Thermodynamics: Symplectic and Contact Geometries

Chen Liang

June 10, 2024

Abstract

We review symplectic and contact geometries and discuss their applications in classical mechanics. We then consider the problem of maximizing entropy subject to constraints to derive the Boltzmann distribution. This leads to a geometric approach to thermodynamics.

Contents

1 Introduction

Symplectic geometry concerns the study of manifolds equipped with a closed non-degenerate 2-form. It is a powerful tool in classical mechanics as it provides a geometric description of the classical phase space. We will show that by incorporating time into the phase space, we can describe a classical system using ideas from contact geometry [2].

Geometric approaches to thermodynamics have been a subject of interest since Gibbs' formulation of thermodynamics. In particular, the fundamental thermodynamic relation can be described by assigning a 1-form, known as a contact form, to the thermodynamic state space [3, 4]. A natural question is whether the geometries of classical mechanics and thermodynamics are related. We will partly answer this question by showing that one can adapt the notion of entropy on a measure space to a notion of entropy on a symplectic manifold. Using the Lagrange multiplier method, we can then derive the Boltzmann distribution as a probability density on the symplectic manifold.

We will begin by providing some fundamental results in symplectic geometry and contact geometry and their applications to classical mechanics in Section 2 and Section 3. Along the way, we will discuss Noether's theorem in the context of both symplectic manifolds and contact manifolds. Section 4 will present a derivation of the Boltzmann distribution from a classical system and discuss its connection to the thermodynamic state space.

2 Symplectic geometry

2.1 Dynamics on symplectic manifolds

Definition 2.1. Let M be a $2n$-dimensional manifold. A *symplectic form* on M is a closed 2-form ω such that the induced map $\tilde{\omega} : TM \to T^*M$ defined by $\tilde{\omega}(v) = \omega(v, \cdot)$ is invertible [7].

Example 2.2. Let N be an n-dimensional manifold. The cotangent bundle T^*N is a symplectic manifold with the canonical symplectic form $\omega \coloneqq d\lambda$ where λ is the Liouville 1-form. The Liouville 1-form is defined at each covector $p \in T^*N$ by $\lambda_p \coloneqq \pi^* p$ where $\pi : T^*N \to N$ is the projection map. In local coordinates (x^i) on N, the Liouville 1-form at each point $p = (x, p_i\, dx^i)$

(Einstein summation convention is used here and throughout) is given by

$$\lambda = p_i \, dx^i \,,$$

where $dx^i = \pi^* \, dx^i$ are 1-forms on T^*N. The canonical symplectic form is then given by

$$\omega = d\lambda = dp_i \wedge dx^i.$$

It is clear that ω is closed by virtue of being exact, and $\tilde{\omega}$ is invertible because for all $A_i \, dx^i + B^i \, dp_i$ we have

$$\omega\left(-B^i \frac{\partial}{\partial x^i} + A_i \frac{\partial}{\partial p_i}, \cdot\right) = A_i \, dx^i + B^i \, dp_i \,.$$

The condition that $\tilde{\omega}$ is invertible is also called non-degeneracy of ω. Below are several equivalent characterizations of non-degeneracy, with the proof of equivalence found in [7].

Proposition 2.3. *Let V be a $2n$-dimensional vector space with $n \geq 1$ and $\omega \in V^* \wedge V^*$ be an antisymmetric bilinear form. The following are equivalent:*

1. *The induced map $\tilde{\omega} : V \to V^*$ defined by $\tilde{\omega}(v) = \omega(v, \cdot)$ is invertible.*
2. *$\omega^n \coloneqq \omega^{\wedge n} \neq 0$.*
3. *There exists a basis $(\alpha_1, \ldots, \alpha_n, \beta^1, \ldots, \beta^n)$ of V^* such that $\omega = \alpha_i \wedge \beta^i$.*

The second characterization establishes that symplectic manifolds are orientable and ω^n serves as a volume form/measure. This fact will come into play when we discuss thermodynamics in Section 4.2. For now, the most important characterization for us is the first one. The following proposition is essentially a direct consequence of the first characterization (see [7]).

Proposition 2.4. *Let (M, ω) be a symplectic manifold. Then, for all 1-forms $\alpha \in \Omega^1(M)$, there exists a unique vector field X such that $\iota_X \omega = \alpha$ where ι_X is the interior product defined by $\iota_X \omega \coloneqq \omega(X, \cdot)$.*

Definition 2.5. Let (M, ω) be a symplectic manifold. A *Hamiltonian vector field* associated to a smooth function $H \in C^\infty(M)$ is the unique vector field X_H such that

$$\iota_{X_H} \omega = -\, dH.$$

Example 2.6. Let $M = T^*N$ be the cotangent bundle of a manifold N. Let (x^i) be local coordinates on N and (p_i) be the dual coordinates on T_x^*N at each point $x \in N$. The Hamiltonian vector field associated to a smooth function $H \in C^\infty(T^*N)$ is given by

$$X_H = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial p_i}\,.$$

An integral curve of X_H is a path $\phi(t) = (x(t), p(t))$ on T^*N that satisfies

$$\dot{x}^i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}\,, \tag{1}$$

which are Hamilton's equations.

Symplectic geometry allows us to easily see that conserved quantities generate symmetries. We will first present a pair of lemmas.

Lemma 2.7 (Cartan's magic formula)**.** *Let X be a vector field and α be an n-form on a manifold, then,*

$$\mathcal{L}_X\alpha = \iota_X\, d\alpha + d\iota_X\alpha.$$

Proof. See [7]. □

Lemma 2.8. *Let (M, ω) be a symplectic manifold and $H \in C^\infty(M)$ be a smooth function, then X_H is a symplectic vector field, meaning that*

$$\mathcal{L}_{X_H}\omega = 0.$$

Proof. Since $d\omega = 0$, we have by Cartan's magic formula

$$\mathcal{L}_{X_H}\omega = d(\iota_{X_H}\omega) = d(-\,dH) = 0.$$

□

Proposition 2.9. *Let (M, ω) be a symplectic manifold and $H \in C^\infty(M)$ be a smooth function. If Q is a conserved quantity in the sense that $X_H Q = 0$ (i.e., it stays constant along the flow of X_H), then X_Q is a symmetry of the system in the sense that X_Q is symplectic and $X_Q H = 0$.*

Proof. By the previous lemma, we know that X_Q is symplectic, so that $\mathcal{L}_{X_Q}\omega = 0$. Direct computation gives

$$X_Q H = \iota_{X_Q}\, dH = -\iota_{X_Q}\iota_{X_H}\omega = \iota_{X_H}\iota_{X_Q}\omega = -\iota_{X_H}\, dQ = -X_H Q\,. \qquad (2)$$

□

Remark 2.10. The quantity in Eq. (2) is called the *Poisson bracket* of Q and H, and it is denoted by $\{Q, H\}$. We have just proved that this bracket is antisymmetric. In fact, it also satisfies the Jacobi identity which implies that $(C^\infty(M), \{\cdot,\cdot\})$ is a Lie algebra (see [7]).

2.2 Noether's theorem on symplectic manifolds

In the previous section, we saw that conserved quantities generate symmetries. Noether's theorem establishes the converse relationship.

We will first outline Noether's theorem in a more general setting and then apply it to symplectic manifolds. Given a functional integral $S[\phi]$, called the action integral, assume that taking arbitrary variation $\delta\phi$ (not necessarily with compact support or with fixed boundaries) gives

$$\delta S[\phi] = \int \left[E_i[\phi]\, \delta\phi^i + \frac{d}{dt}\big(P_i[\phi]\, \delta\phi^i\big) \right] dt, \qquad (3)$$

for some functionals E_i and P_i. When considering compactly supported variations that vanish at the boundary, the second term vanishes, leaving us with

$$\delta S = \int E_i[\phi]\, \delta\phi^i\, dt.$$

By the fundamental lemma of the calculus of variations, $\delta S[\phi, \delta\phi] = 0$ for all compactly supported variations that vanish at the boundary $\delta\phi$ if and only if $E_i[\phi] = 0$. We call the equations $E_i[\phi] = 0$ the equations of motion, and we say that the solutions to the equations are the dynamics generated by the action. A symmetry of the system is a variation R such that $\delta S[\phi, R] = 0$ for all ϕ. However, for simplicity, we will only consider symmetries R that make the integrand in Eq. (3) vanish for all ϕ, that is

$$E_i[\phi] R^i + \frac{d}{dt}\big(P_i[\phi] R^i\big) = 0.$$

When this happens, we see that along any solution $\bar{\phi}$ to the equations of motion, $P_i[\bar{\phi}]R^i$ is a conserved quantity. This underlies Noether's theorem, relating continuous symmetries to conserved quantities.

Now, let us apply Noether's theorem to a symplectic manifold (M, ω). For simplicity, we will assume that ω is exact, meaning $\omega = d\lambda$ for some 1-form λ. In principle, Noether's theorem is local. Since all closed form are locally exact, this is not a particularly restrictive assumption. We will demonstrate that the Hamiltonian flow can be described by minimizing the action functional

$$S[\phi] := \int (\lambda(\dot{\phi}) - H)\, dt.$$

Taking an arbitrary variation $\delta\phi$, we compute in local coordinates

$$\begin{aligned} \int \delta(\lambda_i \dot{\phi}^i)\, dt &= \int \left(\delta\phi^j \frac{\partial \lambda_i}{\partial x^j} \dot{\phi}^i + \lambda_i\, \delta\dot{\phi}^i \right) dt \\ &= \int \left(\delta\phi^j \frac{\partial \lambda_i}{\partial x^j} \dot{\phi}^i - \dot{\phi}^j \frac{\partial \lambda_i}{\partial x^j}\, \delta\phi^i \right) dt + \int \frac{d}{dt}\big(\lambda_i\, \delta\phi^i\big)\, dt \\ &= \int \omega(\delta\phi, \dot{\phi})\, dt + \int \frac{d}{dt}(\lambda(\delta\phi))\, dt \end{aligned} \tag{4}$$

and

$$\int \delta H\, dt = \int dH(\delta\phi)\, dt.$$

Combining these, we have

$$\delta S[\phi] = \int \Big(\omega(\delta\phi, \dot{\phi}) - dH(\delta\phi) \Big)\, dt + \int \frac{d}{dt}(\lambda(\delta\phi))\, dt.$$

We can read off the equations of motion

$$\omega(\cdot, \dot{\phi}) = dH \iff \dot{\phi} = X_H\,,$$

which are Hamilton's equations as in Eq. (1).

A symmetry of this system is a variation R such that

$$\omega(R, \dot{\phi}) - dH(R) + \frac{d}{dt}(\lambda(R)) = 0. \tag{5}$$

Observe that

$$\omega(R, \dot{\phi}) = \iota_{\dot{\phi}} \iota_R\, d\lambda \quad \text{and} \quad \frac{d}{dt}(\lambda(R)) = \iota_{\dot{\phi}}\, d(\iota_R \lambda).$$

By Cartan's magic formula, Eq. (5) simplifies to

$$(\mathcal{L}_R\lambda)(\dot{\phi}) - dH(R) = 0.$$

To make this vanish for all ϕ, we must have[1]

$$\mathcal{L}_R\lambda = 0 \quad \text{and} \quad \mathcal{L}_R H = dH(R) = 0.$$

When R is such a variation, Noether's theorem states that $\lambda(R)$ is a conserved quantity.

3 Contact geometry

3.1 Dynamics on contact manifolds

Definition 3.1. Let M be a $2n+1$-dimensional manifold. A *contact form* on M is a 1-form α such that $d\alpha$ is non-degenerate on $\ker\alpha$. This means that the map $\Phi : \ker\alpha \to (\ker\alpha)^*$ defined by $\Phi(v) = d\alpha(v, \cdot)$ is invertible [7].

Remark 3.2. It is tempting to call the pair (M, α) a contact manifold, but conventionally, a contact manifold refers to odd dimensional manifolds equipped with a non-integrable hyperplane distribution and can be characterized as conformal classes of local contact forms. Since such contact manifolds do not come up in our discussion, we will avoid the terminology of contact manifold entirely. A manifold with a contact form is termed a strict contact manifold.

Example 3.3. Consider $\mathbb{R}^{2n+1}$ as a manifold with coordinates (x^i, p_i, t). Suppose it is equipped with a 1-form $\alpha = H(x,p,t)\,dt - p_i\,dx^i$ where $H \in C^\infty(\mathbb{R}^{2n+1})$, then the kernel of α is spanned by

$$p_i\frac{\partial}{\partial t} + H\frac{\partial}{\partial x^i} \quad \text{and} \quad \frac{\partial}{\partial p_i}.$$

For $A = A^i\big(p_i\frac{\partial}{\partial t} + H\frac{\partial}{\partial x^i}\big) + \tilde{A}_i\frac{\partial}{\partial p_i}$ and $B = B^i\big(p_i\frac{\partial}{\partial t} + H\frac{\partial}{\partial x^i}\big) + \tilde{B}_i\frac{\partial}{\partial p_i}$, we can compute that

$$d\alpha(A,B) = \left(H - p_i\frac{\partial H}{\partial p_i}\right)(-\tilde{A}_iB^i + A^i\tilde{B}_i).$$

[1]Note that ϕ can be reparameterized, so it is similar to saying $ax + b = 0$ for all x implies $a = b = 0$.

Therefore, linear algebra tells us that $d\alpha$ is a contact form if and only if $H - p_i \frac{\partial H}{\partial p_i}$ is nowhere vanishing. In the case of $H = 1$, the contact form $\alpha = dt - p_i\, dx^i$ is known as the standard contact form.

Proposition 3.4. *Let α be a contact form on a manifold M of dimension $2n+1$. Then at each point $x \in M$, the map $\Phi : T_x M \to T_x^* M$ defined by $\Phi(v) = d\alpha(v, \cdot)$ has 1-dimensional kernel.*

Proof. By the definition of a contact form, we know that $\Phi|_{\ker \alpha}$ is injective and hence has rank at least $2n$. It follows that the rank of Φ is exactly $2n$ due to the fact that the determinant of any $(2n+1) \times (2n+1)$ antisymmetric matrix is zero. Therefore, the kernel of Φ has dimension 1. □

The above proposition leads to the existence and uniqueness of the so-called Reeb vector field. The Reeb vector field plays a crucial role in generating the dynamics on contact manifolds.

Definition 3.5. Let α be a contact form on a $2n$-dimensional manifold M. The *Reeb vector field* associated to α is the unique vector field R such that

$$\iota_R\, d\alpha = 0 \quad \text{and} \quad \alpha(R) = 1.$$

Example 3.6. On $\mathbb{R}^{2n+1}$ with coordinates (x^i, p_i, t) and contact form $\alpha = H(x, p, t)\, dt - p_i\, dx^i$, the kernel of $d\alpha$ is spanned by

$$\tilde{R} = \frac{\partial}{\partial t} + \frac{\partial H}{\partial p_i}\frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial p_i}\,.$$

An integral curve of $\tilde{R}$ is a path $\phi(\tau) = (x(\tau), p(\tau), t(\tau))$ such that

$$\dot{t} = 1, \quad \dot{x}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x^i}\,,$$

which again are the Hamilton's equations. By working with time as a coordinate, we can rescale $\tilde{R}$ to get different parameterizations of the same integral curve. The Reeb vector field, for example, will generate the same curve because it is given by

$$R = \left(H - p_i \frac{\partial H}{\partial p_i}\right)^{-1} \tilde{R}.$$

3.2 Noether's theorem on contact manifolds

On a manifold M with a contact form α, we will demonstrate that Reeb dynamics can be described by minimizing the action functional

$$S[\phi] \coloneqq \int_\phi \alpha = \int \alpha(\phi; \dot{\phi})\, d\tau,$$

where $\dot{\phi} = \frac{d\phi}{d\tau}$. Taking the variation amounts to a computation similar to that in Eq. (4), and we have

$$\delta S[\phi] = \int d\alpha(\delta\phi, \dot{\phi})\, d\tau + \int \frac{d}{d\tau}(\lambda(\delta\phi))\, d\tau.$$

The equation of motion is

$$d\alpha(\cdot, \dot{\phi}) = 0,$$

which is determined by Reeb dynamics. Again by a similar computation, a symmetry of the system is a variation A such that

$$\mathcal{L}_A \alpha = 0.$$

In which case, $\alpha(A)$ is a conserved quantity. We can go even further and consider A such that $\mathcal{L}_A\alpha = d\beta$ is exact and not necessarily zero, then we claim that $\alpha(A) - \beta$ is a conserved quantity. Using Cartan's magic formula, we have

$$0 = \mathcal{L}_A\alpha - d\beta = d(\iota_A\alpha - \beta) + \iota_A\, d\alpha$$

which implies that

$$\mathcal{L}_R(\alpha(A) - \beta) = \iota_R(\mathcal{L}_A\alpha - d\beta) - \iota_R\iota_A\, d\alpha = 0.$$

3.3 Comments on odd symplectic manifolds

An interesting property of a contact form α is that $\alpha \wedge (d\alpha)^n$ is a volume form. However, for many classical mechanics applications, all we need is the dynamics determined by the kernel of $d\alpha$. This suggests that we can consider a more general structure called an odd symplectic manifold.

Definition 3.7. Let M be a $2n+1$-dimensional manifold. An *odd symplectic form* on M is a closed 2-form ω such that ω has maximal rank (*i.e.*, $2n$).

Remark 3.8. Sometimes an odd symplectic manifold is defined to have an odd symplectic form and a nowhere vanishing volume form with suitable conditions [5].

An easy consequence of Definition 3.7 is that every contact form α induces an odd symplectic form $d\alpha$, however we can have odd symplectic forms that are not induced by contact forms. In particular, this allows us to relax the condition that $H - p_i \frac{\partial H}{\partial p_i}$ is nowhere vanishing which appeared in Example 3.3

In this contact or odd symplectic formulation of classical mechanics, time is combined with the other coordinates (position and momentum). While Reeb dynamics still provides a classical "causal structure"—where every point in the future and past of a point lies strictly on an integral cruve of the Reeb vector field—there is no longer a canonical choice of spacetime (and momentum) coordinates. In contrast, while the symplectic formulation allows different choices of space and momentum coordinates, time is treated as a universal parameter. Contact or odd symplectic formulation provides us with additional flexibility in choice of space, momentum, and time coordinates. However, on a case by case basis, it also prompts the question of how can we generalize certain ideas that arise in the symplectic formulation and perhaps even depend on a universal time parameter. For instance, in the subsequent section, we are going to investigate the passage from symplectic manifolds to thermodynamics, and we will leave as an open question how to incorporate contact and odd symplectic formulation of classical mechanics.

4 Thermodynamics

4.1 Entropy

On a measure space (X, μ), a probability density ρ is a nonnegative function such that $\int_X \rho\, \mu = 1$. If X is finite and μ is the counting measure, then the *Shannon entropy* of ρ is defined to be

$$S[\rho] \coloneqq -\sum_{x \in X} \rho(x) \log \rho(x),$$

where log is the natural logarithm (though different base can be used and would behave similarly). Many axiomatic characterization theorems of Shannon entropy are available in the literature, see [1] and references therein.

The challenge of generalizing to a continuous space is that some of the axioms for Shannon entropy become hard to formulate. The entropy for a continuous probability density, usually considered in thermodynamics, is called the *differential entropy* and defined as

$$S[\rho] := -\int_X \rho \log \rho \, d\mu.$$

If X has finite measure, then the differential entropy shares some properties with the Shannon entropy such as being maximized by the uniform distribution.

Proposition 4.1. *Suppose (X, μ) is a measure space with finite measure. Then the differential entropy $S[\rho]$ is maximized by the uniform distribution $\rho = 1/\mu(X)$.*

Proof. This is a minimization problem subject to the constraint $\int_X \rho \, d\mu = 1$. We can solve this using Lagrange multipliers. Let

$$L[\rho, \alpha] := -\int_X \rho \log \rho \, d\mu + \alpha\left(\int_X \rho \, d\mu - 1\right),$$

then the variations of L are

$$\delta L = -\int_X (\log \rho + 1 + \alpha)\delta\rho \, d\mu \quad \text{and} \quad \frac{\partial L}{\partial \alpha} = \int_X \rho \, d\mu - 1.$$

Setting both to zero, we have

$$\log \rho + 1 + \alpha = 0 \quad \text{and} \quad \int_X \rho \, d\mu = 1.$$

Hence, the solution is

$$\rho = \frac{\exp(-1-\alpha)}{\int_X \exp(-1-a) \, d\mu} = \frac{1}{\mu(X)}.$$

□

A potential issue with differential entropy is that if the measure μ is dimensionful, then probability density ρ would have the inverse dimension of the measure, which calls into question the meaning of $\log \rho$. Mathematically, this corresponds to the fact that under linear scaling of the measure $\mu' = \lambda\mu$, the uniform distribution will become $\rho' = \rho/\lambda$, and the entropy will transform like the logarithm of probability density:

$$S'[\rho'] = S[\rho] + \log \lambda.$$

If X has finite measure, then we can always normalize the measure to be dimensionless. Alternatively, we would have to introduce some constant of nature (like Planck's constant $\hbar$) to make the measure dimensionless. Due to the relative nature of measurement, a particularly natural solution is to introduce *relative entropy* between two probability densities ρ and σ, which is defined to be

$$S[\rho \mid \sigma] \coloneqq -\int_X \rho \log \frac{\rho}{\sigma}\, d\mu.$$

In general, suppose there is an algebra A acting on the space of probability densities. If the density σ is such that for every other probability density ρ, there exists a unique element $a \in A$ such that $\rho = a\sigma$, then the entropy relative to σ can be defined to be

$$S[\rho \mid \sigma] \coloneqq -\int_X (\log a)\rho\, d\mu$$

assuming that $\log a$ is suitably defined in A (for example, as the inverse of the exponential map defined by the Taylor series). In our example, the algebra can be taken to be the set of functions on X acting by multiplication and σ can be any positive function. This description can be taken further to describe entropy in for example quantum field theory [9].

4.2 Boltzmann distribution

Let (M, ω) be a symplectic manifold of dimension $2n$. Let $H_i \in C^\infty(M)$ be a family of smooth functions indexed by $i \in \Lambda$. If $\{H_i, H_j\} = 0$ for all i, j, then we say that the family (H_i) is a *commuting system of observables*. As an example, we can take $H_0 = 1$ and $H_1 = H$ where H is any smooth function, which we regard as a Hamiltonian. If in addition, $P \in C^\infty(M)$ is a conserved quantity (*i.e.*, $\{H, P\} = 0$), then we can let $H_2 = P$, and we have a commuting system of observables (H_0, H_1, H_2) as another example. Now, using the nowhere vanishing volume form ω^n as a measure, the differential entropy of a nonnegative smooth function $\rho \in C^\infty(M)$ is

$$S[\rho] = -\int_M \rho \log \rho \, \omega^n.$$

Let us now consider the question, given a commuting system of observables (H_i) and real numbers (U_i) indexed again by $i \in \Lambda$, what function ρ maximizes the entropy $S[\rho]$ subject to the constraints

$$\int_M \rho H_i \, \omega^n = U_i \quad \text{for all } i.$$

This is again a constrained optimization problem, which can be solved using Lagrange multipliers. Let β_i be the Lagrange multipliers associated to the constraints. Then, consider

$$\mathcal{L}[\rho, \beta] := -\int_M \rho \log \rho \, \omega^n + \sum_{i \in \Lambda} \beta_i \left(\int_M \rho H_i \, \omega^n - U_i \right).$$

We can compute variations of $\mathcal{L}$

$$\delta \mathcal{L} = -\int_M \left(\log \rho + 1 + \sum_{i \in \Lambda} \beta_i H_i \right) \delta \rho \, \omega^n$$
$$\frac{\partial \mathcal{L}}{\partial \beta_i} = \int_M \rho H_i \, \omega^n - U_i \, .$$

Thus, the solution to the constrained optimization problem is given by

$$\rho = \exp\left(-1 - \sum_{i \in \Lambda} \beta_i H_i \right),$$

where β_i are determined by the constraints

$$\int_M \rho H_i \, \omega^n = U_i \quad \text{for all } i.$$

Example 4.2. We consider the ideal gas of N particles of mass m in a box of volume V. Imposing periodic boundary conditions, we can regard the space as a torus $\mathbb{T}^3$ with volume form d^3x such that

$$\int_{\mathbb{T}^3} d^3x = V.$$

Let $M = T^*(\mathbb{T}^3)^N \cong \mathbb{R}^{3N} \times \mathbb{T}^{3N}$ be the cotangent bundle of $(\mathbb{T}^3)^N$ equipped with the ideal gas Hamiltonian

$$H = \sum_{i=1}^{3N} \frac{p_i^2}{2m} =: \frac{|p|^2}{2m}\,.$$

Considering the commuting system of observables $(1, H)$, we find the Boltzmann distribution

$$\rho = \exp(-1-\alpha-\beta H) = \frac{\exp(-\beta H)}{Z}\,, \tag{6}$$

where (α, β) are the Lagrange multipliers corresponding to $(1, H)$, while $Z \coloneqq \exp(1+\alpha)$. The Lagrange multipliers are subject to the constraints

$$\int_M \rho\,\omega^{3N} = 1 \quad \text{and} \quad \int_M \rho H\,\omega^{3N} = U.$$

The first constraint tells us that

$$\begin{aligned} Z = \int_M \rho\,\omega^{3N} &= V^N \left(\int_{\mathbb{R}} \exp(-\beta p^2/2m)\,dp \right)^{3N} \\ &= V^N \left(\frac{2\pi m}{\beta} \right)^{3N/2}. \end{aligned}$$

Plugging this back into Eq. (6), we obtain the Maxwell-Boltzmann distribution [8] expressed as a distribution on the phase space[2] M,

$$\rho = \frac{1}{V^n} \left(\frac{\beta}{2\pi m} \right)^{3N/2} e^{-\beta |p|^2/2m}\,.$$

The second constraint will allow us to relate the *inverse temperature* β and

[2] The Maxwell-Boltzmann distribution, or the Maxwell distribution, is usually expressed as a velocity distribution. Our distribution is equivalent upon integrating out space directions.

the *internal energy* U. We have

$$\begin{aligned} U &= \frac{V^N}{2mZ}\int_{\mathbb{R}^{3N}} \exp(-\beta|p|^2/2m)|p|^2\, d^{3N}p \\ &= \frac{V^N}{2mZ}\frac{2\pi^{3N/2}}{\Gamma(3N/2)}\int_0^\infty e^{-\beta p^2/2m}p^{3N+1}\, dp \\ &= \frac{V^N}{2mZ}\frac{\pi^{3N/2}}{\Gamma(3N/2)}\left(\frac{2m}{\beta}\right)^{3N/2+1}\Gamma\left(\frac{3N}{2}+1\right) \\ &= \frac{3N}{2}\frac{1}{\beta}\,. \end{aligned}$$

This is the well-known relation between the temperature and the internal energy of an ideal gas.

4.3 Thermodynamic system

A thermodynamic state is said to be an *equilibrium state* if it maximizes the entropy. The Boltzmann distribution from the previous section is an example of an equilibrium state. Evolution of a thermodynamic system along a path of equilibrium states is known as a *quasi-static process*. This is physically interesting because it corresponds well to real systems and mathematically interesting because of its simplicity. For example, if we assume that a classical density distribution is always in the Boltzmann distribution, then the entropy and the Lagrange multipliers β_i can be expressed as functions of the constraints U_i. We can obtain the β_i as functions of U_i by solving

$$\int_M \rho H_i\, \omega^n = U_i \quad \text{for all } i, \tag{7}$$

where ρ is the Boltzmann distribution. The entropy is then directly given by

$$\begin{aligned} S &= \int_M \rho\left(1 + \sum_{i\in\Lambda}\beta_i H_i\right)\omega^n \\ &= 1 + \sum_{i\in\Lambda}\beta_i U_i\,. \end{aligned}$$

Example 4.3. Continuing with the ideal gas example, we have

$$\begin{aligned} S &= \beta H + \log Z \\ &= \frac{3N}{2} + N \log\left(V \left(\frac{4\pi m U}{3N} \right)^{3/2} \right) . \end{aligned}$$

Note that we have been treating the ideal gas particles as distinguishable. If we were to treat them as indistinguishable, we would have to consider the phase space to be a quotient of M by particle exchange. The calculation will be almost identical, but now the Boltzmann distribution becomes $N!\,\rho$, which leads to the entropy

$$S_{\text{identical}} = S - \log N! \, .$$

Using Stirling's approximation $\log N! \sim N \log N - N$ for large N, we have

$$S_{\text{identical}} \sim \frac{5N}{2} + N \log\left(\frac{V}{N} \left(\frac{4\pi m U}{3N} \right)^{3/2} \right) .$$

This is the Sackur-Tetrode equation [8].

Definition 4.4. A *thermodynamic system* is the data of a n-dimensional manifold $\mathcal{U}$ equipped with a smooth real-valued "thermodynamic potential" function $S : \mathcal{U} \to \mathbb{R}$.

Given a thermodynamic system $(\mathcal{U}, S)$ and coordinate variables (U_i) on $\mathcal{U}$, we can define *conjugate variables* (β_i) by

$$\beta_i \coloneqq \frac{\partial S}{\partial U_i} \, .$$

The optimization problem imposes that the conjugate variables β_i will coincide with the Lagrange multipliers satisfying Eq. (7), justifying the reuse of the symbol β_i. However, in a general thermodynamic system, β_i need not stem from Lagrange multipliers. For example, in the ideal gas example, we can consider $S = S(U, V, N)$ as a function of U, V, N.

Sometimes we would like to parameterize the thermodynamic system by the β_i instead of the U_i or even an admixture of them. For this reason, we embed our thermodynamic system into a bigger space $\phi : \mathcal{U} \to \mathcal{T}$ with coordinates (S, U_i, β_i) via the map

$$\phi : (U_i) \mapsto \left(S(U_i), U_i, \frac{\partial S}{\partial U_i} \right).$$

Equipping $\mathcal{T}$ with a contact form $\alpha = dS - \sum_i \beta_i \, dU_i$, we have $\phi^* \alpha = 0$. In the context of contact geometry, this means that $\phi(\mathcal{U})$ is a Legendrian submanifold of $\mathcal{T}$. For further development of thermodynamics in the context of contact geometry, see [3, 6].

Remark 4.5. In a physics context, the U_i are often taken to be *extensive variables*, which are additive under the union of two systems. In the ideal gas example, the energy U, the volume V, the particle number N are extensive variables. The conjugate variables β_i are then called the *intensive variables*. The physical roles of the intensive variables can be illustrated by considering two thermodynamic systems $S_1(U_{1,i})$ and $S_2(U_{2,i})$. Assuming that $U_{1,i} + U_{2,i}$ are conserved, then it can be shown through optimization that the total entropy $S_1(U_{1,i}) + S_2(U_{2,i})$ is maximized when

$$\beta_{1,i} = \frac{\partial S_1}{\partial U_{1,i}} = \frac{\partial S_2}{\partial U_{2,i}} = \beta_{2,i}.$$

The intensive variables characterize whether two systems are in equilibrium with each other, and they reflect the zeroth law of thermodynamics, which states that equilibrium between thermodynamic systems is an equivalence relation[3]. Some examples of intensive variables are temperature, pressure, and chemical potential.

4.4 Open questions

One immediate question is whether the differential entropy or the relative entropy can be characterized by a set of axioms. An alternative phrasing of this question is what physical principles underly the definition of entropy that a thermodynamic system seeks to maximize. Another direction is to unify the origin of the thermodynamic variables. In the ideal gas example,

[3]This means that if two systems are in equilibrium with a third system, then they are in equilibrium with each other.

the energy U, the volume V, and the particle number N are seemingly disparate quantities in the classical system, and yet they ended up being on the same footing in the thermodynamic system. Along this line, we can ask whether the variables conjugate to volume and particle number—which are related to the pressure and the chemical potential—can be derived using Lagrange multipliers and thus lead to a better geometric understanding of thermodynamics.

5 Acknowledgments

I would like to thank Andrew Waldron for his long-term guidance and support throughout this research project. The approach to Noether's theorem in this thesis was influenced by his lectures. I would also like to thank Arthur Jiang for a collaboration in the early stages of this work.

References

[1] J. Aczél, B. Forte, and C. T. Ng. "Why the Shannon and Hartley Entropies Are 'Natural'". In: *Advances in Applied Probability* 6.1 (1974), pp. 131–146. ISSN: 00018678.

[2] V. I. Arnold. New York, NY: Springer New York, 1978. ISBN: 978-1-4757-1693-1. DOI: 10.1007/978-1-4757-1693-1.

[3] V. I. Arnold. "Contact Geometry: The Geometrical Method of Gibbs' Thermodynamics". In: *Proceedings of the Gibbs Symposium*. Yale University, 1989, pp. 163–179.

[4] J. W. Gibbs. *The Collected Works of J. Willard Gibbs: Thermodynamics.* Vol. 1. 1948.

[5] Z. He. "Odd Dimensional Symplectic Manifolds". PhD thesis. MIT, 2010.

[6] A. Jiang. "Contact Geometry abd Thermodynamics". Bachelor's Thesis. University of California, Davis, 2023.

[7] J. M. Lee. *Introduction to Smooth Manifolds*. Springer New York, 2013. ISBN: 978-1-4419-9982-5. DOI: 10.1007/978-1-4419-9982-5.

[8] D. V. Schroeder. *An Introduction to Thermal Physics*. Oxford University Press, Jan. 2021. ISBN: 9780192895547. DOI: 10.1093/oso/9780192895547.001.0001.

[9] E. Witten. "APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory". In: *Reviews of Modern Physics* 90.4 (Oct. 2018). ISSN: 1539-0756. DOI: 10.1103/revmodphys.90.045003.