

CONTENTS

<i>Preface</i>	xi
<i>Acknowledgments</i>	xvii
1 Thermodynamics and Statistical Physics	1
1.1 <i>Basics of Thermodynamics</i>	2
1.2 <i>Thermodynamic Potentials</i>	8
1.3 <i>Microcanonical Distribution</i>	11
1.4 <i>Canonical Distribution and Fluctuations</i>	14
2 Basics of Information Theory	17
2.1 <i>Information as a Choice</i>	18
2.2 <i>Communication Theory</i>	24
2.3 <i>Redundancy in the Alphabet</i>	28
2.4 <i>Mutual Information as a Universal Tool</i>	30
2.5 <i>Channel Capacity</i>	34
2.6 <i>Continuous Case and the Gaussian Channel</i>	38
2.7 <i>Hypothesis Testing and Bayes' Formula</i>	41
2.8 <i>Relative Entropy</i>	44
3 Applications of Information Theory	51
3.1 <i>The Whole Truth and Nothing but the Truth</i>	51
3.2 <i>Exorcising Maxwell's Demon</i>	55
3.3 <i>Information Is Life</i>	59
3.4 <i>Whom to Believe: My Eyes or Myself?</i>	66

viii CONTENTS

3.5	<i>Rate Distortion and the Information Bottleneck</i>	70
3.6	<i>Information Is Money</i>	76
4	New Second Law of Thermodynamics	81
4.1	<i>Stochastic Web Surfing and Google's PageRank</i>	81
4.2	<i>Random Walk and Diffusion</i>	85
4.3	<i>Detailed Balance</i>	87
4.4	<i>Fluctuation Relation and the New Second Law</i>	89
5	Inevitability of Irreversibility	93
5.1	<i>Evolution in the Phase Space</i>	94
5.2	<i>Kinetic Equation and H-Theorem</i>	96
5.3	<i>Phase-Space Mixing and Entropy Growth</i>	101
5.4	<i>Entropy Decrease and Nonequilibrium Fractal Measures</i>	108
5.5	<i>Renormalization Group and the Art of Forgetting</i>	113
6	Fundamental Limits of Uncertainty	120
6.1	<i>Quantum Mechanics and Entropy</i>	121
6.2	<i>Quantum Statistics and the Density Matrix</i>	125
6.3	<i>Entanglement Entropy</i>	128
6.4	<i>Quantum Communications</i>	133
6.5	<i>Conditional Entropy and Teleportation</i>	136
6.6	<i>The Way Out Is via a Black Hole</i>	139
7	Conclusion	143
7.1	<i>Take-Home Lessons</i>	143
7.2	<i>Epilogue</i>	148
Appendix	Extras, Exercises, and Solutions	153
A.1	<i>Formal Structure of Thermodynamics</i>	153
A.2	<i>Central Limit Theorem and Large Deviations</i>	155
A.3	<i>Numbers, Words, and Animal Signals</i>	158

CONTENTS ix

<i>A.4 Ising Model of the Brain</i>	162
<i>A.5 Applying the Infomax Principle</i>	163
<i>A.6 Exploration or Exploitation: Index Strategy</i>	166
<i>A.7 Memory Effects in Particle Collisions</i>	168
<i>A.8 Baker's Map</i>	169
<i>A.9 Multidimensional Renormalization Group</i>	173
<i>A.10 Brownian Motion</i>	177
<i>A.11 Fluctuation Relations in a Multidimensional Case</i>	181
<i>A.12 Quantum Fluctuations and Thermal Noise</i>	183
<i>A.13 Quantum Thermalization</i>	185
<i>A.14 Exercises and Solutions</i>	187
 <i>Index</i>	 207

1

Thermodynamics and Statistical Physics

Our knowledge is always partial. If we study macroscopic systems, some degrees of freedom remain hidden. For small sets of atoms or subatomic particles, their quantum nature prevents us from knowing precise values of their momenta and coordinates simultaneously. We used to believe that we found the way around the partial knowledge in mechanics, electricity, and magnetism, where we have *closed sets of equations describing explicitly known degrees of freedom*. In other words, we learned how to restrict our description only to things that can be considered independent of the unknown within a given accuracy. For example, planets are large complex bodies, and yet the motion of their centers of mass in the limit of large distances satisfies closed equations.¹

Despite the spectacular successes of electromagnetic theory and celestial mechanics, we soon realized how illusory was our belief in the closed description, since we needed to feed it with initial or boundary conditions taken from measurements. Here our knowledge is incomplete because of a finite precision of measurements. This has dramatic consequences when there is instability, so small data uncertainty at a given moment leads to large uncertainty in predicting the future and recovering the past. In a sense, every new decimal in precision is a new degree of freedom for unstable systems (including our solar system).

In this chapter, we deal with *observable manifestations of the hidden degrees of freedom*. While we do not know their state, we do know their nature—whether those degrees of freedom are related to moving particles, spins, bacteria, or

1. The next step—description of a planet rotation—needs to account for many extra degrees of freedom, for instance, oceanic flows (which slow down rotation by tidal forces).

market traders. That means, in particular, that we know the symmetries and conservation laws of the system.

The first two sections present a phenomenological approach called thermodynamics. The last two sections serve as a brief introduction into statistical physics.

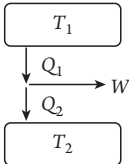
1.1 Basics of Thermodynamics

One can teach monkeys to differentiate; integration requires humans.

—GLEB KOTKIN

For at least a few thousand years, people have been burning things to propel objects. That was put on an industrial scale by the use of steam engines in the mid- to late 1700s. The Industrial Revolution generated a practical need to estimate engine efficiency, which triggered a regular scientific inquiry on general principles governing the conversion of heat into mechanical work. That led to the development of the abstract concept of entropy.

A heat engine works by delivering heat from a reservoir with some temperature T_1 via some system to another reservoir, with T_2 , doing some work in the process. Look under the hood of your car to appreciate the level of abstraction achieved in that definition. The work W is the difference between the heat given by the hot reservoir, Q_1 , and the heat absorbed by the cold one, Q_2 . What is the maximal fraction of heat we can use for work? Carnot in 1824 stated that we cannot make Q_2 arbitrarily small: in all processes, $Q_2/T_2 \geq Q_1/T_1$, so that the efficiency is bounded from above:

$$\frac{W}{Q_1} = \frac{Q_1 - Q_2}{Q_1} \leq 1 - \frac{T_2}{T_1}. \quad (1.1)$$


His elaborate arguments are of only historical interest now. Clausius in 1864 called the ratio Q/T *entropy* (the word starts with *en-*, like *energy*, and ends with *-tropos*, which means “turn” or “way” in Greek). We now interpret the Carnot criterion, saying that the entropy decrease of the hot reservoirs, $\Delta S_1 = Q_1/T_1$, must be less than the entropy increase of the cold one, $\Delta S_2 = Q_2/T_2$. Maximal work is achieved for minimal (zero) total entropy change, $\Delta S_2 = \Delta S_1$.

Just like the path from the Carnot engine to general thermodynamics, we discover the laws of nature by induction: from particular cases to general law and from processes to state functions. The latter step requires integration (to pass, for instance, from the Newton mechanics equations to the Hamiltonian or from thermodynamic equations of state to thermodynamic potentials). It is much easier to differentiate than to integrate, so deduction (or the postulation approach) is usually more pedagogical.² It also provides a good vantage point for generalizations and appeals to our brain, which likes to hypothesize before receiving any data, as we shall see later. In such an approach, one starts by postulating a variational principle for some function of the state of the system. Then one deduces from that principle the laws that govern change when one passes from state to state.

Here we present a deductive description of thermodynamics. *Thermodynamics studies restrictions on the possible macroscopic properties that follow from the fundamental conservation laws.* Therefore, thermodynamics does not predict numerical values but rather sets inequalities and establishes relations among different properties.

A traditional way to start building thermodynamics is to identify a conserved quantity, which can be exchanged but not created. It could be matter, money, energy, etc. For most physical systems, the basic symmetry is the invariance of the fundamental laws with respect to time shifts.³ The evolution of an isolated physical system is usually governed by the Hamiltonian (the energy written in canonical variables), whose time independence means energy conservation. In what follows, the conserved quantity of thermodynamics is called energy and denoted E . We wish to ascribe to the states of the system the values of E . First, we focus on the states independent of how they are prepared; such *equilibrium* states are completely characterized by the *static* values of observable variables.

Passing from state to state involves energy change, which generally consists of two parts: the energy change of visible degrees of freedom (which we call work) and the energy change of hidden degrees of freedom (which we call heat). To be able to measure energy changes in principle, we need adiabatic

2. In science, we strive to get the whole truth at any price. In teaching, we sell its parts at affordable prices.

3. Be careful trying to build thermodynamics for biological or social-economic systems, since generally the laws that govern them are not time invariant. For example, the metabolism of living beings changes with age, and the number of market regulations generally increases (as well as the total money mass, albeit not necessarily in our pockets).

processes, where there is no heat exchange so that all the energy changes are visible (or no under-the-table payments are made). Ascribing to every state its energy (up to an additive constant common for all states) hinges on our ability to relate any two states A and B by an adiabatic process, either $A \rightarrow B$ or $B \rightarrow A$, which allows us to measure the difference in their energies by the work W done by the system. Now, if we encounter a process where the energy change is not equal to the work, we call the difference the heat exchange, δQ :

$$dE = \delta Q - \delta W. \quad (1.2)$$

This statement is known as the first law of thermodynamics. It is nothing but a declaration of our belief in energy conservation: if the visible energy balance does not hold, then the energy of the hidden must change. The energy is a function of the state, so we use the differential, but we use δ for heat and work, which aren't differentials of any function. Heat exchange and work depend on the path taken from A to B ; that is, they refer to particular forms of energy transfer (not energy content). Before the first law was experimentally discovered (Mayer 1842, Joule 1845), heat was believed to be a separate fluid conserved by itself.

The basic problem of thermodynamics is determining the equilibrium state that eventually results after all internal constraints are removed in a closed composite system. The problem is solved with the help of the extremum principle: There exists a quantity S called entropy, which is a function of the parameters of the system. The values assumed by the parameters without an internal constraint maximize the entropy over the manifold of available states (Clausius 1865).

Thermodynamic limit Traditionally, thermodynamics has dealt with extensive parameters whose value grows linearly with the number of degrees of freedom. Additive quantities, like the number of particles N , electric charge, and magnetic moment, are extensive. Energy generally is not additive; that is, the energy of a composite system is not the sum of the parts because of an interaction energy: $E(N_1) + E(N_2) \neq E(N_1 + N_2)$. To treat energy as an additive variable, we assume short-range forces of interaction acting only along the boundary and take the thermodynamic limit $V \rightarrow \infty$. Then one can neglect the interaction energy, which scales as a surface $V^{2/3} \propto N^{2/3}$, in comparison with the additive bulk terms, which scale as $V \propto N$.

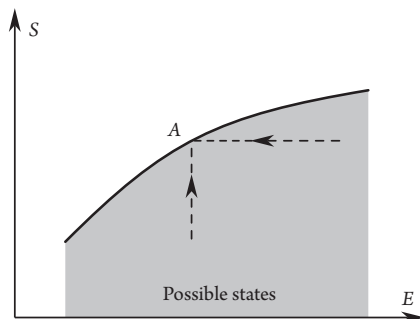
In that limit, thermodynamic entropy is also an extensive variable,⁴ which is a homogeneous first-order function of all the extensive parameters:

$$S(\lambda E, \lambda V, \dots) = \lambda S(E, V, \dots). \quad (1.3)$$

The function $S(E, V, \dots)$, also called a *fundamental relation*, is *everything* one needs to know to solve the basic problem (and others) in thermodynamics.

Of course, (1.3) does not mean that $S(E)$ is a linear function when other parameters are fixed: $S(\lambda E, V, \dots) \neq \lambda S(E, V, \dots)$. On the contrary, the equilibrium curve $S(E)$ must be convex to guarantee the stability of a homogeneous state. Indeed, imagine that a system breaks spontaneously into two halves with a bit different energies, $E + \Delta$ and $E - \Delta$. For equilibration to bring back the homogeneous state, its entropy $2S(E)$ must exceed the sum of the halves: $2S(E) > S(E + \Delta) + S(E - \Delta) \approx 2S(E) + S'' \Delta^2$. That requires $S'' < 0$ (the argument does not work for systems with long-range interaction where energy is nonadditive).

The figure shows the restriction imposed by thermodynamics on possible states: unconstrained equilibrium states are on the curve, while all other states lie below. Convexity guarantees that one can reach state *A* either by maximizing entropy at a given energy or minimizing energy at a given entropy:



Let us complement the visual geometric picture by an analytic relation between the extrema of entropy and energy. We assume the functions $S(E, X)$ and $E(S, X)$ to be continuous differentiable for any other parameter X . An efficient

4. We shall see later that nonextensive parts of entropy are also important for studying interaction and correlations between subsystems.

way to treat partial derivatives of two functions of two variables is to organize them into a 2×2 matrix and use its determinant, called a *Jacobian*:

$$\frac{\partial(u, v)}{\partial(x, y)} \equiv \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial v}{\partial x} \frac{\partial u}{\partial y}.$$

It changes sign upon any interchange of functions or variables. The partial derivative is a Jacobian:

$$\left(\frac{\partial u}{\partial x}\right)_y = \frac{\partial(u, y)}{\partial(x, y)}.$$

Then from

$$\left(\frac{\partial S}{\partial X}\right)_E = \frac{\partial(SE)}{\partial(SX)} = 0$$

follows

$$\left(\frac{\partial E}{\partial X}\right)_S = \frac{\partial(ES)}{\partial(XS)} \frac{\partial(EX)}{\partial(EX)} = -\frac{\partial(ES)}{\partial(EX)} \frac{\partial(EX)}{\partial(SX)} = -\left(\frac{\partial S}{\partial X}\right)_E \left(\frac{\partial E}{\partial S}\right)_X = 0.$$

That means that any entropy extremum is also an energy extremum. Differentiating the last relation one more time, we differentiate only the first factor since it turns into zero at equilibrium:

$$\left(\frac{\partial^2 E}{\partial X^2}\right)_S = -\left(\frac{\partial^2 S}{\partial X^2}\right)_E \left(\frac{\partial E}{\partial S}\right)_X.$$

The equilibrium is an entropy maximum; that is, $(\partial^2 S / \partial X^2)_E$ is negative. Which type of extremum has energy at equilibrium depends on the sign of $(\partial E / \partial S)_X$, which is called temperature; see (1.4) below. When the temperature is positive, as in the figure, the equilibrium is the entropy maximum at a fixed energy or the energy minimum at a fixed entropy—very much like a ball can be defined as the figure of either maximal volume \mathcal{V} for a given surface area \mathcal{A} or minimal area for a given volume. Such analogies create rich connections between thermodynamics and isoperimetric inequalities of the type $\mathcal{A}^d \geq dV^{d-1}\mathcal{V}_0$, where \mathcal{V}_0 is the volume of the unit ball in d dimensions.

The temperature could be negative—an example of a two-level system in section 1.4 shows that $S(E)$ could be nonmonotonic for systems with a finite phase space. Still, for every interval of a definite derivative sign, say, $(\partial E / \partial S)_X > 0$, we can solve $S = S(E, V, \dots)$ uniquely for $E(S, V, \dots)$, which is an equivalent fundamental relation.

Experimentally, one usually measures *changes*, thus finding derivatives. The partial derivatives of an extensive variable with respect to its arguments (which are also extensive parameters) are intensive parameters. In thermodynamics, we have only extensive and intensive variables, because we take the thermodynamic limit $N \rightarrow \infty$, $V \rightarrow \infty$, keeping N/V finite. For energy, one writes

$$\frac{\partial E}{\partial S} \equiv T(S, V, N), \quad \frac{\partial E}{\partial V} \equiv -P(S, V, N) \quad \frac{\partial E}{\partial N} \equiv \mu(S, V, N), \dots \quad (1.4)$$

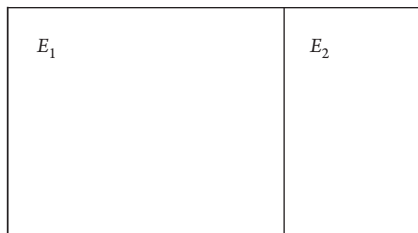
These relations are called the *equations of state*, and they serve as *definitions* for temperature T , pressure P , and chemical potential μ , corresponding to the respective extensive variables S, V, N . Our entropy is dimensionless, so T is assumed to be multiplied by the Boltzmann constant, $k = 1.3 \cdot 10^{-23} \text{J/K}$, and has the same dimensionality as the energy. From (1.4), we write

$$dE(S, V, N) = \delta Q - \delta W = TdS - PdV + \mu dN. \quad (1.5)$$

The extensive parameters V, N describe macroscopic (visible) degrees of freedom. Entropy is responsible for hidden degrees of freedom (i.e., heat). We shall see later that entropy is the missing information, which is thus maximal for hidden degrees of freedom in equilibrium. Temperature is the energetic price of information.

The derivatives (1.4) are taken at equilibrium, where a definite relation exists between variables, for instance, E and S . That means that (1.5) is true only for *quasi-static processes*, i.e., such that the system is close to equilibrium at every point of the process. A process can be considered quasi-static if its typical time of change is larger than the relaxation times (which can be estimated for pressure as L/c , where L is system size and c is sound velocity, and for temperature as L^2/κ , where κ is thermal conductivity). Finite deviations from equilibrium make $dS > \delta Q/T$ because entropy can increase without heat transfer. Only recently have we learned how to measure equilibrium quantities in fast, nonequilibrium processes, as described in section 4.4.

Let us see how the entropy maximum principle solves the basic problem. Consider two simple systems separated by a rigid wall that is impermeable to anything but heat. The whole composite system is closed; that is, $E_1 + E_2 = \text{const}$.



The entropy change under the energy exchange must be nonnegative:

$$dS = \frac{\partial S_1}{\partial E_1} dE_1 + \frac{\partial S_2}{\partial E_2} dE_2 = \frac{dE_1}{T_1} + \frac{dE_2}{T_2} = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dE_1 \geq 0. \quad (1.6)$$

For positive temperature, that means energy flows from the hot subsystem to the cold one: $T_1 > T_2 \Rightarrow dE_1 < 0$. We see that our definition (1.4) agrees with our intuitive notion of temperature. When equilibrium is reached, $dS = 0$, which requires $T_1 = T_2$. If the fundamental relation is known, then so is the function $T(E, V)$. Two equations, $T(E_1, V_1) = T(E_2, V_2)$ and $E_1 + E_2 = \text{const}$, completely determine E_1 and E_2 . In the same way, one can consider a movable wall and get $P_1 = P_2$ in equilibrium. If the wall allows for particle penetration, we get $\mu_1 = \mu_2$ in equilibrium.

Example 1.1: Consider a system that is characterized solely by its energy, which can change between zero and $E_m = N\epsilon$. The equation of state is the energy-temperature relation $E = E_m / (1 + e^{\epsilon/T})$, which tends to $E_m/2$ at $T \gg \epsilon$ and is exponentially small at $T \ll \epsilon$. In section 1.3, we identify this with a set of $N = E_m/\epsilon$ elements with two energy levels, 0 and ϵ . To find the fundamental relation in the entropy representation, we integrate the equation of state:

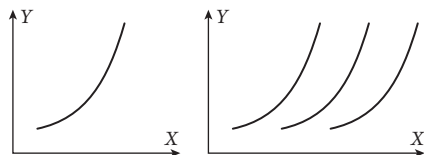
$$\frac{1}{T} = \frac{dS}{dE} = \frac{1}{\epsilon} \ln \frac{E_m - E}{E} \Rightarrow S(E) = N \ln \frac{N}{N - E/\epsilon} + \frac{E}{\epsilon} \ln \frac{N - E/\epsilon}{E/\epsilon}. \quad (1.7)$$

1.2 Thermodynamic Potentials

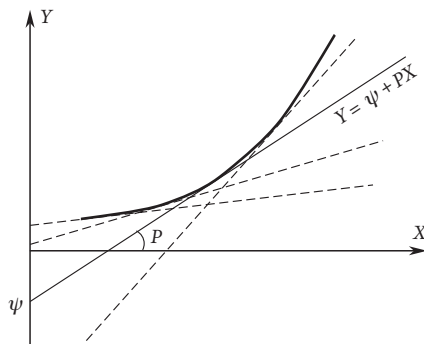
Even though it is always possible to eliminate, say, S from $E = E(S, V, N)$ and $T = T(S, V, N)$, getting $E = E(T, V, N)$, this is *not* a fundamental relation and it does not contain all the information. The point is, $E = E(T, V, N)$ is a partial differential equation (because $T = \partial E / \partial S$), and even if it could

be integrated, the result would contain an undetermined function of V, N . Still, it is easier to measure temperature than entropy, so it is convenient to have a complete formalism with an intensive parameter as an operationally independent variable and an extensive parameter as a derived quantity.

Any function $Y(X)$ defines the curve on the X, Y plane. We want to describe the same curve by some function of $P = \partial Y / \partial X$. It is not enough to eliminate X and consider the function $Y = Y[X(P)] = Y(P)$, because such a function determines the curve $Y = Y(X)$ only up to a shift along X , which changes neither Y nor P :



For example, the function $Y(P) = P^2/4$ corresponds to the whole family $Y = (X + C)^2$, which solves the differential equation $Y = (dY/dX)^2/4$. To pick a single function, we need to nail the curve by fixing the shift along X . For every P , we specify not Y but the position $\psi(P)$, where the straight line tangent to the curve intercepts the y axis: $\psi = Y - PX$:



In this way, we consider the curve $Y(X)$ as the envelope of the family of the tangent lines, each characterized by the slope P and the intercept ψ . The relation between them, $\psi(P) = Y[X(P)] - PX(P)$, completely defines the curve; here one substitutes $X(P)$ found from $P = dY(X)/dX$. The function $\psi(P)$ is called the Legendre transform of $Y(X)$. From $d\psi = -PdX - XdP + dY = -XdP$, one gets $-X = d\psi/dP$ —the inverse transform is the same up to a sign: $Y = \psi + XP$.

The transform is possible when for every X there is one P , that is, $P(X)$ is monotonic and $dP/dX = d^2Y/dX^2 \neq 0$. A sign-definite second derivative means that the function is either concave or convex. This is the second time we have met convexity, which we related above to the stability of a homogeneous state. Convexity and concavity play an important role in our story.

We can now make the Legendre transform of $E(S)$, which replaces the entropy by the temperature as an independent variable: $F = E - TS$ is called free energy. Its differential is as follows: $dF(T, V, N, \dots) = -SdT - PdV + \mu dN + \dots$. The counterpart to $(\partial E/\partial S)_{VN} = T$ is $(\partial F/\partial T)_{VN} = -S$. The free energy is particularly convenient for describing a system in thermal contact with a heat reservoir because the temperature is fixed, and we have one variable less to care about. The maximal work that can be done under a constant temperature (equal to that of the reservoir) is minus the differential of the free energy. Indeed, this is the work done *by the system and the thermal reservoir*. Is that work generally larger or smaller than the work done by the system alone? Let's see. That work is equal to the change in the total energy:

$$d(E + E_r) = dE + T_r dS_r = dE - T_r dS = d(E - T_r S) = d(E - TS) = dF.$$

In other words, the free energy, $F = E - TS$, is that part of the internal energy that is *free* to turn into work; the rest of the energy, TS , we must keep to sustain a constant temperature. The equilibrium state minimizes F —not absolutely, but over the manifold of states with a temperature equal to that of the reservoir. Consider, for instance, minimization of $F(T, V) = E[S(T, V), V] - TS(T, V)$ with respect to volume:

$$\left(\frac{\partial F}{\partial V}\right)_T = \left(\frac{\partial E}{\partial V}\right)_S + \left(\frac{\partial E}{\partial S} - T\right) \frac{\partial S}{\partial V} = \left(\frac{\partial E}{\partial V}\right)_S.$$

The derivatives turn into zero, and E and F reach extrema simultaneously. Also, in the point of an extremum, one gets $(\partial^2 E/\partial V^2)_S = (\partial^2 F/\partial V^2)_T$; i.e., both E and F have the same type of extremum (minimum in a positive-temperature equilibrium).

The system can reach the minimum of the free energy by minimizing energy and maximizing entropy. The former often requires creating some order in the system—for instance, orienting all spins parallel in a magnet or arranging all atoms into a regular crystal. On the contrary, increasing entropy requires disorder. Which of these tendencies wins depends on temperature, setting their relative importance. In later sections, we shall see repeatedly that

looking for a minimum of some free energy is a universal approach, from finding an equilibrium state of a physical system to designing the most optimal algorithm of information processing.

The formal structure of thermodynamics is described in section A.1. Since the Legendre transform is invertible, all thermodynamic potentials are equivalent and contain the same information. The choice of the potential for a given physical situation is that of convenience: we usually take what is fixed as a variable to diminish the number of effective variables.

The next two sections present a brief overview of the classical Boltzmann-Gibbs statistical approach: We introduce microscopic statistical description in the phase space and describe two principal ways (microcanonical and canonical) to derive thermodynamics from statistics.

Example 1.2: Consider a particle in the one-dimensional potential $U(x)$. The force f one needs to apply to keep the particle in the position X is apparently $f(X) = dU(x)/dx$ taken at X . Then $X(f) = dV(f)/df$, where $V(f)$ is minus the Legendre transform of the potential: $V(f) = Xf - U$.

1.3 Microcanonical Distribution

Let us consider a closed system with fixed energy E . Boltzmann conjectured that all microstates with the same energy have equal probability (the ergodic hypothesis). If the number of such states is $\Gamma(E)$, then the *microcanonical probability distribution* is as follows:

$$w_a(E) = 1/\Gamma(E). \quad (1.8)$$

To link statistical physics with thermodynamics, one must define the fundamental relation, i.e., a thermodynamic potential as a function of respective variables. For microcanonical distribution, Boltzmann in 1872 introduced entropy as

$$S(E) = -\ln w_a(E) = \ln \Gamma(E). \quad (1.9)$$

This is one of the most important formulas in physics⁵ (on a par with $f = ma$, $E = mc^2$, and $E = \hbar\omega$).

Noninteracting subsystems are statistically independent. That means that the statistical weight of the composite system is a product—for every state of

5. It is inscribed on Boltzmann's gravestone.

one subsystem, we have all the states of another. If the weight is a product, then the entropy is a sum. For interacting subsystems, this is true only for short-range forces in the thermodynamic limit $N \rightarrow \infty$.

Consider two subsystems, 1 and 2, that can exchange energy. Let's see how statistics solves the basic problem of thermodynamics (to define equilibrium) that we treated in (1.6). Assume that the indeterminacy in the energy of any subsystem Δ is much less than the total energy E . Alternatively, we may presume that the energy could be exchanged by portions Δ . Then

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i) \Gamma_2(E - E_i). \quad (1.10)$$

We denote $\bar{E}_1, \bar{E}_2 = E - \bar{E}_1$ for the values that correspond to the maximal term in the sum (1.10). To find this maximum, we compute the derivative:

$$\frac{\partial \Gamma}{\partial E_i} = \frac{\partial \Gamma_1}{\partial E_i} \Gamma_2 + \frac{\partial \Gamma_2}{\partial E_i} \Gamma_1 = (\Gamma_1 \Gamma_2) \left(\frac{\partial S_1}{\partial E_1} - \frac{\partial S_2}{\partial E_2} \right).$$

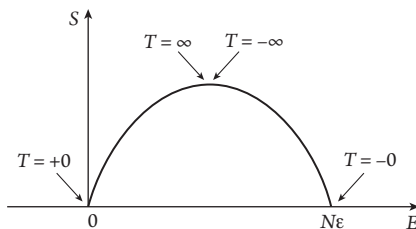
The extremum condition, $(\partial S_1 / \partial E_1)_{\bar{E}_1} = (\partial S_2 / \partial E_2)_{\bar{E}_2}$, corresponds to the thermal equilibrium where the temperatures of the subsystems are equal. The equilibrium is thus where the maximum of probability is. It is obvious that

$$\begin{aligned} \Gamma(\bar{E}_1) \Gamma(\bar{E}_2) &\leq \Gamma(E) \leq \Gamma(\bar{E}_1) \Gamma(\bar{E}_2) E / \Delta \Rightarrow S(E) \\ &= S_1(\bar{E}_1) + S_2(\bar{E}_2) + O(\log N), \end{aligned}$$

where the last term is negligible in the thermodynamic limit.

The same definition of entropy as a logarithm of the number of states is true for any system with a discrete set of states. For example, consider the set of N particles (spins, neurons), each with two energy levels, 0 and ϵ . If the energy of the set is E , then there are $L = E/\epsilon$ upper levels occupied. The statistical weight is determined by the number of ways one can choose L out of N ; that number is denoted C_N^L . This is our first combinatorial computation. Since we treat indistinguishable objects, let us first compute the number of permutations of m things. For each of the m first choices, we have $m - 1$ second choices, $m - 2$ third choices, etc. That means that the total number of permutations is $m(m - 1)(m - 2) \cdots 2 = m!$. To compute the number of ways to choose L out of N , we need to divide the total number of permutations among N by the total number of permutations among L and $N - L$: $\Gamma(N, L) = C_N^L = N! / L!(N - L)!$. We can now define the entropy (i.e., find

the fundamental relation): $S(E, N) = \ln \Gamma$. The entropy is symmetric about $E = N\epsilon/2$ and is zero at $E = 0, N\epsilon$, when either $L! = 1$ or $(N - L)! = 1$; that is, all the particles are in the same state. In the limit, we can use the Stirling formula, $\lim_{N \rightarrow \infty} \ln N! \approx N \ln N$. At the thermodynamic limit $N \gg 1$ and $L \gg 1$, it gives $S(E, N) \approx N \ln[N/(N - L)] + L \ln[(N - L)/L]$, which coincides with (1.7). The entropy as a function of energy is shown in the figure:



The equation of state (temperature-energy relation) is indeed $T^{-1} = \partial S / \partial E \approx \epsilon^{-1} \ln[(N - L)/L]$. We see that, when $E > N\epsilon/2$, the population of the higher level is larger than that of the lower one (inverse population as in a laser) and the temperature is negative. The negative temperature may happen only in systems with the upper limit of energy levels and simply means that, by adding energy beyond some level, we actually decrease the entropy, i.e., the number of accessible states. The example of a negative temperature is to help you disengage from the everyday notion of temperature and get used to the physicist's idea of temperature as the derivative of energy with respect to entropy. Yet it is still worth remembering the unique role played by the particular notion of temperature as mean kinetic energy of the gas molecules in the inductive development of thermodynamics.

Available (nonequilibrium) states lie below the $S(E)$ plot. The entropy maximum corresponds to the energy minimum for positive temperatures and to the energy maximum for negative temperatures. Imagine now that the system with a negative temperature is brought into contact with the thermostat (having a positive temperature). To equilibrate with the thermostat, the system needs to acquire a positive temperature. A glance at the figure shows that our system must move left, that is, give away energy (a laser generates and emits light). If this is done adiabatically slow along the equilibrium curve, the system first decreases the temperature further until it passes through minus infinity right into plus infinity and then down to positive values until it eventually reaches the thermostat's temperature. That is, negative temperatures are

actually “hotter” than positive. If you put your hand on a negative temperature system, you feel heat flowing into you. By itself, though, the system is stable since $\partial^2 S / \partial E^2 = -N / L(N - L)\epsilon^2 < 0$ at any temperature. We stress that there is no volume in $S(E, N)$, which means that we consider only part of the degrees of freedom. Real particles have kinetic energy unbounded from above and can correspond only to positive temperatures since negative temperature and infinite energy give an infinite Gibbs factor $e^{-E/T}$. Assuming detachment between kinetic and internal (electronic, spin, etc.) degrees of freedom is possible when their coupling is weak and only for a finite time.

The derivation of the thermodynamic fundamental relation $S(E, \dots)$ in the microcanonical ensemble is thus via the number of states or phase volume.

Exercise 1.1: Candies and kids.

There are three candies and two systems to distribute them: system 1 contains two boys and system 2 contains three girls. Every boy and girl can have zero, one, two, or three candies with equal probability. Kids are distinguishable, but candies aren't.⁶ What is the most probable number of candies in system 1? What is the average number of candies in system 1? What are the most probable and average numbers of candies in system 2?

1.4 Canonical Distribution and Fluctuations

Let us now discuss the statistical description, which corresponds to the thermodynamic potential of free energy, $F(T)$. Consider a system exchanging energy with a thermostat, which can be thought of as consisting of infinitely many copies of our system—this is the so-called canonical ensemble, characterized by T . Here our system can have any energy, and the question arises, What is the probability of being in a given microstate a with the energy E ? We derive that probability distribution (called canonical) from the microcanonical distribution of the whole system. Since all the states of the thermostat are equally likely to occur, the probability should be directly proportional to the statistical weight of the thermostat $\Gamma_0(E_0 - E)$. Here we assume $E \ll E_0$, expand (in the exponent!) $\Gamma_0(E_0 - E) = \exp[S_0(E_0 - E)] \approx$

6. Exchanging candies between kids leaves the system in the same state; taking candy from one kid and giving it to another brings the system to quite a different state.

$\exp[S_0(E_0) - E/T]$, and obtain

$$w_a(E) = Z^{-1} \exp(-E/T), \quad (1.11)$$

$$Z = \sum_a \exp(-E_a/T). \quad (1.12)$$

Note that there is no trace of the thermostat left except for the temperature. The normalization factor $Z(T, V, N)$ is a sum over all states accessible to the system and is called the partition function.

One again relates statistics and thermodynamics by defining entropy (Gibbs 1878). Recall that, for a closed system, Boltzmann defined entropy as minus the log of probability, $S = -\ln w_a$. There all probabilities were equal. Now we consider a subsystem at a fixed temperature, so that different states have different probabilities and both energy and entropy fluctuate. What should be the thermodynamic entropy: mean entropy, $-\langle \ln w_a \rangle$, or entropy at a mean energy, $-\ln w_a(E)$? They are the same! Indeed, $\ln w_a$ is linear in E_a for the Gibbs distribution, so the entropy at the mean energy is the mean entropy, and we recover the standard thermodynamic relation. Comparing the mean entropy,

$$\begin{aligned} \langle S \rangle &= -\langle \ln w_a \rangle = -\sum w_a \ln w_a = \sum w_a (E_a/T + \ln Z) \\ &= E/T + \ln Z, \end{aligned} \quad (1.13)$$

with the thermodynamic relation for it, $S = (E - F)/T$, we identify

$$F(T) = -T \ln Z(T). \quad (1.14)$$

The log of the probability of the mean energy is indeed the same as the mean log of probability:

$$S(E) = -\ln w_a(E) = -\ln \left[\frac{\exp(-E/T)}{Z} \right] = \frac{E}{T} + \ln Z = \frac{E - F}{T}. \quad (1.15)$$

Even though the Gibbs entropy, $S = -\sum w_a \ln w_a$, is derived here for equilibrium, this definition can be used for any set of probabilities w_a , since it provides a useful measure of our uncertainty about the system, as we shall see in the next chapter, where entropy is a key unlocking many doors (and locking some).

The canonical equilibrium distribution corresponds to the maximum of the Gibbs entropy, $S = -\sum w_a \ln w_a$, under the condition of the given mean energy $\bar{E} = \sum w_a E_a$: Requiring $\partial(S - \beta \bar{E})/\partial w_a = 0$, we obtain (1.11). For

an isolated system with a fixed energy, the entropy maximum corresponds to a uniform microcanonical distribution.

Are canonical and microcanonical statistical descriptions equivalent? Of course not. The descriptions are equivalent only when fluctuations are neglected and consideration is restricted to mean values. That takes place in thermodynamics, where the distributions produce different fundamental relations between the mean values: $S(E)$ for microcanonical, $F(T)$ for canonical. These functions are related by the Legendre transforms. Operationally, how does one check, for instance, the equivalence of canonical and microcanonical energies? One takes an isolated system at a given energy E , measures the derivative $\partial E/\partial S$, then puts it into the thermostat with the temperature equal to $\partial E/\partial S$; the energy now fluctuates, but the *mean* energy must be equal to E (as long as the system is macroscopic and all the interactions are short-range).

As far as fluctuations are concerned, there is a natural hierarchy: microcanonical distribution neglects, and canonical distribution accounts for fluctuations in E . The choice of description is dictated only by convenience in thermodynamics because it treats only mean values. But if we want to describe the whole statistics of the system in a thermostat, we need to use canonical distribution, not microcanonical.

Our subsystem is macroscopic itself, so it has many ways to redistribute the energy E among its degrees of freedom. In other words, it has many microscopic states corresponding to the same total energy of the subsystem. The probability for the subsystem to have a given energy is the probability of the state (1.11) times the number of states, i.e., the statistical weight of the *subsystem*:

$$W(E) = \Gamma(E)w_a(E) = \Gamma(E)Z^{-1} \exp(-E/T). \quad (1.16)$$

The weight $\Gamma(E)$ decreases as $E \rightarrow 0$ and grows as $E \rightarrow \infty$ usually by a power law, but the exponent $\exp(-E/T)$ decays faster than any power. As a result, $W(E)$ is concentrated in a very narrow peak and the energy fluctuations around \bar{E} are very small. For example, for an ideal gas, $W(E) \propto E^{3N/2} \exp(-E/T)$. To conclude, the Gibbs canonical distribution (1.11) tells us that the probability of a given microstate exponentially decays with the energy of the state, while (1.16) tells us that the probability of a given energy has a peak.

INDEX

- active inference, 68
- adiabatic process, 3
- area law, 141
- asymptotic equipartition, 20, 46

- Bayesian approach, 43, 66
- Bekenstein bound, 141
- binocular rivalry, 66
- bit, 18
- blind de-convolution, 165
- box-counting dimension, 86, 111
- brain, 60
- brevity law, 160

- capacity, 35, 63, 145
- Carnot criterion, 2
- central limit theorem, 114, 157
- chain rule, 32, 41
- coarse graining, 101, 106
- cocktail-party problem, 164
- codeword, 25
- complexity, 150
- conditional entropy, 28, 31, 136
- conditional probability, 28, 30, 66
- continuity equation, 86, 95
- convexity, 5, 22, 47, 83, 110, 144, 149
- convolution identity, 180
- critical surface, 175

- density matrix, 125, 201
- detailed balance, 87, 182
- differential entropy, 39

- distortion function, 70
- dynamical chaos, 93, 108

- energy, 3
- entanglement, 126, 131, 205
- entropy, xi, 2, 11, 15
- entropy production, 55, 62, 92, 110, 173
- equation of state, 13, 140, 153, 154
- equilibrium, 54, 95, 99
- equipartition, 178
- erasing information, 58
- ergodic, 107
- Euler equation, 153
- evolution operator, 88

- Fokker-Planck equation, 87, 90, 180
- fractal, 111, 172
- free energy, 10, 46, 56, 60, 91

- general uncertainty relation, 130
- generating function, 114, 156
- generative model, 67
- genome, 60
- geometric mean, 29, 77
- Gibbs distribution, 15, 46, 52, 72, 91
- Gibbs-Duhem relation, 153
- Gittins index, 167

- handicap principle, 162
- heat engine, 2
- Herdan-Heap law, 160

- Hick-Hyman law, 62
holographic principle, 141
horizon, 140
Huffman code, 27
- imaginary time, 184
independent component analysis, 165
infomax principle, 63
information bottleneck, 73
Ising model, 55, 115, 118, 173
isoperimetric inequality, 6, 150
- Jacobian, 6
Jensen inequality, 22, 91
- kinetic equation, 98
kinetic proofreading, 60
Kolmogorov-Sinai entropy, 107
- Landauer principle, 58
large deviation theory, 155, 203
Legendre transform, 9, 11, 156
Liouville theorem, 95, 128
Lyapunov exponent, 105, 112, 170
- Markov chain, 28, 82, 84
mating games, 162
mixed state, 126, 135
mixing, 93, 170
money, 79
monotonicity, 48, 130
mutual information, 47, 57, 99, 117, 129, 145, 164
- natural selection, 61
negative temperature, 13
neuron, 54, 64, 162
non-orthogonal states, 123, 134
- Occam's razor, 42
- partition function, 15, 116
perception, 66
phase transition, 174
phenotype switching, 78
proportional gambling, 76
pure state, 126, 131
- quantum state, 121
quantum superposition, 120
quantum thermalization, 131
quasi-static process, 7
qubit, 123
- random walk, 81, 85, 167
random-typing model, 160
redundancy, 24, 29, 73
relative entropy, 52, 77, 146
renormalization, 116, 148, 176
response function, 63
- Schmidt decomposition, 125
separation of trajectories, 104
spin-block renormalization, 115, 173
statistical weight, 11
Stirling formula, 13
stochastic matrix, 83
sub-additivity of entropy, 33
- temperature, 6
tensor product, 125
thermodynamic limit, 4, 12
thermostat, 14
typical sequence, 19, 36, 134, 145
- uncertainty principle, 122
- Zipf law, 160