the relevant thermodynamic potential, instead of the Gibbs free energy. (See Proposition 10.8. We also refer readers to the discussion for the van der Waals model around Equations (72)–(75) of [Bra19].)

12. Discussion

We would like to mention that we have not touched upon the time evolution or the dynamical point of view of nonequilibrium thermodynamics focusing mostly on the geometric structure and the kinematical aspect thereof. For example, we have not touched upon anything about Boltzmann analysis of involving the evolution equations leading to equilibrium thermodynamics as $t \to \infty$.

The most common form of the Boltzmann equation has the form

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = Q(f, f)$$

in the position-velocity representation by $(x, v) \in \mathbb{R}^K \times \mathbb{R}^K$, where $Q$ is the quadratic Boltzmann collision operator,

$$Q(f, f) = \int_{\mathbb{R}^K} \int_{S^{K-1}} (f' f_s' - f f_s) B(v - v_s, \sigma) \, d\sigma dv_s$$

where $S^{K-1} \subset \mathbb{R}^K$ is the $K - 1$ dimensional unit sphere and $d\sigma$ is its standard volume form. (see [Vil02 Equation (3)] for the meanings of the unexplained notations $f', f_s$, $f_s'$ and the Boltzmann’s collision kernel $B = B(v - v_s, \sigma)$.

There are recent articles by Esen-Grmela-Pavelka [EGP22a, EGP22b] which investigate the role of geometry in statistical mechanics and thermodynamics in the themes somewhat similar to those studied in the present paper. In addition, they unravel some interesting contact geometric aspect and its role in the dynamical perspective. In particular, they put the full kinetic equation, General Equation for Non-Equilibrium Reversible-Irreversible Coupling (GENERIC), involving both Hamiltonian structure and gradient structure into the framework of evolution Hamiltonian dynamics, i.e., dynamics generated by evolution Hamiltonian vector field [SdLLVdD20]. In the course of their investigation, Esen-Grmela-Pavelka also consider some mesoscopic dynamical theory in [EGP22b Section 3.3], a bridge between the statistical mechanics and the thermodynamics, as some sort of reduction, which they call the hierarchy reduction. In [EGP22a], then write down a very general form of the kinetic equation, which they acronym GENERIC: Write $\rho = f \nu_0$ with $f = \frac{\partial \rho}{\partial \nu_0}$. Then GENERIC has the following elegant form

$$\dot{f} = \{f, E\} + \nabla_\beta \Xi |_{\beta = \delta_f} S$$
E is the energy of the system, S is the Boltzmann entropy and \( \Xi = \Xi(f, \beta) \) is a general dissipation potential defined on \( T^*\mathcal{P}(\Gamma) \), which satisfies a list of properties [EGP22b, Equation (9)] so that its evolution leads to a thermodynamic equilibrium as \( t \to \infty \).

It seems that our two-step process of reduction, following our two mottos we take, are related to the Hamiltonian dynamics and the gradient dynamics of their GENERIC framework respectively. It remains to see if the study of the present article would give some new light in the study of the kinetic equation, e.g., of Boltzmann’s equation, and clarifying the relationship between our study and that of [EGP22a, EGP22b] seems to the first step towards that direction.

**Appendix A. Information entropy**

The information entropy is firstly introduced by C. E. Shannon, under the attempt to research the communication in terms of mathematics [SW49]. Consider the probability measure space \((X, \mathcal{B}, \mu)\). We want to construct the function \( I \) that measures the uncertainty of the information saying

“\( x \) is in some ‘event’ \( A \in \mathcal{B} \)”

for any measurable subset \( A \in \mathcal{B} \). The conditions for the information measure \( I : \mathcal{B} \to \mathbb{R} \) should be followings:

1. \( I(A) \) is a function of \( \mu(A) \)
2. \( I(A) \geq 0 \)
3. The smaller \( \mu(A) \) is, the larger \( I(A) \) is.
4. If \( A, B \) are independent, then \( I(A \cap B) = I(A) + I(B) \)

The condition (3) means that the uncertainty of the fact “\( x \in A \)” increases as the size of \( A \) decreases. Intuitively, it reflects conviction that the statement “I am an athlete” gives less information than “I am an NBA player”.

The condition (4) is from the ansatz that the two independent statements connected with ‘and’ give the same information obtained by considering the each statement individually. Roughly speaking, consider the statement “I’m an NBA player and he is an NFL player” This statement just gives no more or less information from the two fact “I’m an NBA player” and “He is an NFL player”.

**Lemma A.1.** *The increasing continuous function \( f : \mathbb{R}_+ \to \mathbb{R} \) satisfying \( f(x) + f(y) = f(xy) \) should be \( f(x) = -\log x \).*

By this lemma, the unique choice of \( I \) satisfying condition (1)–(4) should be \( I(A) = -\log \mu(A) \). Therefore, we define the information measure function.

**Definition A.2.** Let \((X, \mathcal{B}, \mu)\) be a probability space. and \( \mathcal{A} \) be a subalgebra of \( \mathcal{B} \). Then, the information measure \( I_\mu \) is defined by

\[
I_\mu(A) = -\log \mu(A)
\]

and the expected information measure \( \tilde{I}(\mu) \) is defined by

\[
\tilde{I}_\mu(A) = \mu(A)I_\mu(A) = -\mu(A)\log \mu(A)
\]

for all \( A \in \mathcal{A} \).
Suppose that $X$ is a finite set of events. Then, it is easy to define the expected information on probability distribution $\mu$, i.e.

$$S(\mu) = -\sum_{q^i \in X} \mu(q^i) \log \mu(q^i).$$

However, if $X$ is infinite, the summation should become an integral, but the problem is that we do not have any measure on $X$ except the probability measure $\mu$. Also, the logarithm term cannot be defined since $\mu$ does not give any pointwise value without comparing with other reference measure. Therefore, we should consider a reference measure $\nu_0$ in order to define the entropy of a probability [Jay57]. (This is the Liouville measure in the present article.)

**Definition A.3.** Let $(X, B, \nu_0)$ be a measure space and $\rho$ be a probability measure which is absolutely continuous with respect to $\nu_0$. Then, the **relative entropy of $\rho$** is defined by

$$S(\rho) = -\int_X \rho \log \frac{d\rho}{d\nu_0}$$

where $\frac{d\rho}{d\nu_0}$ is the Radon-Nikodym derivative $\rho$.

**APPENDIX B. RELATIVE INFORMATION ENTROPY AND LEGENDRIAN GENERATING FUNCTION**

In this Appendix, we recall the definition of **Legendrian generating function** and explain what we mean by the statement “Relative information entropy in SPS is a generating function of thermodynamic equilibrium”. This framework is motivated by the following Weinstein’s framework he observed on the role of the classical action functional in the representation of Hamiltonian deformations of the zero section of the cotangent bundle $T^*M$. To put our framework in some perspective, we recall Weinstein’s observation.

The standard generating function of a Lagrangian submanifold in the cotangent bundle $T^*B$ is given by a vector bundle $\pi: E \to B$ equipped with a function $S: E \to \mathbb{R}$. We will not repeat its definition in the symplectic context but refer readers to the contact case below, except that we make the following remark.

**Remark B.1.** Traditionally generating function has been a useful tool in the study of symplectic Hamiltonian dynamics in relation to the study of existence question on the periodic orbits and of completely integrable system. More recently it has been used to investigate some homotopical aspect of symplectic topology. (See [Kra17, Kra18] for example.) We anticipate the Legendrian generating functions will play similar role in the study of contact topology.

Laudenbach-Sikorav [LS85] and Sikorav [Sik87] proved that such Lagrangian submanifold admits a generating function on a finite dimensional vector bundle $\pi: E \to B$ by considering an approximation of the classical action functional on the set of piecewise smooth paths which approximates Hamiltonian trajectories issued at the zero section $\sigma_{T^*B}$. The procedure replaces geodesics in Bott’s geodesic approximations of smooth paths [Bot59] by the Hamiltonian paths. Then Weinstein observed that instead of considering this approximation, one may directly put it into the following infinite dimensional framework on the general fibration, not just on the vector bundle.
Lemma B.2 (Weinstein’s lemma, [Wei]). Let $H = H(t,x)$ be a time-dependent Hamiltonian on the cotangent bundle $T^*B$. Then the classical action functional
\[ A_H(\gamma) = \int \gamma^* \theta - H(t, \gamma(t)) \, dt \]
defined on the path space
\[ \mathcal{L}_0(T^*B, 0_{T^*B}) = \{ \gamma : [0,1] \to T^*B \mid \gamma(0) \in 0_{T^*B} \} \]
is a ‘generating function’ of the time-one image $\phi^1_H(0_{T^*B})$ of the zero section under the Hamiltonian flow of $H$.

More precisely, consider the diagram
\[ \mathcal{L}_0(T^*B, 0_{T^*B}) \xrightarrow{A_H} \mathbb{R} \]
\[ \downarrow \pi \circ \text{ev}_1 \]
\[ B \]
where $\text{ev}_1 : \mathcal{L}_0(T^*B, 0_{T^*B}) \to T^*B$ is the evaluation map
\[ \text{ev}_1(c) := c(1), \quad \text{for } c \in \mathcal{L}_0(T^*B, 0_{T^*B}). \]

Then the content and its proof of the above Weinstein’s lemma go as follows:
- First solve the fiberwise (or vertical) critical point equation. This provides the equation of motion, Hamilton’s equation $\dot{x} = X_H(t,x)$.
- Then push forward the image of the differential $dA_H$ by the map $\pi \circ \text{ev}_1$ to the cotangent bundle $T^*B$. This provides the final location of the particle.
- The resulting push-forward image is precisely $\phi^1_H(0_{T^*B})$. This provides the final momentum of the particle.

Now we recall the formal definition of Legendrian generating functions in the context of contact geometry. We refer to [San11] for a good introduction thereto in relation to the purpose of the present section.

Let $B$ be a manifold and $J^1B = T^*B \times \mathbb{R}$ be the one-jet bundle. We equip $J^1B$ with the contact form
\[ \lambda = dz - p dq =: dz - \pi^* \theta \]
where $(q,p,z) =: y$ be the canonical coordinates of $J^1B$. We also denote $x = (q,p) \in T^*B$.

We denote by $\pi : E \to B$ a finite dimensional vector bundle.

**Definition B.3.** We say a Legendrian submanifold $R \subset J^1B$ is generated by a function $S : E \to \mathbb{R}$ if we can express
\[ R = \{ (q,p,z) \mid p = D^h S, \ z = S(q,e), \ d^e S(q,e) = 0 \}. \]

When this holds, we say $S$ is a generating function of the Legendrian submanifold $R$.

Now we apply the above discussions to our diagram (5.4)
\[ \mathcal{M}^B \xrightarrow{S_{red}} \mathbb{R} \]
\[ \pi_f \downarrow \]
\[ 0_{\mathcal{G}_B^*} \]
This diagram is in the same infinite dimensional spirit as Weinstein’s diagram (B.1). Here we break down the procedure of determining the thermodynamic equilibrium and the associated probability distribution:

1. (Finite dimensional reduction) The fiberwise (or vertical) critical point equation provides the equation of motion, which is the constrained extremization problem

\[
d^v(S_{\text{red}}^F)(\rho) = 0.
\]

2. (Generating Lagrangian shadow) Then we push forward the image of the differential \(dS_{\text{red}}^F\) by the map \(\pi_F\) to the cotangent bundle \(T^*g^*_F\).

3. (Generating thermodynamic equilibrium) The canonical lifting to \(J^1g^*_F\) provides a thermodynamic equilibrium state which satisfies the first law of thermodynamics.

4. (Determination of equilibrium probability distribution) The full extremization problem over all observation data which is equivalent to the full critical point equation of \(S_{\text{red}}^F\)

\[
d^v(S_{\text{red}}^F)(\rho) = 0
\]

is now reduced to the finite dimensional problem of solving

\[
D^h(S_{\text{red}}^F)(\rho) = 0 \quad \text{on} \quad R^*F \; ; \; S.
\]

Some thermodynamic interpretation of each step of the above is now in order. What we have shown in Subsection 6 is the statement that the \(F\) iso-data mesoscopic KTPS \(E_{FS}/G\) is the collection of states after performing the finite dimensional reduction of solving \(d^v(S_{\text{red}}^F) = 0\): This corresponds to the above steps (1) and (2).

Step (3) above corresponds to the step of mapping \(E_{FS}/G_{O_S}\) into the finite dimensional TPS \(J^1g^*_F\) in the way that the points in the image \(R_{FS}\) of the map \(pr\) satisfy the first thermodynamic law \(dw - \sum_{i=1}^n p_i dq_i = 0\). By the dimensional reason \(R_{FS}\) becomes a Legendrian submanifold. This is the step (3) above.

When \(R_{FS}\) is holonomic, in which case \(R_{FS} = \text{Image} \; j^1f\) for some smooth function \(f : g^*_F \rightarrow \mathbb{R}\), the finite-dimensional reduction (B.3) becomes a finding a critical point of the relevant thermodynamic potential \(f\). Otherwise one has to solve the extremization problems of a Lipschitz continuous, not necessarily differentiable function, after applying the Maxwell construction and selecting a single-valued branch of the non-holonomic equilibrium state \(R_{FS}\). This process will be explained in more detail with the van der Waals model later in Section 11.

APPENDIX C. GENERAL CONSTRUCTION OF MAXWELL ADJUSTMENT

In this section, we recall the details of construction of Maxwell adjustment from [Oh12] (see also [Arn10] and [AOOdS18]) for a general compact exact Lagrangian submanifold in the cotangent bundle \(T^*N\). It is proved in [Oh97], [AOOdS18] that such a Lagrangian submanifold carries a canonical graph selector \(f_L\) called the basic phase function, and that the structure of its differential \(\sigma_F\), called the Lagrangian selector, is analyzed in [Oh12]. Such an existence of graph selector has been constructed in [OY23] for a general compact Legendrian submanifold contact isotopic to the zero section of one-jet bundle \(O^*\).

This being said we introduce the following general definition motivated by this.

**Definition C.1.** Assume \(N\) is a closed manifold. Let \(R \subset J^1N\) be any compact smooth Legendrian submanifold projecting surjectively to \(N\). We call (a densely defined) single-valued function \(f : N \rightarrow \mathbb{R}\) a graph selector and its differential...
σ := df of $T^*N$ with values lying in $L = \pi_{T^*N}(R)$ its associated Lagrangian selector of $R$.

$R$ carries two projections, the projection of $R$ to $N \times \mathbb{R}$ called the front projection $\pi_{\text{front}}$ and projection of $R$ to $T^*N$, called the Lagrangian projection. We denote by $L$ the image of Lagrangian projection or cotangent projection.

We make the following hypothesis for the further discussion below, which we know holds for the Legendrian submanifold contact isotopic to the zero section of $J^1N$ from [OY23].

**Hypothesis C.2.** We assume that the projection of $R$ to $N \times \mathbb{R}$ is surjective and that $R$ admits a graph selector $f : N \to \mathbb{R}$ and denote by $\sigma := df$ the associated Lagrangian selector.

The differential $\sigma$ is not a continuous map but is a well-defined current. Under this hypothesis, the construction given in [Oh12, Section 3 & 4] immediately generalizes to such Legendrian submanifolds as follows.

For a given graph selector $f : N \to \mathbb{R}$ given by Hypothesis C.2, we denote by $S(\sigma)$ the singular set of the current $\sigma$. Then we consider the open subset of $R$

$$\Sigma_R := \{(q, z) \in J^1N \mid q \in N \setminus S(\sigma_R), p = \sigma_R(q), (q, p, z) \in R\} \subset R.$$  

The projection $\pi_R : R \to N$ restricts to a one-one correspondence on $\Sigma_R \subset R$, and the smooth function

$$f|_{N_0} : N_0 \to \mathbb{R}$$
continuously extends to $\overline{N_0} = N$ where $N_0 = N \setminus S(\sigma)$. By construction, we have the bound

$$|df_R(q)| \leq \max_{x \in R} |p(x)|$$

for any $q \in N_0$, where $x = (q(x), p(x), z(x))$ and the norm $|p(x)|$ is measured by any given Riemannian metric on $N$.

The general structure theorem of the wave front (see [Eli87]) proves that the section $\sigma$ is a differentiable map on a set of full measure for a generic choice of $R$ which is, however, not necessarily continuous.

We note that the singular locus $S(\sigma) \subset N$ is a subset of the bifurcation diagram of the Legendrian submanifold $R$. The bifurcation diagram is the union of the caustic and the Maxwell set where the latter is the set of points of which the different branches of the generating function merge. (See [Giv90] Section 4 for the definition of bifurcation diagram of Lagrangian submanifold $L \subset T^*N$ which can be generalized to the Legendrian submanifold of the type we consider here in general.)

For a generic $R$, $S(\sigma_R)$ is stratified into a finite union of smooth submanifolds

$$\bigcup_{k=1}^n S_k(\sigma_R), \quad S_k(\sigma_R) = \text{Sing}_k(\sigma_R), \quad n = \dim N$$

(see [Arn72, Eli87, Giv90] e.g., for such a result) so that its conormal variety $\nu^*S(\sigma_R)$ can be defined as a finite union of conormals of the corresponding strata. Each stratum $S_k(\sigma_R)$ has codimension $k$ in $N$. The stratum for some $k$ could be empty. In $\dim N = 2$, there are two strata to consider, one $S_1(\sigma_R)$ and the other $S_2(\sigma_R)$.

For $k = 1$, $S_1(\sigma) \subset N$ is a hypersurface and each given point $q \in S_1(\sigma_R)$ has a neighborhood $U(q) \subset N$ such that $U(q) \setminus S_1(\sigma_R)$ has two components. We also
note that $\sigma_R$ carries a natural orientation induced from $N$ by projection when $N$ is orientable. When $N$ is oriented, $S_1(\sigma_R)$ is also orientable as a finite union of smooth hypersurface. We fix any orientation on $S_1(\sigma_R)$.

We denote by $U^\pm(q)$ the closure of each component of $U(q) \setminus S_1(\sigma_R)$ in $U(q)$ respectively. Here we denote by $U^+(q)$ the component whose boundary orientation on $\partial U^+(q)$ coincides with that of the given orientation on $S_1(\sigma_R)$ and by $\partial U^-(q)$ the other one. Then each of $U^\pm(q)$ is an open-closed domain with the same boundary

$$\partial U^\pm(q) = U(q) \cap S_1(\sigma_R).$$

Denote

$$\sigma_R^\pm(q) = \lim_{p^\pm \to q} \sigma_R(p^\pm)$$

(C.2)

obtained by taking the limit on $U^\pm(q)$ respectively. The limits are well-defined from the definition of $\sigma_R$ since $R$ is a smooth manifold.

The following theorem is proved in [Oh12]. See also [Giv90, ZR92] for a related statement.

**Proposition C.3** (Theorem 4.1 [Oh12]). Let $q \in S_1(\sigma_R)$. Then

$$\sigma_R^-(q) - \sigma_R^+(q) \in T_q^* N,$$

is contained in the conormal space $\nu_q^*[S_1(\sigma_R); N] \subset T_q^* N$.

The boundary orientations of the two components arising from that of $\Sigma_R$, which in turn is induced from that of $N$ via the cotangent projection have opposite orientations. We call the one whose projection to $S_1(\sigma_R)$ coinciding with the given orientation the upper branch and the one with the opposite one the lower branch and denote them by

$$\partial^+ \Sigma_R, \partial^- \Sigma_R$$

respectively.

Now let $\ell_q$ be the line segment connecting the two vectors $\sigma_R^\pm(q)$, i.e.,

$$\ell_q : u \in [0, 1] \mapsto \sigma_R^+(q) + u(\sigma_R^-(q) - \sigma_R^+(q)) \subset T_q^* N.$$  (C.3)

This is an affine line in the vector space $T_q^* N$ that is parallel to the conormal space $\nu_q^*[S_1(\sigma_R)]$. Therefore the union

$$\Sigma_R[\cdot;+] := \bigcup_{q \in S_1(\sigma_R)} \ell_q$$  (C.4)

is contained in the translated conormal space

$$\sigma_R^+ + \nu_q^*[S_1(\sigma_R); N]$$  (C.5)

Here the bracket $[\cdot;+]$ stands for the line segment $\ell_q$, and $\nu_q^*[S_1(\sigma_R); N]$ is the conormal bundle of $S_1(\sigma_R)$ in $N$. We would like to point out that since $\sigma_R^+(q) - \sigma_R^-(q) \in \nu_q^*[S_1(\sigma_R); N]$ we have the equality

$$\sigma_R^+(q) + \nu_q^*[S_1(\sigma_R); N] = \sigma_R^-(q) + \nu_q^*[S_1(\sigma_R); N]$$

for all $q \in S_1(\sigma_R)$. Therefore we can simply write (C.5) as

$$\sigma_R + \nu_q^*[S_1(\sigma_R); N]$$  (C.6)

unambiguously. The following definition was called the basic Lagrangian selector chain in [Oh12] as a singular chain.
Definition C.4. We denote by $L^\text{mx}_{\sigma}$ the Lagrangian chain
\[ \Sigma_R \subset T^*N \] (C.7)
with the orientation given as above.

The two components of $\partial \Sigma_R$ associated to each connected component of $S_1(\sigma_R)$ are the graphs of $\sigma_R^\pm$ for the functions $f_R^\pm$ near $S_1(\sigma_R)$. Again we regard $\sigma_R^\pm$ as singular chains.

Note that each connected component of $S_1(\sigma_R)$ gives rise to two components of $\partial \Sigma_R$. We can bridge the ‘cliff’ between the two branches of $\partial \Sigma_R$ over each connected component of $S_1(\sigma_R)$.

Definition C.5 (Cliff wall chain). We define a ‘cliff wall’ chain $\sigma_{R;[-+]}$ whose support is given by the union
\[ \Sigma_{R;[-+]} = \bigcup_{q \in S_1(\sigma_R)} \ell_q \]
Then we define the chain $\Sigma_{R;[-+]}$ similarly as we define $\Sigma_R$ by taking its closure in $T^*N$.

By definition, its tangent space at $x = (q, u)$ has natural identification with
\[ T_x \Sigma_{F;[-+]} \cong \nu_q^* S_1(\sigma_R) \oplus T_q S_1(\sigma_R). \]
Thanks to Proposition C.3, it carries a natural direct sum orientation
\[ o_{\Sigma_{R;[-+]}(q)} = \{ df_R^-(q) - df_R^+(q) \} \oplus o_{S_1(\sigma_R)}(q). \]
Therefore $\Sigma_{R;[-+]}$ carries a natural orientation and defines a singular chain. (Under the natural identification of $T_q N$ with $T_q^*N$ by the dual pairing, which induces an identification
\[ \nu_q^* S_1(\sigma_R) \oplus T_q S_1(\sigma_R) \cong \nu_q S_1(\sigma_R) \oplus T_q S_1(\sigma_R) \]
as an oriented vector space.) Then we have the relation
\[ \partial \Sigma_R = -\partial \Sigma_{R;[-+]} \] (C.8)
along the intersection $\partial \Sigma_R \cap \partial \Sigma_{R;[-+]}$. The above discussion leads us to the following

Proposition C.6 (Maxwell adjustment). Suppose a compact Legendrian submanifold $R \subset J^1N$ satisfies Hypothesis C.2 and let $f : N \rightarrow \mathbb{R}$ be a graph selector. We consider the union
\[ L^\text{mx}_f := \Sigma_R \cup \Sigma_{R;[-+]} \]
Then the set
\[ R^\text{mx}_f := \{(q, p, z) \in J^1N \mid (q, p) \in L^\text{mx}_f, \ z = f(q)\} \]
defines a Legendrian cycle. We call $R^\text{mx}_f$ the Maxwell adjustment of the Legendrian submanifold $R$ associated to $f$.

This finishes our discussion on the Maxwell construction for the general Legendrian submanifold in the one-jet bundle $J^1N$ of general closed manifold $N$. 
References

[AM78] R. Abraham and J. Marsden, *Foundations of mechanics*, Benjamin/Cummings, Reading, MA, 1978, 2nd Edition.

[AOOdS18] L. Amorim, Y.-G. Oh, and J. Oliveira dos Santos, *Exact Lagrangian submanifolds, Lagrangian spectral invariants and Aubry-Mather theory*, Math. Proc. Cambridge Philos. Soc. 165 (2018), no. 3, 411–434.

[Arn72] V. I. Arnold, *Normal forms for functions near degenerate critical points, the weyl groups of $a_k$, $d_k$, $e_k$ and Lagrangian singularities*, Funct. Anal. Appl. 6 (1972), 254–272.

[Arn10] M.-C. Arnaud, *On a theorem due to Birkhoff*, Geom. Funct. Anal. 120 (2010), 1307 – 1316.

[BOdS12] P. Bernard and J. Oliveira dos Santos, *A geometric definition of the Mañé-Mather set and a Theorem of Marie-Claude Arnaud*, Math. Proc. Camb. Phil. Soc. 152 (2012), 167 – 178.

[Bot59] R. Bott, *The stable homotopy of the classical groups*, Ann. of Math. (2) 70 (1959), 313–337.

[Bra19] A. Bravetti, *Contact geometry and thermodynamics*, Int. J. Geom. Methods Mod. Phys. 16, Supp. 1 (2019), 1940003 (51 pages).

[Car09] C. Carathéodory, *Untersuchungen über die Grundlagen der Thermodynamik. (german)*, Math. Annalen 67 (1909), no. 3, 355–386.

[EGP22a] O. Esen, M. Grmela, and M. Pavelka, *On the role of geometry in statistical mechanics and thermodynamics I: geometric perspective*, Journal of Mathematical Physics 63 (2022), no. 12, Paper No. 123305, 22 pp.

[EGP22b] , *On the role of geometry in statistical mechanics and thermodynamics II: thermodynamic perspective*, Journal of Mathematical Physics. 63 (2022), no. 12, Paper No. 122902, 32 pp.

[Eli87] Y. Eliashberg, *A theorem on the structure of wave fronts and its applications*, (Russian) Funktsional. Anal. i Prilazhen 21 (1987), no. 3, 65–72.

[Esp04] Pep Esponol, *Statistical mechanic of coarse-graining*, Springer Verlag, 2004.

[Gib73] Josiah Willard Gibbs, *A method of geometrical representation of the thermodynamic properties of substances by means of surfaces*, Transactions of the Connecticut Academy of Arts and Sciences 2 (1873), 382–404.

[Giv90] A. Givental, *Singualr Lagrangian submanifolds and their Lagrangian mappings*, J. Soviet Math. 52 (1990), no. 4, 3246–3278.

[Grm14] M. Grmela, *Contact geometry of mesoscopic thermodynamics and dynamics*, Entropy 16 (2014), no. 3, 1652–1686.

[Gro86] M. Gromov, *Partial Differential Relations*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 9, Springer-Verlag, Berlin, 1986.

[Her73] M. Hermann, *Geometry, physics, and systems*, Lecture Notes In Pure and Applied Mathematics, M. DEKKER, 1973.

[Jay57] E.T. Jaynes, *Information theory and statistical mechanics*, Phys. Rev. 106 (1957), 620–630.

[Kap] M. Kapranov, *Thermodynamics and the moment map*, preprint, 2011, arXiv:1108.3472.

[KL51] S. Kullback and R. A. Leibler, *On information and sufficiency*, Ann. Math. Statistics 22 (1951), 79–86.

[Kra17] Thomas Kragh, *Homotopy equivalence of nearby lagrangians and the Serre spectral sequence*, Math. Ann. 368 (2017), no. 3–4, 945–970.

[Kra18] , *The Viterbo transfer as a map of spectra*, J. Symplectic Geom. 16 (2018), no. 1, 85–226.

[LS85] F. Laudenbach and J.-C. Sikorav, *Persistance d` ıntersection avec la section nulle au cours d` une isotope hamiltonienne dans un fibre cotangent*, Invent. Math. 82 (1985), no. 2, 349–357.

[Mac03] David J. C. MacKay, *Information theory, inference and learning algorithms*, Cambridge University Press, New York, 2003, xii+628 pp.
[Max67] J. C. Maxwell, *On the dynamic theory of gases*, Philos. Trans. Roy. Soc. London Ser A. 157 (1867), 49–88.

[Max75] ———, *On the dynamical evidence of the molecular constitution of bodies*, NA-TURE 11 (279) (1875), 357–359.

[MNSS90] R. Mrugala, J. D. Nulton, J. C. Schön, and P. Salamon, *Statistical approach to the geometric structure of thermodynamics*, Phys. Rev. A 41 (1990), no. 6, 3156–3160.

[Mru78] R. Mrugala, *Geometrical formulation of equilibrium phenomenological thermody-namics*, Reports on Mathematical Physics 14 (1978), no. 3, 419–427.

[MSBB18] M.J. Moran, H.N. Shapiro, D.D. Boettner, and M.B. Bailey, *Moran’s principles of engineering thermodynamics*, John Wiley & Sons, Limited, 2018.

[MW74] J. Marsden and A. Weinstein, *Reduction of symplectic manifolds with symmetry*, Rep. Mathematical Phys. 5 (1974), no. 1, 121–130.

[MW83] ———, *Coadjoint orbits, vortices and Clebsch variables for incompressible fluids*, Physica D 7 (1983), 305–323.

[Oh97] Y.-G. Oh, *Symplectic topology as the geometry of action functional. I. Relative Floer theory on the cotangent bundle*, J. Differential Geom. 46 (1997), no. 3, 499–577.

[Oh12] ———, *Geometry of generating functions and Lagrangian spectral invariants*, preprint, arXiv:1206.4788, 2012.

[Oh16] ———, *Continuous Hamiltonian dynamics and area-preserving homeomorphism group of D^2*, J. Korean Math. Soc. 53 (2016), no. 4, 795–834.

[Oh21] ———, *Geometry and analysis of contact instantons and entanglement of Legendrian links I*, arXiv:2111.02597, 2021.

[OY23] Yong-Geun Oh and Seungook Yu, *Legendrian spectral invariants on the one-jet bundle via perturbed contact instantons*, preprint, arXiv:2301.06704, 2023.

[San11] S. Sandon, *Contact homology, capacity and non-squeezing in R^{2n} x S^1 via generating functions*, Ann. Inst. Fourier (Grenoble) 61 (2011), 145–185.

[SdLLVdD20] Alexandre Anahory Simoes, Manuel de León, Manuel Lainz Valcázar, and David Martín de Diego, *Contact geometry for simple thermodynamical systems with friction*, Proc. A. 476 (2020), no. 2241.

[Sik87] J.-C. Sikorav, *Problèmes d’intersections et de points fixes en géométrie hamiltonienne*, Comment. Math. Helv. 62 (1987), no. 1, 62–73.

[SW49] C.E. Shannon and W. Weaver, *The mathematical theory of communication*, Illini books, no. V. 1, University of Illinois Press, 1949.

[Vil02] C. Villani, *A review of mathematical topics in collisional kinetic theory*, vol. I, pp. 71–305, North-Holland, Amsterdam, 2002.

[Wei] A. Weinstein, A graduate course in University of California, Berkeley, 1987.

[ZR92] V. Zakalyukin and M. Roberts, *On stable singular Lagrangian varieties*, Funkts. Anal. Prilozhen. 26 (1992), no. 3, 28–34.

Department of Mathematics POSTECH, & Center for Geometry and Physics, Institute for Basic Science (IBS), 77 Cheongam ro, Nam-gu, Pohang-si, Gyeongsangbuk-do, Korea 790-784

Email address: kimchi@postech.ac.kr

Center for Geometry and Physics, Institute for Basic Science (IBS), 77 Cheongam ro, Nam-gu, Pohang-si, Gyeongsangbuk-do, Korea 790-784 & POSTECH, Gyeongsangbuk do, Korea

Email address: yongoh1@postech.ac.kr