

# *Modern Thermodynamics*

---

John Denker



# Contents

<b>0</b>	<b>Introduction</b>	<b>0–1</b>
0.1	Overview . . . . .	0–1
0.2	Availability . . . . .	0–3
0.3	Prerequisites, Goals, and Non-Goals . . . . .	0–3
<b>1</b>	<b>Energy</b>	<b>1–1</b>
1.1	Preliminary Remarks . . . . .	1–1
1.2	Conservation of Energy . . . . .	1–2
1.3	Examples of Energy . . . . .	1–3
1.4	Remark: Recursion . . . . .	1–4
1.5	Energy is Completely Abstract . . . . .	1–5
1.6	Additional Remarks . . . . .	1–6
1.7	Energy versus “Capacity to do Work” or “Available Energy” . . . . .	1–7
1.7.1	Best Case : Non-Thermal Situation . . . . .	1–7
1.7.2	Equation versus Definition . . . . .	1–8
1.7.3	General Case : Some Energy Not Available . . . . .	1–8
1.8	Mutations . . . . .	1–13
1.8.1	Energy . . . . .	1–13
1.8.2	Conservation . . . . .	1–13
1.8.3	Energy Conservation . . . . .	1–13
1.8.4	Internal Energy . . . . .	1–14
1.9	Range of Validity . . . . .	1–14

<b>2</b>	<b>Entropy</b>	<b>2–1</b>
2.1	Paraconservation . . . . .	2–1
2.2	Scenario: Cup Game . . . . .	2–2
2.3	Scenario: Card Game . . . . .	2–3
2.4	Peeking . . . . .	2–5
2.5	Discussion . . . . .	2–6
2.5.1	Connecting Models to Reality . . . . .	2–6
2.5.2	States and Probabilities . . . . .	2–7
2.5.3	Entropy is <i>Not Knowing</i> . . . . .	2–8
2.5.4	Entropy versus Energy . . . . .	2–8
2.5.5	Entropy versus Disorder . . . . .	2–8
2.5.6	False Dichotomy, or Not . . . . .	2–11
2.5.7	$dQ$ , or Not . . . . .	2–11
2.6	Quantifying Entropy . . . . .	2–11
2.7	Microstate versus Macrostate . . . . .	2–14
2.7.1	Surprisal . . . . .	2–14
2.7.2	Contrasts and Consequences . . . . .	2–14
2.8	Entropy of Independent Subsystems . . . . .	2–15
<b>3</b>	<b>Basic Concepts (Zeroth Law)</b>	<b>3–1</b>
<b>4</b>	<b>Low-Temperature Entropy (Alleged Third Law)</b>	<b>4–1</b>
<b>5</b>	<b>The Rest of Physics, Chemistry, etc.</b>	<b>5–1</b>
<b>6</b>	<b>Classical Thermodynamics</b>	<b>6–1</b>
6.1	Overview . . . . .	6–1
6.2	Stirling Engine . . . . .	6–1
6.2.1	Basic Structure and Operations . . . . .	6–1
6.2.2	Energy, Entropy, and Efficiency . . . . .	6–6
6.2.3	Practical Considerations; Temperature Match or Mismatch . . . . .	6–9
6.2.4	Discussion: Reversibility . . . . .	6–10

6.3	All Reversible Heat Engines are Equally Efficient . . . . .	6–10
6.4	Not Everything is a Heat Engine . . . . .	6–11
6.5	Carnot Efficiency Formula . . . . .	6–12
6.5.1	Definition of Heat Engine . . . . .	6–12
6.5.2	Analysis . . . . .	6–13
6.5.3	Discussion . . . . .	6–14
<b>7</b>	<b>Functions of State</b> . . . . .	<b>7–1</b>
7.1	Functions of State : Basic Notions . . . . .	7–1
7.2	Path Independence . . . . .	7–2
7.3	Hess’s Law, Or Not . . . . .	7–3
7.4	Partial Derivatives . . . . .	7–5
7.5	Heat Capacities, Energy Capacity, and Enthalpy Capacity . . . . .	7–8
7.6	$E$ as a Function of Other Variables . . . . .	7–13
7.6.1	$V$ , $S$ , and $h$ . . . . .	7–13
7.6.2	$X$ , $Y$ , $Z$ , and $S$ . . . . .	7–16
7.6.3	$V$ , $S$ , and $N$ . . . . .	7–16
7.6.4	Yet More Variables . . . . .	7–18
7.7	Internal Energy . . . . .	7–18
7.8	Integration . . . . .	7–21
7.9	Advection . . . . .	7–22
7.10	Deciding What’s True . . . . .	7–22
7.11	Deciding What’s Fundamental . . . . .	7–23
<b>8</b>	<b>Thermodynamic Paths and Cycles</b> . . . . .	<b>8–1</b>
8.1	A Path Projected Onto State Space . . . . .	8–1
8.1.1	State Functions . . . . .	8–1
8.1.2	Out-of-State Functions . . . . .	8–4
8.1.3	Converting Out-of-State Functions to State Functions . . . . .	8–5
8.1.4	Reversibility and/or Uniqueness . . . . .	8–6
8.1.5	The Importance of Out-of-State Functions . . . . .	8–6

8.1.6	Heat Content, or Not . . . . .	8-7
8.1.7	Some Mathematical Remarks . . . . .	8-7
8.2	Grady and Ungrady One-Forms . . . . .	8-8
8.3	Abuse of the Notation . . . . .	8-10
8.4	Procedure for Extirpating $dW$ and $dQ$ . . . . .	8-10
8.5	Some Reasons Why $dW$ and $dQ$ Might Be Tempting . . . . .	8-11
8.6	Boundary versus Interior . . . . .	8-14
8.7	The Carnot Cycle . . . . .	8-15
<b>9</b>	<b>Connecting Entropy with Energy</b>	<b>9-1</b>
9.1	The Boltzmann Distribution . . . . .	9-1
9.2	Systems with Subsystems . . . . .	9-1
9.3	Remarks . . . . .	9-3
9.3.1	Predictable Energy is Freely Convertible . . . . .	9-3
9.3.2	Thermodynamic Laws without Temperature . . . . .	9-4
9.3.3	Kinetic and Potential Microscopic Energy . . . . .	9-4
9.3.4	Ideal Gas : Potential Energy as well as Kinetic Energy . . . . .	9-5
9.3.5	Relative Motion versus “Thermal” Energy . . . . .	9-6
9.4	Entropy Without Constant Re-Shuffling . . . . .	9-7
9.5	Units of Entropy . . . . .	9-9
9.6	Probability versus Multiplicity . . . . .	9-10
9.6.1	Exactly Equiprobable . . . . .	9-10
9.6.2	Approximately Equiprobable . . . . .	9-11
9.6.3	Not At All Equiprobable . . . . .	9-13
9.7	Discussion . . . . .	9-15
9.8	Misconceptions about Spreading . . . . .	9-15
9.9	Spreading in Probability Space . . . . .	9-19

<b>10 Additional Fundamental Notions</b>	<b>10–1</b>
10.1 Equilibrium . . . . .	10–1
10.2 Non-Equilibrium; Timescales . . . . .	10–2
10.3 Efficiency; Timescales . . . . .	10–3
10.4 Spontaneity, Irreversibility, Stability, and Equilibrium . . . . .	10–4
10.5 Stability . . . . .	10–4
10.6 Relationship between Static Stability and Damping . . . . .	10–6
10.7 Finite Size Effects . . . . .	10–8
10.8 Words to Live By . . . . .	10–9
<b>11 Experimental Basis</b>	<b>11–1</b>
11.1 Basic Notions of Temperature and Equilibrium . . . . .	11–1
11.2 Exponential Dependence on Energy . . . . .	11–2
11.3 Metastable Systems with a Temperature . . . . .	11–3
11.4 Metastable Systems without a Temperature . . . . .	11–6
11.5 Dissipative Systems . . . . .	11–7
11.5.1 Sudden Piston : Sound . . . . .	11–7
11.5.2 Sudden Piston : State Transitions . . . . .	11–11
11.5.3 Rumford’s Experiment . . . . .	11–12
11.5.4 Simple Example: Decaying Current . . . . .	11–14
11.5.5 Simple Example: Oil Bearing . . . . .	11–15
11.5.6 Misconceptions : Heat . . . . .	11–17
11.5.7 Misconceptions : Work . . . . .	11–17
11.5.8 Remarks . . . . .	11–18
11.6 The Gibbs Gedankenexperiment . . . . .	11–18
11.7 Spin Echo Experiment . . . . .	11–19
11.8 Melting . . . . .	11–19
11.9 Isentropic Expansion and Compression . . . . .	11–20
11.10 Demagnetization Refrigerator . . . . .	11–20
11.11 Thermal Insulation . . . . .	11–21

<b>12 More About Entropy</b>	<b>12–1</b>
12.1 Terminology: Microstate versus Macrostate . . . . .	12–1
12.2 What the Second Law Doesn't Tell You . . . . .	12–2
12.3 Phase Space . . . . .	12–3
12.4 Entropy in a Crystal; Phonons, Electrons, and Spins . . . . .	12–5
12.5 Entropy is Entropy . . . . .	12–6
12.6 Spectator Entropy . . . . .	12–7
12.7 No Secret Entropy, No Hidden Variables . . . . .	12–7
12.8 Entropy is Context Dependent . . . . .	12–9
12.9 Slice Entropy and Conditional Entropy . . . . .	12–10
12.10 Extreme Mixtures . . . . .	12–11
12.10.1 Simple Model System . . . . .	12–12
12.10.2 Two-Sample Model System . . . . .	12–13
12.10.3 Helium versus Snow . . . . .	12–14
12.10.4 Partial Information aka Weak Peek . . . . .	12–15
12.11 Entropy is Not Necessarily Extensive . . . . .	12–15
12.12 Mathematical Properties of the Entropy . . . . .	12–16
12.12.1 Entropy Can Be Infinite . . . . .	12–16
<b>13 Temperature : Definition and Fundamental Properties</b>	<b>13–1</b>
13.1 Example Scenario: Two Subsystems, Same Stuff . . . . .	13–1
13.1.1 Equilibrium is Isothermal . . . . .	13–7
13.2 Remarks about the Simple Special Case . . . . .	13–7
13.3 Two Subsystems, Different Stuff . . . . .	13–8
13.4 Discussion: Constants Drop Out . . . . .	13–12
13.5 Calculations . . . . .	13–12
13.6 Chemical Potential . . . . .	13–14



<b>14 Spontaneity, Reversibility, Equilibrium, Stability, Solubility, etc.</b>	<b>14–1</b>
14.1 Fundamental Notions . . . . .	14–1
14.1.1 Equilibrium . . . . .	14–1
14.1.2 Stability . . . . .	14–1
14.1.3 A First Example: Heat Transfer . . . . .	14–2
14.1.4 Graphical Analysis – One Dimension . . . . .	14–3
14.1.5 Graphical Analysis – Multiple Dimensions . . . . .	14–4
14.1.6 Reduced Dimensionality . . . . .	14–6
14.1.7 General Analysis . . . . .	14–6
14.1.8 What’s Fundamental and What’s Not . . . . .	14–8
14.2 Proxies for Predicting Spontaneity, Reversibility, Equilibrium, etc. . . . .	14–8
14.2.1 Isolated System; Proxy = Entropy . . . . .	14–9
14.2.2 External Damping; Proxy = Energy . . . . .	14–9
14.2.3 Constant $V$ and $T$ ; Proxy = Helmholtz Free Energy . . . . .	14–12
14.2.4 Constant $P$ and $T$ ; Proxy = Gibbs Free Enthalpy . . . . .	14–14
14.3 Discussion: Some Fine Points . . . . .	14–17
14.3.1 Local Conservation . . . . .	14–17
14.3.2 Lemma: Conservation of Enthalpy, Maybe . . . . .	14–17
14.3.3 Energy and Entropy (as opposed to «Heat» . . . . .	14–18
14.3.4 Spontaneity . . . . .	14–18
14.3.5 Conditionally Allowed and Unconditionally Disallowed . . . . .	14–19
14.3.6 Irreversible by State or by Rate . . . . .	14–19
14.4 Temperature and Chemical Potential in the Equilibrium State . . . . .	14–20
14.5 The Approach to Equilibrium . . . . .	14–22
14.5.1 Non-Monotonic Case . . . . .	14–22
14.5.2 Monotonic Case . . . . .	14–23
14.5.3 Approximations and Misconceptions . . . . .	14–23
14.6 Natural Variables, or Not . . . . .	14–25
14.6.1 The “Big Four” Thermodynamic Potentials . . . . .	14–25
14.6.2 A Counterexample: Heat Capacity . . . . .	14–25

14.7	Going to Completion . . . . .	14–26
14.8	Example: Shift of Equilibrium . . . . .	14–27
14.9	Le Châtelier’s «Principle», Or Not . . . . .	14–30
14.10	Appendix: The Cyclic Triple Derivative Rule . . . . .	14–31
14.10.1	Graphical Derivation . . . . .	14–31
14.10.2	Validity is Based on Topology . . . . .	14–33
14.10.3	Analytic Derivation . . . . .	14–34
14.10.4	Independent and Dependent Variables, or Not . . . . .	14–35
14.10.5	Axes, or Not . . . . .	14–35
14.11	Entropy versus “Irreversibility” in Chemistry . . . . .	14–35
<b>15</b>	<b>The “Big Four” Energy-Like State Functions</b>	<b>15–1</b>
15.1	Energy . . . . .	15–1
15.2	Enthalpy . . . . .	15–1
15.2.1	Integration by Parts; $PV$ and its Derivatives . . . . .	15–1
15.2.2	More About $PdV$ versus $VdP$ . . . . .	15–3
15.2.3	Definition of Enthalpy . . . . .	15–6
15.2.4	Enthalpy is a Function of State . . . . .	15–7
15.2.5	Derivatives of the Enthalpy . . . . .	15–8
15.3	Free Energy . . . . .	15–8
15.4	Free Enthalpy . . . . .	15–9
15.5	Thermodynamically Available Energy – Or Not . . . . .	15–9
15.5.1	Overview . . . . .	15–9
15.5.2	A Calculation of “Available” Energy . . . . .	15–12
15.6	Relationships among $E$ , $F$ , $G$ , and $H$ . . . . .	15–14
15.7	Yet More Transformations . . . . .	15–16
15.8	Example: Hydrogen/Oxygen Fuel Cell . . . . .	15–16
15.8.1	Basic Scenario . . . . .	15–16
15.8.2	Enthalpy . . . . .	15–18
15.8.3	Gibbs Free Enthalpy . . . . .	15–19
15.8.4	Discussion: Assumptions . . . . .	15–20
15.8.5	Plain Combustion $\Rightarrow$ Dissipation . . . . .	15–21
15.8.6	Underdetermined . . . . .	15–26
15.8.7	$H$ Stands For Enthalpy – Not «Heat» . . . . .	15–26

<b>16</b>	<b>Adiabatic Processes</b>	<b>16–1</b>
16.1	Various Definitions of “Adiabatic” . . . . .	16–1
16.2	Adiabatic versus Isothermal Expansion . . . . .	16–3
<b>17</b>	<b>Heat</b>	<b>17–1</b>
17.1	Definitions . . . . .	17–1
17.2	Idiomatic Expressions . . . . .	17–5
17.3	Resolving or Avoiding the Ambiguities . . . . .	17–6
<b>18</b>	<b>Work</b>	<b>18–1</b>
18.1	Definitions . . . . .	18–1
18.1.1	Integral versus Differential . . . . .	18–3
18.1.2	Coarse Graining . . . . .	18–3
18.1.3	Local versus Overall . . . . .	18–4
18.2	Energy Flow versus Work . . . . .	18–4
18.3	Remarks . . . . .	18–6
18.4	Hidden Energy . . . . .	18–6
18.5	Pseudowork . . . . .	18–6
<b>19</b>	<b>Cramped versus Uncramped Thermodynamics</b>	<b>19–1</b>
19.1	Overview . . . . .	19–1
19.2	A Closer Look . . . . .	19–4
19.3	Real-World Compound Cramped Systems . . . . .	19–7
19.4	Heat Content, or Not . . . . .	19–7
19.5	No Unique Reversible Path . . . . .	19–9
19.6	Vectors: Direction and Magnitude . . . . .	19–11
19.7	Reversibility . . . . .	19–11
<b>20</b>	<b>Ambiguous Terminology</b>	<b>20–1</b>
20.1	Background . . . . .	20–1
20.2	Overview . . . . .	20–2
20.3	Energy . . . . .	20–2
20.4	Conservation . . . . .	20–3
20.5	Other Ambiguities . . . . .	20–4

<b>21 Thermodynamics, Restricted or Not</b>	<b>21–1</b>
<b>22 The Relevance of Entropy</b>	<b>22–1</b>
<b>23 Equilibrium, Equiprobability, Boltzmann Factors, and Temperature</b>	<b>23–1</b>
23.1 Background and Preview . . . . .	23–1
23.2 Example: $N = 1001$ . . . . .	23–2
23.3 Example: $N = 1002$ . . . . .	23–6
23.4 Example: $N = 4$ . . . . .	23–7
23.5 Role Reversal: $N = 1002$ ; $T_M$ versus $T_\mu$ . . . . .	23–8
23.6 Example: Light Blue . . . . .	23–10
23.7 Discussion . . . . .	23–10
23.8 Relevance . . . . .	23–11
<b>24 Partition Function</b>	<b>24–1</b>
24.1 Basic Properties . . . . .	24–1
24.2 Calculations Using the Partition Function . . . . .	24–2
24.3 Example: Harmonic Oscillator . . . . .	24–4
24.4 Example: Two-State System . . . . .	24–6
24.5 Rescaling the Partition Function . . . . .	24–9
<b>25 Equipartition</b>	<b>25–1</b>
25.1 Generalized Equipartition Theorem . . . . .	25–1
25.2 Corollaries: Power-Law Equipartition . . . . .	25–2
25.3 Harmonic Oscillator, Particle in a Box, and Other Potentials . . . . .	25–2
25.4 Remarks . . . . .	25–3
<b>26 Partition Function: Some Examples</b>	<b>26–1</b>
26.1 Preview: Single Particle in a Box . . . . .	26–1
26.2 Ideal Gas of Point Particles . . . . .	26–2
26.2.1 Distinguishable Particles . . . . .	26–2
26.2.2 Indistinguishable Particles; Delabeling . . . . .	26–3
26.2.3 Mixtures . . . . .	26–3

26.2.4	Energy, Heat Capacity, and Entropy for a Pure Gas . . . . .	26–6
26.2.5	Entropy of a Mixture . . . . .	26–8
26.2.6	Extreme Mixtures . . . . .	26–10
26.2.7	Entropy of the Deal . . . . .	26–10
26.3	Rigid Rotor . . . . .	26–13
26.4	Isentropic Processes . . . . .	26–15
26.5	Polytropic Processes $\cdots$ Gamma etc. . . . .	26–16
26.6	Low Temperature . . . . .	26–21
26.7	Degrees of Freedom, or Not . . . . .	26–21
26.8	Discussion . . . . .	26–22
26.9	Derivation: Particle in a Box . . . . .	26–22
26.10	Area per State in Phase Space . . . . .	26–26
26.10.1	Particle in a Box . . . . .	26–26
26.10.2	Periodic Boundary Conditions . . . . .	26–28
26.10.3	Harmonic Oscillator . . . . .	26–28
26.10.4	Non-Basis States . . . . .	26–28
<b>27</b>	<b>Density Matrices</b>	<b>27–1</b>
<b>28</b>	<b>Summary</b>	<b>28–1</b>
<b>29</b>	<b>About the Book</b>	<b>29–1</b>
<b>30</b>	<b>References</b>	<b>30–1</b>



# Chapter 0

## Introduction

### 0.1 Overview

Real thermodynamics is celebrated for its precision, power, generality, and elegance. However, all too often, students are taught some sort of pseudo-thermodynamics that is infamously confusing, lame, restricted, and ugly. This document is an attempt to do better, i.e. to present the main ideas in a clean, simple, modern way.

The first law of thermodynamics is usually stated in a very unwise form.	We will see how to remedy this.
--	---------------------------------

The second law is usually stated in a very unwise form.	We will see how to remedy this, too.
---	--------------------------------------

The so-called third law is a complete loser. It is beyond repair.	We will see that we can live without it just fine.
---	--

Many of the basic concepts and terminology (including heat, work, adiabatic, etc.) are usually given multiple mutually-inconsistent definitions.	We will see how to avoid the inconsistencies.
--	---

Many people remember the conventional “laws” of thermodynamics by reference to the following joke:<sup>1</sup>

- 0) You have to play the game;
- 1) You can’t win;
- 2) You can’t break even, except on a very cold day; and
- 3) It doesn’t get that cold.

It is not optimal to formulate thermodynamics in terms of a short list of enumerated laws, but if you insist on having such a list, here it is, modernized and clarified as much as possible. The laws appear in the left column, and some comments appear in the right column:

---

<sup>1</sup>This is an elaboration of the jocular laws attributed to C.P. Snow. I haven’t been able to find a more precise citation.

The zeroth law of thermodynamics tries to tell us that certain thermodynamical notions such as “temperature,” “equilibrium,” and “macroscopic state” make sense.	Sometimes these make sense, to a useful approximation ... but not always. See chapter 3.
The first law of thermodynamics states that energy obeys a local conservation law.	This is true and important. See section 1.2.
The second law of thermodynamics states that entropy obeys a local law of paraconservation.	This is true and important. See chapter 2.
There is no third law of thermodynamics.	The conventional so-called third law alleges that the entropy of some things goes to zero as temperature goes to zero. This is never true, except perhaps in a few extraordinary, carefully-engineered situations. It is never important. See chapter 4.

To summarize the situation, we have two laws (#1 and #2) that are very powerful, reliable, and important (but often misstated and/or conflated with other notions) plus a grab-bag of many lesser laws that may or may not be important and indeed are not always true (although sometimes you can make them true by suitable engineering). What’s worse, there are many essential ideas that are not even hinted at in the aforementioned list, as discussed in chapter 5.

We will not confine our discussion to some small number of axiomatic “laws.” We will carefully formulate a first law and a second law, but will leave numerous other ideas un-numbered. The rationale for this is discussed in section 7.10.

The relationship of thermodynamics to other fields is indicated in figure 1. Mechanics and many other fields use the concept of energy, sometimes without worrying very much about entropy. Meanwhile, information theory and many other fields use the concept of entropy, sometimes without worrying very much about energy; for more on this see chapter 22. The hallmark of thermodynamics is that it uses both energy and entropy.

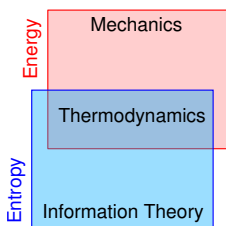


Figure 1: Thermodynamics, Based on Energy and Entropy



## 0.2 Availability

- This document is available in PDF format at <http://www.av8n.com/physics/thermo-laws.pdf>  
You may find this advantageous if your browser has trouble displaying standard HTML math symbols.
- It is also available in HTML format, chapter by chapter. The index is at <http://www.av8n.com/physics/thermo/>

## 0.3 Prerequisites, Goals, and Non-Goals

This section is meant to provide an overview. It mentions the main ideas, leaving the explanations and the details for later. If you want to go directly to the actual explanations, feel free to skip this section.

- (1) There is an important distinction between fallacy and absurdity. An idea that makes wrong predictions every time is absurd, and is not dangerous, because nobody will pay any attention to it. The most dangerous ideas are the ones that are often correct or nearly correct, but then betray you at some critical moment.

Most of the fallacies you see in thermo books are pernicious precisely because they are *not* absurd. They work OK some of the time, especially in simple “textbook” situations ... but alas they do not work in general.

The main goal here is to formulate the subject in a way that is less restricted and less deceptive. This makes it vastly more reliable in real-world situations, and forms a foundation for further learning.

In some cases, key ideas can be reformulated so that they work just as well – *and just as easily* – in simple situations, while working vastly better in more-general situations. In the few remaining cases, we must be content with less-than-general results, but we will make them less deceptive by clarifying their limits of validity.

- (2) We distinguish cramped thermodynamics from uncramped thermodynamics as shown in figure 2.

On the left side of the diagram, the system is constrained to move along the red path, so that there is only one way to get from  $A$  to  $Z$ .

In contrast, on the right side of the diagram, the system can follow any path in the  $(S, T)$  plane, so there are infinitely many ways of getting from  $A$  to  $Z$ , including the simple path  $A \rightarrow Z$  along a contour of constant entropy, as well as more complex paths such as  $A \rightarrow Y \rightarrow Z$  and  $A \rightarrow X \rightarrow Y \rightarrow Z$ . See chapter 19 for more on this.

Indeed, there are infinitely many paths from  $A$  back to  $A$ , such as  $A \rightarrow Y \rightarrow Z \rightarrow A$  and  $A \rightarrow X \rightarrow Y \rightarrow Z \rightarrow A$ . Paths that loop back on themselves like this are called *thermodynamic cycles*. Such a path returns the system to its original state, but generally does not return the surroundings to their original state. This allows us to build heat engines, which take energy from a heat bath and convert it to mechanical work.

There are some simple ideas such as *specific heat capacity* (or molar heat capacity) that can be developed within the limits of cramped thermodynamics, at the high-school level or even the pre-high-school level, and then extended to all of thermodynamics.

Alas there are some other ideas such as “heat content” aka “thermal energy content” that seem attractive in the context of cramped thermodynamics but are extremely deceptive if you try to extend them to uncramped situations.

Even when cramped ideas (such as heat capacity) can be extended, the extension must be done carefully, as you can see from the fact that the energy capacity  $C_V$  is different from the enthalpy capacity  $C_P$ , yet both are widely (if not wisely) called the “heat” capacity.

- (3) Uncramped thermodynamics has a certain irreducible amount of complexity. If you try to simplify it too much, you trivialize the whole subject, and you arrive at a result that wasn’t worth the trouble. When non-experts try to simplify the subject, they all-too-often throw the baby out with the bathwater.
- (4) You can’t do thermodynamics without entropy. Entropy is defined in terms of statistics. As discussed in chapter 2, people who have some grasp of basic probability can understand entropy; those who don’t, can’t. This is part of the price of admission. If you need to brush up on probability, sooner is better than later. A discussion of the basic principles, from a modern viewpoint, can be found in reference 1.

We do not define entropy in terms of energy, nor vice versa. We do not define either of them in terms of temperature. Entropy and energy are well defined even in situations where the temperature is unknown, undefinable, irrelevant, or zero.

- (5) Uncramped thermodynamics is intrinsically multi-dimensional. Even the highly simplified expression  $dE = -P dV + T dS$  involves five variables. To make sense of this requires multi-variable calculus. If you don’t understand how partial derivatives work, you’re not going to get very far.

Furthermore, when using partial derivatives, we must not assume that “variables not mentioned are held constant.” That idea is a dirty trick that may work OK in some simple “textbook” situations, but causes chaos when applied to uncramped thermodynamics, even when applied to something as simple as the ideal gas law, as discussed in reference 2. The fundamental problem is that the various variables are not mutually orthogonal. Indeed, we cannot even

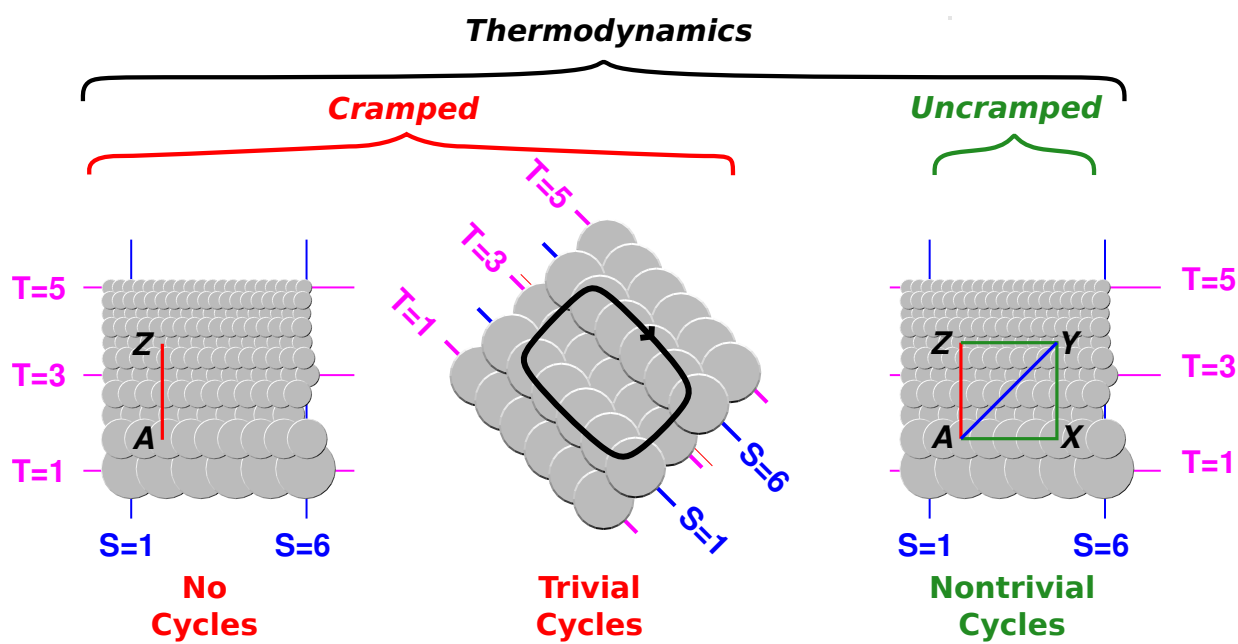


Figure 2: Cramped versus Uncramped Thermodynamics

define what “orthogonal” should mean, because in thermodynamic parameter-space there is no notion of angle and not much notion of length or distance. In other words, there is topology but no geometry, as discussed in section 8.7. This is another reason why thermodynamics is intrinsically and irreducibly complicated.

Uncramped thermodynamics is particularly intolerant of sloppiness, partly because it is so multi-dimensional, and partly because there is no notion of orthogonality. Unfortunately, some thermo books are sloppy in the places where sloppiness is least tolerable.

The usual math-textbook treatment of partial derivatives is dreadful. The standard notation for partial derivatives practically invites misinterpretation.

Some fraction of this mess can be cleaned up just by being careful and not taking shortcuts. Also it may help to *visualize* partial derivatives using the methods presented in reference 3. Even more of the mess can be cleaned up using differential forms, i.e. exterior derivatives and such, as discussed in reference 4. This raises the price of admission somewhat, but not by much, and it’s worth it. Some expressions that seem mysterious in the usual textbook presentation become obviously correct, easy to interpret, and indeed easy to visualize when re-interpreted in terms of gradient vectors. On the other edge of the same sword, some other mysterious expressions are easily seen to be unreliable and highly deceptive.

- (6) If you want to do thermodynamics, beyond a few special cases, you will have to know enough physics to understand what *phase space* is. We have to count states, and the states live in phase space. There are a few exceptions where the states can be counted by other means; these include the spin system discussed in section 11.10, the artificial games discussed in section 2.2 and section 2.3, and some of the more theoretical parts of information theory. Non-exceptions include the more practical parts of information theory; for example, 256-QAM modulation is best understood in terms of phase space. Almost everything dealing with ordinary fluids or chemicals requires counting states in phase space. Sometimes this can be swept under the rug, but it’s still there.

Phase space is well worth learning about. It is relevant to Liouville’s theorem, the fluctuation/dissipation theorem, the optical brightness theorem, the Heisenberg uncertainty principle, and the second law of thermodynamics. It even has application to computer science (symplectic integrators). There are even connections to cryptography (Feistel networks).

- (7) You must appreciate the fact that not every vector field is the gradient of some potential. Many things that non-experts *wish* were gradients are not gradients. You must get your head around this before proceeding. Study Escher’s “Waterfall” as discussed in reference 4 until you understand that the water there has no well-defined height. Even more to the point, study the RHS of figure 8.4 until you understand that there is no well-defined height function, i.e. no well-defined  $Q$  as a function of state. See also section 8.2.

The term “inexact differential” is sometimes used in this connection, but that term is a misnomer, or at best a horribly misleading idiom. We prefer the term *ungrady one-form*. In any case, whenever you encounter a path-dependent integral, you must keep in mind that it is not a potential, i.e. not a function of state. See chapter 19 for more on this.

To say the same thing another way, we will not express the first law as  $dE = dW + dQ$  or anything like that, even though it is traditional in some quarters to do so. For starters, although

such an equation may be meaningful within the narrow context of cramped thermodynamics, it is provably not meaningful for uncramped thermodynamics, as discussed in section 8.2 and chapter 19. It is provably impossible for there to be any  $W$  and/or  $Q$  that satisfy such an equation when thermodynamic cycles are involved.

Even in cramped situations where it might be possible to split  $E$  (and/or  $dE$ ) into a thermal part and a non-thermal part, it is often unnecessary to do so. Often it works just as well (or better!) to use the unsplit energy, making a direct appeal to the conservation law, equation 1.1.

- (8) Almost every newcomer to the field tries to apply ideas of “thermal energy” or “heat content” to uncramped situations. It always *almost* works ... but it never *really* works. See chapter 19 for more on this.
- (9) On the basis of history and etymology, you might think thermodynamics is all about heat, but it’s not. Not anymore. By way of analogy, there was a time when what we now call thermodynamics was all about phlogiston, but it’s not anymore. People wised up. They discovered that one old, imprecise idea (phlogiston) could be and should be replaced two new, precise ideas (oxygen and energy). More recently, it has been discovered that one old, imprecise idea (heat) can be and should be replaced by two new, precise ideas (energy and entropy).

Heat remains central to unsophisticated cramped thermodynamics, but the modern approach to uncramped thermodynamics focuses more on energy and entropy. Energy and entropy are always well defined, even in cases where heat is not.

The idea of entropy is useful in a wide range of situations, some of which do not involve heat or temperature. As shown in figure 1, mechanics involves energy, information theory involves entropy, and thermodynamics involves both energy and entropy.

You can’t do thermodynamics without energy and entropy.

There are multiple mutually-inconsistent definitions of “heat” that are widely used – or you might say wildly used – as discussed in section 17.1. (This is markedly different from the situation with, say, entropy, where there is really only one idea, even if this one idea has multiple corollaries and applications.) There is no consensus as to “the” definition of heat, and no prospect of achieving consensus anytime soon. There is no need to achieve consensus about “heat”, because we already have consensus about entropy and energy, and that suffices quite nicely. Asking students to recite “THE” definition of heat is worse than useless; it rewards rote regurgitation and punishes actual understanding of the subject.

- (10) Our thermodynamics applies to systems of any size, large or small ... not just large systems. This is important, because we don’t want the existence of small systems to create exceptions to the fundamental laws. When we talk about the entropy of a single spin, we are necessarily thinking in terms of an *ensemble* of systems, identically prepared, with one spin per system. The fact that the ensemble is large does not mean that the system itself is large.
- (11) Our thermodynamics is not restricted to the study of ideal gases. Real thermodynamics has a vastly wider range of applicability, as discussed in chapter 22.

- (12) Even in special situations where the notion of “thermal energy” is well defined, we do not pretend that all thermal energy is kinetic; we recognize that random potential energy is important also. See section 9.3.3.

# Chapter 1

## Energy

### 1.1 Preliminary Remarks

Some things in this world are so fundamental that they cannot be defined in terms of anything more fundamental. Examples include:

- Energy, momentum, and mass.
- Geometrical points, lines, and planes.
- Electrical charge.
- Thousands of other things.

Do not place too much emphasis on pithy, dictionary-style definitions. You need to have a vocabulary of many hundreds of words before you can even begin to read the dictionary.

The dictionary approach often leads to circularity. For example, it does no good to define energy in terms of work, define work in terms of force, and then define force in terms of energy.

The general rule is that words acquire meaning from how they are used. For many things, especially including fundamental things, this is the only worthwhile definition you are going to get.

The real definition comes from how the word is used. The dictionary definition is at best secondary, at best an approximation to the real definition.

Words acquire meaning from how they are used.
--

Geometry books often say explicitly that points, lines, and planes are “undefined” terms, but I prefer to say that they are *implicitly* defined. Equivalently, one could say that they are *retroactively* defined, in the sense that they are used before they are defined. They are initially undefined, but

then gradually come to be defined. They are defined by how they are used in the axioms and theorems.

Here is a quote from page 71 of reference 5:

Here and elsewhere in science, as stressed not least by Henri Poincare, that view is out of date which used to say, “Define your terms before you proceed.” All the laws and theories of physics, including the Lorentz force law, have this deep and subtle character, that they both define the concepts they use (here  $B$  and  $E$ ) and make statements about these concepts. Contrariwise, the absence of some body of theory, law, and principle deprives one of the means properly to define or even to use concepts. Any forward step in human knowledge is truly creative in this sense: that theory, concept, law, and method of measurement — forever inseparable — are born into the world in union.

In other words, it is more important to understand what energy *does* than to rote-memorize some dictionary-style definition of what energy «is».

Energy *is* as energy *does*.

We can apply this idea as follows:

The most salient thing that energy *does* is to uphold the local energy-conservation law, equation 1.1.

That means that if we can identify one or more forms of energy, we can identify all the others by seeing how they plug into the energy-conservation law. A catalog of possible starting points and consistency checks is given in equation 1.2 in section 1.3.

## 1.2 Conservation of Energy

The *first law of thermodynamics* states that energy obeys a local conservation law.

By this we mean something very specific:

Any decrease in the amount of energy in a given region of space must be exactly balanced by a simultaneous increase in the amount of energy in an adjacent region of space.

Note the adjectives “simultaneous” and “adjacent.” The laws of physics do not permit energy to disappear now and reappear later. Similarly the laws do not permit energy to disappear from here and reappear at some distant place. Energy is conserved right here, right now.

It is usually possible<sup>1</sup> to observe and measure the physical processes whereby energy is transported from one region to the next. This allows us to express the energy-conservation law as an equation:

$$\begin{array}{lcl} \text{change in energy} & = & \text{net flow of energy} \\ \text{(inside boundary)} & & \text{(inward minus outward across boundary)} \end{array} \quad (1.1)$$

---

<sup>1</sup>Even in cases where measuring the energy flow is not feasible in practice, we assume it is possible in principle.



The word “flow” in this expression has a very precise technical meaning, closely corresponding to one of the meanings it has in everyday life. See reference 6 for the details on this.

There is also a *global* law of conservation of energy: The total energy in the universe cannot change. The local law implies the global law but not conversely. The global law is interesting, but not nearly as useful as the local law, for the following reason: suppose I were to observe that some energy has vanished from my laboratory. It would do me no good to have a global law that asserts that a corresponding amount of energy has appeared “somewhere” else in the universe. There is no way of checking that assertion, so I would not know and not care whether energy was being globally conserved.<sup>2</sup> Also it would be very hard to reconcile a non-local law with the requirements of special relativity.

As discussed in reference 6, there is an important distinction between the notion of *conservation* and the notion of *constancy*. Local conservation of energy says that the energy in a region is constant *except* insofar as energy flows across the boundary.

### 1.3 Examples of Energy

Consider the contrast:

The conservation law presented in section 1.2 does not, by itself, define energy. That’s because there are lots of things that obey the same kind of conservation law. Energy is conserved, momentum is conserved, electric charge is conserved, et cetera.

On the other hand, examples of energy would not, by themselves, define energy.

On the third hand, given the conservation law plus one or more examples of energy, we can achieve a pretty good understanding of energy by a two-step process, as follows:

- 1) Energy includes each of the known examples, such as the things itemized in equation 1.2.
- 2) Energy also includes anything that can be converted to or from previously-known types of energy in accordance with the law of conservation of energy.

For reasons explained in section 1.1, we introduce the terms energy, momentum, and mass as initially-undefined terms. They will gradually acquire meaning from how they are used.

Here are a few well-understood examples of energy. Please don’t be alarmed by the length of the list. You don’t need to understand every item here; indeed if you understand any one item, you can

---

<sup>2</sup>In some special cases, such as Wheeler/Feynman absorber theory, it is possible to make sense of non-local laws, provided we have a non-local conservation law *plus a lot of additional information*. Such theories are unconventional and very advanced, far beyond the scope of this document.

use that as your starting point for the two-step process mentioned above.

$$\text{quantum mechanical energy: } E = \hbar\omega \quad (1.2a)$$

$$\text{relativistic energy in general: } E = \sqrt{(m^2c^4 + p_{xyz}^2c^2)} \quad (1.2b)$$

$$\text{relativistic rest energy: } E_0 = mc^2 \quad (1.2c)$$

$$\text{low-speed kinetic energy: } E_K = \frac{1}{2}p_{xyz} \cdot v \quad (1.2d)$$

$$\text{high-speed kinetic energy: } E_K = p_{xyz} \cdot v \quad (1.2e)$$

$$\text{universal gravitational energy: } E_G = GMm/r \quad (1.2f)$$

$$\text{local gravitational energy: } E_g = mgh \quad (1.2g)$$

$$\text{virtual work: } dE = -F \cdot dx \quad (1.2h)$$

$$\text{Hookean spring energy: } E_{sp} = \frac{1}{2}kx^2 \quad (1.2i)$$

$$\text{capacitive energy: } E_C = \frac{1}{2}CV^2 \quad (1.2j)$$

$$= \frac{1}{2}Q^2/C \quad (1.2k)$$

$$\text{inductive energy: } E_L = \frac{1}{2}LI^2 \quad (1.2l)$$

In particular, if you need a starting-point for your understanding of energy, perhaps the simplest choice is kinetic energy. A fast-moving book has more energy than it would at a lower speed. Some of the examples in equation 1.2 are less fundamental than others. For example, it does no good to define energy via equation 1.2j, if your definition of voltage assumed some prior knowledge of what energy is. Also, equation 1.2c, equation 1.2d and equation 1.2e can all be considered corollaries of equation 1.2b. Still, plenty of the examples are fundamental enough to serve as a starting point. For example:

- If you can define charge, you can calculate the energy via equation 1.2k, by constructing a capacitor of known geometry (and therefore known capacitance). Note that you can measure charge as a multiple of the elementary charge.
- If you can define time, you can calculate the energy via equation 1.2a. Note that SI defines time in terms of cesium hyperfine transitions.
- If you can define mass, you can calculate the energy via equation 1.2c. This is a special case of the more fundamental equation 1.2b. See reference 7 for details on what these equations mean. Note that you can define mass by counting out a mole of  $^{12}\text{C}$  atoms, or go to Paris and use the SI standard kilogram.

The examples that you don't choose as the starting point serve as valuable cross-checks.

We consider things like Planck's constant, Coulomb's constant, and the speed of light to be already known, which makes sense, since they are universal constants. We can use such things freely in our effort to understand how energy behaves.

It must be emphasized that we are talking about the *physics* energy. Do not confuse it with plebeian notions of "available energy" as discussed in section 1.7 and especially section 1.8.1.

## 1.4 Remark: Recursion

The description of energy in section 1.3 is *recursive*. That means we can pull our understanding of energy up by the bootstraps. We can identify new forms of energy as they come along, because

they contribute to the conservation law in the same way as the already-known examples. This is the same basic idea as in reference 8.

Recursive is not the same as circular. A circular argument would be fallacious and useless ... but there are many examples of correct, well-accepted definitions that are recursive. Note one important distinction: Circles never end, whereas a properly-constructed recursion does end.<sup>3</sup> Recursion is very commonly used in mathematics and computer science. For example, it is correct and convenient to define the factorial function so that

$$\begin{array}{ll} \text{factorial}(0) &:= 1 \\ \text{factorial}(N) &:= N \text{ factorial}(N-1) \end{array} \quad \begin{array}{l} \text{and} \\ \text{for all integers } N > 0 \end{array} \quad (1.3)$$

As a more sophisticated example, have you ever wondered how mathematicians define the concept of *integers*? One very common approach is to define the positive integers via the *Peano axioms*. The details aren't important, but the interesting point is that these axioms provide a *recursive* definition ... not circular, just recursive. This is a precise, rigorous, formal definition.

This allows us to make another point: There are a lot of people who are able to count, even though they are not able to provide a concise definition of “integer” – and certainly not able to provide a non-recursive definition. By the same token, there are lots of people who have a rock-solid understanding of how energy behaves, even though they are not able to give a concise and/or non-recursive definition of “energy”.

## 1.5 Energy is Completely Abstract

Energy is an abstraction. This is helpful. It makes things very much simpler. For example, suppose an electron meets a positron. The two of them annihilate each other, and a couple of gamma rays go flying off, with 511 keV of energy apiece. In this situation the number of electrons is not conserved, the number of positrons is not conserved, the number of photons is not conserved, and mass is not conserved. However, energy is conserved. Even though energy cannot exist without being embodied in some sort of field or particle, the point remains that it exists at a different level of abstraction, separate from the field or particle. We can recognize the energy as being the same energy, even after it has been transferred from one particle to another. This is discussed in more detail in reference 9.

Energy is completely abstract. You need to come to terms with this idea, by accumulating experience, by seeing how energy behaves in various situations. As abstractions go, energy is one of the easiest to understand, because it is so precise and well-behaved.

As another example, consider figure 1.1. Initially there is some energy in ball #1. The energy then flows through ball #2, ball #3, and ball #4 without accumulating there. It accumulates in ball #5, which goes flying.

The net effect is that energy has flowed out of ball #1 and flowed into ball #5. Even though the energy is embodied in a completely different ball, we recognize it as the same energy.

Different ball,  
same energy.

---

<sup>3</sup>You can also construct endless recursions, but they are not nearly so useful, especially in the context of recursive definitions.

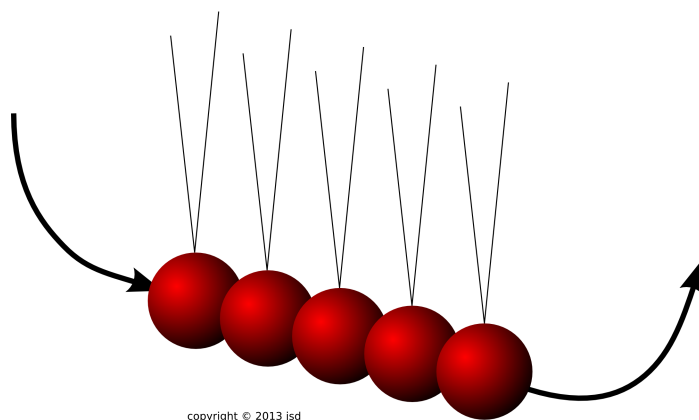


Figure 1.1: Newton's Cradle

## 1.6 Additional Remarks

1. The introductory examples of energy itemized in section 1.3 are only approximate, and are subject to various limitations. For example, the formula  $mgh$  is exceedingly accurate over laboratory length-scales, but is not valid over cosmological length-scales. Similarly the formula  $\frac{1}{2}mv^2$  is exceedingly accurate when speeds are small compared to the speed of light, but not otherwise. These limitations do not interfere with our efforts to understand energy.
2. In non-relativistic physics, energy is a scalar. That means it is not associated with any direction in space. However, in special relativity, energy is not a Lorentz scalar; instead it is recognized as one component of the [energy, momentum] 4-vector, such that energy is associated with the *timelike* direction. For more on this, see reference 10. To say the same thing in other words, the energy is invariant with respect to spacelike rotations, but not invariant with respect to boosts.
3. We will denote the energy by  $E$ . We will denote various sub-categories of energy by putting subscripts on the  $E$ , unless the context makes subscripts unnecessary. Sometimes it is convenient to use  $U$  instead of  $E$  to denote energy, especially in situations where we want to use  $E$  to denote the electric field. Some thermodynamics books state the first law in terms of  $U$ , but it means the same thing as  $E$ . We will use  $E$  throughout this document.
4. Beware of attaching qualifiers to the concept of energy. Note the following contrast:

The symbol  $E$  denotes “the” energy of the system we are considering. If you feel obliged to attach some sort of additional words, you can call  $E$  the “system” energy or the “plain old” energy. This doesn’t change the meaning.

Most other qualifiers change the meaning. There is an important conceptual point here: “The” energy is conserved, but (with rare exceptions) the various sub-categories of energy are not separately conserved. For example, the “available” energy is not necessarily conserved, as discussed in section 1.7.

Associated with the foregoing general conceptual point, here is a specific point of terminology:  $E$  is the plain old total energy, not restricted to “internal” energy or “available” energy.

5. As a related point: If you want to calculate the total energy of the system by summing the various categories of energy, beware that the categories overlap, so you need to be super-careful not to *double count* any of the contributions.
  - For example, suppose we momentarily restrict attention to cramped thermodynamics (such as a heat-capacity experiment), and further suppose we are brave enough to define a notion of “thermal energy” separate from other forms of energy. When adding up the total energy, whatever kinetic energy was counted as part of the so-called “thermal energy” must not be counted again when we calculate the non-thermal kinetic energy, and ditto for the thermal and non-thermal potential energy.
  - Another example that illustrates the same point concerns the *rest energy*,  $E_0$ , which is related to mass via Einstein’s equation<sup>4</sup>  $E_0 = mc^2$ . You can describe the rest energy of a particle in terms of the potential energy and kinetic energy of its internal parts, or in terms of its mass, but you must not add both descriptions together; that would be double-counting.

## 1.7 Energy versus “Capacity to do Work” or “Available Energy”

Non-experts sometimes try to relate energy to “the capacity to do work.” This is never a good idea, for several reasons, as we now discuss.

### 1.7.1 Best Case : Non-Thermal Situation

Consider the following example: We use an ideal battery connected to an ideal motor to raise a weight, doing work against the gravitational field. This is reversible, because we can also operate the motor in reverse, using it as a generator to recharge the battery as we lower the weight.

To analyze such a situation, we don’t need to know anything about thermodynamics. Old-fashioned elementary non-thermal mechanics suffices.

If you do happen to know something about thermodynamics, you can quantify this by saying that the temperature  $T$  is low, and the entropy  $S$  is small, such that any terms involving  $T\Delta S$  are negligible compared to the energy involved.

On the other hand, if you don’t yet know  $T\Delta S$  means, don’t worry about it.

In simple situations such as this, we can define work as  $\Delta E$ . That means energy is related to the ability to do work ... in this simple situation.

---

<sup>4</sup>Einstein intended the familiar expression  $E = mc^2$  to apply only in the rest frame. This is consistent with the modern (post-1908) convention that the mass  $m$  is defined in the rest frame. Calling  $m$  the “rest” mass is redundant but harmless. We write the rest energy as  $E_0$  and write the total energy as  $E$ ; they are not equal except in the rest frame.

### 1.7.2 Equation versus Definition

Even in situations where energy is related to the ability to do work, it is not wise to “define” energy that way, for a number of practical and pedagogical reasons.

Energy is so fundamental that it is not definable in terms of anything more fundamental. You can’t define energy in terms of work unless you already have a solid definition of “work”, and defining work is not particularly easier than defining energy from scratch. It is usually better to start with energy and define work in terms of energy (rather than vice versa), because energy is the more fundamental concept.

### 1.7.3 General Case : Some Energy Not Available

In general, some of the energy of a particular system is available for doing work, and some of it isn’t. The second law of thermodynamics, as discussed in chapter 2, makes it impossible to use all the energy (except in certain very special cases, as discussed in section 1.7.1).

See section 15.5 for more about this.

In this document, the word “energy” refers to the physics energy. However, when business executives and politicians talk about energy, they are generally more concerned about “available energy”, which is an important thing, but it is emphatically not the same as the physics energy. See section 1.8.1 for more about this. It would be a terrible mistake to confuse “available energy” with the physics energy. Alas, this mistake is very common. See section 15.5 for additional discussion of this point.

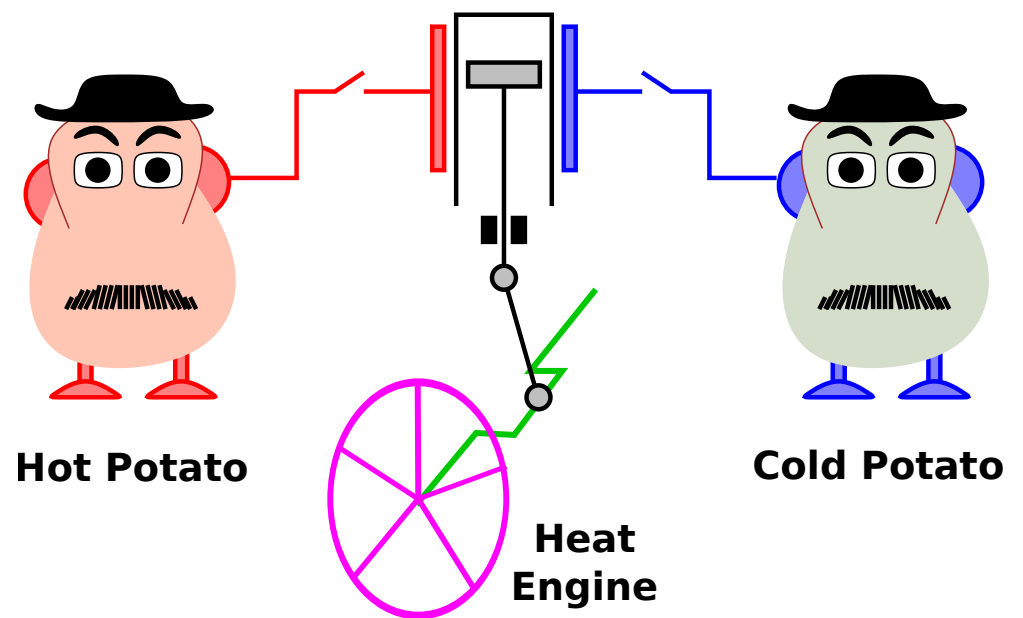
Any attempt to define energy in terms of “capacity to do work” would be inconsistent with thermodynamics, as we see from the following examples:

- |   |   |
|---|---|
| <p>#1: Consider an isolated system containing a hot potato, a cold potato, a tiny heat engine, and nothing else, as illustrated in figure 1.2. This system has some energy and some ability to do work.</p> | <p>#2: Contrast that with a system that is just the same, except that it has two hot potatoes (and no cold potato).</p> |
|---|---|

The second system has *more* energy but *less* ability to do work.

This sheds an interesting side-light on the energy-conservation law. As with most laws of physics, this law, by itself, does not tell you what will happen; it only tells you what *cannot* happen: you cannot have any process that fails to conserve energy. To say the same thing another way: if something is prohibited by the energy-conservation law, the prohibition is absolute, whereas if something is permitted by the energy-conservation law, the permission is conditional, conditioned on compliance with all the other laws of physics. In particular, as discussed in section 9.2, if you want to transfer energy from the “collective” modes of a rapidly-spinning flywheel to the other modes, you have to comply with all the laws, not just conservation of energy. This includes conserving angular momentum. It also includes complying with the second law of thermodynamics.

Let’s be clear: The ability to do work implies energy, but the converse is not true. There are lots of situations where energy *cannot* be used to do work, because of the second law of thermodynamics or some other law.



Equating energy with doable work is just not correct. (In contrast, it might be OK to connect energy with some previously-done work, as opposed to doable work. That is not always convenient or helpful, but at least it doesn't contradict the second law of thermodynamics.)

Some people wonder whether the example given above (the two-potato engine) is invalid because it involves closed systems, not interacting with the surrounding environment. Well, the example is perfectly valid, but to clarify the point we can consider another example (due to Logan McCarty):

- |   |   |
|---|---|
| <p>#1: Consider a system consisting of a room-temperature potato, a cold potato, and a tiny heat engine. This system has some energy and some ability to do work.</p> | <p>#2: Contrast that with a system that is just the same, but except that it has two room-temperature potatoes.</p> |
|---|---|

The second system has *more* energy but *less* ability to do work in the ordinary room-temperature environment.

In some impractical theoretical sense, you might be able to define the energy of a system as the amount of work the system *would* be able to do if it were in contact with an unlimited heat-sink at low temperature (arbitrarily close to absolute zero). That's quite impractical because no such heat-sink is available. If it were available, many of the basic ideas of thermodynamics would become irrelevant.

As yet another example, consider the system shown in figure 1.3. The boundary of the overall "system" is shown as a heavy black line. The system is thermally insulated from its surroundings. The system contains a battery (outlined with a red dashed line) a motor, and a switch. Internal to the battery is a small series resistance  $R_1$  and a large shunt resistance  $R_2$ . The motor drives a thermally-insulated shaft, so that the system can do mechanical work on its surroundings.

- |  |  |
|--|--|
| <p>By closing the switch, we can get the system to perform work on its surroundings by means of the shaft.</p> | <p>On the other hand, if we just wait a moderately long time, the leakage resistor <math>R_2</math> will discharge the battery. This does not change the system's energy (i.e. the energy within the boundary of the system) ... but it greatly decreases the capacity to do work.</p> |
|--|--|

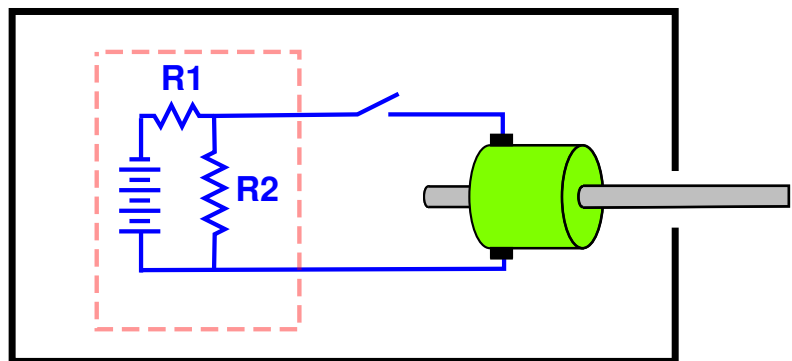
This can be seen as analogous to the NMR  $\tau_2$  process. An analogous mechanical system is discussed in section 11.5.5. All these examples share a common feature, namely a change in entropy with no change in energy.

To remove any vestige of ambiguity, imagine that the system was initially far below ambient temperature, so that the Joule heating in the resistor brings the system closer to ambient temperature. See reference 11 for Joule's classic paper on electrical heating.

To repeat: In real-world situations, energy is not the same as "available energy" i.e. the capacity to do work.

What's worse, any measure of "available" energy is not a function of state. Consider again the two-potato system shown in figure 1.2. Suppose you know the state of the left-side potato, including its energy  $E_1$ , its temperature  $T_1$ , its entropy  $S_1$ , its mass  $m_1$ , its volume  $V_1$ , its free energy  $F_1$ , and its free enthalpy  $G_1$ . That all makes sense so far, because those are all functions of state, determined





by the state of that potato by itself. Alas you don't know what fraction of that potato's energy should be considered thermodynamically "available" energy, and you can't figure it out using only the properties of that potato. In order to figure it out, you would need to know the properties of the other potato as well.

For a homogenous subsystem,  
loosely in contact with the environment,  
its energy is a function of its state.  
Its "available energy" is not.

Every beginner wishes for a state function that specifies the "available energy" content of a system. Alas, wishing does not make it so. No such state function can possibly exist.

(When we say two systems are "loosely in contact" we mean they are neither completely isolated nor completely in equilibrium.)

Also keep in mind that the law of conservation of energy applies to the real energy, not to the "available" energy.

Energy obeys a strict local conservation law.  
"Available energy" does not.

Beware that the misdefinition of energy in terms of "ability to do work" is extremely common. This misdefinition is all the more pernicious because it works OK in simple non-thermodynamical situations, as discussed in section 1.7.1. Many people learn this misdefinition, and some of them have a hard time unlearning it.

## 1.8 Mutations

### 1.8.1 Energy

In physics, there is only one meaning for the term *energy*. For all practical purposes, there is complete agreement among physicists as to what energy is. This stands in dramatic contrast to other terms – such as “heat” – that have a confusing multiplicity of technical meanings even within physics; see section 17.1 for more discussion of “heat”, and see chapter 20 for a more general discussion of ambiguous terminology.

The two main meanings of “energy” are different enough so that the difference is important, yet similar enough to be highly deceptive.

The physics energy is conserved. It is conserved automatically, strictly, and locally, in accordance with equation 1.1.

In ordinary conversation, when people speak of “energy” – even in a somewhat-technical sense – they are usually talking about some kind of “available energy” or “useful energy” or something like that. This is an important concept, but it is very hard to quantify, and it is definitely not equal to the physics energy. Examples include the “Department of Energy” or the “energy industry”.

For the next level of detail on this, see section 20.3.

### 1.8.2 Conservation

In physics, there is almost<sup>5</sup> only one definition of conservation. However, we run into trouble when we consider the plebeian meanings.

The two main meanings of “conservation” are different enough so that the difference is important, yet similar enough to be highly deceptive.

The main physics definition of conservation is synonymous with continuity of flow, as defined in equation 1.1. See reference 6.

The plebeian notion of “conservation” means saving, preserving, not wasting, not dissipating. It definitely is not equivalent to continuity of flow. Example: “conservation” of endangered wildlife.

For the next level of detail on this, see section 20.4.

### 1.8.3 Energy Conservation

Combining the two previous ideas, we see that the simple phrase “energy conservation” is practically begging to be misunderstood. You could suffer from two profound misconceptions in a simple two-word phrase.

---

<sup>5</sup>Beware: You might think the adjective “conservative” refers to the same idea as the noun “conservation,” and this is *almost* true, with one exception, as discussed in section 20.4.

### 1.8.4 Internal Energy

A great many thermodynamics books emphasize the so-called «internal energy», denoted  $U$  or  $E_{\text{in}}$ . I have never found it necessary to make sense of this. Instead I reformulate everything in terms of the plain old energy  $E$  and proceed from there. For details, see section 7.7.

## 1.9 Range of Validity

The law of conservation of energy has been tested and found 100% reliable for all practical purposes, and quite a broad range of impractical purposes besides.

Of course everything has limits. It is not necessary for you to have a very precise notion of the limits of validity of the law of conservation of energy; that is a topic of interest only to a small community of specialists. The purpose of this section is merely to indicate, in general terms, just how remote the limits are from everyday life.

If you aren't interested in details, feel free to skip this section.

Here's the situation:

- For all practical purposes, energy is strictly and locally conserved.
- For all purposes (practical or not), whenever the classical (Newtonian) theory of gravity is an adequate approximation, energy is strictly and locally conserved.
- In special relativity, the [energy,momentum] 4-vector is locally conserved. In any particular frame, each component of the [energy,momentum] 4-vector is separately conserved, and energy is just the timelike component. See reference 6 for a way to visualize continuity of flow in spacetime, in terms of the continuity of world-lines.
- Even in general relativity, even when the Newtonian theory does not apply, there is a well-defined notion of conservation of energy, *provided* we restrict attention to regions of spacetime that are flat or at least asymptotically flat as you go far away. You need to express “flow” in terms of the covariant divergence  $T_{\mu\nu;\nu}$  not the coordinate divergence  $T_{\mu\nu,\nu}$ , as discussed in reference 12. That isn't a special property of energy, but rather a general property of the conservation idea: continuity of flow needs to be expressed in terms of the covariant derivative.
- However, if we take a completely unrestricted view of general relativity, the notion of conservation of energy is problematic. For starters, if the universe has the topology of a torus (such as a donut, or an apple with a wormhole in it), the notion of “energy inside a boundary” is ill-defined, because the fundamental notion of “inside” is ill-defined for any contour that winds through the hole or around the hole, i.e. is not topologically reducible to a point.

## Chapter 2

# Entropy

### 2.1 Paraconservation

The second law of thermodynamics states that entropy obeys a local paraconservation law. That is, entropy is “nearly” conserved. By that we mean something very specific:

$$\begin{array}{lcl} \text{change in entropy} & \geq & \text{net flow of entropy} \\ \text{(inside boundary)} & & \text{(inward minus outward across boundary)} \end{array} \quad (2.1)$$

The structure and meaning of equation 2.1 is very similar to equation 1.1, except that it has an inequality instead of an equality. It tells us that the entropy in a given region can increase, but it cannot decrease except by flowing into adjacent regions.

As usual, the local law implies a corresponding global law, but not conversely; see the discussion at the end of section 1.2.

Entropy is absolutely essential to thermodynamics ... just as essential as energy. The relationship between energy and entropy is diagrammed in figure 1. Section 0.1 discusses the relationship between basic mechanics, information theory, and thermodynamics. Some relevant applications of entropy are discussed in chapter 22.

You can't do thermodynamics without entropy.

Entropy is defined and quantified in terms of probability, as discussed in section 2.6. In some situations, there are important connections between entropy, energy, and temperature ... but these do not define entropy; see section 2.5.7 for more on this. The first law (energy) and the second law (entropy) are logically independent.

If the second law is to mean anything at all, entropy must be well-defined always. Otherwise we could create loopholes in the second law by passing through states where entropy was not defined.

Entropy is related to information. Essentially it is the opposite of information, as we see from the following scenarios.

## 2.2 Scenario: Cup Game

As shown in figure 2.1, suppose we have three blocks and five cups on a table.



Figure 2.1: The Cup Game

To illustrate the idea of entropy, let's play the following game: Phase 0 is the preliminary phase of the game. During phase 0, the dealer hides the blocks under the cups however he likes (randomly or otherwise) and optionally makes an announcement about what he has done. As suggested in the figure, the cups are transparent, so the dealer knows the exact microstate at all times. However, the whole array is behind a screen, so the rest of us don't know anything except what we're told.

Phase 1 is the main phase of the game. During phase 1, we (the players) strive to figure out where the blocks are. We cannot see what's inside the cups, but we are allowed to ask yes/no questions, whereupon the dealer will answer. Our score in the game is determined by the number of questions we ask; each question contributes one bit to our score. The goal is to locate all the blocks using the smallest number of questions.

Since there are five cups and three blocks, we can encode the location of all the blocks using a three-symbol string, such as 122, where the first symbol identifies the cup containing the red block, the second symbol identifies the cup containing the black block, and the third symbol identifies the cup containing the blue block. Each symbol is in the range zero through four inclusive, so we can think of such strings as base-5 numerals, three digits long. There are  $5^3 = 125$  such numerals. (More generally, in a version where there are  $N$  cups and  $B$  blocks, there are  $N^B$  possible microstates.)

1. Example: During phase 0, the dealer announces that all three blocks are under cup #4. Our score is zero; we don't have to ask any questions.
2. Example: During phase 0, the dealer places all the blocks randomly and doesn't announce anything. If we are smart, our score  $S$  is at worst 7 bits (and usually exactly 7 bits). That's because when  $S = 7$  we have  $2^S = 2^7 = 128$ , which is slightly larger than the number of possible states. In the expression  $2^S$ , the base is 2 because we are asking questions with 2 possible answers. Our minimax strategy is simple: we write down all the states in order, from 000 through 444 (base 5) inclusive, and ask questions of the following form: Is the actual state in the first half of the list? Is it in the first or third quarter? Is it in an odd-numbered eighth? After at most seven questions, we know exactly where the correct answer sits in the list.
3. Example: During phase 0, the dealer hides the blocks at random, then makes an announcement that provides partial information, namely that cup #4 happens to be empty. Then (if we follow a sensible minimax strategy) our score will be six bits, since  $2^6 = 64 = 4^3$ .

To calculate what our score will be, we don't need to know anything about energy; all we have to do is count states (specifically, the number of accessible microstates). States are states; they are not energy states.

Superficially speaking, if you wish to make this game seem more thermodynamical, you can assume that the table is horizontal, and the blocks are non-interacting, so that all possible configurations have the same energy. But really, it is easier to just say that over a wide range of energies, energy has got nothing to do with this game.

The point of all this is that we can measure the *entropy* of a given situation according to the number of questions we have to ask to finish the game, starting from the given situation. Each yes/no question contributes one bit to the entropy, assuming the question is well designed. This measurement is indirect, almost backwards, but it works. It is like measuring the volume of an odd-shaped container by quantifying the amount of water it takes to fill it.

The central, crucial idea of entropy is that it measures how much we don't know about the situation. Entropy is *not knowing*.

## 2.3 Scenario: Card Game

Here is a card game that illustrates the same points as the cup game. The only important difference is the size of the state space: roughly eighty million million million million million million million million million states, rather than 125 states. That is, when we move from 5 cups to 52 cards, the state space gets bigger by a factor of  $10^{66}$  or so.

Consider a deck of 52 playing cards. By re-ordering the deck, it is possible to create a large number (52 factorial) of different configurations.

Technical note: There is a separation of variables. We choose to consider only the part of the system that describes the ordering of the cards. We assume these variables are statistically independent of other variables, such as the spatial position and orientation of the cards. This allows us to understand the entropy of this subsystem separately from the rest of the system, for reasons discussed in section 2.8.

Also, unless otherwise stated, we assume the number of cards is fixed at 52 ... although the same principles apply to smaller or larger decks, and sometimes in an introductory situation it is easier to see what is going on if you work with only 8 or 10 cards.

Phase 0 is the preliminary phase of the game. During phase 0, the dealer prepares the deck in a configuration of his choosing, using any combination of deterministic and/or random procedures. He then sets the deck on the table. Finally he makes zero or more announcements about the configuration of the deck.

Phase 1 is the main phase of the game. During phase 1, we (the players) strive to fully describe the configuration, i.e. to determine which card is on top, which card is second, et cetera. We cannot look at the cards, but we can ask yes/no questions, whereupon the dealer will answer. Each question contributes one bit to our score. Our objective is to ascertain the complete configuration using the smallest number of questions. As we shall see, our score is a measure of the entropy of the game.

Note the contrast between microstate and macrostate:

<p>One configuration of the card deck corresponds to one microstate. The microstate does not change during phase 1.</p>	<p>The macrostate is the ensemble of microstates consistent with what we know about the situation. This changes during phase 1. It changes whenever we obtain more information about the situation.</p>
---	---

- (1) Example: The dealer puts the deck in some well-known reference configuration, and announces that fact. Then we don't need to do anything, and our score is zero – a perfect score.
- (2) Example: The dealer puts the deck in the reverse of the reference configuration, and announces that fact. We can easily tell which card is where. We don't need to ask any questions, so our score is again zero.
- (3) Example: The dealer starts with the reference configuration, then “cuts” the deck; that is, he chooses at random one of the 52 possible full-length cyclic permutations, and applies that permutation to the cards. He announces what procedure he has followed, but nothing more.

At this point we know that the deck is in *some* microstate, and the microstate is not changing ... but we don't know *which* microstate. It would be foolish to pretend we know something we don't. If we're going to bet on what happens next, we should calculate our odds based on the ensemble of possibilities, i.e. based on the macrostate.

Our best strategy is as follows: By asking six well-chosen questions, we can find out which card is on top. We can then easily describe every detail of the configuration. Our score is six bits.

- (4) Example: The dealer starts with the standard configuration, cuts it, and then cuts it again. The second cut changes the microstate, but does not change the macrostate. Cutting the deck is, so far as the macrostate is concerned, idempotent; that is,  $N$  cuts are the same as one. It still takes us six questions to figure out the full configuration.

This illustrates that the entropy is a property of the ensemble, i.e. a property of the macrostate, not a property of the microstate. Cutting the deck the second time changed the microstate but did not change the macrostate. See section 2.4 and especially section 2.7.1 for more discussion of this point.

- (5) Example: Same as above, but in addition to announcing the procedure the dealer also announces what card is on top. Our score is zero.
- (6) Example: The dealer shuffles the deck thoroughly. He announces that, and only that. The deck could be in any of the 52 factorial different configurations. If we follow a sensible (minimax) strategy, our score will be 226 bits, since the base-2 logarithm of 52 factorial is approximately 225.581. Since we can't ask fractional questions, we round up to 226.
- (7) Example: The dealer announces that it is equally likely that he has either shuffled the deck completely or left it in the reference configuration. Then our score *on average* is only 114 bits, if we use the following strategy: we start by asking whether the deck is already in the reference configuration. That costs us one question, but half of the time it's the only question we'll need. The other half of the time, we'll need 226 more questions to unshuffle the shuffled deck. The average of 1 and 227 is 114.

Note that we are not depending on any special properties of the “reference” state. For simplicity, we could agree that our reference state is the factory-standard state (cards ordered according to suit and number), but any other agreed-upon state would work just as well. If we know deck is in Moe's favorite state, we can easily rearrange it into Joe's favorite state. Rearranging it from one known state to another known state does not involve any entropy.



## 2.4 Peeking

As a variation on the game described in section 2.3, consider what happens if, at the beginning of phase 1, we are allowed to peek at one of the cards.

In the case of the standard deck, example 1, this doesn't tell us anything we didn't already know, so the entropy remains unchanged.

In the case of the cut deck, example 3, this lowers our score by six bits, from six to zero.

In the case of the shuffled deck, example 6, this lowers our score by six bits, from 226 to 220.

This is worth mentioning because peeking can (and usually does) change the macrostate, but it cannot change the microstate. This stands in contrast to cutting an already-cut deck or shuffling an already-shuffled deck, which changes the microstate but does not change the macrostate. This is a multi-way contrast, which we can summarize as follows:

	Macrostate Changes	Macrostate Doesn't Change
Microstate Usually Changes:	Shuffling a deck that was not already shuffled.	Shuffling an already-shuffled deck, or cutting an already-cut deck.
Microstate Doesn't Change:	Peeking at something that was not already known.	Doing nothing.

This gives us a clearer understanding of what the macrostate is. Essentially the macrostate is the ensemble, in the sense that specifying the ensemble specifies the macrostate and vice versa. Equivalently, we can say that the macrostate is a probability distribution over microstates.

In the simple case where all the microstates are equiprobable, the ensemble is simply the set of all microstates that are consistent with what you know about the system.

In a poker game, there is only one deck of cards. Suppose player Alice has peeked but player Bob has not. Alice and Bob will then play according to very different strategies. They will use different ensembles – different macrostates – when calculating their next move. The deck is the same for both, but the macrostate is not. This idea is discussed in more detail in connection with figure 2.4 in section 2.7.2.

We see that the physical state of the deck does not provide a complete description of the macrostate. The players' knowledge of the situation is also relevant, since it affects how they calculate the probabilities. Remember, as discussed in item 4 and in section 2.7.1, entropy is a property of the macrostate, not a property of the microstate, so peeking can – and usually does – change the entropy.

To repeat: Peeking does not change the microstate, but it can have a large effect on the macrostate. If you don't think peeking changes the ensemble, I look forward to playing poker with you!

## 2.5 Discussion

### 2.5.1 Connecting Models to Reality

The models we have been discussing are in some ways remarkably faithful to reality. In some ways they are not merely models but bona-fide examples of entropy. Even so, the entropy in one situation might or might not convey a good understanding of entropy in another situation. Consider the following contrasts:

#### Small Model Systems

We have been considering systems with only a few bits of entropy (the cup game in section 2.2) or a few hundred bits (the card deck in section 2.3).

Entropy is defined in terms of probability. Entropy is not defined in terms of entropy, nor vice versa. There are plenty of situations where the second law is important, even though the temperature is irrelevant, unknown, and/or undefinable.

#### More Generally

In the physics lab or the chemistry lab, one might be dealing with 100 moles of bits. A factor of  $6 \times 10^{23}$  can have significant practical consequences.

In some (albeit not all) practical situations there is a big emphasis on the relationship between energy and temperature, which is controlled by entropy.

There are some experiments where it is possible to evaluate the entropy two ways, namely by counting the bits and by measuring the energy and temperature. The two ways always agree. Furthermore, the second law allows us to compare the entropy of one object with the entropy of another, which provides strong evidence that all the entropy in the world conforms to the faithful workhorse expression, namely equation 2.2, as we shall see in section 2.6.

We have been asking yes/no questions. Binary questions are required by the rules of the simple games.

The answer to each yes/no question gives us one bit of information, if the two answers (yes and no) are equally likely.

So far, we have only considered scenarios where all accessible microstates are equally probable. In such scenarios, the entropy is the logarithm of the number of accessible microstates.

Other games may allow more general types of questions. For example, consider the three-way measurements in reference 13.

If the two answers are not equally likely, the information will be less.

If the accessible microstates are not equally probable, we need a more sophisticated notion of entropy, as discussed in section 2.6.

Remark on terminology: Any microstates that have appreciable probability are classified as *accessible*. In contrast, microstates that have negligible probability are classified as *inaccessible*. Their probability doesn't have to be zero, just small enough that it can be ignored without causing too much damage.

Thermodynamics is a very practical subject. It is much more common for some probability to be negligible than to be strictly zero.

In order to identify the correct microstate with confidence, we need to ask a sufficient number of questions, so that the information contained in the answers is greater than or equal to the entropy of the game.

### 2.5.2 States and Probabilities

If you want to understand entropy, you must first have at least a modest understanding of basic probability. It's a prerequisite, and there's no way of getting around it. Anyone who knows about probability can learn about entropy. Anyone who doesn't, can't.

Our notion of entropy is completely dependent on having a notion of microstate, and on having a procedure for assigning a probability to each microstate.

In some special cases, the procedure involves little more than counting the accessible (aka "allowed") microstates, as discussed in section 9.6. This type of counting corresponds to a particularly simple, flat probability distribution, which may be a satisfactory approximation in special cases, but is definitely not adequate for the general case.

For simplicity, the cup game and the card game were arranged to embody a clear notion of microstate. That is, the rules of the game specified what situations would be considered the "same" microstate and what would be considered "different" microstates. Note the contrast:

Our discrete games provide a model that is directly and precisely applicable to physical systems where the physics is naturally discrete, such as systems involving only the nonclassical spin of elementary particles. An important example is the demagnetization refrigerator discussed in section 11.10.

For systems involving continuous variables such as position and momentum, counting the states is somewhat trickier ... but it can be done. The correct procedure is discussed in section 12.3.

For additional discussion of the relevance of entropy, as applied to thermodynamics and beyond, see chapter 22.

### 2.5.3 Entropy is *Not Knowing*

The point of all this is that the “score” in these games is an example of entropy. Specifically: at each point in the game, there are two numbers worth keeping track of: the number of questions we have already asked, and the number of questions we must ask to finish the game. The latter is what we call the the entropy of the situation at that point.

Entropy is *not knowing*.  
Entropy measures how much is *not known* about the situation.

Remember that the macrostate is the ensemble of microstates. In the ensemble, probabilities are assigned taking into account what the observer knows about the situation. The entropy is a property of the macrostate.

At each point during the game, the entropy is a property of the macrostate, not of the microstate. The system must be in “some” microstate, but generally we don’t know which microstate, so all our decisions must be based on the macrostate.

The value any given observer assigns to the entropy depends on what *that observer* knows about the situation, not what the dealer knows, or what anybody else knows. This makes the entropy somewhat context-dependent. Indeed, it is somewhat subjective. Some people find this irksome or even shocking, but it is real physics. For a discussion of context-dependent entropy, see section 12.8.

### 2.5.4 Entropy versus Energy

Note that entropy has been defined without reference to temperature and without reference to heat. Room temperature is equivalent to zero temperature for purposes of the cup game and the card game. Arguably, in theory, there is “some” chance that thermal agitation will cause two of the cards to spontaneously hop up and exchange places during the game, but that is really, really negligible.

Non-experts often try to define entropy in terms of energy. This is a mistake. To calculate the entropy, I don’t need to know anything about energy; all I need to know is the probability of each relevant state. See section 2.6 for details on this.

States are states;  
they are not energy states.

Entropy is not defined in terms of energy, nor vice versa.

In some cases, there is a simple mapping that allows us to identify the  $i$ th microstate by means of its energy  $\hat{E}_i$ . It is often convenient to exploit this mapping when it exists, but it does not always exist.

### 2.5.5 Entropy versus Disorder

In pop culture, entropy is often associated with *disorder*. There are even some textbooks that try to explain entropy in terms of disorder. This is not a good idea. It is all the more disruptive because

it is in some sense half true, which means it might pass superficial scrutiny. However, science is not based on half-truths.

Small disorder generally implies small entropy. However, the converse does not hold, not even approximately; A highly-disordered system might or might not have high entropy. The spin echo experiment (section 11.7) suffices as an example of a highly disordered macrostate with relatively low entropy.

Before we go any farther, we should emphasize – again – that entropy is a property of the macrostate, not of the microstate. In contrast, to the extent that “disorder” has any definite meaning at all, it is a property of the microstate. Therefore, whatever the “disorder” is measuring, it isn’t entropy. (A similar microstate versus macrostate argument applies to the “energy dispersal” model of entropy, as discussed in section 9.8.) As a consequence, the usual textbook illustration – contrasting snapshots of orderly and disorderly scenes – cannot be directly interpreted in terms of entropy. To get any scrap of value out of such an illustration, the reader must make a sophisticated leap:

The disorderly snapshot must be interpreted as *representative* of an ensemble with a very great number of similarly-disorderly microstates. The ensemble of disorderly microstates has high entropy. This is a property of the ensemble, not of the depicted microstate or any other microstate.

The orderly snapshot must be interpreted as *representative* of a very small ensemble, namely the ensemble of similarly-orderly microstates. This small ensemble has a small entropy. Again, entropy is a property of the ensemble, not of any particular microstate (except in the extreme case where there is only one microstate in the ensemble, and therefore zero entropy).

To repeat: Entropy is defined as a weighted average over all microstates. Asking about the entropy of a particular microstate (disordered or otherwise) is asking the wrong question. If you know what microstate the system is in, the entropy is zero. Guaranteed.

Note the following contrast:

Entropy is a property of the macrostate. It is defined as an ensemble average.

Disorder, to the extent it can be defined at all, is a property of the microstate. (You might be better off focusing on the surprisal rather than the disorder, as discussed in section 2.7.1.)

The number of orderly microstates is very small compared to the number of disorderly microstates. That’s because when you say the system is “ordered” you are placing constraints on it. Therefore if you know that the system is in one of those orderly microstates, you know the entropy cannot be very large.

The converse does not hold. If you know that the system is in some disorderly microstate, you do *not* know that the entropy is large. Indeed, if you know that the system is in some *particular* disorderly microstate, the entropy is zero. (This is a corollary of the more general proposition that if you know what microstate the system is in, the entropy is zero. it doesn’t matter whether that state “looks” disorderly or not.)

Furthermore, there are additional reasons why the typical text-book illustration of a messy dorm room is not a good model of entropy. For starters, it provides no easy way to define and delimit

the states. Even if we stipulate that the tidy state is unique, we still don't know how many untidy states there are. We don't know to what extent a shirt on the floor "here" is different from a shirt on the floor "there." Since we don't know how many different disorderly states there are, we can't quantify the entropy. (In contrast the games in section 2.2 and section 2.3 included a clear rule for defining and delimiting the states.)

Examples of low entropy and relatively high disorder include, in order of increasing complexity:

1. Quantum mechanical zero-point motion of a particle, such as the electron in a hydrogen atom. The electron is spread out over a nonzero area in phase space. It is spread in a random, disorderly way. However, the electron configuration is in a single, known quantum state, so it has zero entropy.
2. Five coins in a known microstate.

Technical note: There is a separation of variables, analogous to the separation described in section 2.3. We consider only the part of the system that describes whether the coins are in the "heads" or "tails" state. We assume this subsystem is statistically independent of the other variables such as the position of the coins, rotation in the plane, et cetera. This means we can understand the entropy of this subsystem separately from the rest of the system, for reasons discussed in section 2.8.

Randomize the coins by shaking. The entropy at this point is five bits. If you open the box and peek at the coins, the entropy goes to zero. This makes it clear that entropy is a property of the ensemble, not a property of the microstate. Peeking does not change the disorder. Peeking does not change the microstate. However, it can (and usually does) change the entropy. This example has the pedagogical advantage that it is small enough that the entire microstate-space can be explicitly displayed; there are only  $32 = 2^5$  microstates.

3. A deck of cards in a known microstate.

Ordinarily, a well-shuffled deck of cards contains 225.581 bits of entropy, as discussed in section 2.3. On the other hand, if you have peeked at all the cards after they were shuffled, the entropy is now zero, as discussed in section 2.4. Again, this makes it clear that entropy is a property of the ensemble, not a property of the microstate. Peeking does not change the disorder. Peeking does not change the microstate. However, it can (and usually does) change the entropy.

Many tricks of the card-sharp and the "magic show" illusionist depend on a deck of cards arranged to have much disorder but little entropy.

4. In cryptography, suppose we have a brand-new *one time pad* containing a million random hex digits. From our adversary's point of view, this embodies 4,000,000 bits of entropy. If, however, the adversary manages to make a copy of our one time pad, then the entropy of our pad, from his point of view, goes to zero. All of the complexity is still there, all of the disorder is still there, but the entropy is gone.
5. The spin echo experiment involves a highly complicated state that has low entropy. See section 11.7. This is a powerful example, because it involve a macroscopic amount of entropy. For a reasonable-sized sample, the total entropy could be on the order of 1 joule per kelvin, i.e. on the order of  $10^{23}$  bits, not just a few bits or a few hundred bits).

### 2.5.6 False Dichotomy, or Not

There is a long-running holy war between those who try to define entropy in terms of energy, and those who try to define it in terms of disorder. This is based on a grotesquely false dichotomy: If entropy-as-energy is imperfect, then entropy-as-disorder “must” be perfect ... or vice versa. I don’t know whether to laugh or cry when I see this. Actually, both versions are highly imperfect. You might get away with using one or the other in *selected* situations, but not in general.

The right way to define entropy is in terms of probability, as we now discuss. (The various other notions can then be understood as special cases and/or approximations to the true entropy.)

### 2.5.7 $dQ$ , or Not

We do not define entropy via  $dS = dQ/T$  or anything like that, for multiple reasons. For one thing, as discussed in section 8.2, there is no state-function  $Q$  such that  $dQ = TdS$  (with perhaps trivial exceptions). Even more importantly, we need entropy to be well defined even when the temperature is unknown, undefined, irrelevant, or zero.

- In the cup game (section 2.2) and in the card game (section 2.3), the temperature is irrelevant and effectively zero.
- In the three-state system with a population inversion (section 11.4), the temperature is undefined and undefinable.

## 2.6 Quantifying Entropy

The idea of entropy set forth in the preceding examples can be quantified quite precisely. Entropy is defined in terms of probability.<sup>1</sup> For any classical probability distribution  $P$ , we can define its entropy as:

$$S[P] := \sum_i P_i \log(1/P_i) \quad (2.2)$$

where the sum runs over all possible outcomes and  $P_i$  is the probability of the  $i$ th outcome. Here we write  $S[P]$  to make it explicit that  $S$  is a *functional* that depends on  $P$ . For example, if  $P$  is a conditional probability then  $S$  will be a conditional entropy. Beware that people commonly write simply  $S$ , leaving unstated the crucial dependence on  $P$ .

Equation 2.2 is the faithful workhorse formula for calculating the entropy. It ranks slightly below Equation 27.6, which is a more general way of expressing the same idea. It ranks above various less-general formulas that may be useful under more-restrictive conditions (as in section 9.6 for example). See chapter 22 and chapter 27 for more discussion of the relevance and range of validity of this expression.

---

<sup>1</sup>Usually classical probability suffices, as discussed in section 2.6, but if you really want the most general formulation, the quantum statistics version is discussed in chapter 27.

Subject to mild restrictions, equation 2.2 applies to physics as follows: Suppose the system is in a given macrostate, and the macrostate is well described by a distribution  $P$ , where  $P_i$  is the probability that the system is in the  $i$ th microstate. Then we can say  $S$  is the entropy “of the system.”

Expressions of this form date back to Boltzmann (reference 14 and reference 15) and to Gibbs (reference 16). The range of applicability was greatly expanded by Shannon (reference 17). See also equation 27.6.

Beware that uncritical reliance on “the” observed microstate-by-microstate probabilities does not always give a full description of the macrostate, because the  $P_i$  might be correlated with something else (as in section 11.7) or amongst themselves (as in chapter 27). In such cases the unconditional entropy  $S[P]$  will be larger than the conditional entropy  $S[P|Q]$ , and you have to decide which is/are physically relevant.

In the games discussed above, it was convenient to measure entropy in bits, because we were asking yes/no questions. Other units are possible, as discussed in section 9.5.

Figure 2.2 shows the contribution to the entropy from one term in the sum in equation 2.2. Its maximum value is approximately 0.53 bits, attained when  $P(H) = 1/e$ .

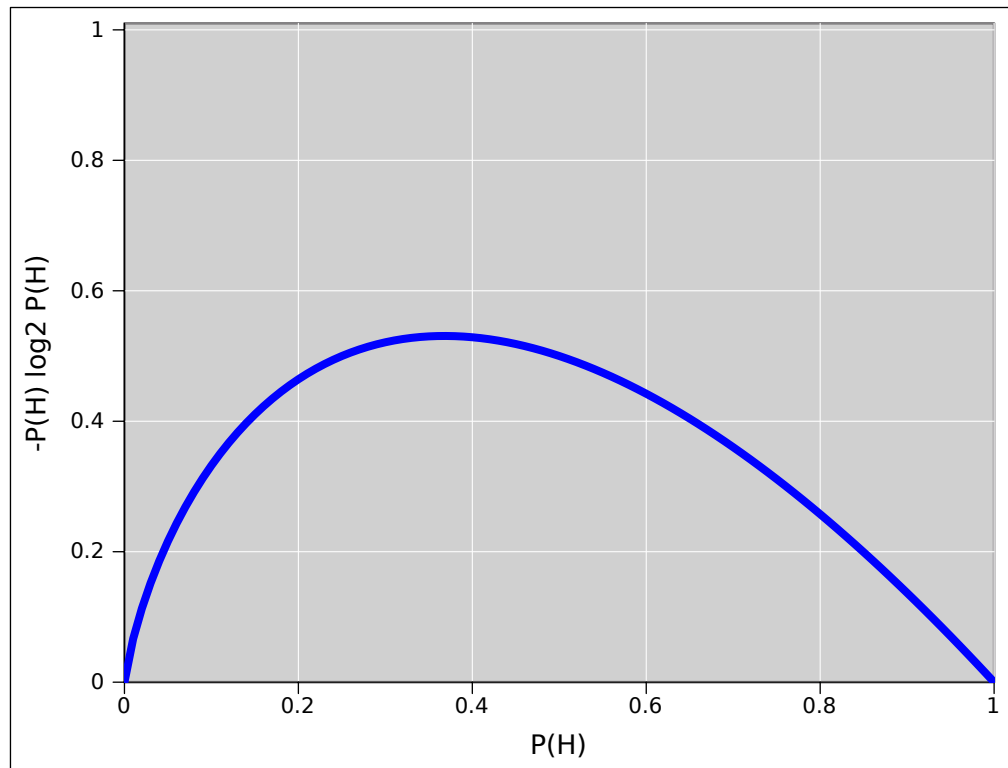


Figure 2.2:  $-P(H) \log_2 P(H)$  – One Term in the Sum

Let us now restrict attention to a system that only has two microstates, such as a coin toss, so there will be exactly two terms in the sum. That means we can identify  $P(H)$  as the probability of the the



“heads” state. The other state, the “tails” state, necessarily has probability  $P(T) \equiv 1 - P(H)$  and that gives us the other term in the sum, as shown by the red curve in figure 2.3. The total entropy is shown by the black curve in figure 2.3. For a two-state system, it is necessarily a symmetric function of  $P(H)$ . Its maximum value is 1 bit, attained when  $P(H) = P(T) = 1/2$ .

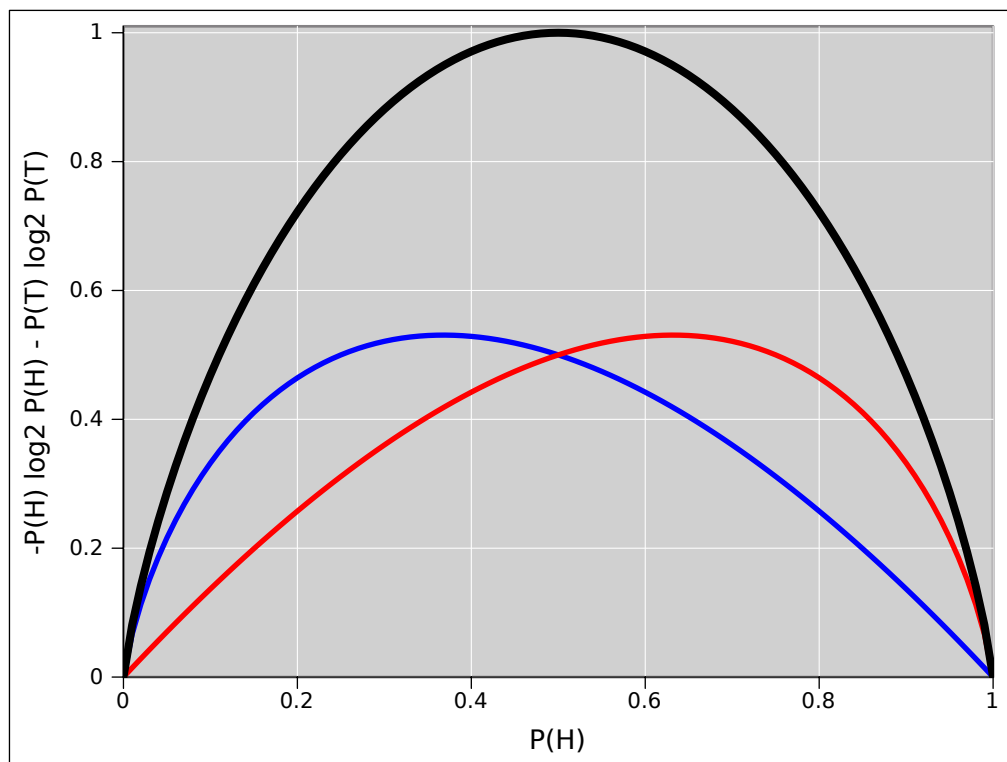


Figure 2.3: Total Entropy – Two-State System

The base of the logarithm in equation 2.2 is chosen according to what units you wish to use for measuring entropy. Alternatively, as discussed in section 9.5, you can fix the base of the logarithm and stick in a prefactor that has some numerical value and some units:

$$S = -k \sum_i P_i \ln P_i \quad (2.3)$$

Entropy itself is conventionally represented by big  $S$ . Meanwhile, molar entropy is conventionally represented by small  $s$  and is the corresponding intensive property.

People commonly think of entropy as being an extensive quantity. This is true to an excellent approximation in typical situations, but there are occasional exceptions. Some exceptions are discussed in section 12.8 and especially section 12.11.

Although it is often convenient to measure molar entropy in units of J/K/mol, other units are allowed, for the same reason that mileage is called mileage even when it is measured in metric units. In particular, sometimes additional insight is gained by measuring molar entropy in units of bits per particle. See section 9.5 for more discussion of units.

When discussing a chemical reaction using a formula such as



it is common to speak of “the entropy of the reaction” but properly it is “the *molar* entropy of the reaction” and should be written  $\Delta s$  or  $\Delta S/N$  (not  $\Delta S$ ). All the other terms in the formula are intensive, so the entropy-related term must be intensive also.

Of particular interest is the *standard* molar entropy,  $s^\circ$  i.e.  $S^\circ/N$ , measured at standard temperature and pressure. The entropy of a gas is strongly dependent on density, as mentioned in section 12.3.

## 2.7 Microstate versus Macrostate

### 2.7.1 Surprisal

If we have a system characterized by a probability distribution  $P$ , the *surprisal* of the  $i$ th microstate is given by

$$s_i := -\log(1/P_i) \quad (2.5)$$

By comparing this with equation 2.2, it is easy to see that the entropy is simply the ensemble average of the surprisal. In particular, it is the *expectation value* of the surprisal. (See equation 27.7 for the fully quantum-mechanical generalization of this idea.)

### 2.7.2 Contrasts and Consequences

Note the following contrast: For any given microstate  $i$  and any given distribution  $P$ :

Surprisal is a property of the microstate $i$ .	Entropy is a property of the distribution $P$ as a whole. It is defined as an ensemble average.
---	---

Entropy is a property of the macrostate  
not of the microstate.

A manifestation of this can be seen in item 4.

When you hear that entropy is a property of «the» distribution, the word «the» should not receive undue emphasis. The complete assertion says that *for any given distribution* the entropy is a property of the distribution, and the second half must not be taken out of context. There are lots of different distributions in this world, and you should not think in terms of «the» one true distribution. Indeed, it is common to deal with two or more distributions at the same time.

Given just a microstate, you do not know what distribution it was drawn from. For example, consider the point at the center of the small black circle in figure 2.4. The point itself has no uncertainty, no width, no error bars, and no entropy. The point could have been drawn from the red distribution or the blue distribution; you have no way of knowing. The entropy of the red distribution is clearly different from the entropy of the blue distribution.

In a card game such as poker or go fish, it is common for different players to use different macrostates (different distributions), even though the microstate (the objective state of the cards) is the same for all, as discussed in section 2.4. Ditto for the game of Clue, or any other game of imperfect information. Similarly, in cryptography the sender and the attacker virtually always use different distributions over plaintexts and keys. The microstate is known to the sender, whereas the attacker presumably has to guess. The microstate is the same for both, but the macrostate is different.

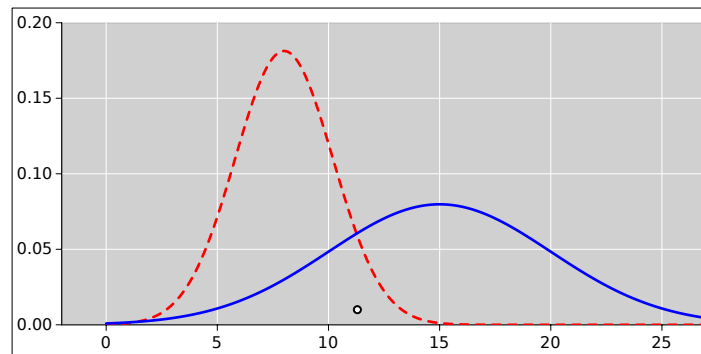


Figure 2.4: One Point, Two Distributions

Another obvious consequence of equation 2.5 is that entropy is not, by itself, the solution to all the world's problems. Entropy measures a particular *average* property of the given distribution. It is easy to find situations where other properties of the distribution are worth knowing. The mean, the standard deviation, the entropy, various Rényi functionals, etc. are just a few of the many properties. (Note that the Rényi functionals (one of which is just the entropy) can be defined even in situations where the mean and standard deviation cannot, e.g. for a distribution over non-numerical symbols.)

For more about the terminology of state, microstate, and macrostate, see section 12.1.

## 2.8 Entropy of Independent Subsystems

Suppose we have subsystem 1 with a set of microstates  $\{(i)\}$  and subsystem 2 with a set of microstates  $\{(j)\}$ . Then in all generality, the microstates of the combined system are given by the direct product of these two sets, namely

$$\{(i)\} \times \{(j)\} = \{(i, j)\} \quad (2.6)$$

where  $(i, j)$  is an ordered pair, which should be a familiar idea and a familiar notation. We use  $\times$  to denote the Cartesian direct product.

We now restrict attention to the less-than-general case where the two subsystems are statistically independent. That means that the probabilities are multiplicative:

$$R(i, j) = P(i)Q(j) \quad (2.7)$$

Let's evaluate the entropy of the combined system:

$$\begin{aligned}
 S[R] &= - \sum_{i,j} R(i,j) \log[R(i,j)] \\
 &= - \sum_{i,j} P(i)Q(j) \log[P(i)Q(j)] \\
 &= - \sum_{i,j} P(i)Q(j) \log[P(i)] - \sum_{i,j} P(i)Q(j) \log[Q(j)] \\
 &= - \sum_j Q(j) \sum_i P(i) \log[P(i)] - \sum_i P(i) \sum_j Q(j) \log[Q(j)] \\
 &= S[P] + S[Q]
 \end{aligned} \tag{2.8}$$

where we have used the fact that the subsystem probabilities are normalized.

So we see that the entropy is additive whenever the probabilities are multiplicative, i.e. whenever the probabilities are independent.

## Chapter 3

# Basic Concepts (Zeroth Law)

There are a bunch of basic notions that are often lumped together and called the zeroth law of thermodynamics. These notions are incomparably less fundamental than the notion of energy (the first law) and entropy (the second law), so despite its name, the zeroth law doesn't deserve priority.

Here are some oft-cited rules, and some comments on each.

We can divide the world into some number of <i>regions</i> that are disjoint from each other.	If there are only two regions, some people like to call one of them “the” system and call the other “the” environment, but usually it is better to consider all regions on an equal footing. Regions are sometimes called <i>systems</i> , or <i>sub-systems</i> , or <i>zones</i> , or <i>parcels</i> . They are sometimes called <i>objects</i> , especially when they are relatively simple.
There is such a thing as <i>thermal equilibrium</i> .	You must not assume that everything is in thermal equilibrium. Thermodynamics and indeed life itself depend on the fact that some regions are out of equilibrium with other regions.
There is such a thing as <i>temperature</i> .	There are innumerable important examples of systems that lack a well-defined temperature, such as the three-state laser discussed in section 11.4.
Whenever any two systems are in equilibrium with each other, and they each have a well-defined temperature, then the two temperatures are the same. See section 10.1 and chapter 23.	This is true and important. (To be precise, we should say they have the same average temperature, since there will be fluctuations, which may be significant for very small systems.)

Equilibrium is *transitive*. That is, if  $A$  is in equilibrium with  $B$  and  $B$  is in equilibrium with  $C$ , then  $A$  is in equilibrium with  $C$ . See chapter 23.

We can establish equilibrium within a system, and equilibrium between selected pairs of systems, without establishing equilibrium between all systems.

If/when we have established equilibrium within a system, a few variables suffice to entirely describe the *thermodynamic state* (i.e. macrostate) of the system.<sup>1</sup> (See section 2.7 and section 12.1 for a discussion of microstate versus macrostate.)

This not always true. To understand how it sometimes goes wrong, we must keep in mind that there are different types of equilibrium. If  $A$  equilibrates with  $B$  by exchange of electrons and  $B$  equilibrates with  $C$  by exchange of ions, interesting things can happen. In particular, we can build a battery. When the battery is sitting there open-circuited, all of the components are essentially in equilibrium ... local pairwise equilibrium ... but the two terminals are not in equilibrium, as you will discover if you wire up a load.

This is an entirely nontrivial statement. Sometimes it takes a good bit of engineering to keep some pairs near equilibrium and other pairs far from equilibrium. See section 11.11.

This is an entirely nontrivial statement, and to make it useful you have to be cagey about what variables you choose; for instance:

- Knowing the temperature and pressure of a parcel of ice gives you more-or-less a complete description of the thermodynamic state of the ice.
- Knowing the temperature and pressure of a parcel of liquid water gives you more-or-less a complete description of the thermodynamic state of the water.
- Meanwhile, in contrast, knowing the temperature and pressure of an ice/water mixture does *not* fully determine the thermodynamic state, because you don't know what fraction is ice and what fraction is water.

In addition to all these thermo-related concepts, we must also comply with all the “non-thermal” laws of physics, including conservation of momentum, conservation of charge, et cetera. This is discussed in chapter 5.

## Chapter 4

# Low-Temperature Entropy (Alleged Third Law)

As mentioned in the introduction, one sometimes hears the assertion that the entropy of a system must go to zero as the temperature goes to zero.

There is no theoretical basis for this assertion, so far as I know – just unsubstantiated opinion.

As for experimental evidence, I know of only one case where (if I work hard enough) I can make this statement true, while there are innumerable cases where it is not true:

- There is such a thing as a *spin glass*. It is a solid, with a spin at every site. At low temperatures, these spins are not lined up; they are highly disordered. And there is a large potential barrier that prevents the spins from flipping. So for all practical purposes, the entropy of these spins is frozen in. The molar entropy involved is substantial, on the order of one J/K/mole. You can calculate the amount of entropy based on measurements of the magnetic properties.
- A chunk of ordinary glass (e.g. window glass) has a considerable amount of frozen-in entropy, due to the disorderly spatial arrangement of the glass molecules. That is, glass is not a perfect crystal. Again, the molar entropy is quite substantial. It can be measured by X-ray scattering and neutron scattering experiments.
- For that matter, it is proverbial that perfect crystals do not occur in nature. This is because it is energetically more favorable for a crystal to grow at a dislocation. Furthermore, the materials from which the crystal was grown will have chemical impurities, not to mention a mixture of isotopes. So any real crystal will have frozen-in nonuniformities. The molar entropy might be rather less than one J/K/mole, but it won't be zero.
- If I wanted to create a sample where the entropy went to zero in the limit of zero temperature, I would proceed as follows: Start with a sample of helium. Cool it to some very low temperature. The superfluid fraction is a single quantum state, so it has zero entropy. But the sample as a whole still has nonzero entropy, because  $^3\text{He}$  is quite soluble in  $^4\text{He}$  (about 6% at zero temperature), and there will always be some  $^3\text{He}$  around. To get rid of that, pump the sample through a superleak, so the  $^3\text{He}$  is left behind. (Call it reverse osmosis if you like.) Repeat this as a function of  $T$ . As  $T$  goes to zero, the superfluid fraction goes to 100% (i.e. the normal-fluid fraction goes to 0%) so the entropy, as far as I know, would go to zero asymptotically.

Note: It is hard to measure the low-temperature entropy by means of elementary thermal measurements, because typically such measurements are insensitive to “spectator entropy” as discussed in section 12.6. So for typical classical thermodynamic purposes, it doesn’t matter whether the entropy goes to zero or not.



## Chapter 5

# The Rest of Physics, Chemistry, etc.

The previous sections have set forth the conventional laws of thermodynamics, cleaned up and modernized as much as possible.

At this point you may be asking, why do these laws call attention to conservation of energy, but not the other great conservation laws (momentum, electrical charge, lepton number, et cetera)? And for that matter, what about all the other physical laws, the ones that aren't expressed as conservation laws? Well, you're right, there are some quite silly inconsistencies here.

The fact of the matter is that in order to do thermo, you need to import a great deal of classical mechanics. You can think of this as the minus-one<sup>th</sup> law of thermodynamics.

- This includes Newton's third law (which is tantamount to conservation of momentum) and Newton's second law, with the associated ideas of force, mass, acceleration, et cetera. Note that the concept of pseudowork, which shows up in some thermodynamic discussions, is more closely related to the momentum and kinetic energy than it is to the total energy.
- In particular, this includes the notion of conservation of energy, which is a well-established part of nonthermal classical mechanics. From this we conclude that the first law of thermodynamics is redundant and should, logically, be left unsaid (although it remains true and important).
- If you are going to apply thermodynamics to a chemical system, you need to import the fundamental notions of chemistry. This includes the notion that atoms exist and are unchanged by ordinary chemical reactions (which merely defines what we mean by a "chemical" as opposed to "nuclear" reaction). This implies about a hundred additional approximate<sup>1</sup> conservation laws, one for each type of atom. The familiar practice of "balancing the reaction equation" is nothing more than an application of these atom-conservation laws.
- If you are going to apply thermodynamics to an electrical or magnetic system, you need to import the laws of electromagnetism, i.e. the Maxwell equations.
- The Maxwell equations imply conservation of charge. This is essential to chemistry, in the context of redox reactions. It means you have to balance the reaction equation with respect to charge. This is in addition to the requirement to balance the reaction equation with respect to atoms.

---

<sup>1</sup>Subject to the approximation that nuclear reactions can be neglected.

Sometimes the process of importing a classical idea into the world of thermodynamics is trivial, and sometimes not. For example:

The law of conservation of momentum would be automatically valid if we applied it by breaking a complex object into its elementary components, applying the law to each component separately, and summing the various contributions. That's fine, but nobody wants to do it that way. In the spirit of thermodynamics, we would prefer a macroscopic law. That is, we would like to be able to measure the overall mass of the object ( $M$ ), measure its average velocity ( $V$ ), and from that compute a macroscopic momentum ( $MV$ ) obeying the law of conservation of momentum. In fact this macroscopic approach works fine, and can fairly easily be proven to be consistent with the microscopic approach. No problem.

The notion of kinetic energy causes trouble when we try to import it. Sometimes you want a microscopic accounting of kinetic energy, and sometimes you want to include only the macroscopic kinetic energy. There is nontrivial ambiguity here, as discussed in section 18.4 and reference 18.

## Chapter 6

# Classical Thermodynamics

### 6.1 Overview

So far we have discussed only a few basic ideas, but already we know enough to handle some interesting applications.

### 6.2 Stirling Engine

#### 6.2.1 Basic Structure and Operations

Let's consider a Stirling engine. This is a type of heat engine; that is, it takes in energy and entropy via one side, dumps out waste energy and entropy via the other side, and performs useful work via a crankshaft.

In operation, the Stirling engine goes through a cycle of four states  $\{A, B, C, D\}$  connected by four legs  $\{AB, BC, CD, DA\}$ . The Stirling cycle is rectangle in  $(T, V)$  space, as shown in figure 6.1.

There are many ways of building a Stirling engine. The ones that are easy to build are not easy to analyze. (Actually all Stirling engines are infamously hard to analyze.) Let's consider an idealized version. One way of carrying out the volume-changes called for in figure 6.1 is shown in figure 6.2. You can see that there are two cylinders: an always-hot cylinder at the left, and an always-cold cylinder at the right. Each cylinder is fitted with a piston. When we speak of "the" volume, we are referring to the combined volume in the two cylinders.

#### Cooling

Let's now examine the required temperature-changes. Figure 6.3 shows a more complete picture of the apparatus. In addition to the two cylinders, there is also a thin tube allowing the working fluid to flow from the hot side to the cold side and back again.

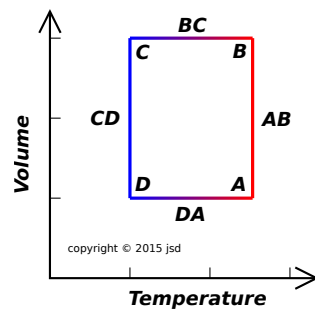


Figure 6.1: Stirling Cycle

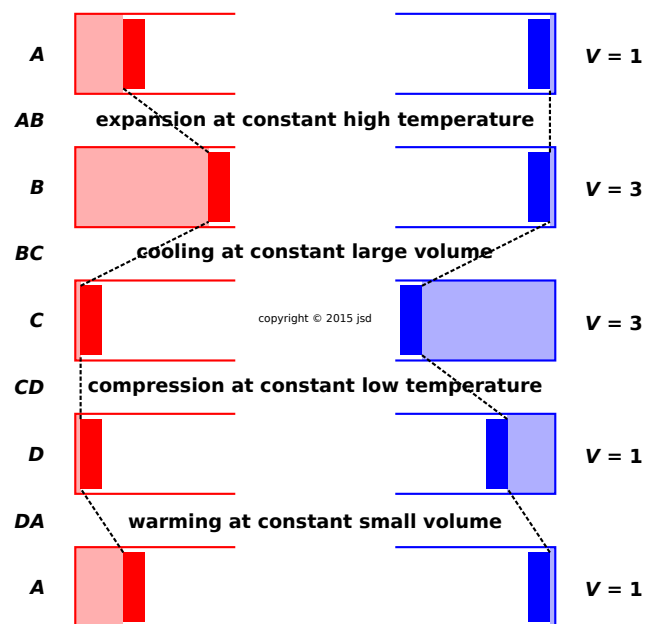


Figure 6.2: Stirling Engine : Volume Changes

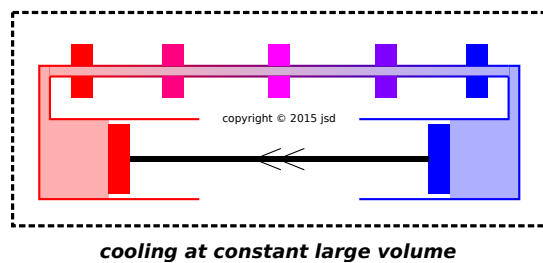


Figure 6.3: Stirling Engine : Cooling Phase

Figure 6.3 shows a snapshot taken during the cooling phase, i.e. the  $BC$  leg. The two pistons move together in lock-step, so the total volume of fluid remains constant. As the fluid flows through the tube, it encounters a series of heat exchangers. Each one is a little cooler than the next. At each point, the fluid is a little bit warmer than the local heat exchanger, so it gives up a little bit of energy and entropy. At each point the temperature difference is small, so the transfer is very nearly reversible. We idealize it as completely reversible. We idealize it as completely reversible.

### Warming

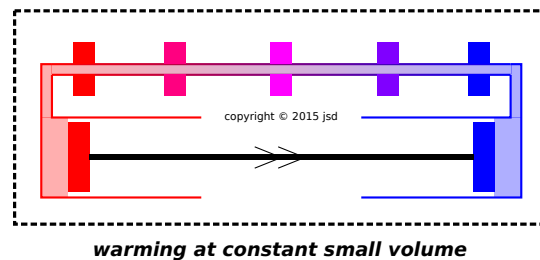


Figure 6.4: Stirling Engine : Warming Phase

Figure 6.4 shows a snapshot taken during the warming phase, i.e. the  $DA$  leg. Once again, the two pistons move together in lock-step, so the total volume of fluid remains constant. The volume here

is less than the volume during the cooling phase, so each piston moves less, but still they force all the volume to flow through the tube. (If you're clever, you can arrange that the pistons move with less velocity during this leg, so that the overall time is the same for both the *DA* leg and the *BC* leg. That is, you can arrange that the flow rate in moles per second is the same, even though the flow rate in cubic meters per second is different. This gives the heat exchangers time to do their job.)

As the fluid flows through the tube, it encounters the same series of heat exchangers, in reverse order. At each point, the fluid is a little bit cooler than the local heat exchanger, so it picks up a little bit of energy and entropy. At each point the temperature difference is small, so the transfer is very nearly reversible. We idealize it as completely reversible.

The same number of moles of fluid underwent the same change of temperature, in the opposite direction, so the energy and entropy involved in the *DA* leg are equal-and-opposite to the energy and entropy involved in the *BC* leg. The molar heat capacity at constant volume is a constant, independent of density, for any ideal gas (or indeed any polytropic gas).

It must be emphasized that during the *BC* leg and also during the *DA* leg, no energy or entropy leaves the system. No energy or entropy crosses the dotted-line boundary shown in the diagram. The heat exchangers arrayed along the thin tube are strictly internal to the system. They do not require fuel or coolant from outside. They are like little cookie jars; you can put cookies in and take cookies out, but they do not produce or consume cookies. At the end of each cycle, they return to their original state.

## Expansion

Figure 6.5 shows the expansion phase. The fluid is all on the hot side of the machine. The piston is retreating, so the fluid expands. If it were allowed to expand at constant entropy, it would cool, but we do not allow that. Instead we supply energy from an external source, to maintain the fluid at a constant temperature as it expands.

## Compression

Figure 6.6 shows the compression phase. The fluid is all on the cold side of the machine. The cold piston is advancing, so the fluid is compressed into a smaller volume. If it were compressed at constant entropy, it would warm up, but we do not allow that. Instead we supply cooling from an external source, to maintain the fluid at a constant temperature during the compression.

### 6.2.2 Energy, Entropy, and Efficiency

It must be emphasized that the source that supplies energy and entropy on the left side of figure 6.5 is in a completely different category from the cookie jars attached to the flow tube. The source provides energy and entropy that flow across the boundary of the system. We imagine that the external high-side heat bath has an infinite heat capacity, so that we can extract energy and entropy from it without changing its temperature.



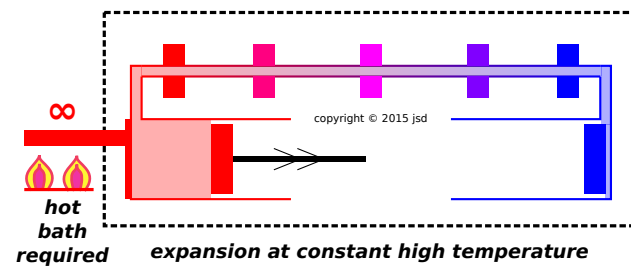


Figure 6.5: Stirling Engine : Expansion Phase

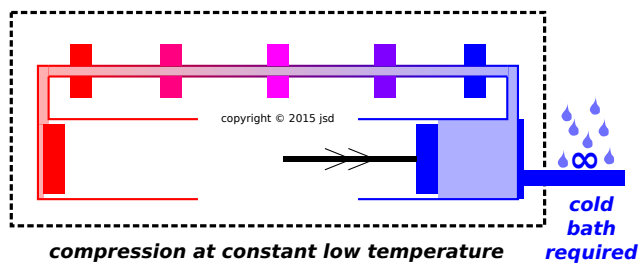


Figure 6.6: Stirling Engine : Compression Phase

Similar words apply to the cooling system on the right side of figure 6.6. This is in a completely different category from the cookie jars, because energy and entropy are crossing the boundary of the system. We imagine that the low-side heat bath has an infinite heat capacity, so that we can dump energy and entropy into it without changing its temperature.

To summarize: Legs  $AB$  and  $CD$  are in one category (energy and entropy crossing the boundary, cumulatively, without end) while  $BC$  and  $DA$  are in another category (energy and entropy being rearranged interneally, with no net change from cycle to cycle).

The diagrams in this section do not show the crankshaft that delivers useful work to the outside. However we can easily enough figure out how much work is involved.

No work is performed on the  $BC$  leg or on the  $DA$  leg. The volume is the same before and after, so there is no  $PdV$  work. We are assuming an ideal engine, so the pressure on both sides is the same. The pressure changes during the  $BC$  leg (and during the  $DA$  leg), but at each instant the pressure at both ends of the machine is the same.

During the  $AB$  leg the volume increases by a certain amount, and then on the  $CD$  leg the volume decreases by the same amount. We know the fluid is hotter on the  $AB$  leg, so it has a higher pressure, so there is a net amount of  $PdV$  work done. For an ideal gas, it is easy to do the integral  $PdV$ . You don't even need to assume a monatomic gas; a polyatomic ideal gas such as air is just as easy to handle. A factor of  $T$  comes out in front of the integral.

So, on a per-cycle basis, the energy that comes in on the left is proportional to  $T_{\text{hot}}$ , while the energy that goes out on the right is proportional to  $T_{\text{cold}}$ .

We define the efficiency of a heat engine as

$$\eta_{\text{he}} := \frac{\text{mechanical energy out}}{\text{energy in}} \quad (\text{for any heat engine}) \quad (6.1)$$

so for this particular heat engine, given all our idealizations, the efficiency is

$$\eta_{\text{he}} := \frac{T_{\text{hot}} - T_{\text{cold}}}{T_{\text{hot}}} \quad (\text{for an ideal Stirling engine}) \quad (6.2)$$

### 6.2.3 Practical Considerations; Temperature Match or Mismatch

Note that each hardware component in a Stirling engine sits at more-or-less steady temperature over time. The working fluid changes temperature significantly, but the hardware components do not. Increasing the thermal contact between the hardware and the fluid (at any given point) makes the engine work better. (On the other hand, we need to minimize thermal conduction between one cookie jar and the next, in the regenerator.)

A gas turbine engine (such as you find on a jet aircraft) is similar, insofar as the fluid changes temperature, while the hardware components do not, to any great extent. Also, the working fluid never comes in contact with hardware that is at a significantly different temperature. It is however in a different subcategory, in that improving the contact between the working fluid and the hardware will not make the engine work better.

This stands in contrast to a Carnot cycle, which depends on heating up and cooling down “the” cylinder that holds the working fluid. This is of course possible in theory, but it is troublesome in

practice. It tends to reduce the power density and/or the efficiency of the engine. James Watt's first steam engine suffered from this problem to a serious degree. It was only later that he figured out the advantages of having a separate condenser.

In a third category we have piston engines of the kind that you find in an ordinary automobile or lawn mower. They have the ugly property that the working fluid comes into contact with hardware at wildly dissimilar temperatures. This limits the efficiency, especially at low speeds. Making stronger contact between the working fluid and the cylinder walls would make the engine work worse. A similar temperature-mismatch problem occurs in the pistons of an external-combustion piston engines, such as piston-type steam engine, because the steam cools as it expands. (Of course a steam engine requires good thermal contact in the boiler and the condenser.)

### 6.2.4 Discussion: Reversibility

As discussed in section 7.1, under mild assumptions we can write

$$dE = -PdV + TdS \quad (6.3)$$

We apply this to the heat baths that attach to the left and right side of our engine. We connect to the heat baths in such a way that we do not affect their pressure or volume, so we have simply:

$$dE = TdS \quad (\text{at constant } V) \quad (6.4)$$

Since for our engine,  $\Delta E$  is proportional to temperature, we see that the amount of entropy we take in from the left is exactly the same as the amount of entropy we dump out on the right. (As always, it is assumed that no entropy flows out via the crankshaft.)

This property – entropy in equals entropy out – is the hallmark of a *reversible* engine. An ideal Stirling engine is reversible.

A heat engine operated in reverse serves as a *heat pump*. A household refrigerator is an example: It pumps energy and entropy from inside the refrigerator to outside. Other heat pumps are used to provide heat for buildings in the winter.

## 6.3 All Reversible Heat Engines are Equally Efficient

We now consider the class of all devices that

1. are heat engines,
2. are reversible, and
3. run off a single hot bath ( $T_{\text{hot}}$ ) and a single cold bath ( $T_{\text{cold}}$ ).

The claim is that all such engines are equally efficient.

The proof is stunningly simple and powerful: If you had two reversible heat engines with different efficiencies, you could hook them in parallel and create a perpetual motion machine!

As a corollary: We know the efficiency of all such devices (for any given pair of temperatures).

The proof is equally simple: Given that all devices that meet our three criteria are equally efficient, if you know the efficiency of one, you know them all. And we do know the efficiency of one, so we in fact we know them all. This is called the *Carnot efficiency* of a heat engine. It is:

$$\eta_{\text{he}} := \frac{T_{\text{hot}} - T_{\text{cold}}}{T_{\text{hot}}} \quad (\text{for any 2-bath reversible heat engine}) \quad (6.5)$$

As long as you are getting your energy from a heat bath, it is not possible to design an engine with greater efficiency than this, no matter how clever you are.

This is the glory of classical thermodynamics. Sadi Carnot figured this out long before there was a modern understanding of what energy is. The result depends on little more than the definition of temperature. In fact, the result can be turned around to serve as an operational definition of temperature: If you have an object at some known reference temperature, and another object at some unknown temperature, you can operate a reversible heat engine between the two, observe the efficiency, and deduce a value for the previously-unknown temperature.

Here's another way of understanding the universal efficiency result. Imagine we have a heat engine that is reversible, but less efficient than we would expect based on equation 6.5. As a heat engine, it is less efficient, but as a heat pump it is more efficient! It is a too-good-to-be-true heat pump. As a heat pump, it takes in too much entropy on the cold side and ejects too little on the hot side. This violates the law of paraconservation of entropy (section 2.1).

## 6.4 Not Everything is a Heat Engine

It must be emphasized that the discussion in section 6.3 applies to *heat engines* and not otherwise. In particular, consider an electrochemical fuel cell. If it's done right, the efficiency of such a thing vastly exceeds  $1 - T_{\text{cold}}/T_{\text{hot}}$ . That's OK, because it's not a heat engine. Even though it takes in fuel and puts out work, it is not a heat engine.

As an even more familiar example, consider a plain old electrochemical battery driving an electric motor. If it's done right, the efficiency of such a thing vastly exceeds  $1 - T_{\text{cold}}/T_{\text{hot}}$ .

It is not worth the trouble to try to understand such things in terms of heat, or in terms of classical thermodynamics generally. I doubt it is even possible.

It is however straightforward to understand such things in terms of modern (post-1898) thermodynamics, i.e. statistical mechanics. Rather than getting bogged down trying to define “heat,” formalize everything in terms of energy and entropy instead. Entropy is defined in terms of probability, as a sum over states. In a battery, the state that stores the energy doesn't even have a temperature, and it isn't in equilibrium with your thermometer.

Let's be clear: For a heat bath, its entropy is proportional to its energy, and the constant of proportionality is the inverse temperature. For a battery, that's just not true. You can stick a lot of energy into a battery without changing its temperature.

## 6.5 Carnot Efficiency Formula

We can derive a famous formula that places limits on the efficiency of any so-called heat-engine.

### 6.5.1 Definition of Heat Engine

Figure 6.7 is a sketch of such an engine. The details don't matter. The key concept is that the heat engine, by definition, has three connections, highlighted by magental labels in the figure.

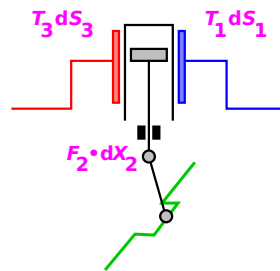


Figure 6.7: Heat Engine

Even though this is called a heat engine, trying to quantify the “heat” is a losing strategy. It is simpler and in every way better to quantify the energy and the entropy.

We could write the change in the energy content of the engine using equation 6.6, including all the terms on the RHS. However, as part of our definition of what we mean by *heat engine*, we require that only the terms shown in magenta are significant.

$$\begin{aligned} dE &= T_3 dS_3 - F_3 \cdot dX_3 + \cdots \quad (\text{connection \#3}) \\ &+ T_2 dS_2 - F_2 \cdot dX_2 + \cdots \quad (\text{connection \#2}) \\ &+ T_1 dS_1 - F_1 \cdot dX_1 + \cdots \quad (\text{connection \#1}) \end{aligned} \quad (6.6)$$

More specifically, we require that connection #3 be 100% thermal, connection #2 be 100% mechanical i.e. nonthermal, and connection #1 be 100% thermal. The engine is designed to *segregate* the thermal energy-transfers from the mechanical energy-transfers.

We have just constructed a theoretical model of an engine. This is a remarkably good model for a wide range of real-world engines. However, it must be emphasized that it does not apply to all engines. In particular, it does not apply to batteries or to electrochemical fuel cells.

We must impose one more restriction: We require that the engine be able to operate in a cycle, such that at the end of the cycle, after doing something useful, all the internal parts of the engine return to their initial state. In particular, we require that the engine not have a “hollow leg” where it can hide unlimited amounts of entropy for unlimited amounts of time. This requirement makes sense, because if entropy could be hidden, it would defeat the spirit of the second law of thermodynamics.

Without loss of generality we assume that  $T_2 \geq T_1$ . There is no loss of generality, because the engine is symmetrical. If necessary, just relabel the connections to *make*  $T_2 \geq T_1$ . Relabeling is just a paperwork exercise, and doesn’t change the physics.

### 6.5.2 Analysis

The engine starts by taking in a certain amount of energy via connection #3. We might hope to convert all of this heat-energy to useful work and ship it out via connection #2, but we can’t do that. The problem is that because connection #3 is a thermal connection, when we took in the energy, we also took in a bunch of entropy, and we need to get rid of that somehow. We can’t get rid of it through connection #2, so the only option is to get rid of it through connection #1.

For simplicity, we assume that  $T_3$  and  $T_1$  are constant throughout the cycle. We also make the rather mild assumption that  $T_1$  is greater than zero. Section 6.5.3 discusses ways to loosen these requirements.

Under these conditions, pushing entropy out through connection #1 costs energy. We call this the *waste heat*:

$$\text{waste heat} = - \int_{\Gamma} T_1 dS_1 \quad (6.7)$$

where the path  $\Gamma$  represents one cycle of operating the engine.

We can now use conservation of energy to calculate the mechanical work done *by* the engine:

$$\text{work out} = \int_{\Gamma} F_2 \cdot dX_2 \quad (6.8a)$$

$$= \int_{\Gamma} T_3 dS_3 + \int_{\Gamma} T_1 dS_1 \quad (6.8b)$$

$$= T_3 \Delta S_3 + T_1 \Delta S_1 \quad (6.8c)$$

where on the last line we have used that fact that the temperatures are unchanging to allow us to do the entropy-integrals. The deltas implicitly depend on the path  $\Gamma$ .

The second law tells us that over the course of a cycle, the entropy going out via connection #1 ( $-\Delta S_1$ ) must be *at least* as much as the entropy coming in via connection #3 ( $+\Delta S_3$ ). For a *reversible* engine, the two are equal, and we can write:

$$\begin{aligned}\Delta S_3 &= +\Delta S_{\blacktriangleright} \\ \Delta S_1 &= -\Delta S_{\blacktriangleright}\end{aligned}\tag{6.9}$$

Where  $\Delta S_{\blacktriangleright}$  is pronounced “delta  $S$  through” and denotes the amount of entropy that flows through the engine, in the course of one cycle. We define the efficiency as

$$\eta := \frac{\text{mechanical transfer out}}{\text{thermal transfer in}}\tag{6.10}$$

Still assuming the temperatures are greater than zero, the denominator in this expression is just  $T_3\Delta S_3$ . Combining results, we obtain:

$$\eta := \frac{T_3\Delta S_3 + T_1\Delta S_1}{T_3\Delta S_3}\tag{6.11}$$

The maximum efficiency is obtained for a thermodynamically reversible engine, in which case

$$\eta_{\text{rev}} = \frac{T_3\Delta S_{\blacktriangleright} - T_1\Delta S_{\blacktriangleright}}{T_3\Delta S_{\blacktriangleright}}\tag{6.12a}$$

$$= \frac{T_3 - T_1}{T_3}\tag{6.12b}$$

where equation 6.12b is the famous Carnot efficiency formula, applicable to a reversible heat engine. The meaning of the formula is perhaps easier to understand by looking at equation 6.12a or equation 6.11, wherein the second term in the numerator is just the waste heat, in accordance with equation 6.7.

Let us now consider an irreversible engine. Dissipation within the engine will create extra entropy, which will have to be pushed out via connection #1. This increases the waste heat, and decreases the efficiency  $\eta$ . In other words, equation 6.12b is the exact efficiency for a reversible heat engine, and an upper bound on the efficiency for any heat engine.

### 6.5.3 Discussion

- (1) If you were wondering whether it is possible to construct even one thermodynamic cycle that complies with all the restrictions in section 6.5.2, fear not. The job can be done using a *Carnot cycle*, as described in section 8.7. On the other hand, as mentioned in that section, the Carnot cycle often gets more attention than it deserves.
- (2) Furthermore, the Carnot efficiency formula often gets more attention than it deserves.
  - For one thing, not all engines are heat engines. The Carnot efficiency formula equation 6.12b applies only to heat engines, as defined by the magenta terms in equation 6.6. As a specific, important counterexample, a battery or an electrochemical fuel cell can have an efficiency enormously greater than what you would guess based on equation 6.12b ... even in situations where the chemicals used to run the fuel cell could have been used to run a heat engine instead.
  - Consider the following scenario: Some guy buys a bunch of coal. He uses it to boil some steam at temperature  $T_3$ . He uses that to drive a steam engine. He uses river water to cool the condenser at temperature  $T_1$ . He manages to operate the engine in such a way



that its efficiency is close to the Carnot efficiency. He is very pleased with himself, and runs spots on television advertising how efficient he is. “I’m as efficient as the laws of physics allow, other things being equal.”

The problem is, he’s asking the wrong question. Rather than asking how well he is doing relative to the Carnot efficiency, he should be asking how well he is doing *relative to the best that could be done using the same amount of fuel*.

Specifically, in this scenario, the guy next door is operating a similar engine, but at a higher temperature  $T_3$ . This allows him to get more power out of his engine. His Carnot efficiency is 30% higher. His *actual* efficiency is only 20% higher, because there are some unfortunate parasitic losses. So this guy is not running as close to the Carnot limit as the previous guy. Still, this guy is doing better *in every way that actually matters*.

[This assumes that at least part of the goal is to minimize the amount of coal consumed (for a given amount of useful work), which makes sense given that coal is a non-renewable resource. Similarly part of the goal is to minimize the amount of  $\text{CO}_2$  dumped into the atmosphere. The competing engines are assumed to have comparable capital cost.]

Suppose you are trying to improve your engine. You are looking for inefficiencies. Let’s be clear: Carnot efficiency tells you one place to look ... but it is absolutely not the only place to look.

- (3) Looking at the structure of the result in equation 6.12b, one can understand why we were tempted to require that  $T_1$  and  $T_3$  have definite, constant values.

However, any heat engine still has a well-defined thermal efficiency, as defined by equation 6.10, even if the temperatures are changing in peculiar ways over the course of the cycle.

Furthermore, with some extra work, you can convince yourself that the integrals in equation 6.8b take the form of a  $\Delta S$  times an *average* temperature. It’s a peculiar type of weighted average. This allows us to interpret the Carnot efficiency formula (equation 6.12b) in a more general, more flexible way.

- (4) As a slightly different line of reasoning that leads to the same conclusion, start with an ordinary Carnot cycle, such as we see in figure 8.5 or figure 8.6. You can imagine covering the entire space with a mosaic of tiny Carnot cycles, as shown in figure 6.8. There is one big cycle made up of nine tiny cycles, such that the left edge of one cycle coincides with the right edge of another, and the top of one coincides with the bottom of another. For an ideal reversible engine, going around each of the nine tiny cycles once is identical to going around the big cycle once, because all of the interior edges cancel.

By selecting a suitable set of the tiny Carnot cycles, you can approximate a wide class of reversible cycles. The efficiency can be calculated from the definition, equation 6.10, by summing the thermal and mechanical energy, summing over all the sub-cycles.

- (5) A hot gas has more entropy than a cool gas. This is partly because the probability distribution is more “spread out” in phase space, more spread out along the momentum direction. It is also partly because the *unit of measure* in phase space is smaller, as discussed in section 12.3.

Because of the higher entropy, you might think the gas has less “available” energy (whatever that means). However, the hot gas contains more energy as well as more entropy, and it is

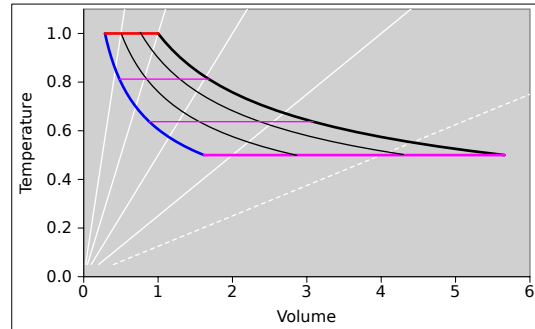


Figure 6.8: Carnot Cycle with Sub-Cycles

easy to find situations where the hot gas is more useful as an energy-source, for instance if the gas is applied to connection #3 in a heat engine. Indeed, in order to increase efficiency in accordance with equation 6.12, engine designers are always looking for ways to make the hot section of the engine run hotter, as hot as possible consistent with reliability and longevity.

Of course, when we examine the cold side of the heat engine, i.e. connection #1, then the reverse is true: The colder the gas, the more valuable it is for producing useful work. This should make it clear that any notion of “available energy” cannot possibly be a function of state. You cannot look at a bottle of gas and determine how much of its energy is “available” – not without knowing a whole lot of additional information about the way in which the gas is to be used. This is discussed in more detail in section 1.7.3.

- (6) If you take equation 6.12b and mindlessly extrapolate it to the case where  $T_1$  is negative, you might hope to achieve an efficiency greater than 100%. However, before doing that note that the defining property of negative temperature is that when entropy goes out via connection #1, energy comes in. Therefore we must include this energy input in the denominator in the definition of efficiency, equation 6.10. When we do that, instead of getting equation 6.11 we get:

$$\begin{aligned}\eta &:= \frac{T_3 \Delta S_3 + T_1 \Delta S_1}{T_3 \Delta S_3 + T_1 \Delta S_1} \\ &= 100\% \quad (\text{always})\end{aligned} \tag{6.13}$$

Both terms in the denominator here are positive; the second term involves a double negative.

We can understand this as follows: Whenever there are two thermal connections, one at a positive temperature and one at a negative temperature, the engine takes in energy via both thermal connections, and sends it all out via the mechanical connection. Therefore the efficiency is always exactly 100%, never more, never less.

Even if the engine is irreversible, such that  $-\Delta S_1$  is greater than  $-\Delta S_3$ , the engine is still 100% efficient in its use of energy, because absolutely all of the energy that is thermally transferred in gets mechanically transferred out.

These conclusions are restricted to a model that assumes the engine has only three connections to the rest of the world. A more realistic engine model would allow for additional connections; see next item.

- (7) A term involving  $T_4\Delta S_4$  could be added to our model, to represent losses due to friction, losses due to heat leaks, et cetera.

$$\begin{aligned}\eta &:= \frac{T_3\Delta S_3 + T_1\Delta S_1 + T_4\Delta S_4}{T_3\Delta S_3 + T_1\Delta S_1} \\ &< 100\% \quad (\text{since } T_4\Delta S_4 < 0)\end{aligned}\tag{6.14}$$

In the real world, the only known ways of producing a heat bath at negative temperature are horrendously inefficient, so the 100% efficiency mentioned in the previous items is nowhere near being relevant to a complete, practical system.



## Chapter 7

# Functions of State

### 7.1 Functions of State : Basic Notions

Terminology: By definition, the term *state function* applies to any measurable quantity that is uniquely determined by the thermodynamic state, i.e. the macrostate.

Terminology: The term *thermodynamic potential* is synonymous with state function. Also the term *function of state* is synonymous with state function.

Example: In an ordinary chunk of metal at equilibrium, state functions include energy ( $E$ ), entropy ( $S$ ), temperature ( $T$ ), molar volume ( $V/N$ ), total mass, speed of sound, et cetera. Some additional important thermodynamic potentials are discussed in chapter 15.

In thermodynamics, we usually arrange for the energy  $E$  to be a function of state. This doesn't tell us anything about  $E$ , but it tells us something about our notion of thermodynamic state. That is, we choose our notion of "state" so that  $E$  will be a function of state.

Similarly, we usually arrange for the entropy  $S$  to be a function of state.

When identifying functions of state, we make no distinction between dependent variables and independent variables. For example, suppose you decide to classify  $V$  and  $T$  as independent variables. That doesn't disqualify them from being functions of state. Calculating  $V$  as a function of  $V$  and  $T$  is not a hard problem. Similarly, calculating  $T$  as a function of  $V$  and  $T$  is not a hard problem. I wish all my problems were so easy.

Counterexample: The microstate is not a function of state (except in rare extreme cases). Knowing the macrostate is not sufficient to tell you the microstate (except in rare extreme cases).

Counterexample: Suppose we have a system containing a constant amount  $\text{H}_2\text{O}$ . Under "most" conditions, specifying the pressure and temperature suffices to specify the thermodynamic state. However, things get ugly if the temperature is equal to the freezing temperature (at the given pressure). Then you don't know how much of the sample is liquid and how much is solid. In such a situation, pressure and temperature do not suffice to specify the thermodynamic state. (In contrast, specifying the pressure and entropy would suffice.)

Note that to be a state function, it has to be a function of the macrostate. This is an idomatic usage of the word “state”.

Something that is a function of the macrostate might or might not make sense as a function of the microstate. Here are some contrasting examples:

<p>The energy is a function of the macrostate, and also a well-behaved function of the microstate. The same can be said for some other quantities including mass, volume, charge, chemical composition, et cetera.</p>	<p>The entropy is a function of the macrostate, but not a function of the microstate. It is defined as an ensemble average. If the macrostate consists of a single microstate, its entropy is zero.</p>
--	---

A crucial prerequisite to idea of “state function” is the idea of “thermodynamic state” i.e. macrostate. Thermodynamics is predicated on the idea that the macrostate can be described by a few variables, such as  $P, V, T$  et cetera. This stands in contrast to describing the microstate, which would require something like  $10^{23}$  variables.

In fluid dynamics, we might divide the fluid into a thousand parcels, each with its own state functions  $P_i, V_i, T_i$  et cetera. That means we have thousands of variables, but that’s still a small number compared to  $10^{23}$ .

Functions of state can be well defined even if the system as a whole is not in equilibrium. For example, the earth’s troposphere is nowhere near thermal equilibrium. (In equilibrium, it would be isothermal, for reasons discussed in section 14.4.) In such a situation, we divide the system into parcels. As long as the parcel unto itself has a well-defined temperature, we can consider temperature to be a function of state, i.e. a function of the state *of that parcel*.

## 7.2 Path Independence

When we say that something is a function of state, we are saying that it does not depend on history; it does not depend on how we got into the given state.

We can apply this idea to *changes* in any function of state. For example, since  $E$  is a function of state, we can write

$$\begin{aligned}\Delta E &= E_{\text{final}} - E_{\text{initial}} \\ &= \text{independent of path}\end{aligned}\tag{7.1}$$

When we say that  $\Delta E$  is independent of path, that mean that  $\Delta E$  is the same, no matter how many steps it takes to get from the initial state to the final state. The path can be simple and direct, or it can involve all sorts of loops and cycles.

As a corollary, if we get from state  $A$  to state  $D$  by two different paths, as shown in figure 7.1, if we add up the changes along each step of each paths, we find that the sum of the changes is independent of paths. That is,

$$\Delta_{AD}(X) = \Delta_{AB}(X) + \Delta_{BC}(X) + \Delta_{CD}(X)\tag{7.2}$$

As usual  $\Delta(X)$  refers to the change in  $X$ . Here  $X$  can any thermodynamic potential.

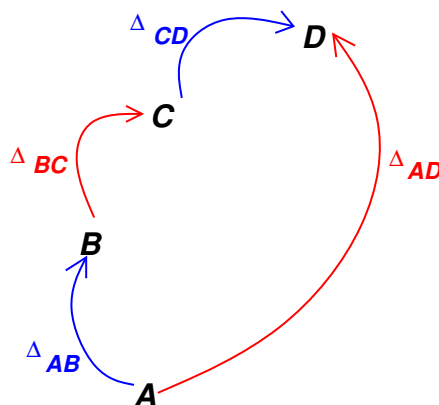


Figure 7.1: Sum of Changes Along Different Paths

The term *sigma-delta* is sometimes used to refer to a sum of changes. Equation 7.2 states that the sigma-delta is independent of path.

It must be emphasized that the principle of the path-independent sigma-delta has got nothing to do with any conservation law. It applies to non-conserved state-functions such as temperature and molar volume just as well as it applies to conserved state-functions such as energy. For example, if the volume  $V$  is a function of state, then:

$$\begin{aligned}\Delta V &= V_{\text{final}} - V_{\text{initial}} \\ &= \text{independent of path}\end{aligned}\tag{7.3}$$

which is true even though  $V$  is obviously not a conserved quantity.

Equation 7.3 looks trivial and usually is trivial. That's because usually you can easily determine the volume of a system, so it's obvious that  $\Delta V$  is independent of path.

The derivation of equation 7.1 is just as trivial as the derivation of equation 7.3, but the applications of equation 7.1 are not entirely trivial. That's because you can't always determine the energy of a system just by looking at it. It may be useful to calculate  $\Delta E$  along one simple path, and then argue that it must be the same along any other path connecting the given initial and final states.

Remark: It is a fairly common mistake for people to say that  $\Delta E$  is a function of state. It's not a function of state; it's a function of *two* states, namely the initial state and the final state, as you can see from the definition:  $\Delta E = E_{\text{final}} - E_{\text{initial}}$ . For more on this, see reference 4. As explained there,

- $\Delta E$  is a scalar but not a function of state.
- $dE$  is a function of state but not a scalar.

### 7.3 Hess's Law, Or Not

Circa 1840, Germain Henri Hess empirically discovered a sum rule for the so-called *heat of reaction*. This is called Hess's Law. Beware that it is not always true, because the heat of reaction is not a

function of state.

A simple counterexample is presented in figure 7.2.

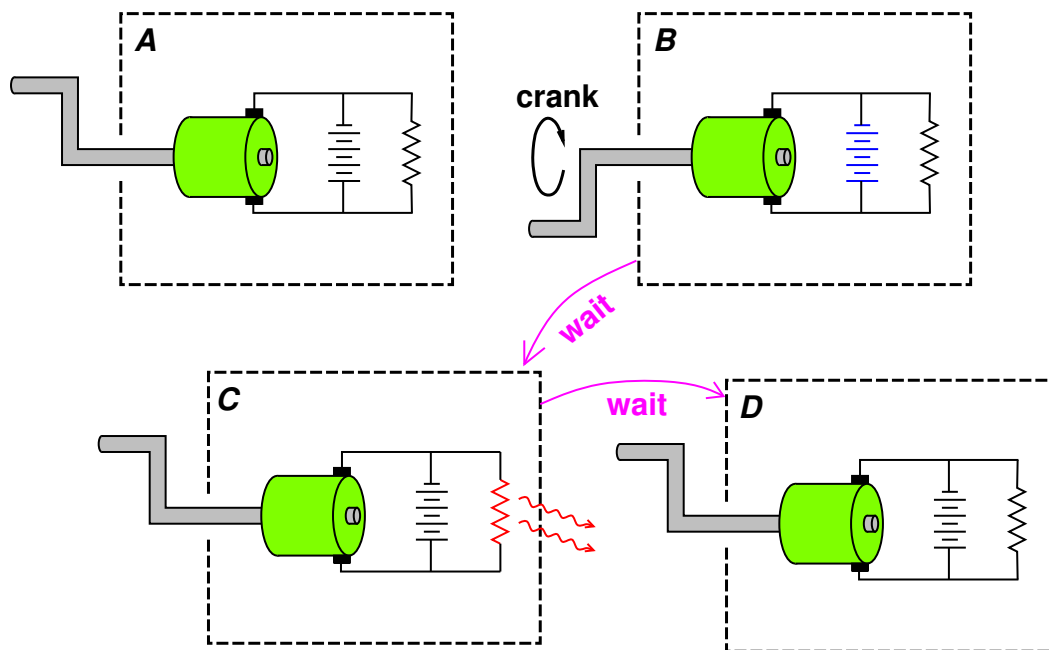


Figure 7.2: Disproof of Hess's Law for Heat

We start in the upper left of the figure. We turn the crank on the generator, which charges the battery. That is, electrochemical reactions take place in the battery. We observe that very little heat is involved in this process. The charged-up battery is shown in blue.

If we stop cranking and wait a while, we notice that this battery has a terrible shelf life. Chemical reactions take place inside the battery that discharge it. This is represented conceptually by a “leakage resistor” internal to the battery. This is represented schematically by an explicit resistor in figure 7.2. In any event, we observe that the battery soon becomes discharged, and becomes warmer. If we wait a little longer, heat flows across the boundary of the system (as shown by the wavy red arrows). Eventually we reach the state shown in the lower right of the diagram, which is identical to the initial state.

There is of course a simpler path for reaching this final state, namely starting at the same initial state and doing nothing ... no cranking, and not even any waiting. This clearly violates Hess's law because the heat of reaction of the discharge process is the dominant contribution along one path, and nothing similar is observed along the other path.

Hess's law in its original form is invalid because heat content is not a state function, and heat of reaction is not the delta of any state function.

Tangential remark: in cramped thermodynamics, a cramped version of Hess's Law is usually valid, because “heat content” is usually a function of state in cramped thermody-



namics. This is a trap for the unwary. This is just one of the many things that are true in cramped thermodynamics but cannot be extended to uncramped thermodynamics.

We can extricate ourselves from this mess by talking about *enthalpy* instead of heat. There is a valid sum rule for the enthalpy of reaction, because enthalpy is a function of state. That is:

$$\begin{aligned}\Delta H &= H_{\text{final}} - H_{\text{initial}} \\ &= \text{independent of path}\end{aligned}\tag{7.4}$$

We emphasize that this does not express conservation of enthalpy. In fact, enthalpy is not always conserved, but equation 7.4 remains true whenever enthalpy is a function of state.

Equation 7.4 could be considered a modernized, “repaired” version of Hess’s law. It is not very important. It does not tell us anything about the enthalpy except that it is a function of state. It is a mistake to focus on applying the sigma-delta idea to enthalpy to the exclusion of the innumerable other state-functions to which the sigma-delta idea applies equally well.

I see no value in learning or teaching any version of Hess’s Law. It is better to simply remember that there is a sigma-delta law for any function of state.

The sigma-delta of any function of state  
is independent of path.

## 7.4 Partial Derivatives

Let’s build up a scenario, based on some universal facts plus some scenario-specific assumptions.

We know that the energy of the system is well defined. Similarly we know the entropy of the system is well defined. These aren’t assumptions. Every system has energy and entropy.

Next, as mentioned in section 7.1, we assume that the system has a well-defined thermodynamic state, i.e. macrostate. This macrostate can be represented as a point in some abstract state-space. At each point in macrostate-space, the macroscopic quantities we are interested in (energy, entropy, pressure, volume, temperature, etc.) take on well-defined values.

We further assume that this macrostate-space has dimensionality  $M$ , and that  $M$  is not very large. (This  $M$  may be larger or smaller than the dimensionality  $D$  of the position-space we live in, namely  $D = 3$ .)

Assuming a well-behaved thermodynamic state is a highly nontrivial assumption.

- As an example where these assumptions are valid, consider the hackneyed example of the ideal gas in equilibrium in a table-top cylinder, where the macrostate is determined by a few variables such as volume, temperature, and number of particles.
- As a more interesting example, consider a heat-flow experiment. We have a metal bar that is kept hot at one end and cold at the other end. This is obviously a non-equilibrium situation, and the heat-flow is obviously irreversible. Yet at each point in the bar, there is a well-defined local temperature, a well-defined local energy density, et cetera. As far as I know, all the assumptions we have made so far hold just fine.

- As a challenging but not hopeless intermediate case, consider a thermal distribution with a few exceptions, as discussed in section 11.3. In this case, our macrostate space must include additional variables to quantify the excitation of the exceptional modes. These variables will show up as additional dimensions in the vector  $V$  or as additional explicit terms in a generalization of equation 7.8.
- As a more problematic example, consider turbulent flow. The motion is chaotic, and the closer you look the more chaos you see. In general, this topic is beyond the scope of this discussion. However, depending on what questions you are asking, it may be possible to average over space and/or average over time so as to establish a well-behaved notion of local temperature in the fluid.
- As an even more problematic example, suppose you have just set off a firecracker inside a cylinder of gas. This is even farther beyond the scope of this discussion. The system will be chaotic and far from equilibrium. It is also nonuniform in space and time, so averaging is problematic (although perhaps not impossible). A great number of modes will be excited. Describing the macrostate of the system will require a tremendous number of variables, so much so that describing the macrostate might be almost as laborious as describing the microstate.

We further assume that the quantities of interest vary *smoothly* from place to place in macrostate-space.

We must be careful how we formalize this “smoothness” idea. By way of analogy, consider a point moving along a great-circle path on a sphere. This path is nice and smooth, by which we mean differentiable. We can get into trouble if we try to describe this path in terms of latitude and longitude, because the coordinate system is singular at the poles. This is a problem with the coordinate system, not with the path itself. To repeat: a great-circle route that passes over the pole is differentiable, but its representation in spherical polar coordinates is not differentiable.

Applying this idea to thermodynamics, consider an ice/water mixture at constant pressure. The temperature is a smooth function of the energy content, whereas the energy-content is not a smooth function of temperature. I recommend thinking in terms of an abstract point moving in macrostate-space. Both  $T$  and  $E$  are well-behaved functions, with definite values at each point in macrostate-space. We get into trouble if we try to parameterize this point using  $T$  as one of the coordinates, but this is a problem with the coordinate representation, not with the abstract space itself.

We will now choose a particular set of variables as a basis for specifying points in macrostate-space. We will use this set for a while, but we are not wedded to it. As one of our variables, we choose  $S$ , the entropy. The remaining variables we will collectively call  $V$ , which is a vector with  $D - 1$  dimensions. In particular, we choose the macroscopic variable  $V$  in such a way that the microscopic energy  $\hat{E}_i$  of the  $i$ th microstate is determined by  $V$ . (For an ideal gas in a box,  $V$  is just the volume of the box.)

Given these rather restrictive assumptions, we can write:

$$dE = \left. \frac{\partial E}{\partial V} \right|_S dV + \left. \frac{\partial E}{\partial S} \right|_V dS \quad (7.5)$$

which is just the chain rule for differentiating a function of two variables. Important generalizations of this equation can be found in section 7.6 and section 18.1.

It is conventional to define the symbols

$$P := - \left. \frac{\partial E}{\partial V} \right|_S \quad (7.6)$$

and

$$\begin{aligned} k\beta &:= \left. \frac{\partial S}{\partial E} \right|_V \\ T &:= 1/k\beta \quad \text{for nonzero } \beta \end{aligned} \quad (7.7)$$

You might say this is just terminology, just a definition of  $T$  ... but we need to be careful because there are also other definitions of  $T$  floating around. For starters, you can compare equation 7.7 with equation 15.10. More importantly, we need to connect this definition of  $T$  to the real physics, and to operational definitions of temperature. There are some basic qualitative properties that temperature should have, as discussed in section 11.1, and we need to show that our definition exhibits these properties. See chapter 13.

Equation 7.7 is certainly not the most general definition of temperature, because of several assumptions that we made in the lead-up to equation 7.5. By way of counterexample, in NMR or ESR, a  $\tau_2$  process changes the entropy without changing the energy. As an even simpler counterexample, internal leakage currents within a thermally-isolated storage battery increase the entropy of the system without changing the energy; see figure 1.3 and section 11.5.5.

Using the symbols we have just defined, we can rewrite equation 7.5 in the following widely-used form:

$$\mathbf{d}E = -P\mathbf{d}V + T\mathbf{d}S \quad (7.8)$$

Again: see equation 7.33, equation 7.34, and section 18.1 for important generalizations of this equation.

Continuing down this road, we can rewrite equation 7.8 as

$$\mathbf{d}E = w + q \quad (7.9)$$

where we choose to define  $w$  and  $q$  as:

$$\begin{aligned} w &:= \left. \frac{\partial E}{\partial V} \right|_S \mathbf{d}V \\ &= -P\mathbf{d}V \end{aligned} \quad (7.10)$$

and

$$\begin{aligned} q &:= \left. \frac{\partial E}{\partial S} \right|_V \mathbf{d}S \\ &= T\mathbf{d}S \end{aligned} \quad (7.11)$$

That’s all fine; it’s just terminology. Note that  $w$  and  $q$  are one-forms, not scalars, as discussed in section 8.2. They are functions of state, i.e. uniquely determined by the thermodynamic state.<sup>1</sup>

Equation 7.9 is fine so long as we don’t misinterpret it. However, beware that equation 7.9 and its precursors are very commonly misinterpreted. In particular, it is tempting to interpret  $w$  as “work” and  $q$  as “heat,” which is either a good idea or a bad idea, depending on which of the various mutually-inconsistent definitions of “work” and “heat” you happen to use. See section 17.1 and section 18.1 for details.

Also: Equation 7.8 is sometimes called the “thermodynamic identity” although that seems like a bit of a misnomer. The only identity involved comes from calculus, not from thermodynamics. We are using a calculus identity to expand the exterior derivative  $\mathbf{d}E$  in terms of some thermodynamically-interesting variables.

Beware that equation 7.8 has got little or nothing to do with the first law of thermodynamics, i.e. with conservation of energy. It has more to do with the fact that  $E$  is a differentiable function of state than the fact that it is conserved. None of the steps used to derive equation 7.8 used the fact that  $E$  is conserved. You could perhaps connect this equation to conservation of energy, but you would have to do a bunch of extra work and bring in a bunch of additional information, including things like the third law of motion, et cetera. To appreciate what I’m saying, it may help to apply the same calculus identity to some *non-conserved* function of state, perhaps the Helmholtz free energy  $F$ . You can go through the same steps and get an equation that is very similar to equation 7.8, as you can see in figure 15.10. If you did not already know what’s conserved and what not, you could not figure it out just by glancing at the structure of these equations.

## 7.5 Heat Capacities, Energy Capacity, and Enthalpy Capacity

Here’s another change of variable that calls attention to some particularly interesting partial derivatives. Now that we have introduced the  $T$  variable, we can write

$$\mathbf{d}E = \left. \frac{\partial E}{\partial V} \right|_T \mathbf{d}V + \left. \frac{\partial E}{\partial T} \right|_V \mathbf{d}T \quad (7.12)$$

assuming things are sufficiently differentiable.

The derivative in the second term on the RHS is conventionally called the *heat capacity at constant volume*. As we shall see in connection with equation 7.19, it is safer to think of this as the *energy capacity*. The definition is:

$$C_V := \left. \frac{\partial E}{\partial T} \right|_V \quad (7.13)$$

again assuming the RHS exists. (This is a nontrivial assumption. By way of counterexample, the RHS does not exist near a first-order phase transition such as the ice/water transition, because

---

<sup>1</sup>In the expression “function of state” or in the equivalent expression “state function,” state always means macrostate. You might think that anything that is a function at all is a function of the microstate, but that’s not true. In particular, entropy is defined as a sum over microstates.

the energy is not differentiable with respect to temperature there. This corresponds roughly to an infinite energy capacity, but it takes some care and some sophistication to quantify what this means. See reference 19.)

The energy capacity in equation 7.13 is an extensive quantity. The corresponding intensive quantities are the specific energy capacity (energy capacity per unit mass) and the molar energy capacity (energy capacity per particle).

The other derivative on the RHS of equation 7.12 doesn't have a name so far as I know. It is identically zero for a table-top sample of ideal gas (but not in general).

The term *isochoric* means “at constant volume,” so  $C_V$  is the isochoric heat capacity ... but more commonly it is just called the “heat capacity at constant volume.”

Using the chain rule, we can find a useful expression for  $C_V$  in terms of entropy:

$$\begin{aligned} C_V &= \frac{\partial E}{\partial S} \frac{\partial S}{\partial T} \quad \text{all at constant } V \\ &= T \left. \frac{\partial S}{\partial T} \right|_V \end{aligned} \quad (7.14)$$

This equation is particularly useful in reverse, as means for measuring changes in entropy. That is, if you know  $C_V$  as a function of temperature, you can divide it by  $T$  and integrate with respect to  $T$  along a contour of constant volume. The relevant formula is:

$$dS = \frac{1}{T} C_V dT \quad \text{at constant } V \quad (7.15)$$

We could have obtained the same result more directly using the often-important fact, from equation 7.8,

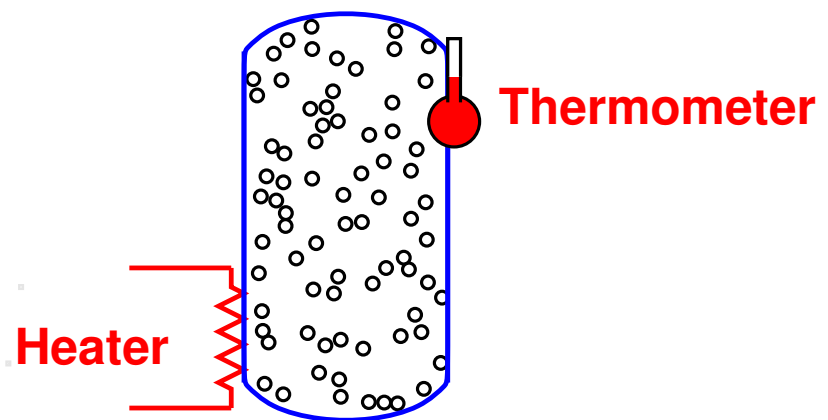
$$dS = \frac{1}{T} dE \quad \text{at constant } V \quad (7.16)$$

and combining it with the definition of  $C_V$  from equation 7.12 and equation 7.13:

$$dE = C_V dT \quad \text{at constant } V \quad (7.17)$$

Equation 7.17 is useful, but there are some pitfalls to beware of. For a given sample, you might think you could ascertain the absolute entropy  $S$  at a given temperature  $T$  by integrating from absolute zero up to  $T$ . Alas nobody has ever achieved absolute zero in practice, and using an approximation of zero K does not necessarily produce a good approximation of the total entropy. There might be a lot of entropy hiding in that last little interval of temperature. Even in theory this procedure is not to be trusted. There are some contributions to the entropy – such as the entropy of mixing – that may be hard to account for in terms of  $dS = dE/T$ . Certainly it would be disastrous to try to “define” entropy in terms of  $dS = dE/T$  or anything like that.

Remark: Equation 7.12 expands the energy in terms of one set of variables, while equation 7.5 expands it in terms of another set of variables. This should suffice to dispel the misconception that  $E$  (or any other thermodynamic potential) is “naturally” a function of one set of variables to the exclusion of other variables. See section 15.6 and reference 3 for more on this.



This concludes our discussion of the constant-volume situation. We now turn our attention to the constant-pressure situation.

Operationally, it is often easier maintain constant ambient pressure than to maintain constant volume. For a gas or liquid, we can measure some sort of “heat capacity” using an apparatus along the lines shown in figure 7.4. That is, we measure the temperature of the sample as a function of the energy put in via the heater. However, this energy is emphatically not the total energy crossing the boundary, because we have not yet accounted for the  $PdV$  work done by the piston as it moves upward (as it must, to maintain constant pressure), doing work against gravity via the weight  $W$ . Therefore the energy of the heater does not measure the change of the real energy  $E$  of the system, but rather of the enthalpy  $H$ , as defined by equation 15.4.

This experiment can be modeled using the equation:

$$dH = \left. \frac{\partial H}{\partial P} \right|_T dP + \left. \frac{\partial H}{\partial T} \right|_P dT \quad (7.18)$$

This is analogous to equation 7.12 ... except that we emphasize that it involves the enthalpy instead of the energy. The second term on the right is conventionally called the *heat capacity at constant pressure*. It is however safer to call it the enthalpy capacity. The definition is:

$$C_P := \left. \frac{\partial H}{\partial T} \right|_P \quad (7.19)$$

Under favorable conditions, the apparatus for measuring  $C_V$  for a chunk of solid substance is particularly simple, because don't need the container and piston shown in figure 7.4; the substance contains itself. We just need to supply thermal insulation. The analysis of the experiment remains the same; in particular we still need to account for the  $PdV$  work done when the sample expands, doing work against the ambient pressure.

The term *isobaric* means “at constant pressure”, so another name for  $C_P$  is the isobaric heat capacity.

In analogy to equation 7.15 we can write

$$dS = \frac{1}{T} C_P dT \quad \text{at constant } P \quad (7.20)$$

which we can obtain using the often-important fact, from equation 15.5,

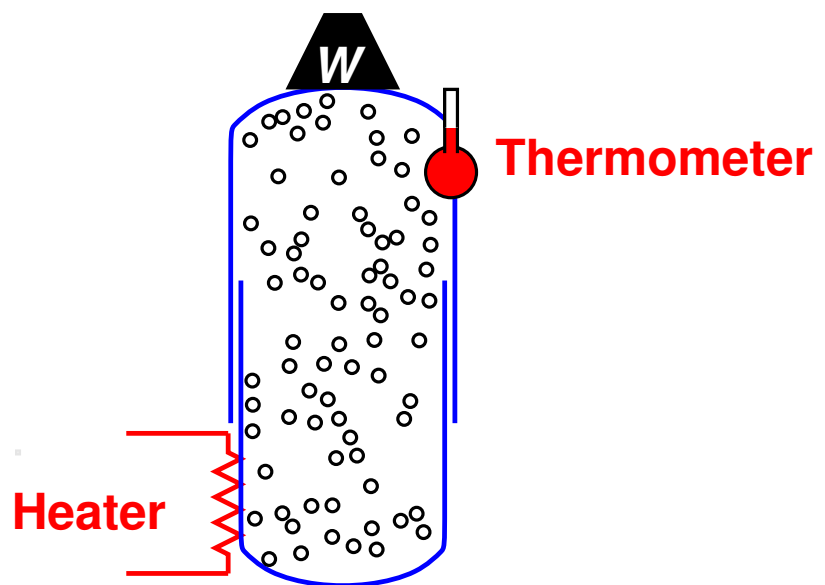
$$dS = \frac{1}{T} dH \quad \text{at constant } P \quad (7.21)$$

and combining it with the definition of  $C_P$  from equation 7.18 and equation 7.19:

$$dH = C_P dT \quad \text{at constant } P \quad (7.22)$$

Collecting results for comparison, we have

$$\begin{aligned} dE &= C_V dT && \text{at constant } V \\ dH &= C_P dT && \text{at constant } P \\ dS &= \frac{1}{T} C_V dT && \text{at constant } V \\ dS &= \frac{1}{T} C_P dT && \text{at constant } P \end{aligned} \quad (7.23)$$





Remark: We see once again that the term “heat” is ambiguous in ways that entropy is not. In the first two rows, the LHS is different, yet both are called “heat”, which seems unwise. In the second two rows, the LHS is the same, and both are called entropy, which is just fine.

Starting with either of the last two lines of equation 7.23 and solving for the heat capacity, we see that we can define a *generalized heat capacity* as:

$$\begin{aligned} C_X &= T \left. \frac{\partial S}{\partial T} \right|_X \\ &= \left. \frac{\partial S}{\partial \ln(T)} \right|_X \end{aligned} \quad (7.24)$$

where  $X$  can be just about anything, including  $X \equiv V$  or  $X \equiv P$ .

Remark: Heat capacity has the same dimensions as entropy.

We see from equation 7.24 that the so-called heat capacity can be thought of as the *entropy capacity* ... especially if you use a logarithmic temperature scale.

Equation 7.24 is useful for many theoretical and analytical purposes, but it does not directly correspond to the way heat capacities are usually measured in practice. The usual procedure is to observe the temperature as a function of energy or enthalpy, and to apply equation 7.13 or equation 7.19.

This supports the point made in section 0.3 and section 17.1, namely that the concept of “heat” is a confusing chimera. It’s part energy and part entropy. It is neither necessary nor possible to have an unambiguous understanding of “heat.” If you understand energy and entropy, you don’t need to worry about heat.

## 7.6 $E$ as a Function of Other Variables

In section 7.4 we temporarily assumed that the energy is known as a function of entropy and volume. This is certainly not the general case. We are not wedded to using  $V$  and  $S$  as coordinates for mapping out the thermodynamic state space (macrostate space).

The point of this section is not to make your life more complicated by presenting lots of additional equations. Instead, the main point is to focus attention on the one thing that really matters, namely conservation of energy. The secondary point is that equations such as equation 7.8, equation 7.27, et cetera fall into a pattern: they express the exterior derivative of  $E$  in terms of the relevant variables, relevant to this-or-that special situation. Once you see the pattern, you realize that the equations are not at all fundamental.

### 7.6.1 $V$ , $S$ , and $h$

Here is a simple generalization that requires a bit of thought. Suppose we have a parcel of fluid with some volume and some entropy. You might think we could write the energy as a function of  $V$  and  $S$ , as in equation 7.25 but that is not always sufficient, as we shall soon see.

$$\begin{aligned} dE &= TdS - PdV \\ &= \text{“heat”} + \text{“work”} \\ &= \text{thermal} + \text{mechanical} \quad \otimes \end{aligned} \quad (7.25)$$

Consider the apparatus shown in figure 7.5. The parcel of fluid is trapped between two pistons in the same cylinder. On the left is the initial situation.

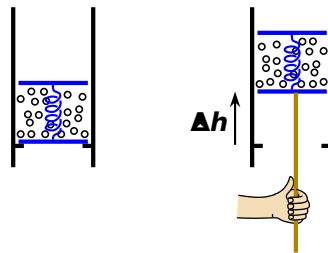


Figure 7.5: Energy Depends on  $V$ ,  $S$ , and  $h$

On the right we see that the parcel has been raised by a distance  $\Delta h$ . It was raised slowly, gently, and reversibly, using a thermally-insulating pushrod. The entropy of the parcel did not change. The volume of the parcel did not change. Overall, no work was done on the parcel.

- Zero overall work is consistent with the fact that at all times while the parcel was being hoisted, the upward force from the pushrod was balanced by the downward force of gravity, i.e. the weight of the parcel of fluid. (We ignore the weight of the pistons and other moving parts.) From the point of view of the guy holding the pushrod, he is doing work ... but from the point of view of the parcel, the pushrod and the gravitational field are doing equal-and-opposite work, so the overall work *on the parcel* is zero.

- Zero overall work is also consistent with the work/KE theorem, since the center-of-mass velocity of the parcel was negligible before, during, and after the hoisting procedure.

The interesting thing is that the gravitational potential energy of the parcel did change.

Nitpickers may argue about whether the gravitational energy is “in” the parcel or “in” the gravitational field, but we don’t care. In any case, the energy is *associated* with the parcel, and we choose to include it in our definition of  $E$ , the energy “of” the parcel. The idea of  $E = mgh$  is perhaps the first and most basic energy-related formula that you ever saw.

The exterior derivative must include terms for each of the relevant variables:

$$dE = \left. \frac{\partial E}{\partial V} \right|_{h,S} dV + \left. \frac{\partial E}{\partial S} \right|_{h,V} dS + \left. \frac{\partial E}{\partial h} \right|_{V,S} dh \quad (7.26)$$

Under mild conditions this simplifies to:

$$\begin{aligned} dE &= TdS - PdV + mgdh \\ &= \text{“heat”} + \text{“work”} + \text{“hoist”} \\ &= \text{thermal} + \text{—mechanical—} \end{aligned} \quad (7.27)$$

where  $mg$  is the weight of the parcel. There are two “mechanical” terms.

Equation 7.27 tells us that we can change the energy of the parcel without doing work on it. This should not come as a surprise; there is a work/kinetic-energy theorem, not a work/total-energy theorem.

Nevertheless this does come as a surprise to some people. Part of the problem is that sometimes people call equation 7.8 “the” first law of thermodynamics. They treat it as “the” fundamental equation, as if it were the 11th commandment. This leads people to think that the only way of changing the energy of the parcel is by doing mechanical work via the  $PdV$  term and/or exchanging heat via the  $TdS$  term. Let’s be clear: the  $mgdh$  term is mechanical, but it is not work (not  $PdV$  work, and not overall work from the parcel’s point of view).

Another part of the problem is that when thinking about thermodynamics, people sometimes think in terms of an oversimplified model system, perhaps a small (“table-top”) sample of ideal gas. They make assumptions on this basis. They equate “mechanical energy transfer” with work. Then, when they try to apply their ideas to the real world, everything goes haywire. For a table-top sample of ideal gas, moving it vertically makes only a small contribution to the energy, negligible in comparison to ordinary changes in pressure or temperature. However, the contribution is not small if you’re looking at a tall column of air in the earth’s atmosphere, or water in the ocean, or the plasma in the sun’s atmosphere, et cetera. Vertical motions can have tremendous implications for temperature, pressure, stability, transport, mixing, et cetera.

A liquid is typically 1000 times denser than a gas at STP. So if you imagine the fluid in figure 7.5 to be a liquid rather than a gas, the  $mgdh$  contribution is 1000 times larger.

In a situation where we know the volume is constant, equation 7.27 simplifies to:

$$dE = TdS + mgdh \quad (7.28)$$

That superficially looks like equation 7.8 (with an uninteresting minus sign). On the RHS we can identify a “thermal” term and a “mechanical” term. However, it is spectacularly different for a non-obvious reason. The reason is that  $V$  and  $h$  enter the equation of state in different ways. Changing  $V$  at constant  $h$  and  $S$  changes the temperature and pressure, whereas changing  $h$  at constant  $V$  and  $S$  does not. For the next level of detail on this, see section 9.3.4.

Some folks try to simplify equation 7.27 by rewriting it in terms of the «internal energy», but I’ve never found this to be worth the trouble. See section 7.7.

### 7.6.2 $X, Y, Z$ , and $S$

Here’s a simple but useful reformulation. It doesn’t involve any new or exciting physics. It’s the same idea as in section 7.4, just with slightly different variables: forces instead of pressure.

Suppose we have a box-shaped parcel of fluid. As it flows along, it might change its size in the  $X$  direction, the  $Y$  direction, or the  $Z$  direction. The volume is  $V = XYZ$ . Instead of equation 7.5 we write:

$$\begin{aligned} \mathbf{d}E &= \left. \frac{\partial E}{\partial X} \right|_{Y,Z,S} \mathbf{d}X + \left. \frac{\partial E}{\partial Y} \right|_{Z,X,S} \mathbf{d}Y + \left. \frac{\partial E}{\partial Z} \right|_{X,Y,S} \mathbf{d}Z + \left. \frac{\partial E}{\partial S} \right|_{X,Y,Z} \mathbf{d}S \\ &= -YZP\mathbf{d}X + -ZXP\mathbf{d}Y - XYP\mathbf{d}Z + T\mathbf{d}S \\ &= -F_X\mathbf{d}X + -F_Y\mathbf{d}Y - F_Z\mathbf{d}Z + T\mathbf{d}S \end{aligned} \quad (7.29)$$

where we define the force  $F_X$  as a directional derivative of the energy:

$$F_X := - \left. \frac{\partial E}{\partial X} \right|_{Y,Z,S} \quad (7.30)$$

and similarly for the forces in the  $Y$  and  $Z$  directions. Compare equation 18.5.

### 7.6.3 $V, S$ , and $N$

Here’s another widely-useful generalization. Sometimes we have a box where the number of particles is not constant. We might be pumping in new particles and/or letting old particles escape. The energy will depend on the number of particles ( $N$ ). The exterior derivative is then:

$$\mathbf{d}E = \left. \frac{\partial E}{\partial N} \right|_{V,S} \mathbf{d}N + \left. \frac{\partial E}{\partial V} \right|_{N,S} \mathbf{d}V + \left. \frac{\partial E}{\partial S} \right|_{N,V} \mathbf{d}S \quad (7.31)$$

For present purposes, we assume there is only one species of particles, not a mixture.

This is a more-general expression; now equation 7.5 can be seen a corollary valid in the special case where  $N$  is constant (so  $\mathbf{d}N = 0$ ).

The conventional pet name for the first derivative on the RHS is *chemical potential*, denoted  $\mu$ . That is:

$$\mu := \left. \frac{\partial E}{\partial N} \right|_{V,S} \quad (7.32)$$

where  $N$  is the number of particles in the system (or subsystem) of interest.

This means we can write:

$$\mathbf{d}E = \mu \mathbf{d}N - P \mathbf{d}V + T \mathbf{d}S \quad (7.33)$$

which is a generalization of equation 7.8.

It is emphatically not mandatory to express  $E$  as a function of  $(V, S)$  or  $(N, V, S)$ . Almost any variables that span the state-space will do, as mentioned in section 15.6 and reference 3.

You should not read too much into the name “chemical” potential. There is not any requirement nor even any connotation that there be any chemical reactions going on.

The defining property of the chemical potential ( $\mu$ ) is that it is conjugate to an increase in number ( $\mathbf{d}N$ ) ... just as the pressure ( $P$ ) is conjugate to a decrease in volume ( $-\mathbf{d}V$ ). Note the contrast: in the scenario described by equation 7.33:

Stepping across a contour of $-\mathbf{d}V$ increases the density (same number in a smaller volume).	Stepping across a contour of $\mathbf{d}N$ increases the density (bigger number in the same volume).
--	--

This can happen if a piston is used to change the volume.	This can happen if particles are carried across the boundary of the system, or if particles are produced within the interior of the system (by splitting dimers or whatever).
---	---

So we see that  $\mathbf{d}N$  and  $\mathbf{d}V$  are two different directions in parameter space. Conceptually and mathematically, we have no basis for declaring them to be “wildly” different directions or only “slightly” different directions; all that matters is that they be different i.e. linearly independent. At the end of the day, we need a sufficient number of linearly independent variables, sufficient to span the parameter space.

Equation 7.33 is a generalization of equation 7.8, but it is not the absolute most-general equation. In fact there is no such thing as the most-general equation; there’s always another generalization you can make. For example, equation 7.33 describes only one species of particle; if there is another species, you will have to define a new variable  $N_2$  to describe it, and add another term involving  $\mathbf{d}N_2$  to the RHS of equation 7.33. Each species will have its own chemical potential. Similarly, if there are significant magnetic interactions, you need to define a variable describing the magnetic field, and add the appropriate term on the RHS of equation 7.33. If you understand the meaning of the equation, such generalizations are routine and straightforward. Again: At the end of the day, any expansion of  $\mathbf{d}E$  needs a sufficient number of linearly independent variables, sufficient to span the relevant parameter space.

For a more formal discussion of using the chain rule to expand differentials in terms of an *arbitrary* number of variables, see reference 3.

### 7.6.4 Yet More Variables

In general, we need even more variables. For example, for a parcel of fluid in a flow reactor, we might have:

$$dE = \sum \mu_i dN_i - PdV + mgdh + mv \cdot dv + TdS + \dots \quad (7.34)$$

where  $N_i$  is the number of molecular entities of the  $i$ th kind,  $m$  is the mass of the parcel,  $g$  is the acceleration of gravity,  $h$  is the height,  $v$  is the velocity, and the ellipsis ( $\dots$ ) represents all the terms that have been left out.

Note that in many cases it is traditional to leave out the ellipsis, recognizing that *no equation* is fully general, and equation 7.33 is merely a corollary of some unstated cosmic generality, valid under the proviso that the omitted terms are unimportant.

Opinions differ, but one common interpretation of equation 7.34 is as follows: the  $TdS$  term can be called the “heat” term, the two terms  $-PdV + mgdh$  can be called “work” terms, the  $\mu_i dN_i$  is neither heat nor work, and I don’t know what to call the  $mv \cdot dv$  term. Obviously the  $mv \cdot dv$  term is important for fluid dynamics, and the  $\mu dN$  term is important for chemistry, so you would risk getting lots of wrong answers if you rashly assumed equation 7.8 were “the” definition of heat and work.

## 7.7 Internal Energy

As foreshadowed in section 1.8.4, a great many thermodynamics books emphasize the so-called «internal energy», denoted  $U$  or  $E_{\text{in}}$ . Mostly they restrict attention to situations where the «internal energy» is identically equal to the plain old energy, so I have to wonder why they bothered to introduce a fancy new concept if they’re not going to use it. In situations where the two concepts are not equivalent, things are even more mysterious. I have never found it necessary to make sense of this. Instead I reformulate everything in terms of the plain old energy  $E$  and proceed from there.

Feel free to skip this section ... but if you’re curious, the «internal energy» is defined as follows:

Suppose we have a smallish parcel of fluid with total mass  $M$  and total momentum  $\Pi$  as measured in the lab frame.<sup>2</sup> Its center of mass is located at position  $R$  in the lab frame. Then we can express the «internal energy» of the parcel as:

$$E_{\text{in}} = E - \frac{\Pi^2}{2M} - \Phi(R) \quad (7.35)$$

where  $\Phi$  is some potential. If it is a gravitational potential then  $\Phi(R) = -Mg \cdot R$ , where  $g$  is a downward-pointing vector.

The «internal energy» is a function of state. As you can see from equation 7.35:

---

<sup>2</sup>We write the momentum as  $\Pi$  instead of the more conventional  $P$ , to avoid a conflict with the pressure ( $P$ ).

- It explicitly excludes the kinetic energy of the parcel as a whole, associated with the center-of-mass motion of the parcel relative to the lab frame. It also explicitly excludes gravitational potential energy and similar potentials associated with “body forces”
- On the other hand, it includes contributions from the kinetic energy of the various particles as they move relative to the center-of-mass frame (not the lab frame). It also includes the potentials that depend on the particles’ positions relative to each other. It also includes each particle’s own intra-particle binding energy, rotational energy, and vibrational energy.

In other words, the «internal energy» can be thought of as the energy of a parcel as observed in a frame that is comoving with and colocated with the parcel’s center of mass. This can be considered a separation of variables. A complete description of the system can be thought of in terms of the variables in the center-of-mass frame plus the “special” variables that describe the location and velocity of the center-of-mass itself (relative to whatever frame we are actually using).

I’ve always found it easier to ignore the «internal energy» and use the plain old energy ( $E$ ) instead, for multiple reasons:

1. The definition of  $E$  is reasonably well known, whereas it’s hard to learn and hard to remember what’s supposed to be included in  $E_{\text{in}}$ .
2. There is a basic conceptual issue: Just because the “special” variables (describing the CM itself) can be separated does not mean they can be ignored, especially if the parcel is changing its velocity, changing its acceleration, changing its altitude, et cetera.
  - If you chose, you can define things so as to exclude  $Mg \cdot R$  from the definition of  $E_{\text{in}}$ , but you cannot exclude it from the actual equation of motion, or from the thermodynamic equation of state. The observed heat capacity of a ten-mile-high column of gas depends on the gravitational interaction. One way or another, you have to account for it.
  - The same goes for kinetic energy. It is well known that if reference frame  $A$  is moving relative to frame  $B$ , switching from one frame to the other can give a wildly different value for the energy, even though the system itself hasn’t changed.
3. The «internal energy» is not conserved (except in special cases). To see this, consider the contrast:

One parcel expands in such a way as to compress a neighboring parcel. $E_{\text{in}}$ is conserved in this special case. So far so good.	One parcel expands in such a way as to <i>hoist</i> a neighboring parcel. It seems to me that $E_{\text{in}}$ is not conserved.
--	---

An even better argument that leads to the same conclusion is based on the situation shown in figure 7.6. There are two jack-in-the-box mechanisms. We focus attention on the blue one, on the left side of the diagram.

- The top row shows the initial situation. Each box is stationary, pushing against its mirror-image partner.
- The bottom row shows the later situation. Each box has pushed off and is now moving away from the centerline.

Figure 7.6: Internal Energy of a Jack-in-the-Box

If you want to make this look more classically thermodynamical, you can replace each spring by some gas molecules under pressure. The idea is the same either way.

It should be obvious by symmetry that no energy crossed the boundary of the blue system. (Some momentum crossed the boundary, but that's the answer to a different question.) No work was done by (or on) the blue system. There was no  $F \cdot dx$ . The pushrod does not move until after contact has ceased and there is no longer any force. At the time and place where there was a nonzero force, there was no displacement. At times and places where there was a nonzero displacement, there was no force.

As a secondary argument leading to the same conclusion, you could equally well replace the red box by a rigid infinitely-massive wall. This version has fewer moving parts but less symmetry. Once again the displacement at the point of contact is zero. Therefore the  $F \cdot dx$  work is zero. The *pseudowork* is nonzero, but the actual thermodynamic work is zero. (Pseudowork is discussed in section 18.5 and reference 18.)

The purpose of this Gedankenexperiment is to make a point about conservation and non-conservation:

The plain old energy $E$ is conserved. Some energy that was stored in the spring has been converted to KE ... more specifically, converted to KE associated with motion of the center of mass.	The «internal energy» is not conserved. The stored energy counts toward $E_{\text{in}}$ , whereas the center-of-mass KE does not.
--	---

It is widely believed<sup>3</sup> that «if no matter or energy crosses the boundary of the system, then the internal energy is constant». First of all, this is not true, as demonstrated by figure 7.6. We can cook up a similar-sounding statement that is actually true, if we define a notion of *super-isolated* system, which cannot exchange matter, energy, momentum, or anything else. Still, even the repaired statement is highly misleading. It is practically begging people to reason by analogy to the plain old energy, which leads to the wrong answer about conservation of «internal energy», as we see from the following table:

	<u>Constant in a super-isolated system</u>	<u>Conserved</u>	
Energy :	yes	yes	
«Internal Energy» :	yes	no	← surprise!

Table 7.1: Constancy versus Conservation

Just because some variable is constant in a super-isolated system does not mean it is conserved.

---

<sup>3</sup>See e.g. Wikipedia, which is a cornucopia of things that are believed to be true without actually being true. It says: «If the containing walls pass neither matter nor energy, the system is said to be isolated. Then its internal energy cannot change.» Reference: [https://en.wikipedia.org/w/index.php?title=Internal\\_energy&oldid=697674365](https://en.wikipedia.org/w/index.php?title=Internal_energy&oldid=697674365)



There are plenty of counterexamples. In special relativity, mass is in this category, as discussed in reference 7. See section 1.2 and reference 6 for a discussion of what we mean by conservation.

4. The definition of «internal energy» seems inconsistent with the usual definition of force, as embodied in the principle of virtual work. Suppose we try to define force as a derivative of the «internal energy». Then if  $Mg \cdot R$  is missing from the «internal energy», then  $Mg$  is missing from the force. That seems like a problem.

The books that glorify  $E_{\text{in}}$  (aka  $U$ ) typically write something like

$$dU = \text{heat} + \text{work} \quad (7.36)$$

and then assert that it expresses conservation of energy. I find this very odd, given that in reality  $U$  is not conserved.

Suggestion: If you are ever tempted to formulate thermodynamics in terms of «internal energy», start by calculating the heat capacity of a ten-mile-high column of air in a standard gravitational field. Hint: As you heat the air, its center of mass goes up, changing the gravitational potential energy, even though the container has not moved. I predict this will make you much less enamored of  $U$ , much less ready to enshrine it as the central focus of thermodynamics.

Here's how I think about it: Very often there are some degrees of freedom that do not equilibrate with others on the timescale of interest. For example:

- I can measure the temperature of a flashlight battery even though the electrical degree of freedom is very far out of equilibrium, i.e. a bazillion  $kT$  higher than would be expected from equipartition. This special mode is protected by conservation of charge in conjunction with well-engineered electrical insulators.
- I can measure the temperature of a car that is moving relative to the lab frame, even though the overall motion of the CM is very far out of equilibrium. This special mode is protected by conservation of momentum in conjunction with well-engineered bearings.

These special modes contribute to the energy in the usual way, even though they do not equilibrate in the usual way. It is necessary to identify them and assign them their own thermodynamic variables. On the other hand, as far as I can tell, it is not necessary or even helpful to define new «energy-like» thermodynamic potentials (such as  $E_{\text{in}}$  aka  $U$ ).

Special variables yes; special potentials no. Note that the «internal energy» as usually defined gives special treatment to the center-of-mass kinetic energy but not to the battery electrical energy. That is yet another indicator that it doesn't really capture the right idea.

## 7.8 Integration

Let's continue to assume that  $T$  and  $P$  are functions of state, and that  $S$  and  $V$  suffice to span the macrostate-space. (This is certainly *not* always a safe assumption, as you can see in e.g. equation 7.27.)

Then, in cases where equation 7.8 is valid, we can integrate both sides to find  $E$ . This gives us an expression for  $E$  as a function of  $V$  and  $S$  alone (plus a constant of integration that has no physical significance). Naturally, this expression is more than sufficient to guarantee that  $E$  is a function of state.

Things are much messier if we try to integrate only one of the terms on the RHS of equation 7.8. Without loss of generality, let's consider the  $T \mathbf{d}S$  term. We integrate  $T \mathbf{d}S$  along some path  $\Gamma$ . Let the endpoints of the path be  $A$  and  $B$ .

It is crucial to keep in mind that the value of the integral depends on the chosen path — not simply on the endpoints. It is OK to write things like

$$s\Delta Q_\Gamma = \int_\Gamma T \mathbf{d}S \quad (7.37)$$

whereas it would be quite unacceptable to replace the path with its endpoints:

$$(\text{anything}) = \int_A^B T \mathbf{d}S \quad (7.38)$$

I recommend writing  $Q_\Gamma$  rather than  $Q$ , to keep the path-dependence completely explicit. This  $Q_\Gamma$  exists only along the low-dimensional subspace defined by the path  $\Gamma$ , and cannot be extended to cover the whole thermodynamic state-space. That's because  $T \mathbf{d}S$  is an ungrady one-form. See section 8.2 for more about this.

## 7.9 Advection

Equation 7.8 is predicated on the assumption that the energy is known as a function  $V$  and  $S$  alone. However, this is not the most general case. As an important generalization, consider the energy budget of a typical automobile. The most-common way of increasing the energy within the system is to transfer fuel (and oxidizer) across the boundary of the system. This is an example of advection of energy. This contributes to  $\mathbf{d}E$ , but is not included in  $P \mathbf{d}V$  or  $T \mathbf{d}S$ . So we should write something like:

$$\mathbf{d}E = -P \mathbf{d}V + T \mathbf{d}S + \text{advection} \quad (7.39)$$

It is possible to quantify the advection mathematically. Simple cases are easy. The general case would lead us into a discussion of fluid dynamics, which is beyond the scope of this document.

## 7.10 Deciding What's True

Having derived results such as equation 7.8 and equation 7.39, we must figure out how to interpret the terms on the RHS. Please consider the following notions and decide which ones are true:

1. Heat is defined to be  $T \mathbf{d}S$  (subject to the usual restrictions, discussed in section 7.4).

2. Heat is defined to be “energy that is transferred from one body to another as the result of a difference in temperature?”
3. The laws of thermodynamics apply even when irreversible processes are occurring.

It turns out that these three notions are mutually contradictory. You have to get rid of one of them, for reasons detailed in section 17.1 and section 8.6.

As a rule, you are allowed to define your terms however you like. However, if you want a term to have a formal, well-defined meaning,

- Each term must be defined only once, and
- You must stick to a well-known unambiguous meaning, and/or clearly explain what definition you are using.

The problem is, many textbooks don’t play by the rules. On some pages they define heat to be  $TdS$ , on some pages they define it to be flow across a boundary, and on some pages they require thermodynamics to apply to irreversible processes.

This is an example of boundary/interior inconsistency, as discussed in section 8.6.

The result is a shell game, or a whack-a-mole game: There’s a serious problem, but nobody can pin down the location of the problem.

This results in endless confusion. Indeed, sometimes it results in holy war between the Little-Endians and the Big-Endians: Each side is 100% convinced that their definition is “right,” and therefore the other side must be “wrong.” (Reference 20.) I will not take sides in this holy war. Viable alternatives include:

1. Pick one definition of heat. Explicitly say which definition you’ve chosen, and use it consistently. Recognize that others may choose differently.
2. Go ahead and use the term informally, with multiple inconsistent meanings, as many experts do. Just don’t pretend you’re being consistent when you’re not. Use other terms and concepts (e.g. energy and entropy) when you need to convey a precise meaning.
3. Avoid using term “heat” any more than necessary. Focus attention on other terms and concepts (e.g. energy and entropy).

For more on this, see the discussion near the end of section 7.11.

## 7.11 Deciding What’s Fundamental

It is not necessarily wise to pick out certain laws and consider them “axioms” of physics. As Feynman has eloquently argued in reference 21, real life is not like high-school geometry, where you were given a handful of axioms and expected to deduce everything from that. In the real world, every fact is linked to many other facts in a grand tapestry. If a hole develops in the tapestry, you can re-weave

it starting from the top of the hole, or the bottom, or either side. That is to say, if you forget one particular fact, you can re-derive it in many different ways.

In this spirit, some folks may wish to consider equation 1.1 and equation 7.9 as being equally axiomatic, or equally non-axiomatic. One can be used to re-derive the other, with the help of other facts, subject to certain limitations.

On the other hand, some facts are more useful than others. Some are absolutely central to our understanding of the world, while others are less so. Some laws are more worth discussing and remembering, while others are less so. Saying that something is true and useful does not make it fundamental; the expression  $1 + 2 + 3 + 4 = 10$  is true and sometimes useful, but it isn't very fundamental, because it lacks generality.

Deciding which laws to emphasize is to some extent a matter of taste, but one ought to consider such factors as simplicity and generality, favoring laws with a large number of predictions and a small number of exceptions.

In my book, energy conservation (equation 1.1) is fundamental. From that, plus a couple of restrictions, we can derive equation 7.9 using calculus. Along the way, the derivation gives us important information about how  $w$  and  $q$  should be interpreted. It's pretty clear what the appropriate restrictions are.

If you try to go the other direction, i.e. from  $w + q$  to conservation of energy, you must start by divining the correct interpretation of  $w$  and  $q$ . The usual "official" interpretations are questionable to say the least, as discussed in section 11.5 and section 8.6. Then you have to posit suitable restrictions and do a little calculus. Finally, if it all works out, you end up with an unnecessarily restrictive version of the local energy-conservation law.

Even in the best case I have to wonder why anyone would bother with the latter approach. I would consider such a derivation as being supporting evidence for the law of local conservation of energy, but not even the best evidence.

I cannot imagine why anyone would want to use equation 7.9 or equation 7.39 as "the" first law of thermodynamics. Instead, I recommend using the local law of conservation of energy ... which is simpler, clearer, more fundamental, more powerful, and more general.

It's not at all clear that thermodynamics should be formulated in quasi-axiomatic terms, but if you insist on having a "first law" it ought to be a simple, direct statement of local conservation of energy. If you insist on having a "second law" it ought to be a simple, direct statement of local paraconservation of entropy.

Another way to judge equation 7.9 is to ask to what extent it describes this-or-that practical device. Two devices of the utmost practical importance are the thermally-insulating pushrod and the ordinary nonmoving heat exchanger. The pushrod transfers energy and momentum (but no entropy) across the boundary, while the heat exchanger transfers energy and entropy (but no momentum) across the boundary.

It is traditional to describe these devices in terms of work and heat, but it is not necessary to do so, and I'm not convinced it's wise. As you saw in the previous paragraph, it is perfectly possible to describe them in terms of energy, momentum, and entropy, which are the true coin of the realm,

the truly primary and fundamental physical quantities. Heat and work are secondary at best (even after you have resolved the nasty inconsistencies discussed in section 7.10 and section 8.6).

Even if/when you *can* resolve  $\mathbf{d}E$  into a  $-P\mathbf{d}V$  term and a  $T\mathbf{d}S$  term, that doesn't mean you *must* do so. In many cases you are better off keeping track of  $E$  by itself, and keeping track of  $S$  by itself. Instead of saying no heat flows down the pushrod, it makes at least as much sense to say that no entropy flows down the pushrod. Keeping track of  $E$  and  $S$  is more fundamental, as you can see from the fact that energy and entropy can be exchanged between systems that don't even have a temperature (section 11.4).

When in doubt, rely on the fundamental laws: conservation of energy, conservation of momentum, paraconservation of entropy, et cetera.



## Chapter 8

# Thermodynamic Paths and Cycles

### 8.1 A Path Projected Onto State Space

#### 8.1.1 State Functions

In this section we will illustrate the ideas using a scenario based on a heat engine. Specifically, it is a mathematical model of a reversible heat engine. For simplicity, we assume the working fluid of the engine is a nondegenerate ideal monatomic gas.

Figure 8.1 shows the thermodynamic state space for our engine. Such a plot is sometimes called the “indicator diagram” for the engine. The vertical white lines are contours of constant pressure. The horizontal white lines are contours of constant volume. The light-blue curves are contours of constant temperature ... and are also contours of constant energy (since the gas is ideal). The dashed light-green curves are contours of constant entropy. Given any arbitrary abstract point  $X$  in this space, we can find the values of all the state variables at that point, including the energy  $E(X)$ , entropy  $S(X)$ , pressure  $P(X)$ , volume  $V(X)$ , temperature  $T(X)$ , et cetera. The spreadsheet used to prepare the figures in this section is cited in reference 22.

That’s nice as far as it goes, but it’s not the whole story. It is also useful to consider the path ( $\Gamma$ ) shown in red in the  $(P, V)$  diagram. We will keep track of what happens as a function of *arc length* ( $\theta$ ) along this path ... not just as a function of state. We note that the path can be broken down into five approximately circular cycles. In this simple case, we choose units of arc length so that each cycle is one unit. Each cycle begins and ends at the point  $X_0$ , which has coordinates  $(P, V) = (2.5, 3)$  in figure 8.1.

Let’s use  $X$  to represent an arbitrary, abstract point in state space, and use  $X_\Gamma(\theta)$  to represent a point *along the specified path*. For simplicity, we write  $X_0 := X_\Gamma(0)$ .

In figure 8.2, we plot several interesting thermodynamic quantities as a function of arc length. Whenever  $\theta$  is an integer, the path crosses through the point  $X_0$ . As a consequence, each of the variables returns exactly to its original value each time  $\theta$  completes a cycle, as you can see in figure 8.2. There have to be definite, reproducible values at the point  $X_0$ , since we are talking about functions of state.

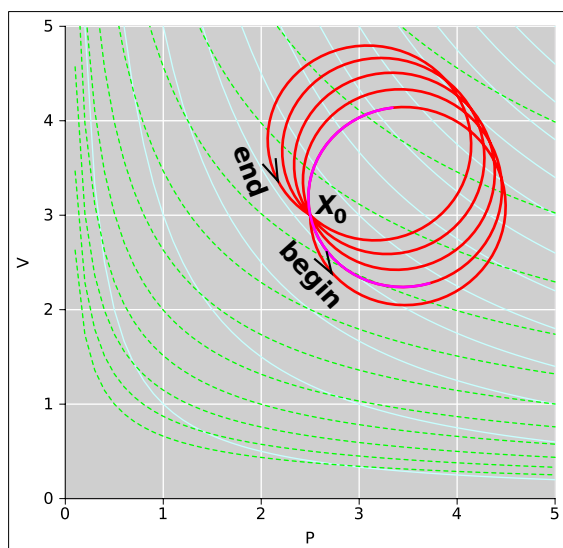


Figure 8.1: A Path Projected Onto State Space

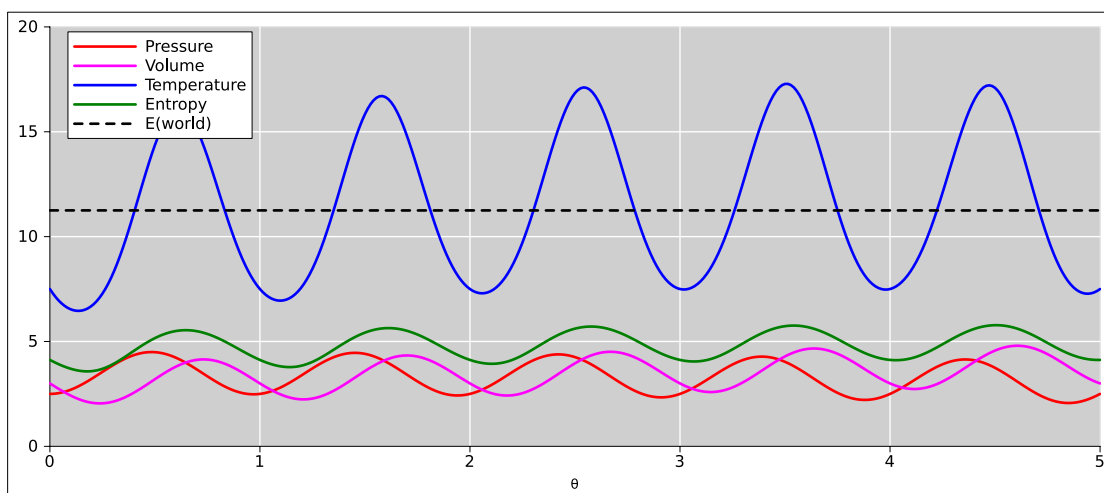


Figure 8.2: State Functions versus Arc Length



In the spreadsheet used to create this figure, the path is chosen to be more-or-less arbitrary function of  $\theta$ , chosen to give us a nice-looking shape in figure 8.1. The path tells us  $P$  and  $V$  as functions of  $\theta$ . Temperature then is calculated as a function of  $P$  and  $V$  using the ideal gas law, equation 26.40. Entropy is calculated as a function of  $V$  and  $T$  using the Sackur-Tetrode law, equation 26.17. The order in which things are calculated is a purely tactical decision. Please do not get the idea that any of these state variables are “natural” variables or “independent” variables.

As discussed in section 7.4, as a corollary of the way we defined temperature and pressure, we obtain equation 7.8, which is reproduced here:

$$dE = -PdV + TdS \quad (8.1)$$

Each of the five variables occurring in equation 8.1 is naturally and quite correctly interpreted as a property of the interior of the engine. For example,  $S$  is the entropy in the interior of the engine, i.e. the entropy of the working fluid. However, in the scenario we are considering, the two terms on the RHS can be re-interpreted as boundary terms, as follows:

- Because the engine is (by hypothesis) in mechanical equilibrium with its surroundings, whenever it changes its volume it exchanges energy with its surroundings.
- Because the engine is (by hypothesis) in thermal equilibrium with its surroundings, and because it is reversible, the only way it can change its entropy is by exchanging entropy with the surroundings.

This re-interpretation is not valid in general, especially if there is any dissipation involved, as discussed in section 8.6. However, if/when we do accept this re-interpretation, equation 8.1 becomes a statement of conservation of energy. The LHS is the change in energy within the engine, and the RHS expresses the mechanical and thermal transfer of energy across the boundary.

Whether or not we re-interpret anything, equation 8.1 is a vector equation, valid point-by-point in state space, independent of path. We can integrate it along any path we like, including our chosen path  $\Gamma$ .

$$\int_{\Gamma} dE = \int_{\Gamma} TdS - \int_{\Gamma} PdV + \text{const} \quad (8.2)$$

As always, “check your work” is good advice. It is an essential part of critical thinking. In this spirit, the spreadsheet carries out the integrals indicated in equation 8.2. The LHS is trivial; it’s just  $E$ , independent of path, and can be calculated exactly. The spreadsheet cannot evaluate integrals on the RHS exactly, but instead approximates each integral by a sum, using a finite stepsize. Rearranging, we get

$$E + \int_{\Gamma} PdV - \int_{\Gamma} TdS = \text{const} \quad (8.3)$$

This quantity on the LHS shown by the black dashed line in figure 8.2. It is constant to better than one part per thousand. This is a nice sensitive check on the accuracy of our numerical integration. It also tells us that our derivation of the ideal gas law and the Sackur-Tetrode law are consistent with our definitions of temperature and pressure.

If we accept the aforementioned re-interpretation, equation 8.3 tells us that our model system conserves energy. It accounts for the energy within the engine plus the energy transferred across the boundary.

### 8.1.2 Out-of-State Functions

Let's give names to the two integrals on the RHS of equation 8.2. We define the *work* done by the engine to be:

$$W_{\Gamma}(\theta) := \int_{\Gamma} P dV \quad (8.4)$$

where the integral runs over all  $\theta'$  on the path  $\Gamma$ , starting at  $\theta' = 0$  and ending at  $\theta' = \theta$ . We call  $W_{\Gamma}$  the *work* done by the engine along the path  $\Gamma$ . This terminology is consistent with ordinary, customary notions of work.

We also define the *heat* absorbed by the engine to be:

$$Q_{\Gamma}(\theta) := \int_{\Gamma} T dS \quad (8.5)$$

For present purposes, we call this the *heat* absorbed by the engine along the specified path. This is consistent with “some” commonly-encountered definitions of heat.<sup>1</sup>

Figure 8.3 includes all the curves from figure 8.2 plus heat (as shown by the brown curve) and work (as shown by the black curve).

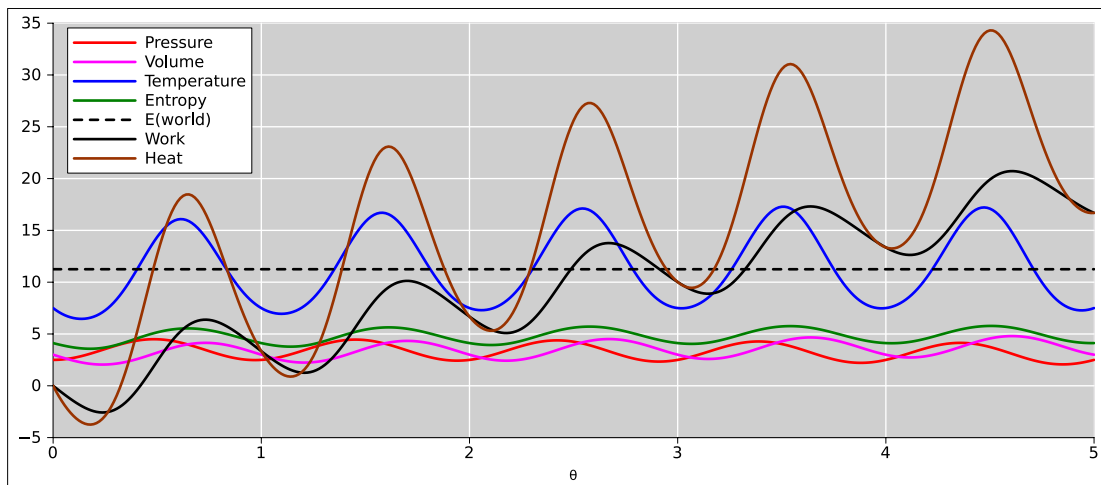


Figure 8.3: State and Out-of-State Functions versus Arc Length

It should be clear in figure 8.3 that the black and brown curves are quite unlike the other curves. Specifically, they are *not* functions of state. They are functions of the arc length  $\theta$  along the path  $\gamma$ , but they are not functions of the thermodynamic state  $X$ . Every time the path returns to the point  $X_0$  in state space,  $W_{\Gamma}(\theta)$  takes on a different value, different from what it was on the previous cycle. Similarly  $Q_{\Gamma}(\theta)$  takes on a different value. You can see this in the figure, especially at places where  $\theta$  takes on integer values. The state of the engine is the same at each of these places, but the heat and work are different.

Heat and work are plotted in figure 8.3, but there is no hope of plotting them on a  $(P, V)$  diagram such as figure 8.1.

<sup>1</sup>... but not all, especially if there is any dissipation involved.

Suppose we continue the operation of this heat engine for more cycles, keeping the cycles always in the vicinity of  $X_0$ .

The state variables such as  $P$ ,  $V$ ,  $T$ , et cetera will remain bounded, returning to their initial values every time the path  $\Gamma$  crosses the initial point  $X_0$ .

At any point  $X$  in state space, if you know  $X$  you know the associated values of the state functions, and the state functions don't care how you got to  $X$ .

The  $W_\Gamma(\theta)$  and  $Q_\Gamma(\theta)$  curves will continue marching off to the northeast, without bound.

At any point  $X$  in state space, even if  $X$  is somewhere on the path  $\Gamma$ , knowing  $X$  is not sufficient to determine out-of-state functions such as  $Q_\Gamma(\theta)$ ,  $W_\Gamma(\theta)$ , or  $\theta$  itself. The out-of-state functions very much depend on how you got to  $X$ .

### 8.1.3 Converting Out-of-State Functions to State Functions

Suppose you start with  $TdS$  and you want to integrate it. There are at least three possibilities:

- You can add something, namely  $-PdV$ , and then integrate. This gives you the energy  $E$ , which is a function of state. The value of the integral depends on the endpoints of the path, but not on other details of the path.
- You can multiply by something, namely  $1/T$ , and then integrate. This gives you the entropy  $S$ , which is a function of state. The value of the integral depends on the endpoints of the path, but not on other details of the path.
- You can integrate  $TdS$  along some chosen path, and accept the fact that the result is not a function of state. The value of the integral depends on every detail of the chosen path.

Please keep in mind that in modern thermodynamics, we define entropy in terms of the probability of the microstates, as discussed in connection with equation 2.2. However, back in pre-modern times, before 1898, people had no understanding of microstates. Quite remarkably, they were able to calculate the entropy anyway, in simple situations, using the following recipe:

- Build a reversible heat engine.
- Measure  $Q_\Gamma(\theta)$ . That is, measure the heat absorbed by the engine as a function of arc length, as the engine goes through its cycles along the path  $\Gamma$ .
- Differentiate with respect to  $\theta$ .
- Multiply by  $1/T$ .
- Integrate with respect to  $\theta$ .

The interesting thing is that even though this recipe depends explicitly on a particular path  $\Gamma$ , the result is path-independent. The result gives you  $S$  as a function of state (plus or minus some constant of integration) – independent of path.

Back in pre-modern times, people not only used this recipe to calculate the entropy, they used it to *define* entropy. It is not a very convenient definition. It is especially inconvenient in situations where dissipation must be taken into account. That's partly because it becomes less than obvious how to define and/or measure  $Q_\Gamma$ .

#### 8.1.4 Reversibility and/or Uniqueness

In our scenario, we choose to operate the heat engine such a way that it moves reversibly along the path  $\Gamma$  in figure 8.1. Every step along the path is done reversibly. There is no law of physics that requires this, but we can make it so (to a good approximation) by suitable engineering.

We can also treat each of the five cycles as a separate path. If we do that, it becomes clear that there are infinitely many paths going through the point  $X_0$ . More generally, the same can be said of any other point in state space: there are infinitely many paths going through that point. As a corollary, given any two points on an indicator diagram, there are infinitely many paths from one to the other. Even if we restrict attention to reversible operations, there are still infinitely many paths from one point to the other.

As a separate matter, we assert that any path through the state-space of the engine can be followed reversibly (to a good approximation) *or irreversibly*. There is no way of looking at a diagram such as figure 8.1 and deciding whether a path is reversible or not. Reversibility depends on variables that are not shown in the diagram. Specifically, when some step along the path is associated with a change in system entropy, the step is either reversible or not, depending on what happens *outside* the system. If there is an equal-and-opposite change in the entropy of the surroundings, the step is reversible, and not otherwise.

I mention this because sometimes people who ought to know better speak of “the” reversible path between two points on a  $(P, V)$  diagram. This is wrong at least twice over.

- There is no such thing as a “reversible” path in a  $(P, V)$  diagram. (You could imagine talking about reversible paths if you constructed a higher-dimensional diagram, including the variable describing not just the heat engine but its surroundings as well, but that’s a different story. That would be a complicated and rather unusual diagram.)
- Even if we restrict attention to reversible operations, there is no such thing as “the” unique path in uncramped thermodynamics. To say the same thing the other way, if the system is so cramped that there is only one path from one  $(P, V)$  point to another, it is obviously impossible to build a heat engine.

#### 8.1.5 The Importance of Out-of-State Functions

As we shall see, state functions are more elegant than out-of-state functions. They are vastly easier to work with. Usually, anything that can be expressed in terms of state functions should be.

However, out-of-state functions remain importance. The entire *raison d'être* of a heat engine is to convert heat into work, so it's nice to have a good way to quantify what we mean by “heat” and “work.”

An electric power company provides each customer with a watt-hour meter, which measures the electrical work done. The customer's bill largely depends on the amount of work done, so this is an example of an out-of-state function with tremendous economic significance.

Also note that in pre-modern thermodynamics, “heat” was super-important, because it was part of the recipe for defining entropy, as discussed in section 8.1.3.

### 8.1.6 Heat Content, or Not

In cramped thermodynamics, heat content is a state function. You can heat up a bottle of milk and let it cool down, in such a way that the energy depends on the temperature and the heat capacity, and that's all there is to it.

However, in uncramped thermodynamics, there is no such thing as heat content. You can define  $Q_\Gamma$  as a function of arc length along a specified path, but it is not a state function and it cannot be made into a state function.

More-or-less everybody who studies thermodynamics gets to the point where they try to invent a heat-like quantity that is a function of state, but they never succeed. For more on this, see section 8.2 and chapter 17.

### 8.1.7 Some Mathematical Remarks

If you ever have something that looks like “heat” and you want to differentiate it, refer back to the definition in equation 8.5 and write the derivative as  $TdS$ . That is elegant, because as discussed in section 8.2 and elsewhere,  $T$  is a state function  $S$  is a state function,  $dS$  is a state function, and the product  $TdS$  is a state function. This stands in contrast to  $Q_\Gamma$  which is not a state function, and  $dQ$  which does not even exist.

It is possible to differentiate  $Q_\Gamma$  if you are very very careful. It is almost never worth the trouble – because you are better off with  $TdS$  – but it can be done. As a corollary of equation 8.1, equation 8.4, and equation 8.5, we can write:

$$\frac{d}{d\theta}E(X_\Gamma) = \frac{d}{d\theta}Q_\Gamma - \frac{d}{d\theta}W_\Gamma \quad (8.6)$$

which is of course shorthand for

$$\frac{d}{d\theta}E(X_\Gamma(\theta)) = \frac{d}{d\theta}Q_\Gamma(\theta) - \frac{d}{d\theta}W_\Gamma(\theta) \quad (8.7)$$

Whereas equation 8.1 is a vector equation involving gradient vectors, equation 8.6 is a scalar equation involving total derivatives with respect to  $\theta$ . It can be understood as a projection of equation 8.1, obtained by projecting the gradient vectors onto the direction specified by the path  $\Gamma$ .

On the LHS of equation 8.6, we might be able to consider the energy  $E$  to be a function of state, especially if we can write as  $E(X)$ , which does not depend on the path  $\Gamma$ . However, if we write

it as  $E(X_\Gamma)$ , the energy depends indirectly on the path. In particular, the derivative on the LHS of equation 8.6 very much depends on the path. Considered as a derivative in state space, it is a directional derivative, in the direction specified by the path.

Here is where non-experts go off the rails: In equation 8.6, it is tempting (but wrong) to “simplify” the equation by dropping the “ $d\theta$ ” that appears in every term. Similarly it is tempting (but wrong) to “simplify” the equation by dropping the  $\Gamma$  that appears in every term. The fact of the matter is that these are directional derivatives, and the direction matters a great deal. If you leave out the direction-specifiers, the derivatives simply do not exist.

## 8.2 Grady and Ungrady One-Forms

Sometimes people who are trying to write equation 8.6 instead write something like:

$$dE = dQ - dW \quad (\text{allegedly}) \quad (8.8)$$

which is deplorable.

Using the language of differential forms, the situation can be understood as follows:

- $E$  is a scalar state-function.
- $V$  is a scalar state-function.
- $S$  is a scalar state-function.
- $P$  is a scalar state-function.
- $T$  is a scalar state-function.
- $\Delta E := E_2 - E_1$  is a scalar function of *two* states.
- $\Delta S := S_2 - S_1$  is a scalar function of *two* states.
- $\Delta V := V_2 - V_1$  is a scalar function of *two* states.
- $dE$  is a grady one-form state-function.
- $dS$  is a grady one-form state-function.
- $dV$  is a grady one-form state-function.
- $w := PdV$  is in general an ungrady one-form state-function.
- $q := TdS$  is in general an ungrady one-form state-function.
- There is in general no state-function  $W$  such that  $w = dW$ .
- There is in general no state-function  $Q$  such that  $q = dQ$ .

where in the last four items, we have to say “in general” because exceptions can occur in peculiar situations, mainly cramped situations where it is not possible to construct a heat engine. Such situations are very unlike the general case, and not worth much discussion beyond what was said in conjunction with equation 7.37. When we say something is a state-function we mean it is a function of the thermodynamic state. The last two items follow immediately from the definition of grady versus ungrady.

Beware that in one dimension,  $\Delta S$  is rather similar to  $dS$ . However, that’s not true in higher dimensions ... and uncramped thermodynamics is intrinsically multi-dimensional. If you have experience in one dimension, do not let it mislead you. Recognize the fact that  $\Delta S$  is a scalar, whereas  $dS$  is a vector.

Figure 8.4 shows the difference between a grady one-form and an ungrady one-form.

As you can see in on the left side of the figure, the quantity  $dS$  is grady. If you integrate clockwise around the loop as shown, the net number of upward steps is zero. This is related to the fact that we can assign an unambiguous height ( $S$ ) to each point in  $(T, S)$  space.

In contrast, as you can see on the right side of the diagram, the quantity  $TdS$  is not grady. If you integrate clockwise around the loop as shown, there are considerably more upward steps than downward steps. There is no hope of assigning a height “ $Q$ ” to points in  $(T, S)$  space.

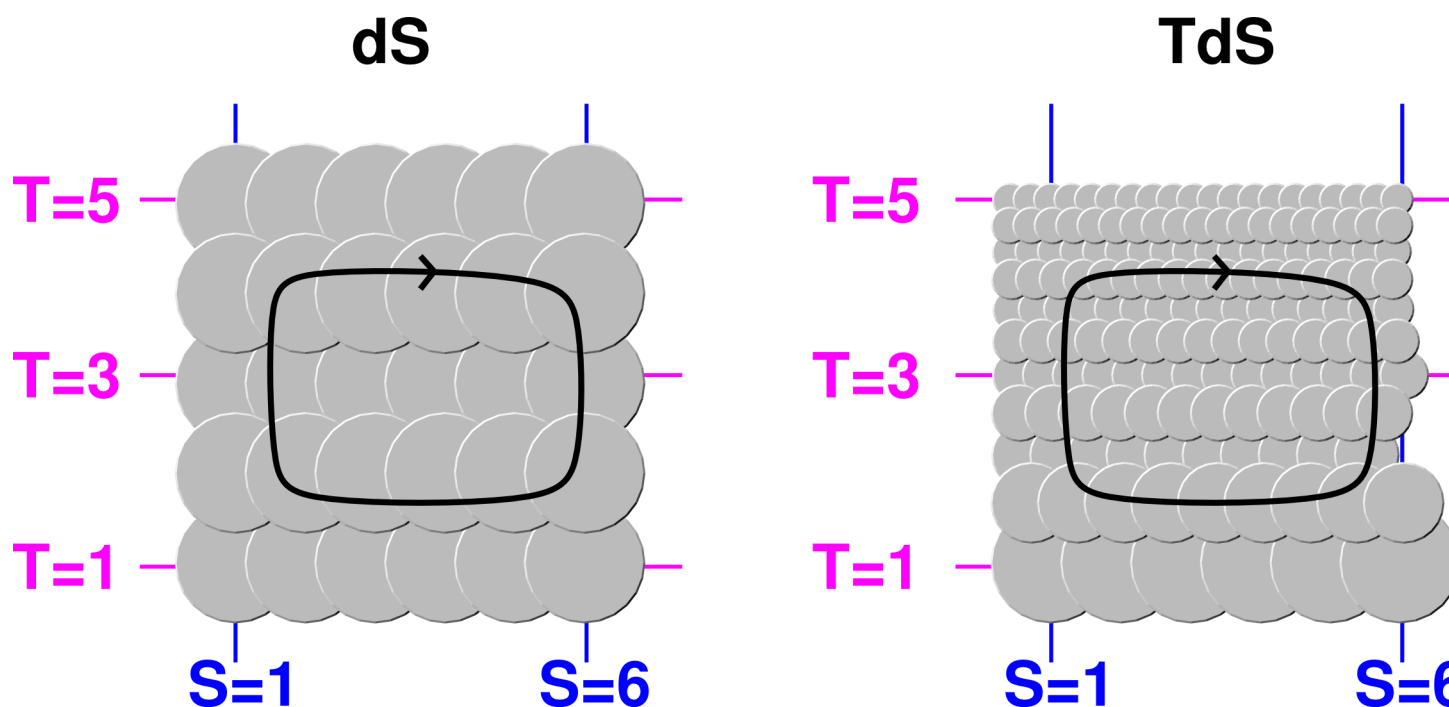


Figure 8.4:  $dS$  is Grady,  $TdS$  is Not

Some additional diagrams showing the relationship between certain grady and ungrady one-forms, see section 15.2.2. For details on the properties of one-forms in general, see reference 4 and perhaps reference 23.

Be warned that in the mathematical literature, what we are calling ungrady one-forms are called “inexact” one-forms. The two terms are entirely synonymous. A one-form is called “exact” if and only if it is the gradient of something. We avoid the terms “exact” and “inexact” because they are too easily misunderstood. In particular, in this context,

- exact is not even remotely the same as accurate.
- inexact is not even remotely the same as inaccurate.
- inexact does not mean “plus or minus something.”

- exact just means grady. An exact one-form is the gradient of some potential.

The difference between grady and ungrady has important consequences for practical situations such as heat engines. Even if we restrict attention to reversible situations, we still cannot think of  $Q$  as a function of state, for the following reasons: You can define any number of functions  $Q_1, Q_2, \dots$  by integrating  $TdS$  along some paths  $\Gamma_1, \Gamma_2, \dots$  of your choosing. Each such  $Q_i$  can be interpreted as the total heat that has flowed into the system *along the specified path*. As an example, let's choose  $\Gamma_6$  to be the path that a heat engine follows as it goes around a complete cycle – a reversible cycle, perhaps something like a Carnot cycle, as discussed in section 8.7. Let  $Q_6(N)$  be the value of  $Q_6$  at the end of the  $N$ th cycle. We see that even after specifying the path,  $Q_6$  is still not a state function, because at the end of each cycle, all the state functions return to their initial values, whereas  $Q_6(N)$  grows linearly with  $N$ . This proves that in any situation where you can build a heat engine,  $q$  is not equal to  $d(\text{anything})$ .

### 8.3 Abuse of the Notation

Suppose there are two people, namely WAYNE and DWAYNE. There is no special relationship between them. In particular, we interpret DWAYNE as a simple six-letter name, not as  $D(\text{WAYNE})$  i.e. not as the derivative of WAYNE.

Some people try to use the same approach to supposedly define  $dQ$  to be a “two-letter name” that represents  $TdS$  – supposedly without implying that  $dQ$  is the derivative of anything. That is emphatically not acceptable. That would be a terrible abuse of the notation.

In accordance with almost-universally accepted convention,  $d$  is an operator, and  $dQ$  denotes the operator  $d$  applied to the variable  $Q$ . If you give it any other interpretation, you are going to confuse yourself and everybody else.

The point remains that in thermodynamics, there does not exist any  $Q$  such that  $dQ = TdS$  (except perhaps in trivial cases). Wishing for such a  $Q$  does not make it so. See chapter 19 for more on this.

### 8.4 Procedure for Extirpating $dW$ and $dQ$

Constructive suggestion: If you are reading a book that uses  $dW$  and  $dQ$ , you can repair it using the following simple procedure:

- For reversible processes, it's easy: Every time you see  $dQ$ , cross it out and write  $TdS$  instead. Every time you see  $dW$ , cross it out and write  $PdV$  or  $-PdV$  instead. The choice of sign depends on convention. It should be easy to determine which convention the book is using.
- For irreversible processes, much more effort is required. Classical thermodynamics books like to say that for an irreversible process « $TdS$  is greater than  $dQ$ ». In this case, you can't simply



replace  $\mathbf{d}Q$  by  $T \mathbf{d}S$  because  $\mathbf{d}Q$  (to the extent that it means anything at all) sometimes does not account for the entire  $T \mathbf{d}S$ . In this context, it probably involves only the entropy that flowed in across the boundary – not the entropy that was created from scratch. So the rule in this context is to cross out  $\mathbf{d}Q$  and replace it by  $T \mathbf{d}S_{\text{transferred}}$ .

As for the idea that  $T \mathbf{d}S > T \mathbf{d}S_{\text{transferred}}$  for an irreversible process, we cannot accept that at face value. For one thing, we would have problems at negative temperatures. We can fix that by getting rid of the  $T$  on both sides of the equation. Another problem is that according to the modern interpretation of the symbols,  $\mathbf{d}S$  is a vector, and it is not possible to define a “greater-than” relation involving vectors. That is to say, vectors are not well ordered. We can fix this by integrating. The relevant equation is:

$$\begin{aligned} \int_{\Gamma} \mathbf{d}S &= \int_{\Gamma} (\mathbf{d}S_{\text{transferred}} + \mathbf{d}S_{\text{created}}) \\ &> \int_{\Gamma} \mathbf{d}S_{\text{transferred}} \end{aligned} \quad (8.9)$$

for some definite path  $\Gamma$ . We need  $\Gamma$  to specify the “forward” direction of the transformation; otherwise the inequality wouldn’t mean anything. We have an inequality, not an equality, because we are considering an irreversible process.

At the end of the day, we find that the assertion that « $T \mathbf{d}S$  is greater than  $\mathbf{d}Q$ » is just a complicated and defective way of saying that the irreversible process created some entropy from scratch.

Note: The underlying idea is that for an irreversible process, entropy is not conserved, so we don’t have *continuity of flow*. Therefore the classical approach was a bad idea to begin with, because it tried to define entropy in terms of heat divided by temperature, and tried to define heat in terms of *flow*. That was a bad idea on practical grounds and pedagogical grounds, in the case where entropy is being created from scratch rather than flowing. It was a bad idea on conceptual grounds, even before it was expressed using symbols such as  $\mathbf{d}Q$  that don’t make sense on mathematical grounds.

Beware: The classical thermo books are inconsistent. Even within a single book, even within a single chapter, sometimes they use  $\mathbf{d}Q$  to mean the entire  $T \mathbf{d}S$  and sometimes only the  $T \mathbf{d}S_{\text{transferred}}$ .

## 8.5 Some Reasons Why $\mathbf{d}W$ and $\mathbf{d}Q$ Might Be Tempting

It is remarkable that people are fond of writing things like  $\mathbf{d}Q \dots$  even in cases where it does not exist. (The remarks in this section apply equally well to  $\mathbf{d}W$  and similar monstrosities.)

Even people who know it is wrong do it anyway. They call  $\mathbf{d}Q$  an “inexact differential” and sometimes put a slash through the  $\mathbf{d}$  to call attention to this. The problem is, neither  $\mathbf{d}Q$  nor  $\mathbf{d}Q$  is a differential at all. Yes,  $T \mathbf{d}S$  is an ungrady one-form or (equivalently) an inexact one-form, but no, it is not properly called an inexact differential, since it is generally not a differential at all. It is not the derivative of anything.

One wonders how such a bizarre tangle of contradictions could arise, and how it could persist. I hypothesize part of the problem is a too-narrow interpretation of the traditional notation for

integrals. Most mathematics books say that every integral should be written in the form

$$\int (\text{integrand}) \mathbf{d}(\text{something}) \quad (8.10)$$

where the  $\mathbf{d}$  is alleged to be merely part of the notation – an obligatory and purely mechanical part of the notation – and the integrand is considered to be separate from the  $\mathbf{d}(\text{something})$ .

However, it doesn't have to be that way. If you think about a simple scalar integral from the Lebesgue point of view (as opposed to the Riemann point of view), you realize that what is indispensable is a *weighting function*. Specifically:  $\mathbf{d}(\text{something})$  is a perfectly fine, normal type of weighting function, but not the only possible type of weighting function.

In an ordinary one-dimensional integral, we are integrating along a path, which in the simplest case is just an interval on the number line. Each element of the path is a little pointy vector, and the weighing function needs to map that pointy vector to a number. Any one-form will do, grady or otherwise. The grady one-forms can be written as  $\mathbf{d}(\text{something})$ , while the ungrady ones cannot.

For purposes of discussion, in the rest of this section we will put square brackets around the weighting function, to make it easy to recognize even if it takes a somewhat unfamiliar form. As a simple example, a typical integral can be written as:

$$\int_{\Gamma} (\text{integrand}) [(\text{weight})] \quad (8.11)$$

where  $\Gamma$  is the domain to be integrated over, and the weight is typically something like  $\mathbf{d}x$ .

As a more intricate example, in two dimensions the moment of inertia of an object  $\Omega$  is:

$$I := \int_{\Omega} r^2 [\mathbf{d}m] \quad (8.12)$$

where the weight is  $\mathbf{d}m$ . As usual,  $r$  denotes distance and  $m$  denotes mass. The integral runs over all elements of the object, and we can think of  $\mathbf{d}m$  as an operator that tells us the mass of each such element. To my way of thinking, this is the definition of moment of inertia: a sum of  $r^2$ , summed over all elements of mass in the object.

The previous expression can be expanded as:

$$I = \int_{\Omega} r^2 [\rho(x, y) \mathbf{d}x \mathbf{d}y] \quad (8.13)$$

where the weighting function is same as before, just rewritten in terms of the density,  $\rho$ .

Things begin to get interesting if we rewrite that as:

$$I = \int_{\Omega} r^2 \rho(x, y) [\mathbf{d}x \mathbf{d}y] \quad (8.14)$$

where  $\rho$  is no longer part of the weight but has become part of the integrand. We see that the distinction between the integrand and the weight is becoming a bit vague. Exploiting this vagueness in the other direction, we can write:

$$\begin{aligned} I &= \int_{\Omega} [r^2 \mathbf{d}m] \\ &= \int_{\Omega} [r^2 \rho(x, y) \mathbf{d}x \mathbf{d}y] \end{aligned} \quad (8.15)$$

which tells us that the distinction between integrand and weighting function is completely meaningless. Henceforth I will treat everything inside the integral on the same footing. The integrand and weight together will be called the *argument*<sup>2</sup> of the integral.

Using an example from thermodynamics, we can write

$$\begin{aligned} Q_\Gamma &= \int_\Gamma T[\mathbf{d}S] \\ &= \int_\Gamma [T \mathbf{d}S] \\ &= \int_\Gamma [q] \end{aligned} \tag{8.16}$$

where  $\Gamma$  is some path through thermodynamic state-space, and where  $q$  is an ungrady one-form, defined as  $q := T\mathbf{d}S$ .

It must be emphasized that these integrals *must not* be written as  $\int[\mathbf{d}Q]$  nor as  $\int[\mathbf{d}q]$ . This is because the argument in equation 8.16 is an ungrady one-form, and therefore cannot be equal to  $\mathbf{d}(\text{anything})$ .

There is no problem with using  $T\mathbf{d}S$  as the weighting function in an integral. The only problem comes when you try to write  $T\mathbf{d}S$  as  $\mathbf{d}(\text{something})$  or  $\mathbf{d}(\text{something})$ :

- Yes,  $T\mathbf{d}S$  is a weighting function.
- Yes, it is a one-form.
- No, it is not a grady one-form.
- No, it is not  $\mathbf{d}(\text{anything})$ .

I realize an expression like  $\int[q]$  will come as a shock to some people, but I think it expresses the correct ideas. It's a whole lot more expressive and more correct than trying to write  $T\mathbf{d}S$  as  $\mathbf{d}(\text{something})$  or  $\mathbf{d}(\text{something})$ .

Once you understand the ideas, the square brackets used in this section no longer serve any important purpose. Feel free to omit them if you wish.

There is a proverb that says if the only tool you have is a hammer, everything begins to look like a nail. The point is that even though a hammer is the ideal tool for pounding nails, it is suboptimal for many other purposes. Analogously, the traditional notation  $\int \cdots \mathbf{d}x$  is ideal for some purposes, but not for all. Specifically: sometimes it is OK to have no explicit  $\mathbf{d}$  inside the integral.

There are only two things that are required: the integral must have a domain to be integrated over, and it must have some sort of argument. The argument must be an operator, which operates on an element of the domain to produce something (usually a number or a vector) that can be summed by the integral.

A one-form certainly suffices to serve as an argument (when elements of the domain are pointy vectors). Indeed, some math books introduce the notion of one-forms by *defining* them to be operators of the sort we need. That is, the space of one-forms is defined as an operator space, consisting of the operators that map column vectors to scalars. (So once again we see that one-forms correspond to row vectors, assuming pointy vectors correspond to column vectors). Using these operators does not require taking a dot product. (You don't need a dot product unless you

---

<sup>2</sup>This corresponds to saying that  $\theta$  is the argument of the cosine in the expression  $\cos(\theta)$ .

want to multiply two column vectors.) The operation of applying a row vector to a column vector to produce a scalar is called a *contraction*, not a dot product.

It is interesting to note that an ordinary summation of the form  $\sum_i F_i$  corresponds exactly to a Lebesgue integral using a measure that assigns unit weight to each integer ( $i$ ) in the domain. No explicit  $\mathbf{d}$  is needed when doing this “integral.” The idea of “weighting function” is closely analogous to the idea of “measure” in Lebesgue integrals, but not exactly the same. We must resist the temptation to use the two terms interchangeably. In particular, a measure is by definition a scalar, but sometimes (such as when integrating along a curve) it is important to use a weighting function that is a vector.

People heretofore have interpreted  $\mathbf{d}$  in several ways: as a differential operator (with the power, among other things, to produce one-forms from scalars), as an infinitesimal step in some direction, and as the marker for the weighting function in an integral. The more I think about it, the more convinced I am that the differential operator interpretation is far and away the most advantageous. The other interpretations of  $\mathbf{d}$  can be seen as mere approximations of the operator interpretation. The approximations work OK in elementary situations, but produce profound misconceptions and contradictions when applied to more general situations ... such as thermodynamics.

In contrast, note that in section 17.1, I do not take such a hard line about the multiple incompatible definitions of heat. I don’t label any of them as right or wrong. Rather, I recognize that each of them *in isolation* has some merit, and it is only when you put them together that conflicts arise.

Bottom line: There are two really simple ideas here: (1)  $\mathbf{d}$  always means exterior derivative. The exterior derivative of any scalar-valued function is a vector. It is a one-form, not a pointy vector. In particular it is always a grady one-form. (2) An integral needs to have a weighting function, which is not necessarily of the form  $\mathbf{d}(\text{something})$ .

## 8.6 Boundary versus Interior

We now discuss two related notions:

- The flow of something across the boundary of the region.
- The change in the amount of something inside the region.

When we consider a conserved quantity such as energy, momentum, or charge, these two notions stand in a one-to-one relationship. In general, though, these two notions are not equivalent.

In particular, consider equation 7.39, which is restated here:

$$\mathbf{d}E = -PdV + T\mathbf{d}S + \text{advection} \quad (8.17)$$

Although officially  $\mathbf{d}E$  represents the change in energy in the interior of the region, we are free to interpret it as the flow of energy across the boundary. This works because  $E$  is a conserved quantity.

The advection term is explicitly a boundary-flow term.

It is extremely tempting to interpret the two remaining terms as boundary-flow terms also ... but this is not correct!

Officially  $PdV$  describes a property of the interior of the region. Ditto for  $TdS$ . Neither of these can be converted to a boundary-flow notion, because neither of them represents a conserved quantity. In particular,  $PdV$  energy can turn into  $TdS$  energy entirely within the interior of the region, without any boundary being involved.

Let's be clear: boundary-flow ideas are elegant, powerful, and widely useful. Please don't think I am saying anything bad about boundary-flow ideas. I am just saying that the  $PdV$  and  $TdS$  terms *do not represent* flows across a boundary.

Misinterpreting  $TdS$  as a boundary term is a ghastly mistake. It is more-or-less tantamount to assuming that heat is a conserved quantity unto itself. It would set science back over 200 years, back to the “caloric” theory.

Once these mistakes have been pointed out, they seem obvious, easy to spot, and easy to avoid. But beware: mistakes of this type are extremely prevalent in introductory-level thermodynamics books.

## 8.7 The Carnot Cycle

A Carnot cycle is not the only possible thermodynamic cycle, or even the only reversible cycle, but it does have special historical and pedagogical significance. Because it is easy to analyze, it gets more attention than it really deserves, which falls into the category of “looking under the lamp-post.” The Carnot cycle involves only relatively simple operations, namely *isothermal* expansion and contraction and *thermally isolated* expansion and contraction. It is fairly easy to imagine carrying out such operations without introducing too much dissipation.

The topic of “Carnot cycle” is only tangentially related to the topic of “Carnot efficiency.” The topics are related insofar as the efficiency of the Carnot cycle is particularly easy to calculate. However, the definition of efficiency applies to a far broader class of cycles, as discussed in section 6.5.

In this section, we do a preliminary analysis of the Carnot cycle. We carry out the analysis using two slightly different viewpoints, running both analyses in parallel. In figure 8.5 and figure 8.6 the physics is the same; only the presentation is different. The primary difference is the choice of axes:  $(V, T)$  versus  $(P, T)$ . Every other difference can be seen as a consequence of this choice.

In figure 8.5:

When the engine is doing work we go clockwise around the cycle.

In figure 8.6:

When the engine is doing work we go counter-clockwise around the cycle.

In other words, the sequence for positive work being done by the engine is this: red, black, magenta, blue.

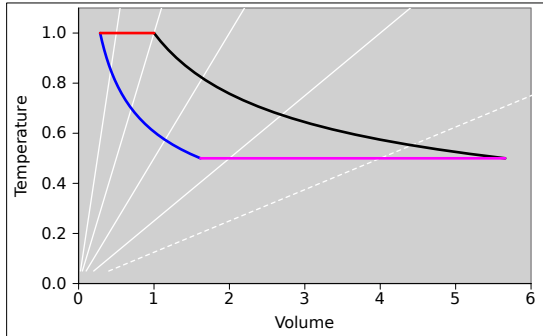


Figure 8.5: Carnot Cycle : T versus V

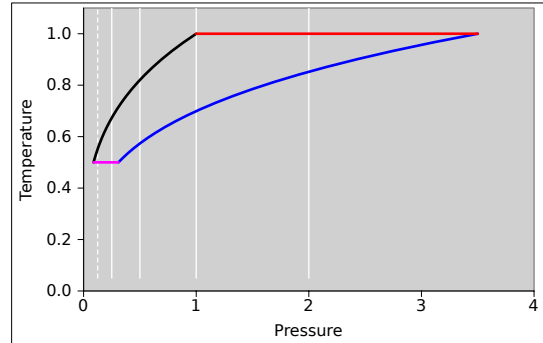


Figure 8.6: Carnot Cycle : T versus P

In all of these figures, the contours of constant pressure are shown in white. The pressure values form a geometric sequence:  $\{1/8, 1/4, 1/2, 1, 2\}$ . In figure 8.6 this is obvious, but in the other figures you will just have to remember it. The  $P = 1/8$  contour is shown as a dotted line, to make it easy to recognize. The  $P = 1$  contour goes through the corner where red meets black.

There are at least a dozen different ways of plotting this sort of data. Another version is shown in figure 8.7. It is somewhat remarkable that multiple figures portraying the same physics would look so different at first glance. I diagrammed the situation in multiple ways partly to show the difference, and partly to make – again – the point that any of the thermodynamic state-functions can be drawn as a function of any reasonable set of variables.

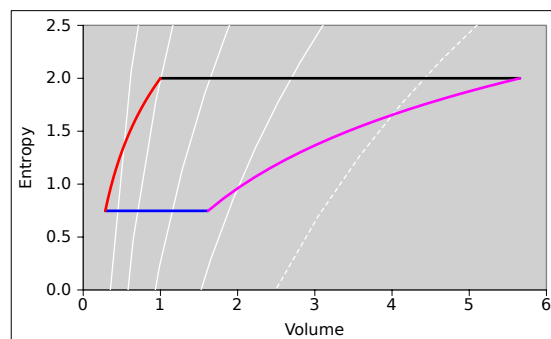


Figure 8.7: Carnot Cycle : S versus V

The spreadsheet used to compute these diagrams is cited in reference 24.

It's interesting that we can replot the data in such a way as to change the apparent shape of the diagram ... without changing the meaning. This clarifies something that was mentioned in goal 5 in section 0.3: In thermodynamic state-space, there is a topology but not a geometry. We can measure the distance (in units of  $S$ ) between contours of constant  $S$ , but we cannot compare that to any distance in any other direction.

In thermodynamics, there are often more variables than one can easily keep track of. Here we focus attention on  $T$ ,  $V$ ,  $P$ , and  $S$ . There are plenty of other variables (such as  $E$ ) that we will barely

even mention. Note that the number of variables far exceeds the number of dimensions. We say this is a two-dimensional situation, because it can be projected onto a two-dimensional diagram without losing any information.

It takes some practice to get the hang of interpreting these diagrams.

- In figure 8.5:
  - Contours of constant temperature are horizontal lines. Examples include the red and magenta path-segments.
  - Contours of constant volume are vertical lines. These are not explicitly shown, but you can easily figure out what they would look like.
  - Contours of constant pressure are sloping lines, shown in white.
  - Contours of constant entropy are curves. Examples include the blue and black path-segments.
- In figure 8.6:
  - Contours of constant temperature are horizontal lines. Examples include the red and magenta path-segments.
  - Contours of constant volume are not shown. (They would be sloping lines, forming a starburst pattern similar to the white lines in figure 8.5.)
  - Contours of constant pressure are vertical lines, shown in white.
  - Contours of constant entropy are curves. Examples include the blue and black path-segments.
- In figure 8.7:
  - Contours of constant temperature are curves. Examples include the red and magenta path-segments.
  - Contours of constant volume are vertical lines. These are not explicitly shown, but you can easily figure out what they would look like.
  - Contours of constant pressure are curves, shown in white.
  - Contours of constant entropy are horizontal lines. Examples include the blue and black path-segments.

The Carnot cycle has four phases.

1. Isothermal expansion: We connect the cylinder to the high-temperature heat bath. We gently expand the gas at constant high temperature. This is shown in red in the figures.

The amount of expansion is a design choice. In the example I arbitrarily chose a volume ratio of 3.5:1 for this phase of the expansion. (This is not the so-called “compression ratio” of the overall engine, since we have so far considered only one of the two compression phases.)

2. Isentropic expansion: We disconnect the cylinder. We expand the gas some more, gently, at constant entropy. This is shown in black in the figures.

We continue the expansion until the temperature of the gas reaches the temperature of the low-temperature heat bath.

The amount of expansion required to achieve this depends on the gamma of the gas and (obviously) on the ratio of temperatures. For the example I assumed a 2:1 temperature ratio, which calls for a 5.7:1 expansion ratio during this phase. In general, for an adiabatic expansion, if you know the temperature ratio you can calculate the expansion ratio for this phase via

$$\frac{V_1}{V_2} = \left( \frac{T_1}{T_2} \right)^{-1/(\gamma-1)} \quad (8.18)$$

For the figures I used the gamma value appropriate for air and similar diatomic gases, namely  $\gamma = 7/5 = 1.4$ . A monatomic gas would need a significantly lower compression ratio for any given temperature ratio.

3. Isothermal compression: We connect the cylinder to the low-temperature heat bath. We gently compress the gas at constant low temperature. This is shown in magenta in the figures.
4. Isentropic compression: We disconnect the cylinder from both heat baths. We compress the gas some more, gently, along a contour of constant entropy. This is shown in blue in the figures.

During the two expansion phases, the gas does work on the flywheel. During the two compression phases, the flywheel needs to do work on the gas. However, the compression-work is strictly less in magnitude than the expansion-work, so during the cycle net energy is deposited in the flywheel.

To understand this at the next level of detail, recall that mechanical work is  $-P dV$ . Now the integral of  $dV$  around the cycle is just  $\Delta V$  which is zero. But what about the integral of  $P dV$ ? Consider for example the step in  $V$  that goes from  $V = 3$  to  $V = 4$  along the black curve, and the corresponding step that goes from  $V = 4$  to  $V = 3$  along the magenta curve. The work done during the expansion phase (black) will be strictly greater than the work done during the recompression phase (magenta) because the pressure is higher. You can make the same argument for every piece of  $V$  in the whole  $V$ -space: For every  $\Delta V$  in one of the expansion phases there will be a corresponding  $\Delta V$  directly below it in one of the compression phases. For each of these pairs, expansion-work will be strictly greater than the compression-work because the pressure will be higher.

You can infer the higher pressure from the white contours of constant pressure, or you can just observe that the pressure must be higher because the temperature is higher.



## Chapter 9

# Connecting Entropy with Energy

### 9.1 The Boltzmann Distribution

For reasons discussed in chapter 23, whenever a system is in thermal equilibrium, the energy is distributed among the microstates according to a very special probability distribution, namely the Boltzmann distribution. That is, the probability of finding the system in microstate  $i$  is given by:

$$P_i = e^{-\hat{E}_i/kT} \quad \dots \quad \text{for a thermal distribution} \quad (9.1)$$

where  $\hat{E}_i$  is the energy of the  $i$ th microstate, and  $kT$  is the temperature measured in energy units. That is, plain  $T$  is the temperature, and  $k$  is Boltzmann's constant, which is just the conversion factor from temperature units to energy units.

Figure 9.1 shows this distribution graphically.

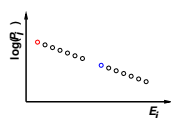


Figure 9.1: An Equilibrium Distribution

Evidence in favor of equation 9.1 is discussed in section 11.2.

### 9.2 Systems with Subsystems

When thinking about equation 9.1 and figure 9.1 it is important to realize that things don't have to be that way. There are other possibilities. Indeed, a theory of thermodynamics that assumed that everything in sight was always in equilibrium at temperature  $T$  would be not worth the trouble. For starters, it is impossible to build a heat engine unless the hot reservoir is not in equilibrium with the cold reservoir.

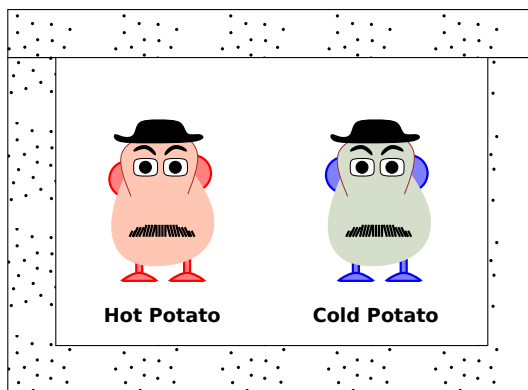


Figure 9.2: A System Consisting of Two Potatoes

We start by considering the system shown in figure 9.2, namely a styrofoam box containing a hot potato and a cold potato. (This is a simplified version of figure 1.2.)

In situations like this, we can make good progress if we divide the system into *subsystems*. Here subsystem *A* is the red potato, and subsystem *B* is the blue potato. Each subsystem has a well defined temperature, but initially the system as a whole does not.

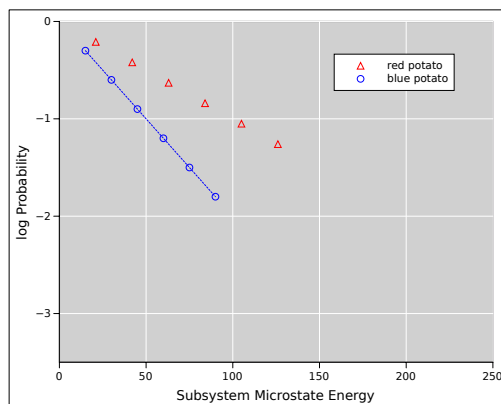


Figure 9.3: Probability versus Energy, Non-Equilibrium

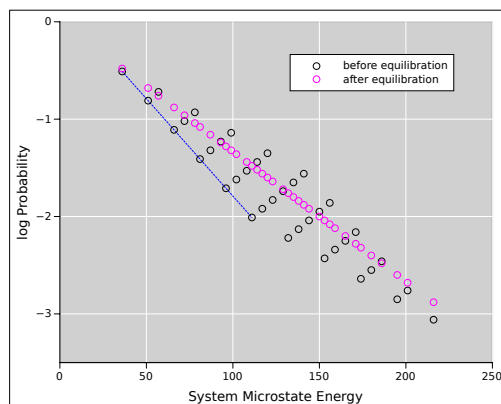


Figure 9.4: Probability versus Energy, System as a Whole

If we wait long enough, the two potatoes will come into equilibrium with each other, and at this point the system as a whole will have a well defined temperature. However, we are not required to wait for this to happen.

## 9.3 Remarks

### 9.3.1 Predictable Energy is Freely Convertible

The difference between random energy and predictable energy has many consequences. The most important consequence is that the predictable energy can be freely converted to and from other

forms, such as gravitational potential energy, chemical energy, electrical energy, et cetera. In many cases, these conversions can be carried out with very high efficiency. In some other cases, though, the laws of thermodynamics place severe restrictions on the efficiency with which conversions can be carried out, depending on to what extent the energy distribution deviates from the Boltzmann distribution.

### 9.3.2 Thermodynamic Laws without Temperature

Ironically, the first law of thermodynamics (equation 1.1) does not depend on temperature. Energy is well-defined and is conserved, no matter what. It doesn't matter whether the system is hot or cold or whether it even has a temperature at all.

Even more ironically, the second law of thermodynamics (equation 2.1) doesn't depend on temperature, either. Entropy is well-defined and is paraconserved no matter what. It doesn't matter whether the system is hot or cold or whether it even has a temperature at all.

(This state of affairs is ironic because thermodynamics is commonly defined to be the science of heat and temperature, as you might have expected from the name: *thermodynamics*. Yet in our modernized and rationalized thermodynamics, the two most central, fundamental ideas – energy and entropy – are defined without reference to heat or temperature.)

Of course there are many important situations that do involve temperature. Most of the common, every-day applications of thermodynamics involve temperature – but you should not think of temperature as the essence of thermodynamics. Rather, it is a secondary concept which is defined (if and when it even exists) in terms of energy and entropy.

### 9.3.3 Kinetic and Potential Microscopic Energy

You may have heard the term “kinetic theory.” In particular, the thermodynamics of ideal gases is commonly called the kinetic theory of gases. However, you should be careful, because “kinetic theory” is restricted to ideal gases (indeed to a subset of ideal gases) ... while thermodynamics applies to innumerable other things. Don't fall into the trap of thinking that there is such a thing as “thermal energy” and that this so-called “thermal energy” is necessarily *kinetic* energy. In almost all systems, including solids, liquids, non-ideal gases, and even some ideal gases, the energy is a mixture of kinetic and potential energy. (Furthermore, in any non-cramped situation, i.e. in any situation where it is possible to build a heat engine, it is impossible in principle to define any such thing as “thermal energy.”) In any case, it is safer and in all ways better to say *thermodynamics* or *statistical mechanics* instead of “kinetic theory.”

In typical systems, potential energy and kinetic energy play parallel roles:

- To visualize microscopic potential energy, imagine that the atoms in a crystal lattice are held in place by springs. Roughly speaking, half of these springs have positive potential energy because they are extended relative to their resting-length, while the other half have positive potential energy because they are compressed relative to their resting-length. They've all got energy, but you can't easily harness it because you don't know which ones are compressed and which ones are extended.

- To visualize microscopic kinetic energy, imagine that half the atoms have a leftward velocity and half have a rightward velocity. They all have kinetic energy, but you can't easily harness it because you don't know which ones are moving leftward and which are moving rightward.

In fact, for an ordinary crystal such as quartz or sodium chloride, almost exactly half of the heat capacity is due to potential energy, and half to kinetic energy. It's easy to see why that must be: The heat capacity is well described in terms of thermal phonons in the crystal. Each phonon mode is a harmonic<sup>1</sup> oscillator. In each cycle of any harmonic oscillator, the energy changes from kinetic to potential and back again. The kinetic energy goes like  $\sin^2(\text{phase})$  and the potential energy goes like  $\cos^2(\text{phase})$ , so on average each of those is half of the total energy.

Not all kinetic energy contributes to the heat capacity.  
Not all of the heat capacity comes from kinetic energy.

A table-top sample of ideal gas is a special case, where *all* the energy is kinetic energy. This is very atypical of thermodynamics in general. Table-top ideal gases are very commonly used as an illustration of thermodynamic ideas, which becomes a problem when the example is overused so heavily as to create the misimpression that thermodynamics deals only with kinetic energy.

You could argue that in many familiar systems, the *temperature* is closely related to random kinetic energy ... but temperature is not the same thing as so-called “heat” or “thermal energy.” Furthermore, there are other systems, such as spin systems, where the temperature is not related to the random kinetic energy.

All in all, it seems quite unwise to define heat or even temperature in terms of kinetic energy.

This discussion continues in section 9.3.4.

### 9.3.4 Ideal Gas : Potential Energy as well as Kinetic Energy

We have seen that for an ideal gas, there is a one-to-one correspondence between the temperature and the kinetic energy of the gas particles. However, that does *not* mean that there is a one-to-one correspondence between kinetic energy and heat energy. (In this context, heat energy refers to whatever is measured by a heat capacity experiment.)

To illustrate this point, let's consider a sample of pure monatomic nonrelativistic nondegenerate ideal gas in a tall cylinder of horizontal radius  $r$  and vertical height  $h$  at temperature  $T$ . The pressure measured at the bottom of the cylinder is  $P$ . Each particle in the gas has mass  $m$ . We wish to know the heat capacity per particle at constant volume, i.e.  $C_V/N$ .

At this point you may already have in mind an answer, a simple answer, a well-known answer, independent of  $r$ ,  $h$ ,  $m$ ,  $P$ ,  $T$ , and  $N$ . But wait, there's more to the story: The point of this exercise is that  $h$  is not small. In particular,  $m|g|h$  is not small compared to  $kT$ , where  $g$  is the acceleration of gravity. For simplicity, you are encouraged to start by considering the limit where  $h$  goes to infinity, in which case the exact value of  $h$  no longer matters. Gravity holds virtually all the gas near the bottom of the cylinder, whenever  $h \gg kT/m|g|$ .

---

<sup>1</sup>Anharmonicity can cause the average KE to be not exactly equal to the average PE, but for a crystal well below its melting point, the thermal phonon modes are not significantly anharmonic.

- Later, if you want to come back and work the problem a second time, with no restrictions on  $h$ , that's interesting and not very difficult.
- Also if you want to generalize to a polyatomic gas, that's also worth doing. Section 26.4 continues the discussion of the energy of an ideal gas, including equation 26.46 and equation 26.47 which related energy to pressure, volume, and temperature.

You will discover that a distinctly nontrivial contribution to the heat capacity comes from the potential energy of the ideal gas. When you heat it up, the gas column expands, lifting its center of mass, doing work against gravity. (Of course, as always, there will be a contribution from the kinetic energy.)

For particles the size of atoms, the length-scale  $kT/m|g|$  is on the order of several kilometers, so the cylinder we are considering is much too big to fit on a table top. I often use the restrictive term “*table-top*” as a shorthand way of asserting that  $m|g|h$  is small compared to  $kT$ .

So, this reinforces the points made in section 9.3.3. We conclude that in general, heat energy is not just kinetic energy.

Beware that this tall cylinder is not a good model for the earth's atmosphere. For one thing, the atmosphere is not isothermal. For another thing, if you are going to take the limit as  $h$  goes to infinity, you can't use a cylinder; you need something more like a cone, spreading out as it goes up, to account for the spherical geometry.

### 9.3.5 Relative Motion versus “Thermal” Energy

Over the years, lots of people have noticed that you can always split the kinetic energy of a complex object into the KE of the center-of-mass motion plus the KE of the relative motion (i.e. the motion of the components relative to the center of mass).

Also a lot of people have tried (with mixed success) to split the energy of an object into a “thermal” piece and a “non-thermal” piece.

It is an all-too-common mistake to think that the overall/relative split is the same as the non-thermal/thermal split. Beware: they're not the same. Definitely not. See section 7.7 for more on this.

First of all, the microscopic energy is not restricted to being kinetic energy, as discussed in section 9.3.3. So trying to understand the thermal/non-thermal split in terms of kinetic energy is guaranteed to fail. Using the work/KE theorem (reference 18) to connect work (via KE) to the thermal/nonthermal split is guaranteed to fail for the same reason.

Secondly, a standard counterexample uses flywheels, as discussed in section 18.4. You can impart macroscopic, non-Locrian KE to the flywheels without imparting center-of-mass KE or any kind of potential energy ... and without imparting any kind of Locrian energy (either kinetic or potential).

The whole idea of “thermal energy” is problematic, and in many cases impossible to define, as discussed in chapter 19. If you find yourself worrying about the exact definition of “thermal energy”;

it means you're trying to solve the wrong problem. Find a way to reformulate the problem in terms of energy and entropy.

Center-of-mass motion is an example *but not the only example* of low-entropy energy. The motion of the flywheels is one perfectly good example of low-entropy energy. Several other examples are listed in section 11.3.

A macroscopic object has something like  $10^{23}$  modes. The center-of-mass motion is just one of these modes. The motion of counter-rotating flywheels is another mode. These are slightly special, but not very special. A mode to which we can apply a conservation law, such as conservation of momentum, or conservation of angular momentum, might require a little bit of special treatment, but usually not much ... and there aren't very many such modes.

Sometimes on account of conservation laws, and sometimes for other reasons as discussed in section 11.11 it may be possible for a few modes of the system to be strongly coupled to the outside (and weakly coupled to the rest of the system), while the remaining  $10^{23}$  modes are more strongly coupled to each other than they are to the outside. It is these issues of coupling-strength that determine which modes are in equilibrium and which (if any) are far from equilibrium. This is consistent with our definition of equilibrium (section 10.1).

Thermodynamics treats all the equilibrated modes on an equal footing. One manifestation of this can be seen in equation 9.1, where each state contributes one term to the sum ... and addition is commutative.

There will never be an axiom that says such-and-such mode is always in equilibrium or always not; the answer is sensitive to how you engineer the couplings.

## 9.4 Entropy Without Constant Re-Shuffling

It is a common mistake to visualize entropy as a highly dynamic process, whereby the system is constantly flipping from one microstate to another. This may be a consequence of the fallacy discussed in section 9.3.5 (mistaking the thermal/nonthermal distinction for the kinetic/potential distinction) ... or it may have other roots; I'm not sure.

In any case, the fact is that re-shuffling is not an essential part of the entropy picture.

An understanding of this point proceeds directly from fundamental notions of probability and statistics.

By way of illustration, consider one hand in a game of draw poker.

- A) The deck is shuffled and hands are dealt in the usual way.
- B) In preparation for the first round of betting, you look at your hand and discover that you've got the infamous "inside straight." Other players raise the stakes, and when it's your turn to bet you drop out, saying to yourself "if this had been an outside straight the probability would have been twice as favorable"

- C) The other players, curiously enough, stand pat, and after the hand is over you get a chance to flip through the deck and see the card you would have drawn.

Let's more closely examine step (B). At this point you have to make a decision based on probability. The deck, as it sits there, is not constantly re-arranging itself, yet you are somehow able to think about the probability that the card you draw will complete your inside straight.

The deck, as it sits there during step (B), is not flipping from one microstate to another. It is in *some* microstate, and staying in that microstate. At this stage you don't know what microstate that happens to be. Later, at step (C), long after the hand is over, you might get a chance to find out the exact microstate, but right now at step (B) you are forced to make a decision based only on the probability.

The same ideas apply to the entropy of a roomful of air, or any other thermodynamic system. At any given instant, the air is in *some* microstate with 100% probability; you just don't know what microstate that happens to be. If you did know, the entropy would be zero ... but you don't know. You don't need to take any sort of time-average to realize that you don't know the microstate.

The bottom line is that the essence of entropy is the same as the essence of probability in general: The essential idea is that you don't know the microstate. Constant re-arrangement is not essential.

This leaves us with the question of whether re-arrangement is *ever* important. Of course the deck needs to be shuffled at step (A). Not constantly re-shuffled, just shuffled the once.

Again, the same ideas apply to the entropy of a roomful of air. If you did somehow obtain knowledge of the microstate, you might be interested in the timescale over which the system re-arranges itself, making your erstwhile knowledge obsolete and thereby returning the system to a high-entropy condition.

The crucial point remains: the *process* whereby knowledge is lost and entropy is created is not part of the definition of entropy, and need not be considered when you evaluate the entropy. If you walk into a room for the first time, the re-arrangement rate is not your concern. You don't know the microstate of this room, and that's all there is to the story. You don't care how quickly (if at all) one unknown microstate turns into another.

If you don't like the poker analogy, we can use a cryptology analogy instead. Yes, physics, poker, and cryptology are all the same when it comes to this. Statistics is statistics.

If I've intercepted just one cryptotext from the opposition and I'm trying to crack it, on some level what matters is whether or not I know their session key. It doesn't matter whether that session key is 10 microseconds old, or 10 minutes old, or 10 days old. If I don't have any information about it, I don't have any information about it, and that's all that need be said.

On the other hand, if I've intercepted a stream of messages and extracted partial information from them (via a partial break of the cryptosystem), the opposition would be well advised to "re-shuffle the deck" i.e. choose new session keys on a timescale fast compared to my ability to extract information about them.

Applying these ideas to a roomful of air: Typical sorts of measurements give us only a pathetically small amount of partial information about the microstate. So it really doesn't matter whether the



air re-arranges itself super-frequently or super-infrequently. We don't have any significant amount of information about the microstate, and that's all there is to the story.

Reference 25 presents a simulation that demonstrates the points discussed in this subsection.

## 9.5 Units of Entropy

Before we go any farther, convince yourself that

$$\begin{aligned}\log_{10}(x) &= \frac{\ln(x)}{\ln(10)} \\ &\approx 0.434294 \ln(x)\end{aligned}\tag{9.2}$$

and in general, multiplying a logarithm by some positive number corresponds to changing the base of the logarithm.

In the formula for entropy, equation 2.2, the base of the logarithm has intentionally been left unspecified. You get to choose a convenient base. This is the same thing as choosing what units will be used for measuring the entropy.

Some people prefer to express the units by choosing the base of the logarithm, while others prefer to stick with natural logarithms and express the units more directly, using an expression of the form:

$$S[P] := k \sum_i P_i \ln(1/P_i)\tag{9.3}$$

In this expression we stipulated  $e$  as the base of the logarithm. Whereas equation 2.2 we could choose the base of the logarithm, in equation 9.3 we get to choose the numerical value and units for  $k$ . This is a superficially different solution to the same problem. Reasonable choices include:

$$\begin{aligned}k &= 1.380649 \times 10^{-23} && \text{J/K} \\ &= 1/\ln(3) && \text{trit} \\ &= 1 && \text{nat} \\ &= 1/\ln(2) && \text{bit}\end{aligned}\tag{9.4}$$

It must be emphasized that all these expressions are mathematically equivalent. In each case, the numerical part of  $k$  balances the units of  $k$ , so that the meaning remains unchanged. In some cases it is convenient to absorb the numerical part of  $k$  into the base of the logarithm:

$$\begin{aligned}S[P] &= k \sum_i P_i \ln(1/P_i) && \text{(in general)} \\ k &= 1.380649 \times 10^{-23} \sum_i P_i \ln(1/P_i) && \text{J/K (SI)} \\ &= \sum_i P_i \log_{\Omega}(1/P_i) && \text{J/K (SI)} \\ &= \sum_i P_i \log_3(1/P_i) && \text{trit} \\ &= \sum_i P_i \log_e(1/P_i) && \text{nat} \\ &= \sum_i P_i \log_2(1/P_i) && \text{bit}\end{aligned}\tag{9.5}$$

where the third line uses the remarkably huge base  $\Omega = \exp(1.24 \times 10^{22})$ . When dealing with smallish amounts of entropy, units of bits are conventional and often convenient. When dealing with large amounts of entropy, units of J/K are conventional and often convenient. These are related as follows:

$$\begin{aligned}1\text{J/K} &= 1.04 \times 10^{23} \text{bits} \\ 1\text{bit} &= 9.57 \times 10^{-24} \text{J/K}\end{aligned}\tag{9.6}$$

A convenient unit for *molar entropy* is Joules per Kelvin per mole:

$$\begin{aligned} 1\text{J/K/mol} &= 0.17\text{bit/particle} \\ 1\text{bit/particle} &= 5.76\text{J/K/mol} = R \ln(2) \end{aligned} \quad (9.7)$$

Values in this range (on the order of one bit per particle) are very commonly encountered.

If you are wondering whether equation 9.7 is OK from a dimensional-analysis point of view, fear not. Temperature units are closely related to energy units. Specifically, energy is extensive and measured in J, while temperature is intensive and measured in K. Therefore combinations such as (J/K/mol) are dimensionless units. A glance at the dimensions of the ideal gas law should suffice to remind you of this if you ever forget.

See reference 26 for more about dimensionless units.

## 9.6 Probability versus Multiplicity

Let us spend a few paragraphs discussing a *strict* notion of multiplicity, and then move on to a more nuanced notion. (We also discuss the relationship between an equiprobable distribution and a microcanonical ensemble.)

### 9.6.1 Exactly Equiprobable

Suppose we have a system where a certain set of states<sup>2</sup> (called the “accessible” states) are equiprobable, i.e.  $P_i = 1/W$  for some constant  $W$ . Furthermore, all remaining states are “inaccessible” which means they all have  $P_i = 0$ . The constant  $W$  is called the *multiplicity*.

Note: Terminology: The  $W$  denoting multiplicity in this section is unrelated to the  $W$  denoting work elsewhere in this document. Both usages of  $W$  are common in the literature. It is almost always obvious from context which meaning is intended, so there isn’t a serious problem. Some of the literature uses  $\Omega$  to denote multiplicity.

The probability per state is necessarily the reciprocal of the number of accessible states, since (in accordance with the usual definition of “probability”) we want our probabilities to be normalized:  $\sum P_i = 1$ .

In this less-than-general case, the entropy (as given by equation 2.2) reduces to

$$S = \log W \quad (\text{provided the microstates are equiprobable}) \quad (9.8)$$

As usual, you can choose the base of the logarithm according to what units you prefer for measuring entropy: bits, nats, trits, J/K, or whatever. Equivalently, you can fix the base of the logarithm and express the units by means of a factor of  $k$  out front, as discussed in section 9.5:

$$S = k \ln W \quad (\text{provided the microstates are equiprobable}) \quad (9.9)$$

---

<sup>2</sup>Here “states” means “microstates”.

This equation is prominently featured on Boltzmann’s tombstone. However, I’m pretty sure (a) he didn’t put it there, (b) Boltzmann was not the one who originated or emphasized this formula (Planck was), and (c) Boltzmann was well aware that this is not the most general expression for the entropy. I mention this because a lot of people who ought to know better take equation 9.9 as THE unassailable definition of entropy, and sometimes they cite Boltzmann’s tombstone as if it were the ultimate authority on the subject.

In any case, (d) even if Boltzmann had endorsed equation 9.9, appeal to authority is not an acceptable substitute for scientific evidence and logical reasoning. We know more now than we knew in 1898, and we are allowed to change our minds about things ... although in this case it is not necessary. Equation 2.2 has been the faithful workhorse formula for a very long time.

There are various ways a system could wind up with equiprobable states:

- Consider a well-shuffled deck of cards. This is an example of a non-thermal system where the states are equiprobable.
- Consider a thermal system that is thermally isolated so that all accessible microstates have the same energy. This is called a *microcanonical* system. It will have equiprobable states in accordance with the equipartition law, equation 9.1.

Consider two blocks of copper that are identical except that one of them has more energy than the other. They are thermally isolated from each other and from everything else. The higher-energy block will have a greater number of accessible states, i.e. a higher multiplicity. In this way you can, if you wish, define a notion of multiplicity *as a function of energy level*.

On the other hand, you must not get the idea that multiplicity is a monotone function of energy or vice versa. Such an idea would be quite incorrect when applied to a spin system.

**Terminology:** By definition, a *level* is a group of microstates. An energy level is a group of microstates all with the same energy (or nearly the same energy, relative to other energy-scales in the problem). By connotation, usually when people speak of a level they mean energy level.

### 9.6.2 Approximately Equiprobable

We now introduce a notion of “approximate” equiprobability and “approximate” multiplicity by reference to the example in the following table:

Level	# microstates in level	Probability of microstate	Probability of level	Entropy (in bits)
1	2	0.01	0.020	0.133
2	979	0.001	0.989	9.757
3	1,000,000	1E-09	0.001	0.030
<b>Total:</b>	<b>1,000,981</b>		<b>1.000</b>	<b>9.919</b>

The system in this example 1,000,981 microstates, which we have grouped into three levels. There are a million states in level 3, each of which occurs with probability one in a billion, so the probability of observing *some* state from this level is one in a thousand. There are only two microstates in level 1, each of which is observed with a vastly larger probability, namely one in a hundred. Level 2 is baby-bear just right. It has a moderate number of states, each with a moderate probability ... with the remarkable property that on a level-by-level basis, this level dominates the probability distribution. The probability of observing some microstate from level 2 is nearly 100%.

The bottom line is that the entropy of this distribution is 9.919 bits, which is 99.53% of the entropy you would have if all the probability were tied up in 1000 microstates with probability 0.001 each.

Beware of some overloaded terminology:

<p>In the table, the column we have labeled “# microstates in level” is conventionally called the <i>multiplicity</i> of the level.</p>	<p>If we apply the <math>S = \log(W)</math> formula in reverse, we find that our example distribution has a multiplicity of <math>W = 2^S = 2^{9.919} = 968</math>; this is the effective multiplicity of the distribution as a whole.</p>
---	--

So we see that the effective multiplicity of the distribution is dominated by the multiplicity of level 2. The other levels contribute very little to the entropy.

You have to be careful how you describe the microstates in level 2. Level 2 is the most probable level (on a level-by-level basis), but its microstates are not the most probable microstates (on a microstate-by-microstate basis).

In the strict notion of multiplicity, all the states that were not part of the dominant level were declared “inaccessible”; but alas this terminology becomes hopelessly tangled when we progress to the nuanced notion of multiplicity. In the table, the states in level 3 are high-energy states, and it might be OK to say that they are energetically inaccessible, or “almost” inaccessible. It might be superficially tempting to label level 1 as also inaccessible, but that would not be correct. The states in level 1 are perfectly accessible; their only problem is that they are few in number.

I don’t know how to handle “accessibility” except to avoid the term, and to speak instead of “dominant” levels and “negligible” levels.

<p>A system that is thermally isolated so that all microstates have the same energy is called microcanonical.</p>	<p>In contrast, an object in contact with a constant-temperature heat bath is called canonical (not microcanonical). Furthermore, a system that can exchange particles with a reservoir, as described by a chemical potential, is called grand canonical (not microcanonical or canonical).</p>
---	---

<p>The strict definition of multiplicity applies directly to microcanonical ensembles and other strictly equiprobable distributions. Equation 9.8 applies exactly to such systems.</p>	<p>Equation 9.8 does not apply exactly to canonical or grand-canonical systems, and may not apply even approximately. The correct <i>thermal</i> probability distribution is shown in figure 9.1.</p>
--	---

There exist intermediate cases, which are common and often important. In a canonical or grand-canonical thermal system, we can get into a situation where the notion of multiplicity is a good approximation – not exact, but good enough. This can happen if the energy distribution is so strongly peaked near the most-probable energy that the entropy is very nearly what you would get in the strictly-equiprobable case. This can be roughly understood in terms of the behavior of Gaussians. If we combine  $N$  small Gaussians to make one big Gaussian, the absolute width scales like  $\sqrt{N}$  and the relative width scales like  $\sqrt{N}/N$ . The latter is small when  $N$  is large.

One should not attach too much importance to the tradeoff in the table above, namely the tradeoff between multiplicity (increasing as we move down the table) and per-microstate probability (decreasing as we move down the table). It is tempting to assume all thermal systems must involve a similar tradeoff, but they do not. In particular, at negative temperatures (as discussed in reference 27), it is quite possible for the lower-energy microstates to outnumber the higher-energy microstates, so that both multiplicity *and* per-microstate probability are decreasing as we move down the table toward higher energy.

You may reasonably ask whether such a system might be unstable, i.e. whether the entire system might spontaneously move toward the high-energy high-probability high-multiplicity state. The answer is that such a move cannot happen because it would not conserve energy. In a thermally-isolated system, if half of the system moved to higher energy, you would have to “borrow” that energy from the other half, which would then move to lower energy, lower multiplicity, and lower probability per microstate. The overall probability of the system depends on the probability of the two halves taken jointly, and this joint probability would be unfavorable. If you want to get technical about it, stability does not depend on the increase or decrease of multiplicity as a function of energy, but rather on the *convexity* which measures what happens if you borrow energy from one subsystem and lend it to another.

### 9.6.3 Not At All Equiprobable

Consider the probability distribution shown in figure 9.5. There is one microstate with probability  $1/2$ , another with probability  $1/4$ , another with probability  $1/8$ , et cetera. Each microstate is represented by a sector in the diagram, and the area of the sector is proportional to the microstate’s probability.

Some information about these microstates can be found in the following table.

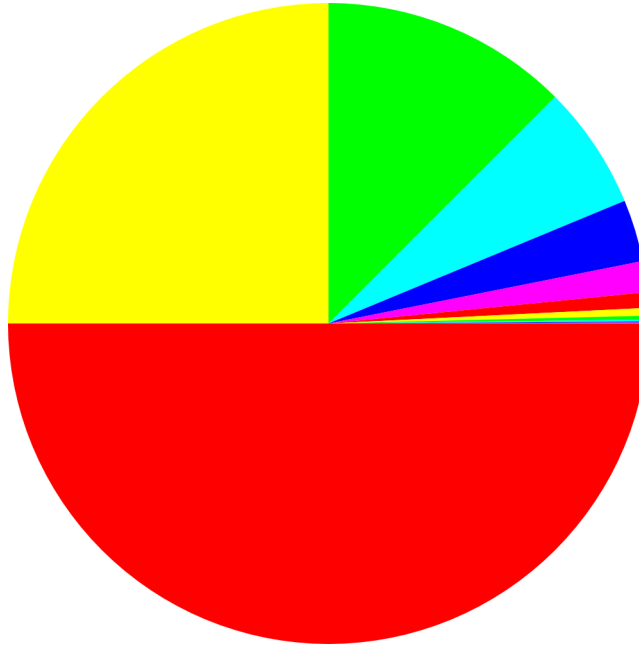


Figure 9.5: Exponential Distribution

State#	Probability	Suprise Value / bits
1	0.5	1
2	0.25	2
3	0.125	3
4	0.0625	4
5	0.03125	5
6	0.015625	6
7	0.0078125	7
8	0.00390625	8
9	0.001953125	9
10	0.0009765625	10
...	et cetera	...

The total probability adds up to 1, as you can verify by summing the numbers in the middle column. The total entropy is 2, as you can verify by summing the surprisals *weighted by the corresponding probabilities*. The total number of states is infinite, and the multiplicity  $W$  is infinite. Note that

$$\begin{aligned}
 S &= 2 \\
 k \ln W &= \infty \\
 S &\neq k \ln W
 \end{aligned}
 \tag{9.10}$$

which means that equation 9.9 definitely fails to work for this distribution. It fails by quite a large margin.

## 9.7 Discussion

Some people are inordinately fond of equation 9.8 or equivalently equation 9.9. They are tempted to take it as the definition of entropy, and sometimes offer outrageously unscientific arguments in its support. But the fact remains that Equation 2.2 is an incomparably more general, more reliable expression, while equation 9.9 is a special case, a less-than-general corollary, a sometimes-acceptable approximation.

Specific reasons why you should *not* consider equation 9.8 to be axiomatic include:

1. Theory says that you cannot exactly reconcile a Boltzmann probability distribution with an equiprobable distribution.
2. In practice, equation 9.8 is usually not an acceptable approximation for small systems. Thermodynamics applies to small systems, but equation 9.8 usually does not.
3. For large systems, even though equation 9.8 commonly leads to valid approximations for first-order quantities (e.g. energy, entropy, temperature, and pressure) ... it does not lead to valid results for second-order quantities such as fluctuations (energy fluctuations, temperature fluctuations, et cetera).

For a thermal distribution, the probability of a microstate is given by equation 9.1. So, even within the restricted realm of thermal distributions, equation 9.9 does not cover all the bases; it applies if and only if all the accessible microstates have the same energy. It is possible to arrange for this to be true, by constraining all accessible microstates to have the same energy. That is, it is possible to create a microcanonical system by isolating or insulating and sealing the system so that no energy can enter or leave. This can be done, but it places drastic restrictions on the sort of systems we can analyze.

- Two of the four phases of the Carnot cycle are carried out at constant temperature, not constant energy. The system is in contact with a heat bath, not isolated or insulated. A theory of “thermodynamics” without heat engines would be pretty lame.
- A great many chemistry-lab recipes call for the system to be held at constant temperature while the reaction proceeds. Vastly fewer call for the system to be held in a thermally-insulated flask while the reaction proceeds. A theory of “thermodynamics” incapable of describing typical laboratory procedures would be pretty lame.
- Even if the overall system is insulated, we often arrange it so that various subsystems within the system are mutually in equilibrium. For example, if there is liquid in a flask, we expect the left half of the liquid to be in thermal equilibrium with the right half, especially if we stir things. But remember, equilibrium involves having a shared temperature. The left half is *not* thermally insulated from the right half; energy is exchanged between the two halves. The microstates of the left half are not equiprobable. A theory of “thermodynamics” incapable of describing thermal equilibrium would be pretty lame.

## 9.8 Misconceptions about Spreading

This section exists mainly to dispel a misconception. If you do not suffer from this particular misconception

Non-experts sometimes get the idea that whenever something is more dispersed – more spread out in position – its entropy must be higher. This is a mistake.

There is a fundamental conceptual problem, which should be obvious from the fact that the degree of dispersal, insofar as it can be defined at all, is a property of the microstate + whereas entropy is a property of the macrostate as a whole, i.e. a property of the ensemble, i.e. a property of the probability distribution as a whole, as discussed in section 2.4 and especially section 2.7.1. A similar microstate versus macrostate argument applies to the “disorder” model of entropy, as discussed in section 2.5.5. In any case, whatever “dispersal” is measuring, it’s not entropy.

Yes, you can find selected scenarios where a gas expands and *does* gain entropy (such as isothermal expansion, or diffusive mixing as discussed in section 11.6) ... but there are also scenarios where a gas expands but does *not* gain entropy (reversible thermally-isolated expansion). Indeed there are scenarios where a gas gains entropy by becoming *less* spread out, as we now discuss:

Consider a closed system consisting of a tall column of gas in a gravitational field, at a uniform not-too-high temperature such that  $kT < m|g|H$ . Start from a situation where the density<sup>3</sup> is uniform, independent of height, as in figure 9.6a. This is not the equilibrium distribution.

As the system evolves toward equilibrium, irreversibly, its entropy will increase. At equilibrium, the density will be greater toward the bottom and lesser toward the top, as shown in figure 9.6b. Furthermore, the equilibrium situation does not exhibit even dispersal of energy. The kinetic energy per particle is evenly dispersed, but the potential energy per particle and the total energy per particle are markedly dependent on height.

There are theorems about what does get uniformly distributed, as discussed in chapter 25. Neither density nor energy is the right answer.

As another example, consider two counter-rotating flywheels. In particular, imagine that these flywheels are annular in shape, i.e. hoops, as shown in figure 9.7, so that to a good approximation, all the mass is at the rim, and every bit of mass is moving at the same speed. Also imagine that they are stacked on the same axis. Now let the two wheels rub together, so that friction causes them to slow down and heat up. Entropy has been produced, but the energy has not become more spread-out in space. To a first approximation, the energy was everywhere to begin with and everywhere afterward, so there is no change.

If we look more closely, we find that as the entropy increased, the energy dispersal actually *decreased* slightly. That is, the energy became slightly *less* evenly distributed in space. Under the initial conditions, the macroscopic rotational mechanical energy was evenly distributed, and the microscopic forms of energy were evenly distributed on a macroscopic scale, plus or minus small local thermal fluctuations. Afterward, the all the energy is in the microscopic forms. It is still evenly distributed on a macroscopic scale, plus or minus thermal fluctuations, but the thermal fluctuations are now *larger* because the temperature is higher. Let’s be clear: If we ignore thermal fluctuations, the increase in entropy was accompanied by no change in the spatial distribution of energy, while if we include the fluctuations, the increase in entropy was accompanied by *less* even dispersal of the energy.

---

<sup>3</sup>If you want to be quantitative about the density, consider a smallish parcel of gas at a particular height.



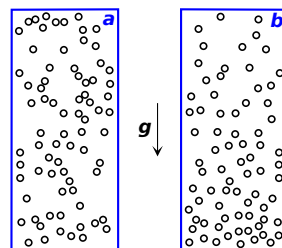


Figure 9.6: Equilibrium is Less Evenly Dispersed

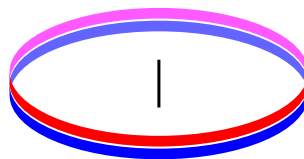


Figure 9.7: Grinding Wheels

Here's yet another nail in the coffin of the «dispersal» model of entropy. Consider a thermally isolated system consisting of gas in a piston pushing up on a mass, subject to gravity, as shown in figure 9.8. Engineer it to make dissipation negligible. Let the mass oscillate up and down, reversibly. The matter and energy become repeatedly more and less «disperse», with no change in entropy.

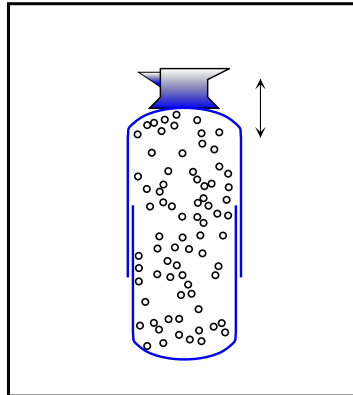


Figure 9.8: Matter and Energy Dispersing and Un-Dispersing

Here's another reason why any attempt to define entropy in terms of “energy dispersal” or the like is Dead on Arrival: Entropy is defined in terms of probability, and applies to systems where the energy is zero, irrelevant, and/or undefinable.

As previously observed, states are states; they are not necessarily energy states.

## 9.9 Spreading in Probability Space

You can salvage the idea of spreading if you apply it to the spreading of probability in an abstract probability-space (not the spreading of energy of any kind, and not spreading in ordinary position-space).

This is not the recommended way to introduce the idea of entropy, or to define entropy. It is better to intr

If you insist on using the idea of spreading, here’s an example that illustrates the idea and can be analyzed in detail. Figure 9.9 shows two blocks under three transparent cups. In the first scenario, the blocks are “concentrated” in the 00 state. In the probability histogram below the cups, there is unit probability (shown in magenta) in the 00 slot, and zero probability in the other slots, so  $p \log(1/p)$  is zero everywhere. That means the entropy is zero.

In the next scenario, the blocks are spread out in position, but since we know exactly what state they are in, all the probability is in the 02 slot. That means  $p \log(1/p)$  is zero everywhere, and the entropy is still zero.

In the third scenario, the system is in some randomly chosen state, namely the 21 state, which is as disordered and as random as any state can be, yet since we know what state it is,  $p \log(1/p)$  is zero everywhere, and the entropy is zero.

The fourth scenario is derived from the third scenario, except that the cups are behind a screen. We can’t see the blocks right now, but we remember where they are. The entropy remains zero.

Finally, in the fifth scenario, we simply don’t know what state the blocks are in. The blocks are behind a screen, and have been shuffled since the last time we looked. We have some vague notion that on average, there is 2/3rds of a block under each cup, but that is only an average over many states. The probability histogram shows there is a 1-out-of-9 chance for the system to be in any of the 9 possible states, so  $\sum p \log(1/p) = \log(9)$ .

One point to be made here is that entropy is not defined in terms of particles that are spread out (“dispersed”) in position-space, but rather in terms of probability that is spread out in state-space. This is quite an important distinction. For more details on this, including an interactive simulation, see reference 25.

Entropy involves probability spread out in state-space,

- not necessarily anything spread out in position-space,
- not necessarily particles spread out in any space,
- not necessarily energy spread out in any space.

To use NMR language, entropy is produced on a timescale  $\tau_2$ , while energy-changes take place on a timescale  $\tau_1$ . There are systems where  $\tau_1$  is huugely longer than  $\tau_2$ . See also section 11.5.5 and figure 1.3. (If this paragraph doesn’t mean anything to you, don’t worry about it.)

As a way of reinforcing this point, consider a system of spins such as discussed in section 11.10. The spins change orientation, but they don’t change position at all. Their positions are locked to the crystal lattice. The notion of entropy doesn’t require any notion of position; as long as we have

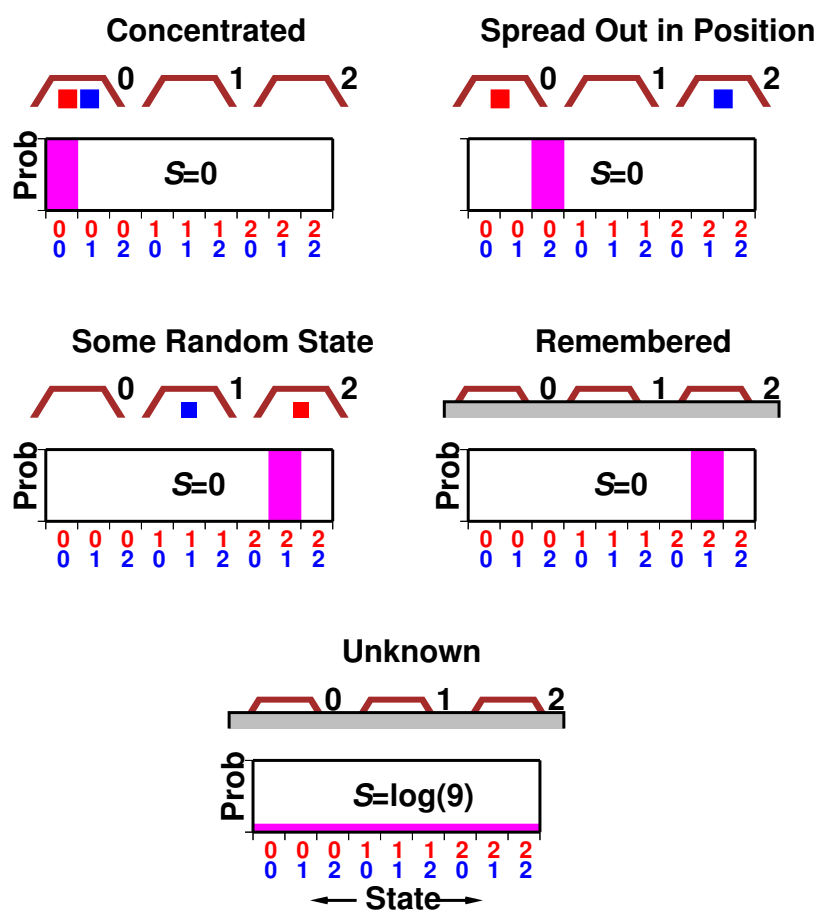


Figure 9.9: Spreading vs. Randomness vs. Uncertainty

states, and a probability of occupying each state, then we have a well-defined notion of entropy. High entropy means the probability is spread out over many states in state-space.

State-space can sometimes be rather hard to visualize. As mentioned in section 2.3, a well-shuffled card deck has nearly  $2^{226}$  bits of entropy ... which is a stupendous number. If you consider the states of gas molecules in a liter of air, the number of states is even larger – far, far beyond what most people can visualize. If you try to histogram these states, you have an unmanageable number of slots (in contrast to the 9 slots in figure 9.9) with usually a very small probability in each slot.

Another point to be made in connection with figure 9.9 concerns the relationship between observing and stirring (aka mixing, aka shuffling). Here's the rule:

	not looking	looking
not stirring	entropy constant	entropy decreasing (aa)
stirring	entropy increasing (aa)	indeterminate change in entropy

where (aa) means almost always; we have to say (aa) because entropy can't be increased by stirring if it is already at its maximum possible value, and it can't be decreased by looking if it is already zero. Note that if you're not looking, lack of stirring does not cause an increase in entropy. By the same token, if you're not stirring, lack of looking does not cause a decrease in entropy. If you are stirring and looking simultaneously, there is a contest between the two processes; the entropy might decrease or might increase, depending on which process is more effective.

The simulation in reference 25 serves to underline these points.

Last but not least, it must be emphasized that spreading of probability in probability-space is dramatically different from spreading energy (or anything else) in ordinary position-space. For one thing, these two spaces don't even have the same size. Suppose you have a crystal with a million evenly-spaced copper atoms. We consider the magnetic energy of the nuclear spins. Each nucleus can have anywhere from zero to four units of energy. Suppose the total energy is two million units, which is what we would expect for such a system at high temperature.

- If you are just dispersing the energy in position-space, there are a million different locations in this space. That's just barely a seven-digit number. You just assign one one-millionth of the energy to each one. That is, you give the atoms two units of energy apiece. It's all very simple ... but wrong.
- In probability space, the total number of states is 4 to the millionth power. That's a number with 600,000 digits. At high temperature, the probability is spread evenly over all of these states.

Let's be clear: a number with 600,000 digits is very much larger than a number with six or seven digits. If you imagine spreading the energy in position-space, it gives entirely the wrong picture. The physics cares about spreading the probability in a completely different space, a very much larger space. The probability is spread very much more thinly.



## Chapter 10

# Additional Fundamental Notions

### 10.1 Equilibrium

Feynman defined equilibrium to be “when all the fast things have happened but the slow things have not” (reference 28). That statement pokes fun at the arbitrariness of the split between “fast” and “slow” – but at the same time it is 100% correct and insightful. There *is* an element of arbitrariness in our notion of equilibrium. Note the following contrast:

Over an ultra-long timescale, a diamond will turn into graphite.	Over an ultra-short timescale, you can have non-equilibrium distributions of phonons rattling around inside a diamond crystal, such that it doesn't make sense to talk about the temperature thereof.
--	---

Usually thermodynamics deals with the intermediate timescale, long after the phonons have become thermalized but long before the diamond turns into graphite. During this intermediate timescale it makes sense to talk about the temperature, as well as other thermodynamic properties such as volume, density, entropy, et cetera.

One should neither assume that equilibrium exists, nor that it doesn't.

- Diamond has a vast, clear-cut separation between the slow timescale and the fast timescale. Most intro-level textbook thermodynamics deal only with systems that have a clean separation.
- In the real world, one often encounters cases where the separation of timescales is not so clean, and an element of arbitrariness is involved. The laws of thermodynamics can still be applied, but more effort and more care is required. See section 11.3 for a discussion.

The word *equilibrium* is quite ancient. The word has the same stem as the name of the constellation “Libra” — the scale. The type of scale in question is the two-pan balance shown in figure 10.1, which has been in use for at least 7000 years.

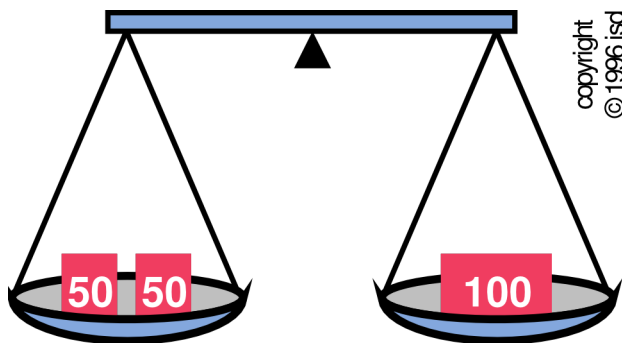


Figure 10.1: Equilibrium — Forces in Balance

The notion of equilibrium originated in mechanics, long before thermodynamics came along. The compound word “equilibrium” translates literally as “equal balance” and means just that: everything in balance. In the context of mechanics, it means there are no unbalanced forces, as illustrated in the top half of figure 10.2.

Our definition of equilibrium applies to infinitely large systems, to microscopic systems, and to everything in between. This is important because in finite systems, there will be fluctuations *even at equilibrium*. See section 10.7 for a discussion of fluctuations and other finite-size effects.

## 10.2 Non-Equilibrium; Timescales

The idea of equilibrium is one of the foundation-stones of thermodynamics ... but any worthwhile theory of thermodynamics must also be able to deal with non-equilibrium situations.

Consider for example the familiar Carnot heat engine: It depends on having two heat reservoirs at two different temperatures. There is a well-known and easily-proved theorem (section 14.4) that says at equilibrium, everything must be at the same temperature. Heat bath #1 may be internally in equilibrium with itself at temperature  $T_1$ , and heat bath may be internally in equilibrium with itself at temperature  $T_2$ , but the two baths cannot be in equilibrium with each other.

So we must modify Feynman’s idea. We need to identify a *timescale of interest* such that all the fast things have happened and the slow things have not. This timescale must be long enough so that certain things we want to be in equilibrium have come into equilibrium, yet short enough so that things we want to be in non-equilibrium remain in non-equilibrium.

Here’s another everyday example where non-equilibrium is important: sound. As you know, in a sound wave there will be some points where the air is compressed and other points, a half-wavelength away, where the air is expanded. For ordinary audible sound, this expansion occurs *isentropically* not isothermally. If you analyze the physics of sound using the isothermal compressibility instead of the isentropic compressibility, you will get the wrong answer. Among other things, your prediction



for the speed of sound will be incorrect. This is an easy mistake to make; Isaac Newton made this mistake the first time he analyzed the physics of sound.

Again we invoke the theorem that says in equilibrium, the whole system must be at the same temperature. Since the sound wave is not isothermal, and cannot even be satisfactorily approximated as isothermal, we conclude that any worthwhile theory of thermodynamics must include non-equilibrium thermodynamics.

For a propagating wave, the time (i.e. period) scales like the distance (i.e. wavelength). In contrast, for diffusion and thermal conductivity, the time scales like distance *squared*. That means that for ultrasound, at high frequencies, a major contribution to the attenuation of the sound wave is thermal conduction between the high-temperature regions (wave crests) and the low-temperature regions (wave troughs). If you go even farther down this road, toward high thermal conductivity and short wavelength, you can get into a regime where sound is well approximated as isothermal. Both the isothermal limit and the isentropic limit have relatively low attenuation; the intermediate case has relatively high attenuation.

### 10.3 Efficiency; Timescales

Questions of efficiency are central to thermodynamics, and have been since Day One (reference 29).

For example in figure 1.3, if we try to extract energy from the battery very quickly, using a very low impedance motor, there will be a huge amount of power dissipated inside the battery, due to the voltage drop across the internal series resistor  $R_1$ . On the other hand, if we try to extract energy from the battery very slowly, most of the energy will be dissipated inside the battery via the shunt resistor  $R_2$  before we have a chance to extract it. So efficiency requires a timescale that is not too fast and not too slow.

Another example is the familiar internal combustion engine. It has a certain tach at which it works most efficiently. The engine is always nonideal because some of the heat of combustion leaks across the boundary into the cylinder block. Any energy that goes into heating up the cylinder block is unavailable for doing  $PdV$  work. This nonideality becomes more serious when the engine is turning over slowly. On the other edge of the same sword, when the engine is turning over all quickly, there are all sorts of losses due to friction in the gas, friction between the mechanical parts, et cetera. These losses increase faster than linearly as the tach goes up.

If you have gas in a cylinder with a piston and compress it slowly, you can (probably) treat the process as reversible. On the other hand, if you move the piston suddenly, it will stir the gas. This can be understood macroscopically in terms of sound radiated into the gas, followed by frictional dissipation of the sound wave (section 11.5.1). It can also be understood microscopically in terms of time-dependent perturbation theory; a sudden movement of the piston causes microstate transitions that would otherwise not have occurred (section 11.5.2).

Timescales matter.

## 10.4 Spontaneity, Irreversibility, Stability, and Equilibrium

Another of the great achievements of thermodynamics is the ability to understand what processes occur spontaneously (and therefore irreversibly) and what processes are reversible (and therefore non-spontaneous). The topic of spontaneity, reversibility, stability, and thermodynamic equilibrium is discussed in depth in chapter 14.

Any theory of thermodynamics that considers only reversible processes – or which formulates its basic laws and concepts in terms of reversible processes – is severely crippled.

If you want to derive the rules that govern spontaneity and irreversibility, as is done in chapter 14, you need to consider perturbations away from equilibrium. If you assume that the perturbed states are in equilibrium, the derivation is guaranteed to give the wrong answer.

In any reversible process, entropy is a conserved quantity. In the real world, entropy is not a conserved quantity.

If you start with a reversible-only equilibrium-only (ROEO) theory of thermodynamics and try to extend it to cover real-world situations, it causes serious conceptual difficulties. For example, consider an irreversible process that creates entropy from scratch in the interior of a thermally-isolated region. Then imagine trying to model it using ROEO ideas. You could try to replace the created entropy by entropy that flowed in from some fake entropy reservoir, but that would just muddy up the already-muddy definition of heat. Does the entropy from the fake entropy reservoir count as “heat”? The question is unanswerable. The “yes” answer is unphysical since it violates the requirement that the system is thermally isolated. The “no” answer violates the basic conservation laws.

Additional examples of irreversible processes that deserve our attention are discussed in sections 10.3, 11.5.1, 11.5.3, 11.5.5, and 11.6.

Any theory of reversible-only equilibrium-only thermodynamics is dead on arrival.

ROEO = DoA
------------

## 10.5 Stability

The basic ideas of stability and equilibrium are illustrated in figure 10.2. (A more quantitative discussion of stability, equilibrium, spontaneity, reversibility, etc. can be found in chapter 14.)

We can understand stability as follows: Suppose we have two copies (two instances) of the same system. Suppose the initial condition of instance *A* is slightly different from the initial condition of instance *B*. If the subsequent behavior of the two copies remains closely similar, we say the system is stable.

More specifically, we define stability as follows: If the difference in the behavior is *proportionate* to the difference in initial conditions, we say the system is stable. Otherwise it is unstable. This notion of stability was formalized by Lyapunov in the 1880s, although it was understood in less-formal ways long before then.

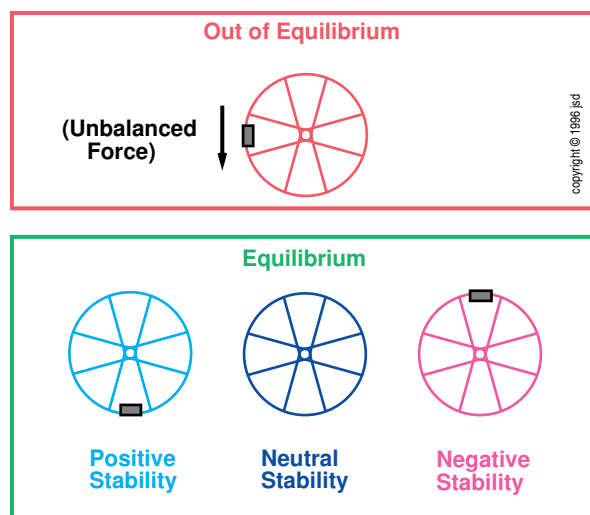


Figure 10.2: Equilibrium and Stability

For a mechanical system, such as in figure 10.2, we can look into the workings of how equilibrium is achieved. In particular,

- Consider the wheel with the weight at the bottom, as shown at the lower left in figure 10.2. Suppose instance *A* starts out in equilibrium and remains in equilibrium. Instance *B* starts out at the same position, but is not in equilibrium because of an additional, external, applied force. The applied force gives rise to a displacement, and the displacement gives rise to an internal force that opposes the applied force, eventually canceling the applied force altogether. Such a system has positive static stability.

Note that the equilibrium position in system *B* is shifted relative to the equilibrium position in system *A*. Stability does not require the system to return to its original position. It only requires that the response be *proportionate* to the disturbance.

- Now consider the perfectly balanced wheel. An applied force gives rise to a displacement, and the displacement gives rise to no force whatsoever. Such a system has zero static stability. If rather than applying a force, we simply move this system to a new position, it will be at equilibrium at the new position. There will be infinitely many equilibrium positions.
- Now consider the wheel with the weight at the top. An applied force gives rise to a displacement. The displacement gives rise to an internal force in the same direction, amplifying the effect of the applied force. We say such a system has negative static stability.

For a non-mechanical system, such as a chemical reaction system, corresponding ideas apply, although you have to work harder to define the notions that correspond to displacement, applied force, restoring force, et cetera.

A system with positive static stability will be stable in the overall sense, unless there is a lot of negative damping or something peculiar like that.

Note that a system can be stable with respect to one kind of disturbance but unstable with respect to another. As a simple example, consider the perfectly balanced wheel, with no damping.

- If system “A” is in equilibrium at rest at position zero, and system “B” is the same except at a different initial position, then the long-term difference in position is proportionate to the disturbance. The system is stable.
- If system “A” is in equilibrium at rest at position zero, and system “B” is the same except with a different initial velocity, then the long-term difference in position is unbounded, completely disproportionate to the disturbance. The system is unstable.

To determine stability, normally you need to consider all the dynamical variables. In the previous example, the long-term velocity difference is bounded, but that doesn’t mean the system is stable, because the long-term position is unbounded.

Properly speaking, a system with zero stability can be called “neutrally unstable.” More loosely speaking, sometimes a system with zero stability is called “neutrally stable,” although that is a misnomer. A so-called “neutrally stable” system is not stable, just as “zero money” is not the same as “money.”

Tangential remark: In chemistry class you may have heard of “Le Châtelier’s principle.” Ever since Le Châtelier’s day there have been two versions of the “principle,” neither of which can be taken seriously, for reasons discussed in section 14.9.

## 10.6 Relationship between Static Stability and Damping

To reiterate: Stability means that two systems that start out with similar initial conditions will follow similar trajectories. Sometimes to avoid confusion, we call this the “overall” stability or the “plain old” stability ... but mostly we just call it the stability.

Meanwhile, *static stability* arises from a force that depends on *position* of the system. In contrast, *damping* refers to a force that depends on the *velocity*.

The term “dynamic stability” is confusing. Sometimes it refers to damping, and sometimes it refers to the plain old stability, i.e. the overall stability. The ambiguity is semi-understandable and usually harmless, because the only way a system can have positive static stability and negative overall stability is by having negative damping.

Static stability can be positive, zero, or negative; damping can also be positive, zero, or negative. A dynamical system can display any combination of these two properties — nine possibilities in all, as shown in figure 10.3. In the top row, the bicycle wheel is dipped in molasses, which provides damping. In the middle row, there is no damping. In the bottom row, you can imagine there is some hypothetical “anti-molasses” that provides negative damping.

The five possibilities in the bottom row and the rightmost column have negative overall stability, as indicated by the pale-red shaded region. The three possibilities nearest the upper-left corner have positive overall stability, as indicated by the pale-blue shaded region. The middle possibility (no

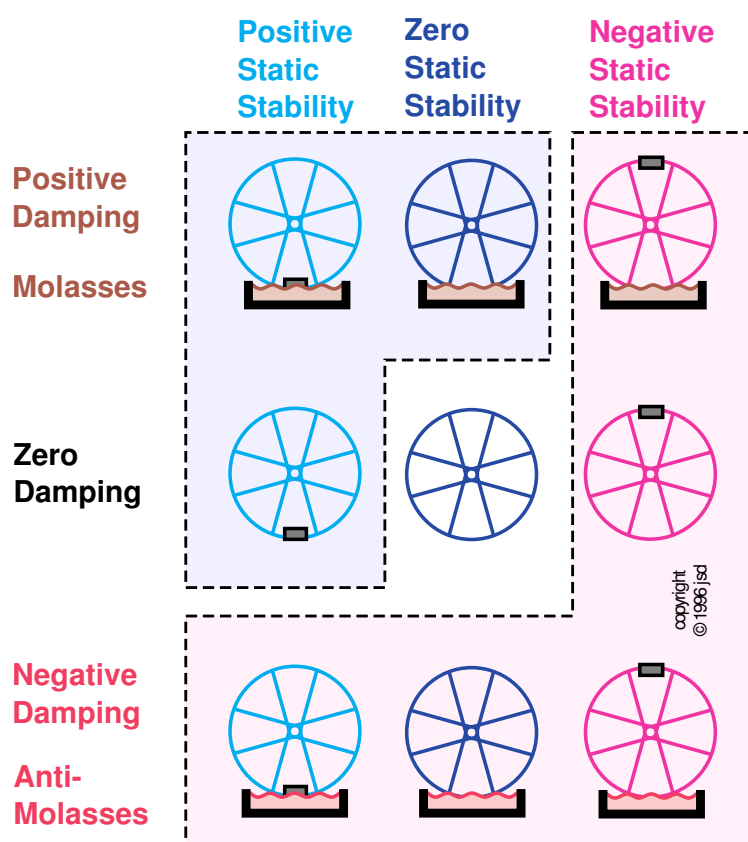


Figure 10.3: Stability and Damping — Possible Combinations

static stability and no damping) is stable with respect to some disturbances (such as a change in initial position) but unstable with respect to others (such as a change in initial velocity).

By the way: Damping should be called “damping” not “dampening” — if you start talking about a “dampener” people will think you want to moisten the system.

## 10.7 Finite Size Effects

As we shall discuss, finite size effects can be categorized as follows (although there is considerable overlap among the categories):

- Boundary effects;
- Quantization effects;
- Mean free path effects;
- Transport and dissipation effects; and
- Fluctuations.

We shall see that:

1. In microscopic systems, finite-size effects dominate.
2. In moderately-large systems, finite-size effects lead to smallish correction terms.
3. In infinite systems, finite-size effects are negligible.

Let’s start with an example: The usual elementary analysis of sound in air considers only *adiabatic* changes in pressure and density. Such an analysis leads to a wave equation that is non-dissipative. In reality, we know that there is some dissipation. Physically the dissipation is related to transport of energy from place to place by thermal conduction. The amount of transport depends on wavelength, and is negligible in the *hydrodynamic limit*, which in this case means the limit of very long wavelengths.

We can come to the same conclusion by looking at things another way. The usual elementary analysis treats the air in the *continuum limit*, imagining that the gas consists of an infinite number density of particles each having infinitesimal size and infinitesimal mean free path. That’s tantamount to having no particles at all; the air is approximated as a continuous fluid. In such a fluid, sound would travel without dissipation.

So we have a macroscopic view of the situation (in terms of nonzero conductivity) and a microscopic view of the situation (in terms of quantized atoms with a nonzero mean free path). These two views of the situation are equivalent, because thermal conductivity is proportional to mean free path (for any given heat capacity and given temperature).

In any case, we can quantify the situation by considering the ratio of the wavelength  $\Lambda$  to the mean free path  $\lambda$ . Indeed we can think in terms of a Taylor series in powers of  $\lambda/\Lambda$ .

- The zeroth-order term is all that survives in the hydrodynamic limit (or the continuum limit). This is the only term that is considered in the usual elementary analysis.
- The first-order terms allow us to understand additional details of the situation, such as the fact that the sound wave undergoes some damping.

Let us now discuss fluctuations.

As an example, in a system at equilibrium, the pressure as measured by a very large piston will be essentially constant. Meanwhile, the pressure as measured by a very small piston will fluctuate. These pressure fluctuations are closely related to the celebrated Brownian motion.

Fluctuations are the rule, whenever you look closely enough and/or look at a small enough subsystem. There will be temperature fluctuations, density fluctuations, entropy fluctuations, et cetera.

We remark in passing that the dissipation of sound waves is intimately connected to the fluctuations in pressure. They are connected by the **fluctuation / dissipation theorem**, which is a corollary of the second law of thermodynamics.

There is magnificent discussion of fluctuations in Feynman volume I chapter 46 (“Ratchet and Pawl”). See reference 8.

As another example, consider *shot noise*. That is: in a small-sized electronic circuit, there will be fluctuations in the current, because the current is not carried by a continuous fluid but rather by electrons which have a quantized charge.

Let us now discuss boundary terms.

If you change the volume of a sample of compressible liquid, there is a well-known  $P dV$  contribution to the energy, where  $P$  is the pressure and  $V$  is the volume. There is also a  $\tau dA$  contribution, where  $\tau$  is the surface tension and  $A$  is the area.

A simple scaling argument proves that for very large systems, the  $P dV$  term dominates, whereas for very small systems the  $\tau dA$  term dominates. For moderately large systems, we can start with the  $P dV$  term and then consider the  $\tau dA$  term as a smallish correction term.

## 10.8 Words to Live By

1. Temperature is not energy.
2. Rate is not equilibrium.
3. Entropy is not energy.
4. «Heat» is not the same as «heat».

In more detail:

1. Temperature and energy follow completely different scaling laws: one them is intensive, while the other is extensive. As a familiar example, the sparks that fly from a grinding wheel have a very high temperature, but very little energy.

Just because a reaction proceeds faster at high temperature does not mean it is exothermic. As a familiar example, the combustion of coal is famously exothermic, yet it proceeds much faster at elevated temperature.

2. As a familiar example, catalysis can change the rate of reaction by many orders of magnitude, but it never changes the equilibrium point.

Temperature is not the same as catalysis, insofar as sometimes it changes the equilibrium point. However, you can't infer the equilibrium point or the energy balance just by casual observation of the temperature.

3. Equilibrium is determined more directly by entropy than by energy. Therefore the fact that you can dehydrate something at temperatures above 100 C in a dessicator and rehydrate it by adding a huge excess of water below 100 C tells you practically nothing about the enthalpies of formation.
4. The nice thing about defining «heat» is that there are so many definitions to choose from. When people say let's «heat» the sample, they might be talking about temperature. When they say the sample gives off «heat», they might be talking about energy.



# Chapter 11

## Experimental Basis

In science, questions are not decided by taking votes, or by seeing who argues the loudest or the longest. Scientific questions are decided by a careful combination of experiments and reasoning. So here are some epochal experiments that form the starting point for the reasoning presented here, and illustrate why certain other approaches are unsatisfactory.

### 11.1 Basic Notions of Temperature and Equilibrium

Make a bunch of thermometers. Calibrate them, to make sure they agree with one another. Use thermometers to measure each of the objects mentioned below.

- Temperature is an intensive quantity. That means that if you have a large parcel of fluid with a definite temperature, and you divide it into two equal sub-parcels, each sub-parcel has (to an excellent approximation) the same temperature you started with.

In contrast, energy and entropy are extensive quantities. Each sub-parcel has half as much energy and half as much entropy as the original large parcel.

The terms intensive and extensive are a shorthand way of expressing simple scaling properties. Any extensive property scales like the first power of any other extensive property, so if you know any extensive property you can recognize all the others by their scaling behavior. Meanwhile, an intensive property scales like the zeroth power of any extensive property.

- At thermal equilibrium, things have the same temperature, no matter what they are made of. To demonstrate this, take two objects that start out at different temperatures. Put them in a box together. Observe that they end up at the same temperature.

(This can be seen to be related to the previous point, if we consider two bodies that are simply parts of a larger body.)

- Take two objects that start out at the same temperature. Put them in a box together. Observe that they never (if left alone) end up at different temperatures. You can build a machine, called a refrigerator or a heat pump, that will cool off one object while heating up the other,

but all such machines require an energy input, so they are irrelevant to any discussion of equilibrium.

- The molar volume of an ideal gas is proportional to absolute temperature.
- The efficiency of an ideal heat engine depends on absolute temperature.
- Temperature can be connected to a derivative of energy with respect to entropy, as in equation 7.7, as further discussed in chapter 13. Another equation that involves temperature is the Boltzmann distribution, i.e. the equipartition law, equation 24.7.

## 11.2 Exponential Dependence on Energy

Here is a collection of observed phenomena that tend to support equation 9.1.

- There is a wide (but not infinitely wide) class of chemical reactions where the rate of reaction depends exponentially on inverse temperature according to the Arrhenius rate equation:

$$\text{rate} = Ae^{-E_a/kT} \quad (11.1)$$

where  $E_a$  is called the *activation energy* and the prefactor  $A$  is called the *attempt frequency*. The idea here is that the reaction pathway has a potential barrier of height  $E_a$  and the rate depends on thermal activation over the barrier. In the independent-particle approximation, we expect that thermal agitation will randomly give an exponentially small fraction of the particles an energy greater than  $E_a$  in accordance with equation 9.1.

Of course there are many examples where equation 11.1 would not be expected to apply. For instance, the flow of gas through a pipe (under the influence of specified upstream and downstream pressures) is not a thermally activated process, and does not exhibit an exponential dependence on inverse temperature.

- In a wide class of materials, the strength of the NMR signal closely follows the Curie law over a range of many orders of magnitude. That is, the strength is proportional to  $1/T$ . This is exactly what we would expect from treating each individual nucleus as a system unto itself (while treating everything else as the “environment” aka “heat bath”) and assigning probabilities to its individual microstates in accordance with equation 9.1.
- The density of saturated water vapor (i.e. the density of gaseous  $\text{H}_2\text{O}$  in equilibrium with liquid  $\text{H}_2\text{O}$ ) is rather accurately an exponential function of inverse temperature. This is what we would expect from equation 9.1, if we once again make the independent-particle approximation and say that particles in the liquid are in a low-energy state while particles in the vapor are in a high-energy state.

### 11.3 Metastable Systems with a Temperature

Consider an ordinary electrical battery. This is an example of a system where *most* of the modes are characterized by well-defined temperature, but there are also a few exceptional modes. Often such systems have an energy that is higher than you might have guessed based on the temperature and entropy, which makes them useful repositories of “available” energy.

Figure 11.1 shows the probability of the various microstates of the battery, when it is discharged (on the left) or charged (on the right). Rather than labeling the states by the subscript  $i$  as we have done in the past, we label them using a pair of subscripts  $i, j$ , where  $i$  takes on the values 0 and 1 meaning discharged and charged respectively, and  $j$  runs over the thermal phonon modes that we normally think of as embodying the heat capacity of an object.

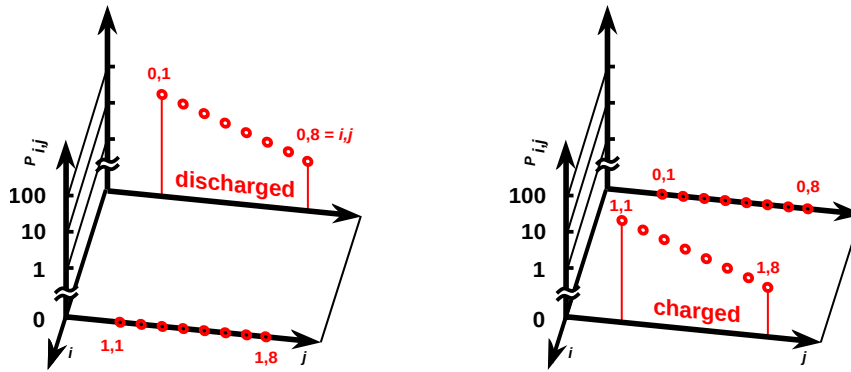


Figure 11.1: Probability of Microstates for a Battery

Keep in mind that probabilities such as  $P_{i,j}$  are defined with respect to some ensemble. For the discharged battery at temperature  $T$ , all members of the ensemble are in contact with a heat bath at temperature  $T$ . That means the thermal phonon modes can exchange energy with the heat bath, and different members of the ensemble will have different amounts of energy, leading to the probabilistic distribution of energies shown on the left side of figure 11.1. The members of the ensemble are not able to exchange electrical charge with the heat bath (or with anything else), so that the eight microstates corresponding to the charged macrostate have zero probability. It doesn't

matter what energy they would have, because they are not accessible. They are not in equilibrium with the other states.

Meanwhile, on the right side of the figure, the battery is in the charged state. The eight microstates corresponding to the discharged macrostate have zero probability, while the eight microstates corresponding to the charged macrostate have a probability distribution of the expected Boltzmann form.

Comparing the left side with the right side of figure 11.1, we see that the two batteries have the same temperature. That is, the slope of  $\log(P_{i,j})$  versus  $E_{i,j}$  – for the modes that are actually able to contribute to the heat capacity – is the same for the two batteries.

You may be wondering how we can reconcile the following four facts: (a) The two batteries have the same temperature  $T$ , (b) the accessible states of the two batteries have different energies, indeed every accessible state of the charged battery has a higher energy than any accessible state of the discharged battery, (c) corresponding accessible states of the two batteries have the same probabilities, and (d) both batteries obey the Boltzmann law,  $P_{i,j}$  proportional to  $\exp(-E_{i,j}/kT)$ . The answer is that there is a bit of a swindle regarding the meaning of “proportional.” The discharged battery has one proportionality constant, while the charged battery has another. For details on this, see section 24.1.

Here is a list of systems that display this sort of separation between thermal modes and nonthermal modes:

- an electrochemical storage battery that may have a shelf life of ten months or ten years.
- a flywheel that may keep spinning for one second or one hour or one day.
- a large piece of metal that rings like a bell, i.e. with a high excitation in one of its mechanical resonance modes.
- a capacitor that may hold its charge for hours or days.
- a fluid-dynamic excitation such as the wingtip vortices trailing behind an airplane.
- a weight-driven cuckoo clock that may go a day or a week between windings.
- a spring-driven clock that may go a day or a week or a year between windings.
- a microwave oven that puts potato-molecules into an excited state.
- a metastable chemical species such as  $\text{H}_2\text{O}_2$  or TNT. If left to themselves, they will decompose quickly or slowly, depending on temperature, catalysis, and other details.
- a classical Carnot-cycle heat engine. If you operate it too quickly, there will be nonidealities because the parts of the cycle that are supposed to be isothermal won’t be (i.e. the working fluid won’t be in good thermal contact with the heat bath). On the other hand, if you operate it too slowly, there will be nonidealities due to parasitic thermal conduction through structures such as the pushrod that connects the piston to the load. You cannot *assume* or *postulate* that there is a nice big separation between the too-slow timescale and the too-fast timescale; if you need a big separation you must arrange for it by careful engineering.

(Section 11.4 takes another look at metastable systems.)

There are good reasons why we might want to apply thermodynamics to systems such as these. For instance, the Clausius-Clapeyron equation can tell us interesting things about a voltaic cell.

Also, just analyzing such a system as a Gedankenexperiment helps us understand a thing or two about what we ought to mean by “equilibrium,” “temperature,” “heat,” and “work.”

In equilibrium, the “accessible” states are supposed to be occupied in accordance with the Boltzmann distribution law (equation 9.1).

An example is depicted in figure 11.1, which is a scatter plot of  $P_{i,j}$  versus  $E_{i,j}$ .

As mentioned in section 10.1, Feynman defined equilibrium to be “when all the fast things have happened but the slow things have not” (reference 28). The examples listed at the beginning of this section all share the property of having *two* timescales and therefore two notions of equilibrium. If you “charge up” such a system you create a Boltzmann distribution with exceptions. There are not just a few exceptions as in figure 11.3, but huge classes of exceptions, i.e. huge classes of microstates that are (in the short run, at least) inaccessible. If you revisit the system on longer and longer timescales, eventually the energy may become dissipated into the previously-inaccessible states. For example, the battery may self-discharge via some parasitic internal conduction path.

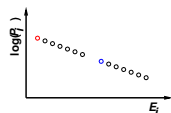


Figure 11.2: An Equilibrium Distribution

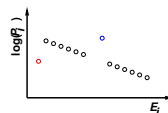


Figure 11.3: An Equilibrium Distribution with Exceptions

The idea of temperature is valid even on the shorter timescale. In practice, I can measure the temperature of a battery or a flywheel without waiting for it to run down. I can measure the temperature of a bottle of  $\text{H}_2\text{O}_2$  without waiting for it to decompose.

This proves that in some cases of interest, we cannot write the system energy  $E$  as a function of the macroscopic thermodynamic variables  $V$  and  $S$ . Remember,  $V$  determines the spacing between energy levels (which is the same in both figures) and  $S$  tells us *something* about the occupation of those levels, but alas  $S$  does not tell us everything we need to know. An elementary example of this can be seen by comparing figure 9.1 with figure 11.3, where we have the same  $V$ , the same  $S$ , and different  $E$ . So we must not assume  $E = E(V, S)$ . A more spectacular example of this can be seen by comparing the two halves of figure 11.1.

Occasionally somebody tries to argue that the laws of thermodynamics do not apply to figure 11.3 or figure 11.1, on the grounds that thermodynamics requires strict adherence to the Boltzmann exponential law. This is a bogus argument for several reasons. First of all, strict adherence to the Boltzmann exponential law would imply that *everything in sight* was at the same temperature. That means we can’t have a heat engine, which depends on having two heat reservoirs at different temperatures. A theory of pseudo-thermodynamics that cannot handle exceptions to the Boltzmann exponential law is useless.

So we must allow some exceptions to the Boltzmann exponential law ... maybe not every imaginable exception, but some exceptions. A good criterion for deciding what sort of exceptions to allow is to ask whether it is operationally possible to measure the temperature. For example, in the case of a storage battery, it is operationally straightforward to design a thermometer that is electrically insulated from the exceptional mode, but thermally well connected to the thermal modes.

Perhaps the most important point is that equation 1.1 and equation 2.1 apply directly, without

modification, to the situations listed at the beginning of this section. So from this point of view, these situations are not “exceptional” at all.

The examples listed at the beginning of this section raise some other basic questions. Suppose I stir a large tub of water. Have I done work on it ( $w$ ) or have I heated it ( $q$ )? If the question is answerable at all, the answer must depend on timescales and other details. A big vortex can be considered a single mode with a huge amount of energy, i.e. a huge exception to the Boltzmann distribution. But if you wait long enough the vortex dies out and you’re left with just an equilibrium distribution. Whether you consider this sort of dissipation to be  $q$  and/or heat is yet another question. (See section 7.10 and especially section 17.1 for a discussion of what is meant by “heat”)

In cases where the system’s internal “spin-down” time is short to all other timescales of interest, we get plain old dissipative systems. Additional examples include:

- The Rumford experiment (section 11.5.3).
- Shear in a viscous fluid (section 11.5.5).
- A block sliding down an inclined plane, under the influence of sliding friction.
- The brake shoes on a car.
- et cetera.

## 11.4 Metastable Systems without a Temperature

An interesting example is:

- a three-state laser, in which there is a population inversion.

In this case, it’s not clear how to measure the temperature or even define the temperature of the spin system. Remember that in equilibrium, states are supposed to be occupied with probability proportional to the Boltzmann factor,  $P_i \propto \exp(-\hat{E}_i/kT)$ . However, the middle microstate is more highly occupied than the microstates on either side, as depicted in figure 11.4. This situation is clearly not describable by any exponential, since exponentials are monotone.

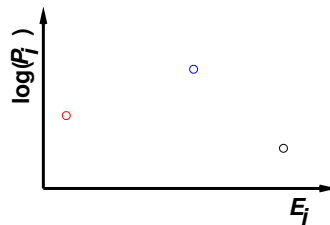


Figure 11.4: Three-State System without a Temperature

We cannot use the ideas discussed in section 11.3 to assign a temperature to such a system, because it has so few states that we can’t figure out which ones are the thermal “background” and which ones are the “exceptions”

Such a system does have an entropy – even though it doesn’t have a temperature, even though it is metastable, and even though it is grossly out of equilibrium. It is absolutely crucial that the system have a well-defined entropy, for reasons suggested by figure 11.5. That is, suppose

the system starts out in equilibrium, with a well-defined entropy  $S(1)$ . It then passes through in intermediate state that is out of equilibrium, and ends up in an equilibrium state with entropy  $S(3)$ . The law of paraconservation of entropy is meaningless unless we can somehow relate  $S(3)$  to  $S(1)$ . The only reasonable way that can happen is if the intermediate state has a well-defined entropy. The intermediate state typically does not have a temperature, but it does have a well-defined entropy.

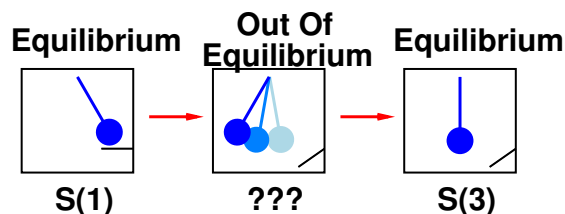


Figure 11.5: Non-Equilibrium: Well-Defined Entropy

## 11.5 Dissipative Systems

### 11.5.1 Sudden Piston : Sound

Consider the apparatus shown in figure 11.6. You can consider it a two-sided piston.

Equivalently you can consider it a loudspeaker in an unusual *full* enclosure. (Loudspeakers are normally only half-enclosed.) It is roughly like two unported speaker enclosures face to face, completely enclosing the speaker driver that sits near the top center, shown in red. The interior of the apparatus is divided into two regions, 1 and 2, with time-averaged properties  $(E_1, S_1, T_1, P_1, V_1)$  and  $(E_2, S_2, T_2, P_2, V_2)$  et cetera. When the driver (aka piston) moves to the right, it increase volume  $V_1$  and decreases volume  $V_2$ . The box as a whole is thermally isolated / insulated / whatever. That is to say, no entropy crosses the boundary. No energy crosses the boundary except for the electricity feeding the speaker.

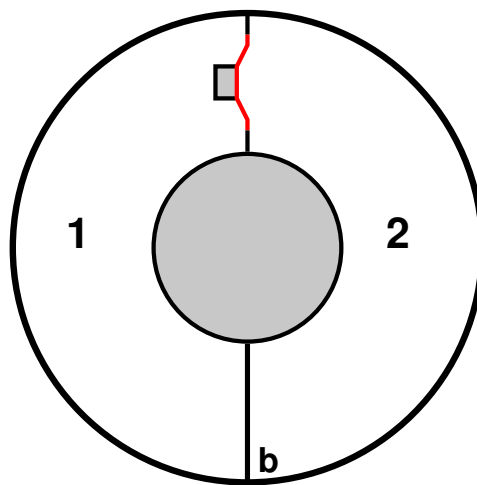
You could build a simplified rectangular version of this apparatus for a few dollars. It is considerably easier to build and operate than Rumford's cannon-boring apparatus (section 11.5.3).

We will be primarily interested in a burst of oscillatory motion. That is, the piston is initially at rest, then oscillates for a while, and then returns to rest at the original position.

When the piston moves, it does  $F \cdot d\mathbf{x}$  work against the gas. There are two contributions. Firstly, the piston does work against the gas in each compartment. If  $P_1 = P_2$  this contribution vanishes to first order in  $dV$ . Secondly, the piston does work against the pressure in the sound field.

The work done against the average pressure averages to zero over the course of one cycle of the oscillatory motion ... but the work against the radiation field does *not* average to zero. The  $dV$  is oscillatory but the field pressure is oscillatory too, and the product is positive on average.

The acoustic energy radiated into the gas is in the short term not in thermal equilibrium with the gas. In the longer term, the sound waves are damped i.e. dissipated by internal friction and also by thermal conductivity, at a rate that depends on the frequency and wavelength.





What we put in is  $F \cdot \mathbf{dx}$  (call it “work” if you wish) and what we get out in the long run is an increase in the energy and entropy of the gas (call it “heat” if you wish).

It must be emphasized that whenever there is appreciable energy in the sound field, it is not possible to write  $E_1$  as a function of  $V_1$  and  $S_1$  alone, or indeed to write  $E_1$  as a function of any two variables whatsoever. In general, the sound creates a pressure  $P(r)$  that varies from place to place as a function of the position-vector  $r$ . That’s why we call it a sound *field*; it’s a scalar field, not a simple scalar.

As a consequence, when there is appreciable energy in the sound field, it is seriously incorrect to expand  $\mathbf{d}E = T \mathbf{d}S - P \mathbf{d}V$ . The correct expansion necessarily has additional terms on the RHS. Sometimes you can analyze the sound field in terms of its normal modes, and in some simple cases most of the sound energy resides in only a few of the modes, in which case you need only a few additional variables. In general, though, the pressure can vary from place to place in an arbitrarily complicated way, and you would need an arbitrarily large number of additional variables. This takes us temporarily outside the scope of ordinary thermodynamics, which requires us to describe the macrostate as a function of some reasonably small number of macroscopic variables. The total energy, total entropy, and total volume are still perfectly well defined, but they do not suffice to give a complete description of what is going on. After we stop driving the piston, the sound waves will eventually dissipate, whereupon we will once again be able to describe the system in terms of a few macroscopic variables.

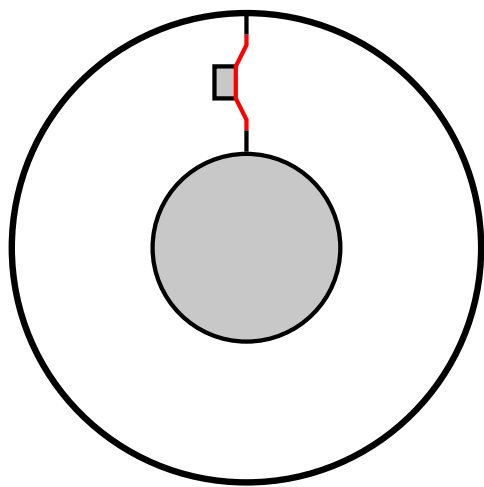
If the piston moves slowly, very little sound will be radiated and the process will be essentially isentropic and reversible. On the other hand, if the piston moves quickly, there will be lots of sound, lots of dissipation, and lots of newly created entropy. This supports the point made in section 10.2: timescales matter.

At no time is any entropy transferred across the boundary of the region. The increase in entropy of the region is due to new entropy, created from scratch in the interior of the region.

If you want to ensure the gas exerts zero average force on the piston, you can cut a small hole in the baffle near point  $b$ . Then the only work the piston can do on the gas is work against the sound pressure field. There is no longer any important distinction between region 1 and region 2.

You can even remove the baffle entirely, resulting in the “racetrack” apparatus shown in figure 11.7.

The kinetic energy of the piston is hardly worth worrying about. When we say it takes more work to move the piston rapidly than slowly, the interesting part is the work done on the gas, not the work done to accelerate the piston. Consider a very low-mass piston if that helps. Besides, whatever KE goes into the piston is recovered at the end of each cycle. Furthermore, it is trivial to calculate the  $F \cdot \mathbf{dx}$  of the piston *excluding* whatever force is necessary to accelerate the piston. Let’s assume the experimenter is clever enough to apply this trivial correction, so that we know, moment by moment, how much  $F \cdot \mathbf{dx}$  “work” is being done on the gas. This is entirely conventional; the conventional pressures  $P_1$  and  $P_2$  are associated with the forces  $F_1$  and  $F_2$  on the faces of the piston *facing the gas*, not the force  $F_d$  that is driving the piston. To relate  $F_d$  to  $F_1$  and  $F_2$  you would need to consider the mass of the piston, but if you formulate the problem in terms of  $F_1 \cdot \mathbf{dx}$  and  $F_2 \cdot \mathbf{dx}$ , as you should, questions of piston mass and piston KE should hardly even arise.



### 11.5.2 Sudden Piston : State Transitions

Let's forget about all the complexities of the sound field discussed in section 11.5.1. Instead let's take the quantum mechanical approach. Let's simplify the gas down to a single particle, the familiar particle in a box, and see what happens.

As usual, we assume the box is rigid and thermally isolated / insulated / whatever. No entropy flows across the boundary of the box. Also, no energy flows across the boundary except for the work done by the piston.

Since we are interested in entropy, it will not suffice to talk about “the” quantum state of the particle. The entropy of any particular quantum state (microstate) is zero. We can however represent the thermodynamic state (macrostate) using a density matrix  $\rho$ . For some background on density matrices in the context of thermodynamics, see chapter 27.

The entropy is given by equation 27.6. which is *the* gold-standard most-general definition of entropy; in the classical limit it reduces to the familiar workhorse expression equation 2.2

For simplicity we consider the case where the initial state is a pure state, i.e. a single microstate. That means the initial entropy is zero, as you can easily verify. Hint: equation 27.6 is particularly easy to evaluate in a basis where  $\rho$  is diagonal.

Next we perturb our particle-in-a-box by moving one wall of the box inward. We temporarily assume this is done in such a way that the particle ends up in the “same” microstate. That is, the final state is identical to the original quantum state except for the shorter wavelength as required to fit into the smaller box. It is a straightforward yet useful exercise to show that this does  $P dV$  “work” on the particle. The KE of the new state is higher than the KE of the old state.

Now the fun begins. We retract the previous assumption about the final state; instead we calculate the final macrostate using perturbation theory. In accordance with Fermi's golden rule we calculate the *overlap integral* between the original quantum state (original wavelength) and each of the possible final quantum states (slightly shorter wavelength).

Each member of the original set of basis wavefunctions is orthogonal to the other members. The same goes for the final set of basis wavefunctions. However, each final basis wavefunction is only *approximately* orthogonal to the various original basis wavefunctions. So the previous assumption that the particle would wind up in the corresponding state is provably not quite true; when we do the overlap integrals there is always some probability of transition to nearby states.

It is straightforward to show that if the perturbation is slow and gradual, the corresponding state gets the lion's share of the probability. Conversely, if the perturbation is large and sudden, there will be lots of state transitions. The final state will not be a pure quantum state. It will be a mixture. The entropy will be nonzero, i.e. greater than the initial entropy.

To summarize:	slow and gradual	$\implies$	isentropic, non dissipative
	sudden	$\implies$	dissipative

So we are on solid grounds when we say that in a thermally isolated cylinder, a gradual movement of the piston is isentropic, while a sudden movement of the piston is dissipative. Saying that the system is adiabatic in the sense of thermally insulated does *not* suffice to make it adiabatic in the sense of isentropic.

Note that in the quantum mechanics literature the slow and gradual case is conventionally called the “adiabatic” approximation in contrast to the “sudden” approximation. These terms are quite firmly established ... even though it conflicts with the also well-established convention in other branches of physics where “adiabatic” means thermally insulated; see next message.

There is a nice introduction to the idea of “radiation resistance” in reference 8 chapter 32.

### 11.5.3 Rumford’s Experiment

Benjamin Thompson (Count Rumford) did some experiments that were published in 1798. Before that time, people had more-or-less assumed that “heat” by itself was conserved. Rumford totally demolished this notion, by demonstrating that unlimited amounts of “heat” could be produced by nonthermal mechanical means. Note that in this context, the terms “thermal energy”, “heat content”, and “caloric” are all more-or-less synonymous ... and I write each of them in scare quotes.

From the pedagogical point of view Rumford’s paper is an optimal starting point; the examples in section 11.5.1 and section 11.5.2 are probably better. For one thing, a microscopic understanding of sound and state-transitions in a gas is easier than a microscopic understanding of metal-on-metal friction.

Once you have a decent understanding of the modern ideas, you would do well to read Rumford’s original paper, reference 30. The paper is of great historical importance. It is easy to read, informative, and entertaining. On the other hand, beware that it contains at least one major error, plus one trap for the unwary:

- The title is wrong and the bottom-line conclusion is wrong, for reasons discussed in section 9.3.3.
- The analysis depends on a cramped notion of “heat content” that is more-or-less OK in this particular context, but is not clearly defined, and but would cause serious problems if you tried to extend it to uncramped thermodynamics.

The main point of the paper is that “heat” is not conserved. This point remains true and important. The fact that the paper has a couple of bugs does not detract from this point.

You should reflect on how something can provide valuable (indeed epochal) information and still not be 100% correct.

All too often, the history of science is presented as monotonic “progress” building one pure “success” upon another, but this is not how things really work. In fact there is a lot of back-tracking out of dead ends. Real science and real life are like football, in the sense that any play that advances the ball 50 or 60 yards it is a major accomplishment, even if you get knocked out of bounds before reaching the ultimate goal. Winning is important, but you don’t need to win the entire game, single handedly, the first time you get your hands on the ball.

Rumford guessed that all the heat capacity was associated with “motion” – because he couldn’t imagine anything else. It was a process-of-elimination argument, and he blew it. This is understandable, given what he had to work with.

A hundred years later, guys like Einstein and Debye were able to cobble up a theory of heat capacity based on the atomic model. We know from this model that the heat capacity of solids is half kinetic and half potential. Rumford didn’t stand much of a chance of figuring this out.

It is possible to analyze Rumford’s experiment without introducing the notion of “heat content.” It suffices to keep track of the energy and the entropy. The energy can be quantified by using the first law of thermodynamics, i.e. the conservation of energy. We designate the cannon plus the water bath as the “system” of interest. We know how much energy was pushed into the system, pushed across the boundary of the system, in the form of macroscopic mechanical work. We can quantify the entropy by means of equation 7.21, i.e.  $dS = (1/T)dE$  at constant pressure. Energy and entropy are functions of state, even in situations where “heat content” is not.

Heat is a concept rooted in cramped thermodynamics, and causes serious trouble if you try to extend it to uncramped thermodynamics. Rumford got away with it, in this particular context, because he stayed within the bounds of cramped thermodynamics. Specifically, he did everything at constant pressure. He used the heat capacity of water *at constant pressure* as his operational definition of heat content.

To say the same thing the other way, if he had strayed off the contour of constant  $P$ , perhaps by making little cycles in the  $PV$  plane, using the water as the working fluid in a heat engine, any notion of “heat content” would have gone out the window. There would have been an unholy mixture of  $C_P$  and  $C_V$ , and the “heat content” would have not been a function of state, and everybody would have been sucked down the rabbit-hole into crazy-nonsense land.

We note in passing that it would be impossible to reconcile Rumford’s notion of “heat” with the various other notions listed in section 17.1 and section 18.1. For example: work is being done in terms of energy flowing across the boundary, but no work is being done in terms of the work/KE theorem, since the cannon is not accelerating.

For more about the difficulties in applying the work/KE theorem to thermodynamic questions, see reference 18.

We can begin to understand the microscopics of sliding friction using many of the same ideas as in section 11.5.1. Let’s model friction in terms of *asperities* on each metal surface. Each of the asperities sticks and lets go, sticks and lets go. When it lets go it wiggles and radiates ultrasound into the bulk of the metal. This produces in the short term a nonequilibrium state due to the sound waves, but before long the sound field dissipates, depositing energy and creating entropy in the metal.

Again, if you think in terms only of the (average force) dot (average  $dx$ ) you will never understand friction or dissipation. You need to model many little contributions of the form (short term force) dot (short term  $dx$ ) and then add up all the contributions. This is where you see the work being done against the radiation field.

At ordinary temperatures (not too hot and not too cold) most of the heat capacity in a solid is associated with the phonons. Other phenomena associated with friction, including deformation and abrasion of the materials, are only very indirectly connected to heating. Simply breaking a bunch of bonds, as in cleaving a crystal, does not produce much in the way of entropy or heat. At some point, if you want to understand heat, you need to couple to the phonons.

#### 11.5.4 Simple Example: Decaying Current

Suppose we have some current  $I$  flowing in a wire loop, as shown in figure 11.8. The current will gradually decay, on a timescale given by  $L/R$ , i.e. the inductance divided by the resistance.

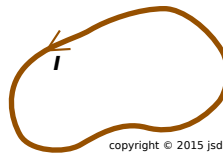


Figure 11.8: Current in a Wire Loop

The temperature of the wire will increase, and the entropy of the wire will increase, even though no energy is being transferred (thermally or otherwise) across the boundary of the system.

(Even if you consider an imaginary boundary between the conduction electrons and the rest of the metal, you cannot possibly use any notion of energy flowing across the boundary to explain the fact that *both* subsystems heat up.)

A decaying current of water in an annular trough can be used to make the same point.

### 11.5.5 Simple Example: Oil Bearing

Here is a modified version of Rumford’s experiment, more suitable for quantitative analysis. Note that reference 31 carries out a similar analysis and reaches many of the same conclusions. Also note that this can be considered a macroscopic mechanical analog of the NMR  $\tau_2$  process, where there is a change in entropy with no change in energy. See also figure 1.3.

Suppose we have an oil bearing as shown in figure 11.9. It consists of an upper plate and a lower plate, with a thin layer of oil between them. Each plate is a narrow annulus of radius  $R$ . The lower plate is held stationary. The upper plate rotates under the influence of a force  $F$ , applied via a handle as shown. The upper plate is kept coaxial with the lower plate by a force of constraint, not shown. The two forces combine to create a pure torque,  $\tau = F/R$ . The applied torque  $\tau$  is balanced in the long run by a frictional torque  $\tau'$ ; specifically

$$\langle \tau \rangle = \langle \tau' \rangle \quad (11.2)$$

where  $\langle \dots \rangle$  denotes a time-average. As another way of saying the same thing, in the long run the upper plate settles down to a more-or-less steady velocity.

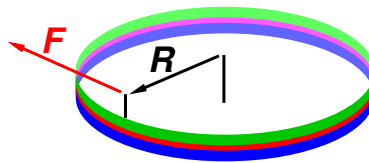


Figure 11.9: Oil Bearing

We arrange that the system as a whole is thermally insulated from the environment, to a sufficient approximation. This includes arranging that the handle is thermally insulating. In practice this isn’t difficult.

We also arrange that the plates are somewhat thermally insulating, so that heat in the oil doesn’t immediately leak into the plates.

Viscous dissipation in the oil causes the oil to heat up. To a good approximation this is the only form of dissipation we must consider.

In an infinitesimal period of time, the handle moves through a distance  $d\mathbf{x}$  or equivalently through an angle  $d\theta = d\mathbf{x}/R$ . We consider the driving force  $F$  to be a controlled variable. We consider  $\theta$  to be an observable dependent variable. The relative motion of the plates sets up a steady shearing motion within the oil. We assume the oil forms a sufficiently thin layer and has sufficiently high viscosity that the flow is laminar (i.e. non-turbulent) everywhere. We say the fluid has a very low Reynolds number (but if you don't know what that means, don't worry about it). The point is that the velocity of the oil follows the simple pattern shown by the red arrows in figure 11.10.

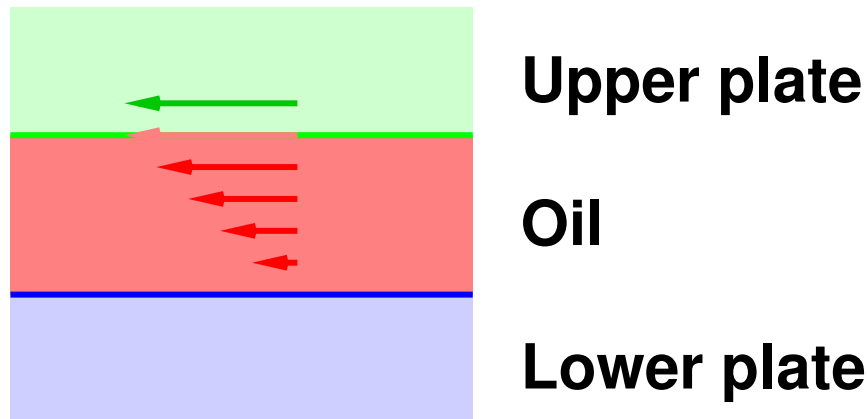


Figure 11.10: Shear: Velocity Field in the Oil

The local work done on the handle by the driving force is  $w = Fd\mathbf{x}$  or equivalently  $w = \tau d\theta$ . This tells us how much energy is flowing across the boundary of the system. From now on we stop talking about work, and instead talk about energy, confident that energy is conserved.

We can keep track of the energy-content of the system by integrating the energy inputs. Similarly, given the initial entropy and the heat capacity of the materials, we can predict the entropy at all times<sup>1</sup> by integrating equation 7.14. Also given the initial temperature and heat capacity, we can predict the temperature at all times by integrating equation 7.13. We can then measure the temperature and compare it with the prediction.

We can understand the situation in terms of equation 1.1. Energy  $\tau d\theta$  comes in via the handle. This energy cannot be stored as potential energy within the system. This energy also cannot be stored as macroscopic or mesoscopic kinetic energy within the system, since at each point the velocity is essentially constant. By a process of elimination we conclude that this energy accumulates inside the system in microscopic form.

This gives us a reasonably complete description of the thermodynamics of the oil bearing.

This example is simple, but helps make a very important point. If you base your thermodynamics on wrong foundations, it will get wrong answers, including the misconceptions discussed in section 11.5.6 and section 11.5.7.

<sup>1</sup>If the flow pattern were turbulent, calculating the entropy would entail practical as well as conceptual difficulties.



Some people who use wrong foundations try to hide from the resulting problems narrowing their definition of “thermodynamics” so severely that it has nothing to say – right or wrong – about dissipative systems. Making no predictions is a big improvement over making wrong predictions ... but still it is a terrible price to pay. Real thermodynamics has tremendous power and generality. Real thermodynamics applies just fine to dissipative systems. See chapter 21 for more on this.

### 11.5.6 Misconceptions : Heat

There are several correct ways of analyzing the oil-bearing system, one of which was presented in section 11.5.5. In addition, there are innumerable many incorrect ways of analyzing things. We cannot list all possible misconceptions, let alone discuss them all. However, it seems worthwhile to point out some of the most prevalent pitfalls.

You may have been taught to think of heating in terms of the “thermal” transfer of energy across a boundary. If you’re going to use that definition, you must keep in mind that it is not equivalent to the  $TdS$  definition. In other words, in section 17.1, definition #5 is sometimes very different from definition #2. The decaying current in section 11.5.4 and the oil-bearing example in section 11.5.5 clearly demonstrates this difference.

Among other things, this difference can be seen as another instance of boundary/interior inconsistency, as discussed in section 8.6. Specifically:

No heat is flowing into the oil. The oil is hotter than its surroundings, so if there is any heat-flow at all, it flows <i>outward</i> from the oil.	The $TdS/dt$ is strongly positive. The entropy of the oil is steadily increasing.
--	---

Another point that can be made using this example is that the laws of thermodynamics apply just fine to dissipative systems. Viscous damping has a number of pedagogical advantages relative to (say) the sliding friction in Rumford’s cannon-boring experiment. It’s clear where the dissipation is occurring, and it’s clear that the dissipation does not prevent us from assigning a well-behaved temperature to each part of the apparatus. Viscous dissipation is more-or-less ideal in the sense that it does not depend on submicroscopic nonidealities such as the asperities that are commonly used to explain solid-on-solid sliding friction.

### 11.5.7 Misconceptions : Work

We now discuss some common misconceptions about work.

Work is susceptible to boundary/interior inconsistencies for some of the same reasons that heat is.

You may have been taught to think of work as an energy transfer across a boundary. That’s one of the definitions of work discussed in section 18.1. It’s often useful, and is harmless *provided* you don’t confuse it with the other definition, namely  $PdV$ .

Work-flow is the “work” that shows up in the principle of virtual work (reference 32), e.g. when we want to calculate the force on the handle of the oil bearing.	Work- $PdV$ is the “work” that shows up in the work/KE theorem.
---	---

### 11.5.8 Remarks

This discussion has shed some light on how equation 7.5 can and cannot be interpreted.

- Sometimes the terms on the RHS are well-defined and can be interpreted as “work” and “heat.”
- Sometimes the terms on the RHS are well-defined but do not correspond to conventional notions of “work” and “heat.”
- Sometimes the terms on the RHS are not even well-defined, i.e. the derivatives do not exist.

In all cases, the equation should not be considered the first law of thermodynamics, because it is inelegant and in every way inferior to a simple, direct statement of local conservation of energy.

## 11.6 The Gibbs Gedankenexperiment

As shown in figure 11.11, suppose we have two moderate-sized containers connected by a valve. Initially the valve is closed. We fill one container with an ideal gas, and fill the other container with a different ideal gas, at the same temperature and pressure. When we open the valve, the gases will begin to mix. The temperature and pressure will remain unchanged, but there will be an irreversible increase in entropy. After mixing is complete, the molar entropy will have increased by  $R \ln 2$ .

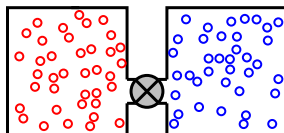


Figure 11.11: The Gibbs Gedankenexperiment

As Gibbs observed,<sup>2</sup> the  $R \ln 2$  result is independent of the choice of gases, “... except that the gases which are mixed must be of different kinds. If we should bring into contact two masses of the same kind of gas, they would also mix, but there would be no increase of entropy.”

There is no way to explain this in terms of 19th-century physics. The explanation depends on quantum mechanics. It has to do with the fact that one helium atom is identical (absolutely totally identical) with another helium atom.

Also consider the following contrast:

In figure 11.11, the pressure on both sides of the valve is the same. There is no net driving force. The process proceeds by diffusion, not by macroscopic flow.

This contrasts with the scenario where we have gas on one side of the partition, but vacuum on the other side. This is dramatically different, because in this scenario there is a perfectly good 17th-century dynamic (not thermodynamic) explanation for why the gas expands: there is a pressure difference, which drives a *flow* of fluid.

---

<sup>2</sup>Quoted in reference 33.

Entropy drives the process. There is no hope of extracting energy from the diffusive mixing process.	Energy drives the process. We could extract some of this energy by replacing the valve by a turbine.
--	--

The timescale for free expansion is roughly  $L/c$ , where  $L$  is the size of the apparatus, and  $c$  is the speed of sound. The timescale for diffusion is slower by a huge factor, namely by a factor of  $L/\lambda$ , where  $\lambda$  is the mean free path in the gas.

Pedagogical note: The experiment in figure 11.11 is not very exciting to watch. Here's an alternative: Put a drop or two of food coloring in a beaker of still water. The color will spread throughout the container, but only rather slowly. This allows students to visualize a process driven by entropy, not energy.

Actually, it is likely that most of the color-spreading that you see is due to convection, not diffusion. To minimize convection, try putting the water in a tall, narrow glass cylinder, and putting it under a Bell jar to protect it from drafts. Then the spreading will take a very long time indeed.

Beware: Diffusion experiments of this sort are tremendously valuable if explained properly ... but they are horribly vulnerable to misinterpretation if not explained properly, for reasons discussed in section 9.9.

For a discussion of the microscopic theory behind the Gibbs mixing experiments, see section 26.2.

## 11.7 Spin Echo Experiment

It is possible to set up an experimental situation where there are a bunch of nuclei whose spins appear to be oriented completely at random, like a well-shuffled set of cards. However, if I let you in on the secret of how the system was prepared, you can, by using a certain sequence of Nuclear Magnetic Resonance (NMR) pulses, get all the spins to line up – evidently a very low-entropy configuration.

The trick is that there is a lot of information in the lattice surrounding the nuclei, something like  $10^{23}$  bits of information. I don't need to communicate all this information to you explicitly; I just need to let you in on the secret of how to use this information to untangle the spins.

The ramifications and implications of this are discussed in section 12.8.

## 11.8 Melting

Take a pot of ice water. Add energy to it via friction, à la Rumford, as described in section 11.5.3. The added energy will cause the ice to melt. The temperature of the ice water will *not* increase, not until all the ice is gone.

This illustrates the fact that temperature is not the same as thermal energy. It focuses our attention on the entropy. A gram of liquid water has more entropy than a gram of ice. So at any given temperature, a gram of water has more energy than a gram of ice.

The following experiment makes an interesting contrast.

## 11.9 Isentropic Expansion and Compression

Take an ideal gas acted upon by a piston. For simplicity, assume a nonrelativistic nondegenerate ideal gas, and assume the sample is small on the scale of  $kT/mg$ . Assume everything is thermally insulated, so that no energy enters or leaves the system via thermal conduction. Gently retract the piston, allowing the gas to expand. The gas cools as it expands. In the expanded state,

- The gas has essentially the same entropy, if the expansion was done gently enough.
- The gas has a lower temperature.
- The gas has less energy, by some amount  $\Delta E$ .

Before the expansion, the energy in question ( $\Delta E$ ) was in microscopic Locrian form, within the gas.	After the expansion, this energy is in macroscopic non-Locrian form, within the mechanism that moves the piston.
--	--

This scenario illustrates some of the differences between temperature and entropy, and some of the differences between energy and entropy.

Remember, the second law of thermodynamics says that the entropy obeys a local law of paraconservation. Be careful not to misquote this law.

It doesn't say that the temperature can't decrease. It doesn't say that the so-called "thermal energy" can't decrease.	It says the entropy can't decrease in any given region of space, except by flowing into adjacent regions.
--	---

Energy is conserved. That is, it cannot increase or decrease except by flowing into adjacent regions. (You should not imagine that there is any law that says "thermal energy" by itself is conserved.)

If you gently push the piston back in, compressing the gas, the temperature will go back up.

Isentropic compression is an increase in temperature at constant entropy. Melting (section 11.8) is an increase in entropy at constant temperature. These are two radically different ways of increasing the energy.

## 11.10 Demagnetization Refrigerator

Obtain (or construct) a simple magnetic compass. It is essentially a bar magnet that is free to pivot. Attach it to the middle of a board. By placing some small magnets elsewhere on the board, you should (with a little experimentation) be able to null out the earth's field and any other stray fields, so that the needle rotates freely. If you don't have convenient physical access to the needle, you can set it spinning using another magnet.

Note: The earth's field is small, so it doesn't take much to null it out. You can make your own not-very-strong magnets by starting with a piece of steel (perhaps a sewing pin) and magnetizing it.

Once you have a freely-moving needle, you can imagine that if it were smaller and more finely balanced, thermal agitation would cause it to rotate randomly back and forth forever.

Now hold another bar magnet close enough to ruin the free rotation, forcing the spinner to align with the imposed field.

This is a passable pedagogical model of the guts of a demagnetization refrigerator. Such devices are routinely used to produce exceedingly low temperatures, within a millionth of a degree of absolute zero. Copper nuclei can be used as the spinners.

- At low temperatures in a high magnetic field, the copper nuclei have only one accessible state. This corresponds to zero molar entropy.
- At high temperatures in a low magnetic field, the nuclei are free and have 4 equiprobable states. The 4 arises because both kinds of naturally-occurring copper nuclei have total spin  $I = 3/2$ , so there are four possible  $m_I$  values, namely  $+3/2$ ,  $+1/2$ ,  $-1/2$ , and  $-3/2$ . This corresponds to a molar entropy of  $s = R \ln(4)$ . All the other quantum numbers are irrelevant; the four  $m_I$  states are the only accessible states.

The compass is not a perfect model of the copper nucleus, insofar as it has more than four states when it is spinning freely. However, if you use your imagination, you can pretend there are only four states. When a strong field is applied, only one of these states remains accessible.

It is worthwhile to compare theory to experiment:

<p>These values for the molar entropy <math>s</math> have a firm theoretical basis. They require little more than counting. We count microstates and apply the definition of entropy. Then we obtain <math>\Delta s</math> by simple subtraction.</p>	<p>Meanwhile, <math>\Delta s</math> can also be obtained experimentally, by observing the classical macroscopic thermodynamic behavior of the refrigerator.</p>
---	---

Both ways of obtaining  $\Delta s$  give the same answer. What a coincidence! This answers the question about how to connect microscopic state-counting to macroscopic thermal behavior. The Shannon entropy is not merely analogous to the thermodynamic entropy; it *is* the thermodynamic entropy.

Spin entropy is discussed further in section 12.4.

Tangential remark: There are some efforts devoted to using demagnetization to produce refrigeration under conditions that are not quite so extreme; see e.g. reference 34.

## 11.11 Thermal Insulation

As a practical technical matter, it is often possible to have a high degree of thermal insulation between some objects, while other objects are in vastly better thermal contact.

For example, if we push on an object using a thermally-insulating stick, we can transfer energy to the object, without transferring much entropy. In contrast, if we push on a hot object using a non-insulating stick, even though we impart energy to one or two of the object's modes by pushing, the object could be losing energy overall, via thermal conduction through the stick.

Similarly, if you try to build a piece of thermodynamic apparatus, such as an automobile engine, it is essential that some parts reach thermal equilibrium reasonably quickly, and it is equally essential that other parts do not reach equilibrium on the same timescale.

## Chapter 12

# More About Entropy

### 12.1 Terminology: Microstate versus Macrostate

Beware: In the thermodynamics literature, the word “state” is used with two inconsistent meanings. It could mean either microstate or macrostate.

In a system such as the deck of cards discussed in section 2.3, the microstate is specified by saying exactly which card is on top, exactly which card is in the second position, et cetera.

In that system, the macrostate is the ensemble of all card decks consistent with what we know about the situation.

In a system such as a cylinder of gas, a microstate is a single fully-specified quantum state of the gas.

For such a gas, the macrostate is specified by macroscopic variables such as the temperature, density, and pressure.

In general, a macrostate is an equivalence class, i.e. a set containing some number of microstates (usually many, many microstates).

In the context of quantum mechanics, state always means microstate.

In the context of classical thermodynamics, state always means macrostate, for instance in the expression “function of state”.

The idea of microstate and the idea of macrostate are both quite useful. The problem arises when people use the word “state” as shorthand for one or both. You can get away with state=microstate in introductory quantum mechanics (no thermo), and you can get away with state=macrostate in introductory classical thermo (no quantum mechanics) ... but there is a nasty collision as soon as you start doing statistical mechanics, which sits astride the interface between QM and thermo.

In this document, the rule is that state means microstate, unless the context requires otherwise.

When we mean macrostate, we explicitly say macrostate or thermodynamic state. The idiomatic expression “function of state” necessarily refers to macrostate.

The relationship between microstate and macrostate, and their relationship to entropy, is discussed in section 2.7 and section 12.1.

Also, chapter 20 is a tangentially-related discussion of other inconsistent terminology.

## 12.2 What the Second Law Doesn't Tell You

Remember that the entropy is a property of the macrostate, not of any particular microstate. The macrostate is an ensemble of identically-prepared systems.

Therefore, if you are studying a single system, the second law doesn't necessarily tell you what that system is going to do. It tells you what an ensemble of such systems would do, but that's not necessarily the same thing. At this point there are several options. The ensemble average is the strongly recommended option.

1. For a large system, you may be able to use your imagination to divide the system into many pieces, so that you have an ensemble of pieces. Each of the  $N$  pieces is in equilibrium with a heat bath consisting of the other  $N - 1$  pieces. You can take the average over pieces and (maybe) substitute it for the ensemble average.
  - You must assume the system is homogeneous, or do a lot of extra work to account for any inhomogeneity.
  - This throws the baby out with the bathwater if you are interested in small systems.
2. Similarly, you can take the time-average and (maybe) substitute it for the ensemble average. The Poincaré recurrence theorem states that under some conditions, the time-average will give you the same answer as the ensemble average. Obviously this doesn't apply to time-dependent systems.
  - You must assume the system is ergodic and time-invariant.
  - This throws the baby out with the bathwater if you are interested in the time dependence.
3. We want the foundations of thermodynamics to apply to all systems, including small, time-dependent systems. Therefore have defined entropy in terms of an ensemble average. We do not depend on any time-averaging, we do not depend on ergodicity, we do not depend on any large- $N$  limit, and we do not depend on averaging over subsystems.

Note that energy is well defined for a single microstate but entropy is not. Entropy is a property of the macrostate. You may wish for more information about the microstate, but you won't get it, not from the second law anyway.

Given a probability distribution:

You can find the mean of the distribution.

However, the mean does not tell you everything there is to know.



You can find the standard deviation of the distribution.

However, the standard deviation does not tell you everything there is to know.

You can find the entropy of the distribution.

However, the entropy does not tell you everything there is to know.

Even those three things together do not tell you everything there is to know.

Note that distribution = ensemble = macrostate. The literature uses three words that refer to the same concept. That's annoying, but it's better than the other way around. (Using one word for three different concepts is a recipe for disaster.)

Suppose we have some source distribution, namely a distribution over some  $N$ -dimensional vector  $X$ . This could represent the positions and momenta of  $N/6$  atoms, or it could represent something else. Now suppose we draw one point from this distribution – i.e. we select one vector from the ensemble. We call that the sample. We can easily evaluate the sample-mean, which is just equal to the  $X$ -value of this point. We do not expect the sample-mean to be equal to the mean of the source distribution. It's probably close, but it's not the same.

Similarly we do not expect the sample-entropy to be the same as the source-distribution-entropy. Forsooth, the entropy of the sample is zero!

Given a larger sample with many, many points, we expect the sample-entropy to converge to the source-distribution-entropy, but the convergence is rather slow.

The discrepancy between the sample-mean and the source-dist-mean is not what people conventionally think of as a thermodynamic fluctuation. It's just sampling error. Ditto for the sample-entropy versus the source-dist-entropy. Fluctuations are dynamic, whereas you get sampling error even when sampling a distribution that has no dynamics at all. For example, given an urn containing colored marbles, the marbles are not fluctuating ... but different samples will contain different colors, in general.

For more about the crucial distinction between a distribution and a point drawn from that distribution, see reference 1, especially the section on sampling.

## 12.3 Phase Space

As mentioned in section 2.5.2, our notion of entropy is completely dependent on having a notion of microstate, and on having a procedure for assigning probability to microstates.

For systems where the relevant variables are naturally discrete, this is no problem. See section 2.2 and section 2.3 for examples involving symbols, and section 11.10 for an example involving real thermal physics.

We now discuss the procedure for dealing with continuous variables. In particular, we focus attention on the position and momentum variables.

It turns out that we must account for position and momentum *jointly*, not separately. That makes a lot of sense, as you can see by considering a harmonic oscillator with period  $\tau$ : If you know the oscillator's position at time  $t$ , you know its momentum at time  $t + \tau/4$  and vice versa.

Figure 12.1 shows how this works, in the semi-classical approximation. There is an abstract space called *phase space*. For each position variable  $q$  there is a momentum variable  $p$ . (In the language of classical mechanics, we say  $p$  and  $q$  are dynamically conjugate, but if you don't know what that means, don't worry about it.)

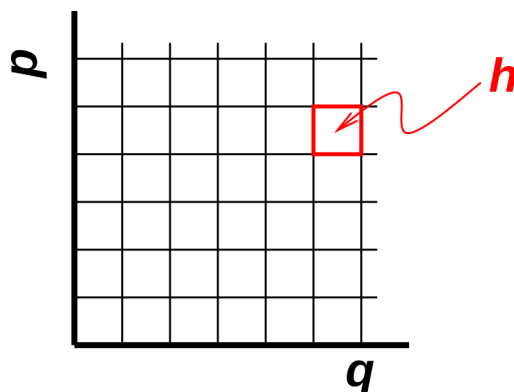


Figure 12.1: Phase Space

Area in phase space is called *action*. We divide phase space into cells of size  $h$ , where  $h$  is Planck's constant, also known as the *quantum of action*. A system has zero entropy if it can be described as sitting in a single cell in phase space. If we don't know exactly where the system sits, so that it must be described as a probability distribution in phase space, it will have some correspondingly greater entropy.

If you are wondering why each state has area  $h$ , as opposed to some other amount of area, see section 26.10.

If there are  $M$  independent position variables, there will be  $M$  momentum variables, and each microstate will be associated with a  $2M$ -dimensional cell of size  $h^M$ .

Using the phase-space idea, we can already understand, qualitatively, the entropy of an ideal gas in simple situations:

- If we keep the volume constant and increase the temperature, the entropy goes up. The spread in position stays the same, but the spread in momentum increases.
- If we keep the temperature constant and increase the volume, the entropy goes up. The spread in momentum stays the same, but the spread in position increases.

For a non-classical variable such as spin angular momentum, we don't need to worry about conjugate variables. The spin is already discrete i.e. quantized, so we know how to count states ... and it already has the right dimensions, since angular momentum has the same dimensions as action.

In chapter 2, we introduced entropy by discussing systems with only discrete states, namely rearrangements of a deck of cards. We now consider a continuous system, such as a collection of free particles. The same ideas apply.

For each continuous variable, you can divide the phase space into cells of size  $h$  and then see which cells are occupied. In classical thermodynamics, there is no way to know the value of  $h$ ; it is just an arbitrary constant. Changing the value of  $h$  changes the amount of entropy by an additive constant. But really there is no such arbitrariness, because “classical thermodynamics” is a contradiction in terms. There is no fully self-consistent classical thermodynamics. In modern physics, we definitely know the value of  $h$ , Planck’s constant. Therefore we have an absolute scale for measuring entropy.

As derived in section 26.2, there exists an explicit, easy-to-remember formula for the molar entropy of a monatomic three-dimensional ideal gas, namely the Sackur-Tetrode formula:

$$\frac{S/N}{k} = \ln\left(\frac{V/N}{\Lambda^3}\right) + \frac{5}{2} \quad (12.1)$$

where  $S/N$  is the molar entropy,  $V/N$  is the molar volume, and  $\Lambda$  is the thermal de Broglie length, i.e.

$$\Lambda := \sqrt{\left(\frac{2\pi\hbar^2}{mkT}\right)} \quad (12.2)$$

and if you plug this  $\Lambda$  into the Sackur-Tetrode formula you find the previously-advertised dependence on  $h^3$ .

You can see directly from equation 26.17 that the more spread out the gas is, the greater its molar entropy. Divide space into cells of size  $\Lambda^3$ , count how many cells there are per particle, and then take the logarithm.

The thermal de Broglie length  $\Lambda$  is very commonly called the thermal de Broglie *wavelength*, but this is something of a misnomer, because  $\Lambda$  shows up in a wide variety of fundamental expressions, usually having nothing to do with wavelength. This is discussed in more detail in reference 35.

## 12.4 Entropy in a Crystal; Phonons, Electrons, and Spins

Imagine a crystal of pure copper, containing only the  $^{63}\text{Cu}$  isotope. Under ordinary desktop conditions, most of the microscopic energy in the crystal takes the form of random potential and kinetic energy associated with vibrations of the atoms relative to their nominal positions in the lattice. We can find “normal modes” for these vibrations. This is the same idea as finding the normal modes for two coupled oscillators, except that this time we’ve got something like  $10^{23}$  coupled oscillators. There will be three normal modes per atom in the crystal. Each mode will be occupied by some number of phonons.

At ordinary temperatures, almost all modes will be in their ground state. Some of the low-lying modes will have a fair number of phonons in them, but this contributes only modestly to the entropy. When you add it all up, the crystal has about 6 bits per atom of entropy in the thermal phonons at room temperature. This depends strongly on the temperature, so if you cool the system, you quickly get into the regime where thermal phonon system contains much less than one bit of entropy per atom.

There is, however, more to the story. The copper crystal also contains conduction electrons. They are mostly in a low-entropy state, because of the exclusion principle, but still they manage to contribute a little bit to the entropy, about 1% as much as the thermal phonons at room temperature.

A third contribution comes from the fact that each  $^{63}\text{Cu}$  nucleus can be in one of four different spin states:  $+3/2$ ,  $+1/2$ ,  $-1/2$ , or  $-3/2$ . Mathematically, it's just like flipping two coins, or rolling a four-sided die. The spin system contains two bits of entropy per atom under ordinary conditions.

You can easily make a model system that has four states per particle. The most elegant way might be to carve some tetrahedral dice ... but it's easier and just as effective to use four-sided "bones," that is, parallelepipeds that are roughly 1cm by 1cm by 3 or 4 cm long. Make them long enough and/or round off the ends so that they never settle on the ends. Color the four long sides four different colors. A collection of such bones is profoundly analogous to a collection of copper nuclei. The which-way-is-up variable contributes two bits of entropy per bone, while the nuclear spin contributes two bits of entropy per atom.

In everyday situations, you don't care about this extra entropy in the spin system. It just goes along for the ride. This is an instance of *spectator* entropy, as discussed in section 12.6.

However, if you subject the crystal to a whopping big magnetic field (many teslas) and get things really cold (a few millikelvins), you can get the nuclear spins to line up. Each nucleus is like a little bar magnet, so it tends to align itself with the applied field, and at low-enough temperature the thermal agitation can no longer overcome this tendency.

Let's look at the cooling process, in a high magnetic field. We start at room temperature. The spins are completely random. If we cool things a little bit, the spins are still completely random. The spins have no effect on the observable properties such as heat capacity.

As the cooling continues, there will come a point where the spins start to line up. At this point the spin-entropy becomes important. It is no longer just going along for the ride. You will observe a contribution to the heat capacity whenever the crystal unloads some entropy.

You can also use copper nuclei to make a refrigerator for reaching very cold temperatures, as discussed in section 11.10.

## 12.5 Entropy is Entropy

Some people who ought to know better try to argue that there is more than one kind of entropy.

Sometimes they try to make one or more of the following distinctions:

Shannon entropy.

Thermodynamic entropy.

Entropy of abstract symbols.

Entropy of physical systems.

Entropy as given by equation 2.2 or equation 27.6.

Entropy defined in terms of energy and temperature.

Small systems: 3 blocks with  $5^3$  states, or 52 cards with  $52!$  states

Large systems:  $10^{25}$  copper nuclei with  $4^{10^{25}}$  states.

It must be emphasized that none of these distinctions have any value.

For starters, having two types of entropy would require two different paraconservation laws, one for each type. Also, if there exist any cases where there is some possibility of converting one type of entropy to the other, we would be back to having one overall paraconservation law, and the two type-by-type laws would be seen as mere approximations.

Also note that there are plenty of systems where there are two ways of evaluating the entropy. The copper nuclei described in section 11.10 have a maximum molar entropy of  $R \ln(4)$ . This value can be obtained in the obvious way by counting states, just as we did for the small, symbol-based systems in chapter 2. This is the same value that is obtained by macroscopic measurements of energy and temperature. What a coincidence!

Let's be clear: The demagnetization refrigerator counts *both* as a small, symbol-based system and as a large, thermal system. Additional examples are mentioned in chapter 22.

## 12.6 Spectator Entropy

Suppose we define a bogus pseudo-entropy  $S'$  as

$$S' := S + K \tag{12.3}$$

for some arbitrary constant  $K$ . It turns out that in some (but not all!) situations, you may not be sensitive to the difference between  $S'$  and  $S$ .

For example, suppose you are measuring the heat capacity. That has the same units as entropy, and is in fact closely related to the entropy. But we can see from equation 7.14 that the heat capacity is not sensitive to the difference between  $S'$  and  $S$ , because the derivative on the RHS annihilates additive constants.

Similarly, suppose you want to know whether a certain chemical reaction will proceed spontaneously or not. That depends on the difference between the initial state and the final state, that is, *differences* in energy and *differences* in entropy. So once again, additive constants will drop out.

There are many standard reference books that purport to tabulate the entropy of various chemical compounds ... but if you read the fine print you will discover that they are really tabulating the pseudo-entropy  $S'$  not the true entropy  $S$ . In particular, the tabulated numbers typically do not include the contribution from the nuclear spin-entropy, nor the contribution from mixing the various isotopes that make up each element. They can more-or-less get away with this because under ordinary chem-lab conditions those contributions are just additive constants.

However, you must not let down your guard. Just because you can get away with using  $S'$  instead of  $S$  in a few simple situations does not mean you can get away with it in general. As discussed in section 12.7 there is a correct value for  $S$  and there are plenty of cases where the correct value is needed.

## 12.7 No Secret Entropy, No Hidden Variables

Suppose we want to find the value of the true entropy,  $S$ . We account for the thermal phonons, and the electrons, and the nuclear spins. We even account for isotopes, chemical impurities, and structural defects in the crystal. But ... how do we know when to stop? How do we know if/when we've found all the entropy? In section 12.6 we saw how some of the entropy could silently go along for the ride, as a spectator, under certain conditions. Is there some additional entropy lurking here or there? Could there be hitherto-unimagined quantum numbers that couple to hitherto-unimagined fields?

The answer is no. According to all indications, there is no secret entropy. At any temperature below several thousand degrees, electrons, atomic nuclei, and all other subatomic particles can be described by their motion (position and momentum) and by their spin, but that's it, that's a complete description. Atoms, molecules, and all larger structures can be completely described by what their constituent particles are doing.

In classical mechanics, there could have been an arbitrary amount of secret entropy, but in the real world, governed by quantum mechanics, the answer is no.

We have a firm experimental basis for this conclusion. According to the laws of quantum mechanics, the scattering of indistinguishable particles is different from the scattering of distinguishable particles.

Therefore let's consider a low-energy proton/proton scattering experiment. We arrange that the protons are not distinguishable on the basis of position, or on any basis other than spin. That is, the protons are indistinguishable if and only if they have the same spin.

Next we randomize the spins, so that for each proton, each of the two spin states is equally likely. Our ignorance of the spin state contributes exactly 1 bit per particle to the entropy.

Now, to make things interesting, suppose that in addition to the aforementioned 1 bit of spin-entropy, each proton had 17 bits of "secret entropy", in whatever form you can imagine. That would mean that there would be  $2^{17}$  different *distinguishable* types of proton. If you pick protons at random, they would almost certainly be distinguishable, whether or not their spins were aligned, and you would almost never observe like-spin scattering to be different from unlike-spin scattering.

Such scattering experiments have been conducted with electrons, protons, various heavier nuclei, and sometimes entire atoms. There has never been any indication of any secret entropy.

The thermodynamics of chemical reactions tells us that larger structures can be described in terms of their constituents with no surprises.

The existence of superfluidity is further evidence that we can correctly account for entropy. All the atoms in the superfluid phase are described by a single quantum wavefunction. The entropy per atom is zero; otherwise it wouldn't be a superfluid. Superfluid  $^4\text{He}$  depends on the fact that all  $^4\text{He}$  atoms are absolutely totally indistinguishable – not distinguishable on the basis of position, spin, or any other quantum numbers. This is what we expected, based on two-particle scattering experiments, but the existence of superfluidity reassures us that we haven't overlooked anything when going from two particles to  $10^{23}$  particles.

Superfluidity occurs because certain identical-particle effects are cumulative and therefore have a spectacular effect on the entire fluid. Similar macroscopic identical-particle effects have been directly observed in  $^3\text{He}$ , spin-polarized monatomic hydrogen, sodium atomic gas, and other systems.

It might also be remarked that the existence of superconductors, semiconductors, metals, molecular bonds, and the periodic table of elements is strong evidence that electrons have no secret entropy. The existence of lasers is strong evidence that photons have no secret entropy.

I can't *prove* that no hitherto-secret entropy will ever be discovered. We might discover a new atom tomorrow, called *loonium*, which is exactly the same as helium except that for some reason it always obeys the distinguishable-particle scattering law when scattering against helium. This wouldn't be the end of the world; we would just postulate a new quantum number and use it to distinguish the two types of atom. All I can say is that loonium must be exceedingly rare; otherwise it would have been noticed.

Reminder: The foregoing discussion applies to “secret entropy” that might exist at room temperature or below, in analogy to spin entropy. In contrast we are not talking about the plethora of quantum numbers that are known to come into play at *higher* energies, but are all in their ground state under ordinary room-temperature conditions.

## 12.8 Entropy is Context Dependent

Consider 100 decks of cards. The first one is randomly shuffled. It has an entropy of just under 226 bits. All the rest are ordered the same way as the first. If you give me any one of the decks *in isolation*, it will take me 226 yes/no questions to figure out how to return the deck to standard order. But after I've seen any one of the decks, I know the exact microstate of every other deck without asking additional questions. The other 99 decks contain zero additional entropy.

In a situation like this, it's hard to consider entropy to be a state variable. In particular, the entropy density will not be an intensive property.

I know this sounds creepy, but it's real physics. Creepy situations like this do not *usually* occur in physical systems, but sometimes they do. Examples include:

- The spin-echo experiment (section 11.7) is the perfect example of this.
- Small thermodynamic systems, including Maxwell demons and Szilárd engines, are also excellent examples.
- There are many magic tricks that involve a deck of cards that is (or appears to be) completely disordered, yet important details of the configuration are known to the magician.
- Similarly, in cryptology, a string of symbols that is well encrypted will pass any standard test for randomness, and is therefore completely unpredictable to most parties ... yet it is highly predictable to parties who hold the key.

In an ordinary ideal gas, you can pretty much assume the entropy density is a well-behaved intensive property – but don't completely let down your guard, or you'll be badly fooled by the spin-echo setup.

A related issue concerns the dependence of entropy on the choice of observer. Entropy is not simply a property of a system, but rather a property of the system *and* the description thereof. This was mentioned in passing near the end of chapter 2.

Let's be clear: As a matter of principle, two different observers will in general assign two different values to “the” entropy.

This is easy to express in mathematical terms. The trustworthy workhorse formula for entropy is equation 2.2. If we have a conditional probability, things get slightly more complicated, as discussed in section 12.9.

*Human* observers are so grossly dissipative and usually “know” so little that it is academic to worry about the thermodynamics of *human* “knowledge.” However, the issue takes on new life when we consider tiny, highly-optimized robot measuring devices – Maxwell demons and the like.

For microscopic systems, it is for sure possible for different observers to report different values of “the” entropy (depending on what each observer knows about the system). The discrepancy can be a large percentage of the total.

By way of analogy, you know that different observers report different values of “the” kinetic energy (depending on the velocity of the observer), and this hasn't caused the world to end.

For macroscopic systems ( $10^{23}$  particles or thereabouts) it is uncommon for one observer to know  $10^{23}$  things that the other observer doesn't ... but even this is possible. The spin echo experiment is a celebrated example, as discussed in section 11.7.

Regardless of the size of the system, it is often illuminating to consider a complete thermodynamic *cycle*, such that all participants are returned to the same state at the end of the cycle. This de-emphasizes what the observers “know” and instead focuses attention on how they “learn” ... *and how they forget*. In more technical terms: this focuses attention on the observation/measurement process, which is crucial if you want a deep understanding of what entropy is and where it comes from. See reference 36 and reference 37.

In particular, at some point in each cycle the observer will have to forget previous information, to make room for the new information. This forgetting expels entropy, and at temperature  $T$  it dissipates energy  $TS$ .

To repeat: When evaluating “the” entropy, it is necessary to account for the information in the observer-system. In a closed cycle, this focuses attention on the observation and measurement process. If you don't do this, you will get the wrong answer every time when analyzing spin echo systems, Maxwell demons, Szilárd engines, reversible computers, et cetera.

## 12.9 Slice Entropy and Conditional Entropy

Suppose we have the following joint probability:

$$P[i, j] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (12.4)$$



From that we can form the marginal probabilities such as  $u = a + b$  and  $v = c + d$ :

$$P[i, j] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \vdots \quad \begin{bmatrix} u \\ v \end{bmatrix} \\ \dots \\ \begin{bmatrix} x & y \end{bmatrix} \quad (12.5)$$

To allow for the possibility that the probabilities are not normalized, we define

$$\begin{aligned} z &= u + v \\ &= x + y \end{aligned} \quad (12.6)$$

The conditional probabilities (conditioned on row number) are:

$$P[@i, j] = \begin{bmatrix} a/u & b/u \\ c/v & d/v \end{bmatrix} \quad (12.7)$$

Notation: The probability is conditioned on the variable marked with the “@” sign, as discussed in reference 1. Briefly, we can define this notation in terms of the older “|” bar notation:

$$\begin{aligned} P[@i, j](y, x) &\equiv P[j, @i](x, y) \\ &\equiv P[j|i](x, y) \end{aligned} \quad (12.8)$$

Then for each row  $y$ , we can define the *row entropy*:

$$S'[P[@i, j]](y) := -\sum_x P[@i, j](y, x) \log(P[@i, j](y, x)) \quad (12.9)$$

So in our example, the row entropy of the first row is:

$$S'[P[@i, j]](1) := -a/u \log(a/u) - b/u \log(b/u) \quad (12.10)$$

An analogous expression exists for the column entropy. The term *slice entropy* is meant to encompass row entropy, column entropy, and all other ways of slicing up the joint distribution. If the probability is a function of  $N$  variables, there will be  $2^N - 2$  ways of slicing things.

We can now define the full-blown conditional entropy as the weighted sum of the row entropies:

$$S[P[@i, j]] := \sum_y S'[P[@i, j]](y) P[i](y) \quad (12.11)$$

So in our example, the conditional entropy is:

$$\begin{aligned} S[P[@i, j]] &:= -a/z \log(a/u) - b/z \log(b/u) \\ &\quad -c/z \log(c/v) - d/z \log(d/v) \end{aligned} \quad (12.12)$$

Note that the row entropy is a function of row number, but the full conditional entropy is not. Both of them are, of course, a functional of the probability distribution.

## 12.10 Extreme Mixtures

In this section, we consider three possibilities:

- We could have a pure system. A good example is  $^4\text{He}$  gas. All  $^4\text{He}$  atoms are identical, in the strict quantum-mechanical sense.

- We could have an ordinary mixture with a small number of components, small compared to the number of particles in the sample. A good example is  $^3\text{He}$  gas. Each  $^3\text{He}$  atom has two possible states (associated with the orientation of its nuclear spin).
- We could have an extreme mixture, where the components of the mixture are drawn from some “universal set” such that the number of possibilities exceeds the number of particles in our sample, and – crucially – we don’t know which of the possible particles actually appear in our sample. We call this *snow*, since snowflakes are proverbially all different.

### 12.10.1 Simple Model System

Calculating the entropy of an extreme mixture is a bit tricky. It may help to use playing cards as a model system. There are two possible scenarios:

**Scenario 1A :** Suppose we deal out three cards, and then look at them to see WHICH cards they are. At this point, the entropy of the hand is zero, since we know the exact microstate.

Next we shuffle the cards within the hand, so that we no longer know the order of the cards. Now there is 2.6 bits of entropy, i.e.  $\log_2(3!)$ .

Note: In this section, all numerical values for the entropy are rounded to the nearest 0.1 bit.

Let’s call this the *entropy of permutation* of the hand. For some purposes we might consider this “the” entropy, but not for all purposes, as we shall see. As another way of saying the same thing: we have just calculated a *row entropy* (as defined in section 12.9 – conditioned on knowing WHICH three cards are in the hand).

**Scenario 1B:** I shuffle a new deck and deal three cards at random. Unlike in the previous scenario, we do not peek to see WHICH cards are in the hand. This hand contains 17 bits of entropy, i.e.  $\log_2(52 * 51 * 50)$ . This is the sum of

- 14.4 bits associated with WHICH cards got dealt, which we call *the entropy of the deal*, which we denote  $S_D$ ; and
- the aforementioned 2.6 bits representing the row entropy, conditioned on the deal, which we denote  $S_{|D}$ .

In this example,  $S_{|D}$  is just the entropy of permutation within the hand. (For another system, such as a sample of gas, we would need a more complicated expression for  $S_{|D}$ .)

Very commonly, the entropy of the deal is enormous compared to the entropy of permutation within the hand.

In general, the total system entropy is

$$S = S_D + S_{|D} \quad (12.13)$$

Reminder: The entropy of the deal, by definition, does not include the entropy of permutation within the sample.

Under mild conditions, the number of possible deals will be:

$$\begin{aligned} S_D &= \log(W_D) \\ &= \log(\text{MchooseN}) \\ &= \log\left(\frac{M!}{(M-N)!N!}\right) \end{aligned} \quad (12.14)$$

where  $W_D$  is the multiplicity of the deal,  $M$  is the number of particles in the universal set, and  $N$  is the number of particles that were selected (dealt) into the actual sample. Here we are assuming that the particles, although distinguishable, all behave very nearly the same, so that they have equal probability of being selected.

When the number of particles in the universal set (cards in the deck) is very large compared to the number of particles actually in the sample, equation 12.14 simplifies to:

$$S_D \approx \log\left(\frac{M^N}{N!}\right) \quad (\text{for } M \gg N) \quad (12.15)$$

### 12.10.2 Two-Sample Model System

We now consider a two-handed version of the same game. Once again, there are two scenarios, depending on whether we peek or not.

**Scenario 2A:** We deal two hands. We look to see which cards are in each hand. Then we shuffle each hand separately. The entropy of the each hand is 2.6 bits. This is the entropy of permutation within each hand. The two shuffles are statistically independent, so the entropy is additive. Therefore the entropy of the system as a whole is 5.2 bits.

Next, we put the two hands together and shuffle the six cards. Now the system entropy is 9.5 bits, i.e.  $\log_2(6!)$ .

The main result here is that the system entropy is increased by mixing the two subsystems. It has gone up from 5.2 to 9.5, for a gain of 4.3 bits. This newly created entropy, called the entropy of mixing, can be considered a miniature entropy of the deal, created by dealing two hands from the six-card mini-deck.

We now turn to an ugly and messy side-issue, which you can skip if you want. If we re-separate these card into two hands of three, it is not easy to decide how the system entropy gets distributed. Each player, considering his hand alone, evaluates the entropy as 6.9 bits, i.e.  $\log_2(6 * 5 * 4)$ , since he doesn't know which three of the six cards he's got, or what order they're in. That leaves the familiar 2.6 bits in the other player's hand.

This is another example of context-dependent entropy, as discussed in section 12.8. By symmetry, each player evaluates the entropy of his own hand as 6.9 bits and the entropy of the other player's hand as 2.6 bits. Each of these evaluations is consistent with itself, but not consistent with the other evaluation.

**Scenario 2B:** The same as above, but we don't peek so we don't know WHICH cards are in either hand.

Immediately after dealing the cards, the system entropy is 33.8 bits, i.e.  $\log_2(52 * 51 * 50 * 49 * 48 * 47)$ . If the cards were dealt in the usual way, we can say this entropy is distributed as

follows: There is 16.9 bits in the first hand, i.e.  $\log_2(52 * 50 * 48)$ , plus 16.8 bits in the other hand, i.e.  $\log_2(51 * 49 * 47)$ .

Next, we put the two hands together and shuffle the six cards. This does not increase the entropy. Before the shuffling, we knew nothing about which six cards were involved and knew nothing about the order, and after the shuffling we don't know any more or less. The system entropy is still 33.8 bits.

The main result is that in this scenario, mixing does not increase the entropy.

Once again we have a messy and ugly side issue: We now re-separate the cards into two hands of three. It is, alas, tricky to decide how the entropy is distributed. It *might* make sense to distribute the entropy evenly, by symmetry. Or it *might* make sense to say it was distributed as before, namely 16.9 plus 16.8. Or it might make sense for one player to decide that in the absence of information about the other player's hand, he should use the maximally-asymmetric worst-case estimate, namely 17 bits in his hand, i.e.  $\log_2(51 * 51 * 50)$ , plus 16.8 bits in the other hand i.e.  $\log_2(49 * 48 * 47)$ .

Note the contrast:

In scenario 2A, mixing increased the entropy. This can be explained by saying that we zeroed out the entropy of the deal by peeking, and then re-created a little bit of entropy of the deal by shuffling the six-card mini-deck and re-dealing.

In scenario 2B, mixing did not increase the entropy. The system already had the maximal amount of entropy of the deal, so shuffling the mini-deck and re-dealing could not possibly increase the entropy.

### 12.10.3 Helium versus Snow

We now consider a contrast along a different axis, namely the contrast between a pure substance and an extreme mixture:

For a sample of helium, the entropy of the deal is zero. That's because all  $^4\text{He}$  atoms are identical. One sample of  $N$  atoms of helium is just like any other sample of  $N$  atoms of helium.

For a sample of gas where the particles are all different, like the proverbial snowflakes, the total entropy necessarily includes the entropy of the deal. If you have  $N$  snowflakes, in general it matters WHICH  $N$  snowflakes they are. If you have two such samples, each individual flake is different, and each sample as a whole is different. The two samples may be equivalent as to various macroscopic average properties, but from the point of view of the foundations of thermodynamics, the two samples are very different. They are different in a way that two samples of helium are not.

If we perform a Gibbs-type mixing experiment with helium, mixing does not increase the entropy. This makes sense, because all helium atoms are identical.

If

we perform a Gibbs-type mixing experiment with snow, there are two possibilities. In scenario 2A

(as discussed in section 12.10.2), mixing increases the entropy. In this scenario the entropy of the deal had been removed, and mixing re-created a certain amount of entropy of the deal.

Meanwhile, in scenario *2B*, mixing does not increase the entropy.

Note that the non-increase of entropy for mixing snow in scenario *2B* is not equivalent to the non-increase for mixing helium; they occur for different reasons.

#### 12.10.4 Partial Information aka Weak Peek

There is yet another scenario that is worth considering, namely where we have only partial information about the particles in each sample. Roughly speaking, imagine mixing red snow with blue snow, or mixing chocolate milk with regular milk.

Again we use playing cards as our model system.

**Scenario 2C :** We deal three black cards to one player, and deal three red cards to the other player. The entropy of each hand is a little over 13.9 bits, i.e.  $\log_2(26 * 25 * 24)$  ... of which 11.3 bits is the entropy of the deal and 2.6 bits is the familiar entropy of permutation within the hand. The hands are statistically independent, so the system entropy is just twice that, namely 27.9 bits, which breaks down as 22.7 (deal) plus 5.2 (two permutations).

The information about the coloration of each hand can be considered partial information. This information could come from any of several possible sources, including a weak peek, i.e. a peek that ascertains the color without ascertaining the exact suit or rank.

If we put the two hands together and shuffle the six cards, the system entropy increases to 32.2 bits, which breaks down as 22.7 for the entropy of the original deal (unchanged), plus 9.5 for the entropy of permutation of the six cards, i.e.  $\log_2(6!)$ . So the entropy-gain due to mixing is 4.3 bits.

Not coincidentally, this is the exact-same entropy gain as we saw in scenario *2A*. The change is the same, even though the system entropy of the new scenario (*2C*) is greater. In both of these scenarios, the entropy of mixing is associated with not knowing which of the two original hands contributed a given card to the final hand.

Again it is tricky to distribute the system entropy. If each player adopts the worst-case maximum entropy strategy, he will attribute 17.0 bits of entropy to his hand, i.e.  $\log_2(52 * 51 * 50)$ , since he really has no idea what cards are in his hand. This leaves 15.2 bits in the other hand. This is another situation where each player's evaluation is internally consistent, but inconsistent with the other player's evaluation.

### 12.11 Entropy is Not Necessarily Extensive

People commonly think of entropy as being an extensive quantity. This is true to a good approximation in many situations, but not all.

As an extreme example of non-extensive entropy, consider the situation described at the beginning of section 12.8, where we have one randomly-shuffled deck plus some number of clones. Whichever one of these decks we look at first will look completely random, but the total entropy of the system does not grow in proportion to the number of decks; indeed it does not grow at all as we add more decks.

Even for something as simple as a pure monatomic ideal gas, the entropy is “almost” extensive, but not quite, as discussed in connection with equation 26.17. Similarly, for a gas mixture, see equation 26.21.

The idea that entropy might be non-extensive should not come as any big surprise. The energy of a liquid or solid is not exactly extensive either, because of things like surface tension and surface reconstruction.

Note that in a Gibbs-type mixing experiment, if you start out with helium on one side and neon on the other, the entropy of the mixture is not the sum of the two original entropies. Not even close. This is because there is a large entropy of mixing. In this situation we might say the entropy is grossly non-extensive (if it even makes sense to talk about extensivity in such a situation).

On the other hand, if we start out with a mixture of helium and neon on one side, and the same mixture on the other side, when we pull out the partition, there is “almost” no change in entropy, in accordance with equation 26.21.

Similarly, even for an extreme mixture, i.e. snow on one side and snow on the other, the entropy is extensive (or nearly so) ... provided we account for the entropy of the deal. Beware that the entropy of the deal is often neglected, even though it is enormous.

## 12.12 Mathematical Properties of the Entropy

### 12.12.1 Entropy Can Be Infinite

Let’s see what it takes to construct a probability distribution that has infinite entropy.

For simplicity, let’s consider a discrete distribution, with discrete microstates ... such as a distribution over the integers. (A continuous distribution would make the problem more complicated but not in any way more interesting.)

There are two ways to proceed. Let’s deal with the simple, not-so-interesting case first. Imagine tossing a very large number of coins. The entropy per coin ( $s$ ) is intensive and is the same everywhere, namely 1 bit per coin. The total entropy ( $S$ ) is extensive, and grows in proportion to the number of coins. This situation is easy to understand in the case of a large but finite number of coins, but things go haywire if we ask about a truly infinite number of coins, because then there would be an infinite number of microstates with zero probability apiece. Zero probability is not good.

We can apply the same logic to the physics of a large chunk of metal. The molar entropy,  $s$ , is intensive and is the same everywhere. The total entropy,  $S$ , is extensive and grows in proportion to the size of the chunk. Again, this is easy to understand for a large but finite chunk, but things

to haywire if we ask about a truly infinite chunk, because then every microstate would have zero probability.

So let's not approach the problem that way. Instead we insist on a well behaved probability distribution, such that

$$p_i > 0 \quad \text{for all integers } i \geq j \quad (12.16a)$$

$$\sum p_i = 1 \quad (12.16b)$$

where  $j$  is some convenient starting-point. The goal of infinite entropy means that

$$\sum p_i \ln(1/p_i) = \infty \quad (12.17)$$

Finding a suitable distribution is going to be a challenge, because we want the series in equation 12.16b to converge, but we want the closely-related series in equation 12.17 to diverge. This is possible, but it means that one series will “just barely” converge while the other will “just barely” diverge.

Let's see whether the probability defined by equation 12.18 does what we want.

$$\begin{aligned} q_i &= \frac{1}{i (\ln i)^h} & \text{for all } i \geq 3 \\ Z &= \sum q_i \\ p_i &= \frac{q_i}{Z} \end{aligned} \quad (12.18)$$

As an aid to understanding what's going on, consider the series in equation 12.19. Using the integral test, you can easily show that the series converges for all  $h$  strictly greater than 1, and diverges to  $+\infty$  otherwise.

$$Z = \sum_{i=3}^{\infty} \frac{1}{i (\ln i)^h} \quad (12.19)$$

Returning to equation 12.18, let's assume  $h$  is greater than one. For definiteness, you can take  $h = 1.5$  if you wish. This guarantees that the probability is well-behaved as specified in equation 12.16.

We can calculate a lower bound on the entropy as follows:

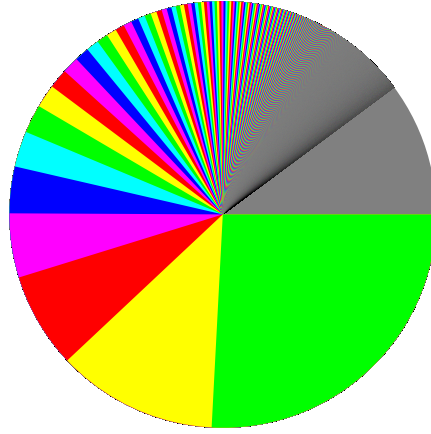
$$\begin{aligned} S &= \sum p_i \ln(1/p_i) \\ &= \sum \frac{\ln(Z) + \ln(i) + h \ln(\ln i)}{Z i (\ln i)^h} \\ &> \sum \frac{\ln(Z) + \ln(i)}{Z i (\ln i)^h} \\ &> \text{const} + \sum \frac{1}{Z i (\ln i)^{(h-1)}} \\ &= \infty \end{aligned} \quad (12.20)$$

provided  $h > 1$  and  $h \leq 2$

The pie chart for this distribution is shown (somewhat imperfectly) in figure 12.2. You can see that the series converges quite slowly. In fact, the first 1000 terms cover only about 86% of the total pie when  $h = 2.0$ . Smaller values of  $h$  give even slower convergence.

I can't think of any physics situations where there are countably many states, each with a positive amount of probability ... but such situations are routine in data compression, communications, and cryptography. It's an argument for having variable-length codewords. There are infinitely many different messages that could be sent. A few of them are very common and should be encoded with short codewords, while most of the rest are very very unlikely, and should be encoded with much longer codewords.

This is a way of driving home the point that entropy is a property of the distribution.

Figure 12.2: Distribution with Infinite Entropy ( $h = 2.0$ )

- You can talk about the mean of the distribution.
- You can talk about the second moment of the distribution.
- You can talk about the fourth moment of the distribution.
- You can talk about the *entropy* of the distribution.
- Et cetera.

The idea of entropy applies to any distribution, not just thermal-equilibrium distributions.

Here is another example of a countable, discrete distribution with infinite entropy:

$$\begin{aligned}
 q_i &= \frac{1}{\ln(i)} - \frac{1}{\ln(i+1)} && \text{for all } i \geq 3 \\
 Z &= \sum q_i \\
 &= \frac{1}{\ln(3)} \\
 p_i &= \frac{q_i}{Z}
 \end{aligned}
 \tag{12.21}$$

This has the advantage that we can easily calculate the numerical value of  $Z$ . If you approximate this distribution using a Taylor series, you can see that for large  $i$ , this distribution behaves similarly to the  $h = 2$  case of equation 12.18. This distribution is discussed with more formality and more details in reference 38.



## Chapter 13

# Temperature : Definition and Fundamental Properties

The first goal for today is to define what we mean by temperature. In the process, we will discover that *equilibrium is isothermal*. That is, when two objects are in contact (so that they can exchange energy), then in equilibrium they have the same temperature. This is one of the big reasons why people care about temperature. It could be argued that the notion of temperature is constructed and defined in such a way as to guarantee that equilibrium is isothermal.

We follow the same approach as reference 39. See especially figure 3 therein. We shall use these ideas in section 14.4 and elsewhere.

### 13.1 Example Scenario: Two Subsystems, Same Stuff

In our first scenario, suppose we have two subsystems.

- Subsystem #1 has  $N_1 = 48$  spins.
- Subsystem #2 has  $N_2 = 24$  spins.

Each spin has  $-1$  units of energy when it is in the DOWN state and  $+1$  units of energy when it is in the UP state.

The two subsystems are able to exchange energy (and spin angular momentum) with each other, but the system as a whole is isolated from the rest of the universe.

Within limits, we can set up initial conditions with a specified amount of energy in each subsystem, namely  $E_1$  and  $E_2$ . We can then calculate the entropy in each subsystem, namely  $S_1$  and  $S_2$ . By considering various ways of distributing the energy and entropy, we can figure out which distribution corresponds to thermal equilibrium.

In particular, the subsystems reach equilibrium by exchanging energy with each other, under conditions of constant total energy  $E = E_1 + E_2$ . Therefore, at any given value of  $E$ , we can keep track

of the equilibration process as a function of  $E_1$ . We can calculate  $E_2$  as a function of  $E$  and  $E_1$ . Then we can calculate everything else we need to know as a function of  $E_1$  and  $E_2$ .

This equilibration process is diagrammed in figure 13.1, for the case where  $E = -60$ . Subsequent figures show the same thing for other amounts of energy. We pick eight different values for the total system energy, and calculate everything else accordingly. The method of calculation is discussed in section 13.5.

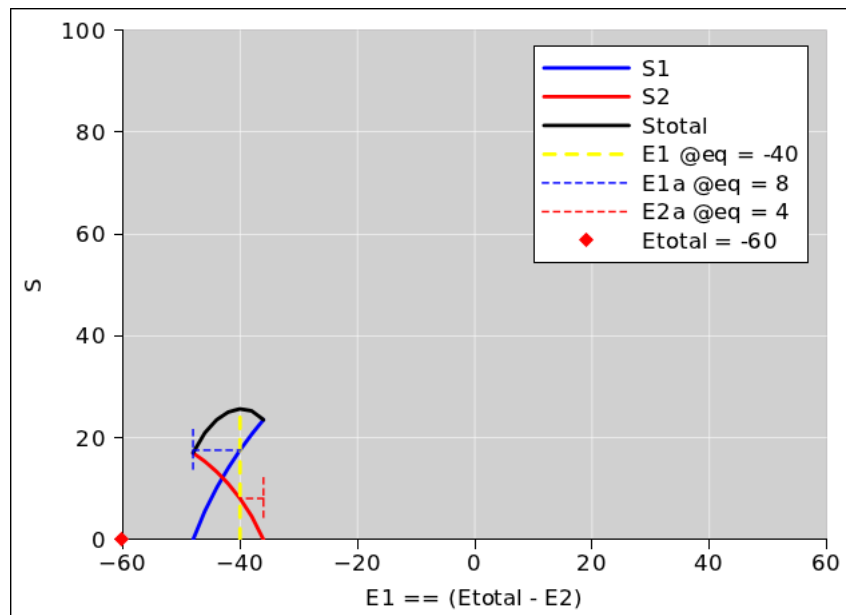
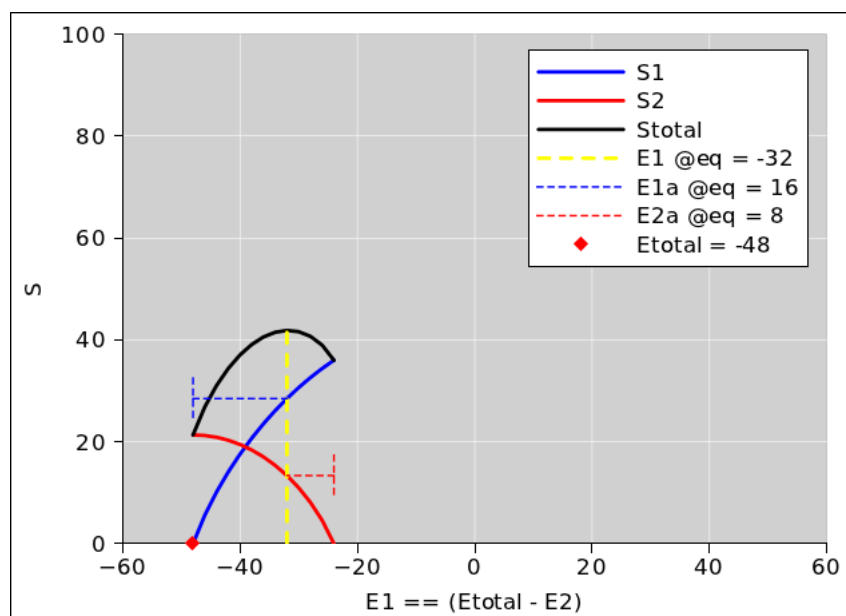
In each of the figures, there is quite a bit of information:

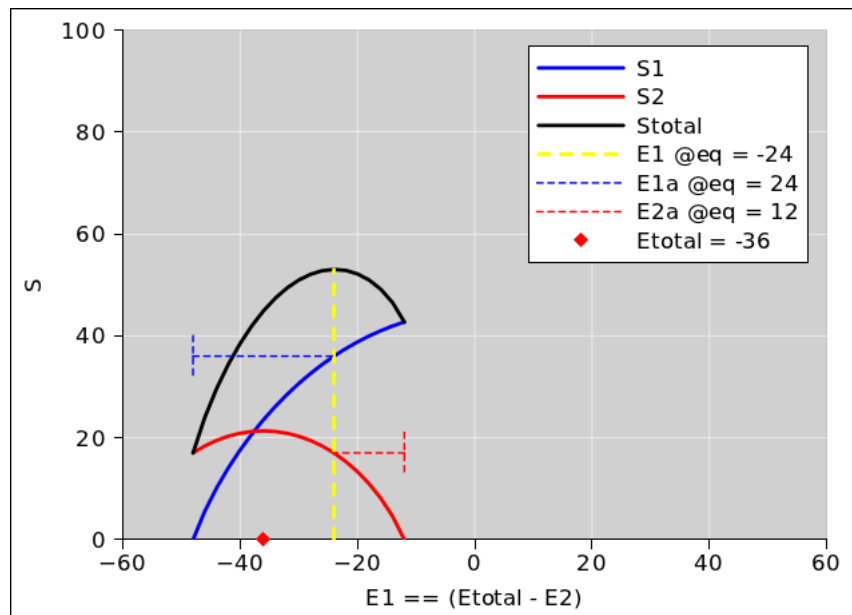
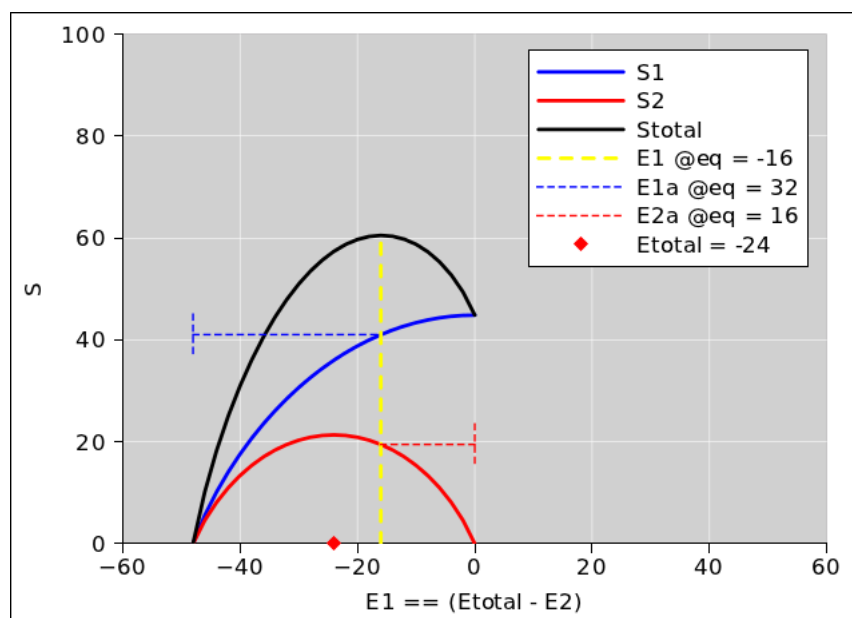
- The blue curve indicates  $S_1$ , the entropy of system #1, as a function of  $E_1$ .
- The red curve indicates  $S_2$ , the entropy of system #2, as a function of  $E_1$ . You might have expected  $S_2$  to be plotted as a function of  $E_2$ , but for present purposes it is more useful to plot everything on a common abscissa, and we choose  $E_1$  for this.
- The black curve indicates the entropy of the system as a whole,  $S = S_1 + S_2$ , plotted as a function of  $E_1$ .
- The equilibrium situation is indicated by a vertical dashed yellow line. You can see that the total entropy is maximal at this point, which corresponds to  $dS/dE = 0$ .
- The horizontal dashed blue line indicates  $E_{1a}$ , the amount of energy *above* the ground state in subsystem #1. The “a” in  $E_{1a}$  stands for “above.” Note that  $E_1$  and  $E_{1a}$  increase left-to-right in the diagram, in the conventional way.
- The horizontal dashed red line indicates  $E_{2a}$ , the amount of energy *above* the ground state in subsystem #2. Note that  $E_2$  and  $E_{2a}$  increase right-to-left in the diagram, which makes sense in this case, even though it is unconventional.

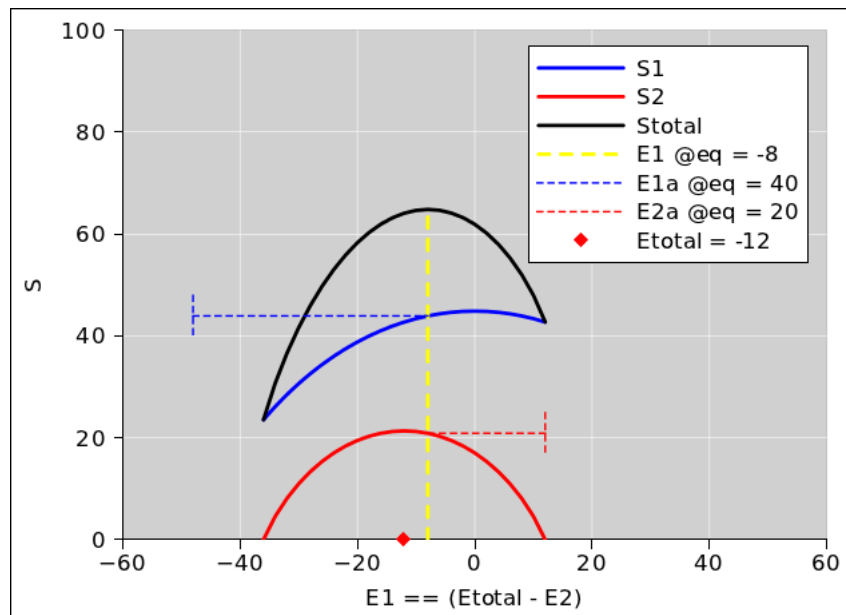
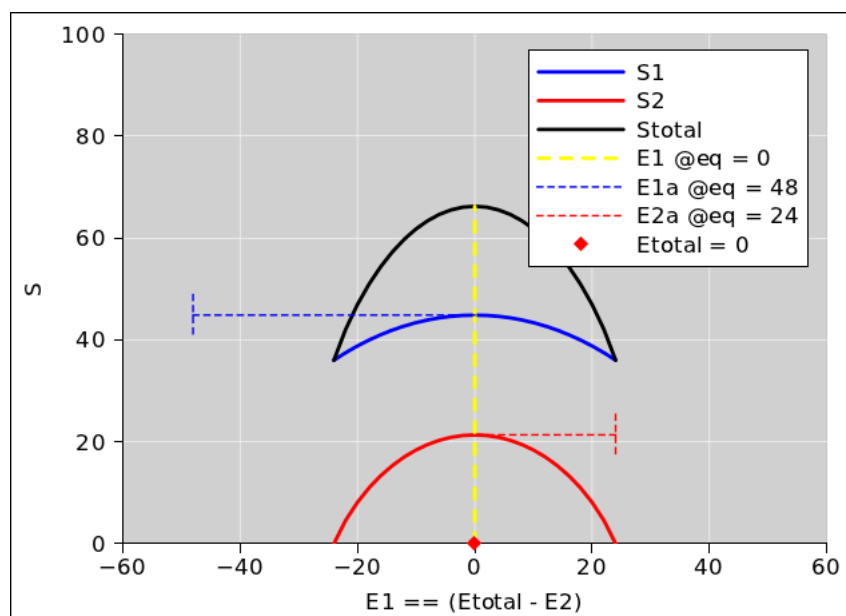
It is important to notice that the red curve plus the blue curve add up to make the black curve, everywhere. Therefore the slope of the red curve plus the slope of the blue curve add up to make the slope of the black curve, everywhere. At equilibrium, the slope of the black curve is zero, so the slope of the other two curves must be equal and opposite. You can see this in the graphs, at the places where the curves cross the yellow line.

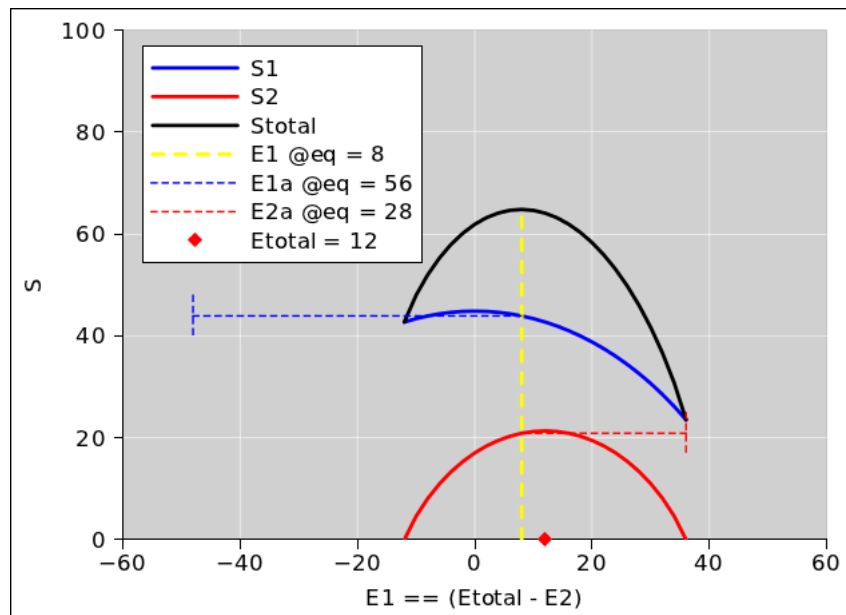
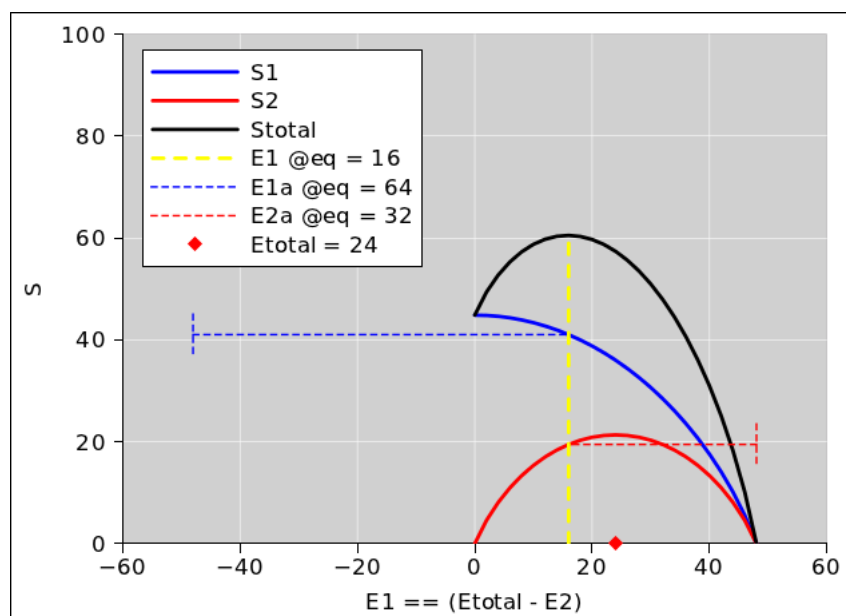
For the blue curve, the slope is:

$$\beta_1 := \left. \frac{\partial S_1}{\partial E_1} \right|_{N,V} \quad (13.1)$$

Figure 13.1: Entropy versus  $E_1$ , at constant  $E = -60$ Figure 13.2: Entropy versus  $E_1$ , at constant  $E = -48$

Figure 13.3: Entropy versus  $E_1$ , at constant  $E = -36$ Figure 13.4: Entropy versus  $E_1$ , at constant  $E = -24$

Figure 13.5: Entropy versus  $E_1$ , at constant  $E = -12$ Figure 13.6: Entropy versus  $E_1$ , at constant  $E = 0$

Figure 13.7: Entropy versus  $E_1$ , at constant  $E = +12$ Figure 13.8: Entropy versus  $E_1$ , at constant  $E = +24$

which depends only on properties of subsystem #1. As plotted in the figure, the slope of the red curve is  $\frac{\partial S_2}{\partial E_1}$ , which is somewhat interesting, but in the long run we will be much better off if we focus attention on something that depends only on properties of subsystem #2, namely:

$$\beta_2 := \left. \frac{\partial S_2}{\partial E_2} \right|_{N,V} \quad (13.2)$$

We now make use of the fact that the system is isolated, so that  $dE = 0$ , and make use of conservation of energy, so that  $E = E_1 + E_2$ . Plugging this in to the definition of  $\beta_2$ , we find

$$\begin{aligned} \beta_2 &:= \left. \frac{\partial S_2}{\partial E_2} \right|_{N,V} \\ &= - \left. \frac{\partial S_2}{\partial E_1} \right|_{N,V} \\ &= \text{negative slope of red curve} \end{aligned} \quad (13.3)$$

Therefore, when we observe that the slope of the blue curve and the slope of the red curve are equal and opposite, it tells us that  $\beta_1$  and  $\beta_2$  are just plain equal (and not opposite).

These quantities are so important that the already has a conventional name:  $\beta_1$  is the *inverse temperature* of subsystem #1. Similarly  $\beta_2$  is the *inverse temperature* of subsystem #2.

### 13.1.1 Equilibrium is Isothermal

The punch line is that when two subsystems have reached equilibrium by exchanging energy, they will be at the same temperature. We have just explained why this must be true, as a consequence of the definition of temperature, the definition of equilibrium, the law of conservation of energy, and the fact that the system is isolated from the rest of the world.

Equilibrium is isothermal.

Experts note: The black curve measures  $\partial S/\partial E_1$  not  $\partial S/\partial E$ , so it cannot serve as a definition of temperature. Not even close. If we want to ascertain the temperature of the system, it usually suffices to measure the temperature of some subsystem. This is the operational approach, and it almost always makes sense, although it can get you into trouble in a few extreme cases. Hint: make sure there are at least two subsystems, each of which is big enough to serve as a heat sink for the other.

## 13.2 Remarks about the Simple Special Case

The remarks in this section apply only in the special case where one subsystem is twice as large as the other, and the two subsystems are made of the same kind of stuff. (The case where they are made of different stuff is more interesting, as discussed in section 13.3.)

You can see from the length of the horizontal dashed lines that at equilibrium, the blue subsystem has 2/3rds of the energy while the red subsystem has 1/3rd of the energy. This makes sense, since the blue system is twice as large, and the two subsystems are made of the same kind of stuff.

Meanwhile, you can see from the height of the horizontal dashed lines that at equilibrium, the blue subsystem has 2/3rds of the entropy while the red subsystem has 1/3rd of the entropy. Again this is unsurprising.

The vertical “tail” on the blue dashed line serves to indicate the  $E_1$  value that corresponds to  $E_{1a} = 0$ . Similarly, the vertical “tail” on the red dashed line serves to indicate the  $E_1$  value that corresponds to  $E_{2a} = 0$ .

Also, turning attention to  $E_1$  rather than  $E_{1a}$ , you can see from the position of the yellow dashed line that  $E_1$  is 2/3rds of the total  $E$ , as shown by the red diamond, which represents total  $E$  even though it is plotted on the nominal  $E_1$  axis.

### 13.3 Two Subsystems, Different Stuff

In this section we consider a new scenario. It is the same as the previous scenario, except that we imagine that the spins in the red subsystem have only half as much magnetic moment (or are sitting in half as much magnetic field). That means that the amount of energy that flips one spin in the blue subsystem will now flip *two* spins in the red subsystem.

We also double the number of spins in red subsystem, so that its maximum energy and minimum energy are the same as before.

Even though the red system cannot hold any more energy than it could before, it is now markedly more effective at attracting energy.

In this scenario, the two subsystems do not share the energy in a simple 1/3rd, 2/3rds fashion. At low temperatures, the red system is much more aggressive than the blue system, and soaks up more than its “share” of the energy, more than you would have predicted based on its physical size or on the maximal amount of energy it could hold.

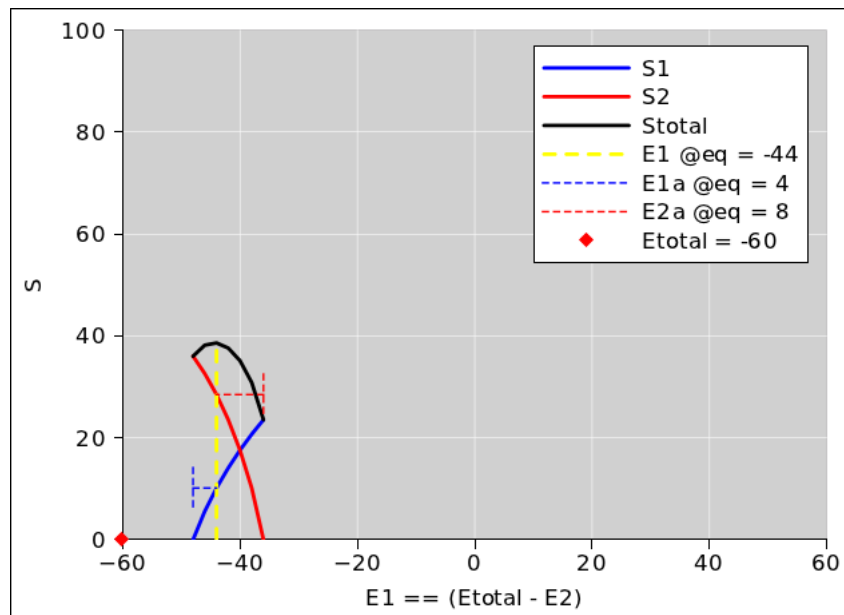
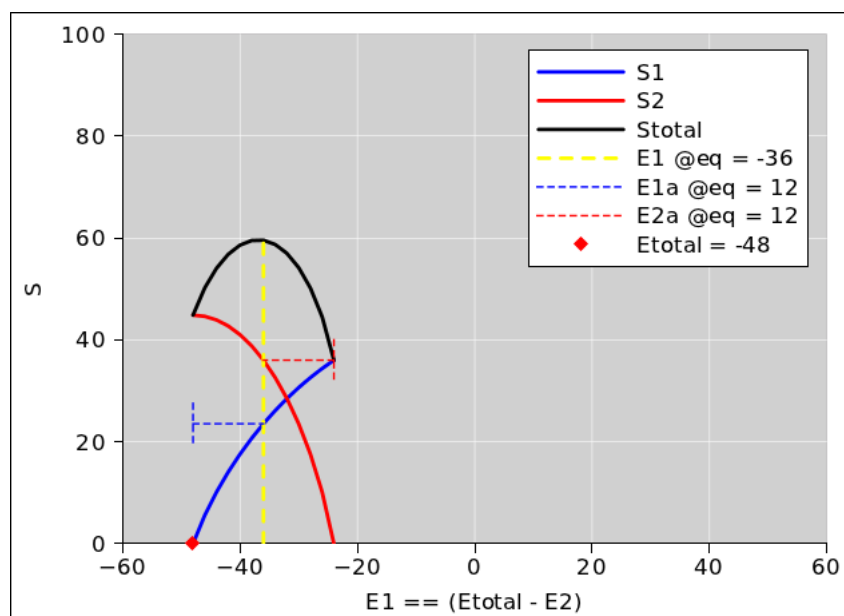
This illustrates an important point: All microstates are equally probable.

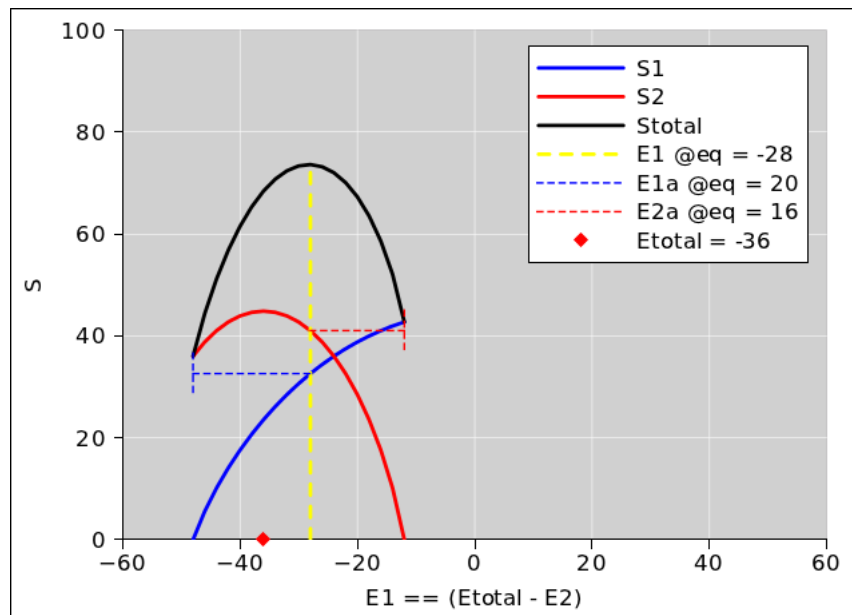
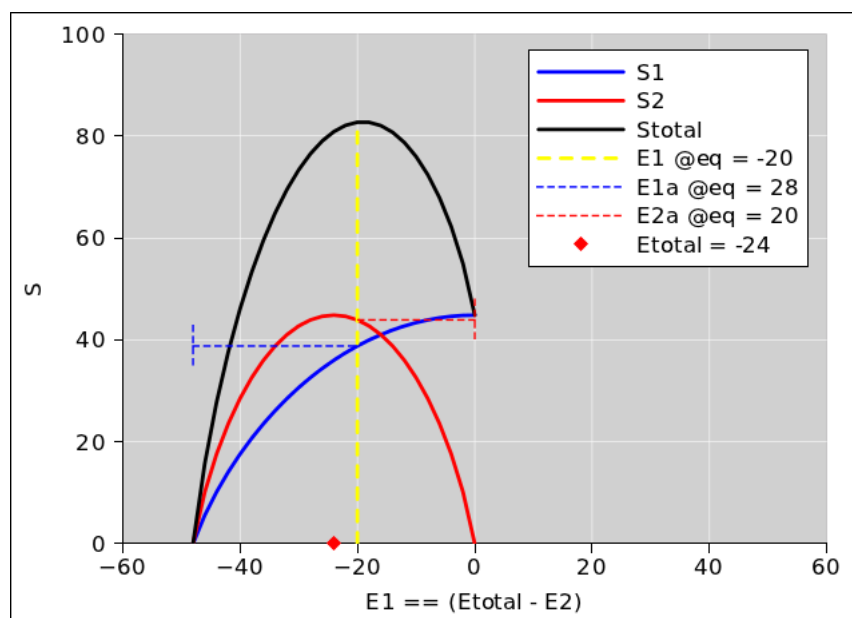
This stands in contrast to an oft-mentioned notion, namely the so-called principle of “equipartition of energy.” It is simply not correct to think that energy is equally distributed per unit mass or per unit volume or per atom or per spin. The fact is that probability (not energy) is what is getting distributed, and it gets distributed over microstates.

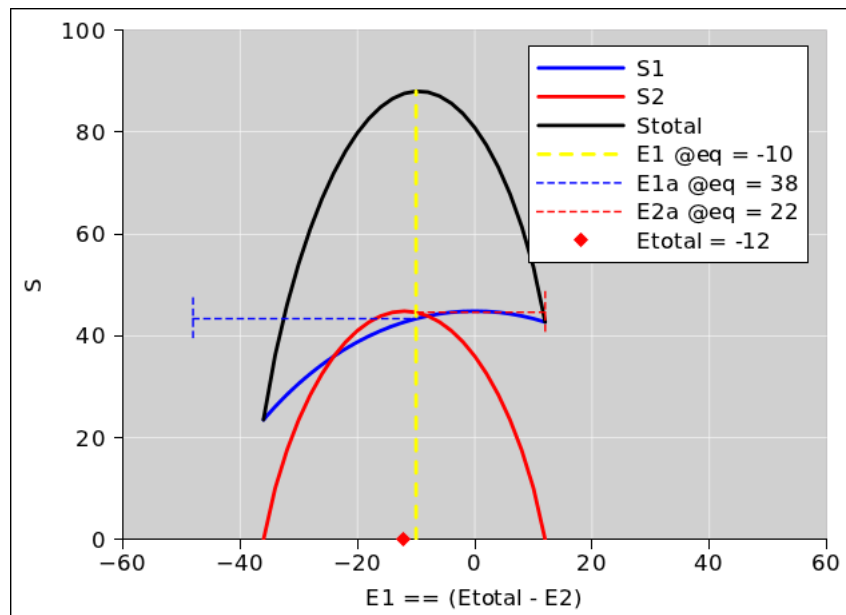
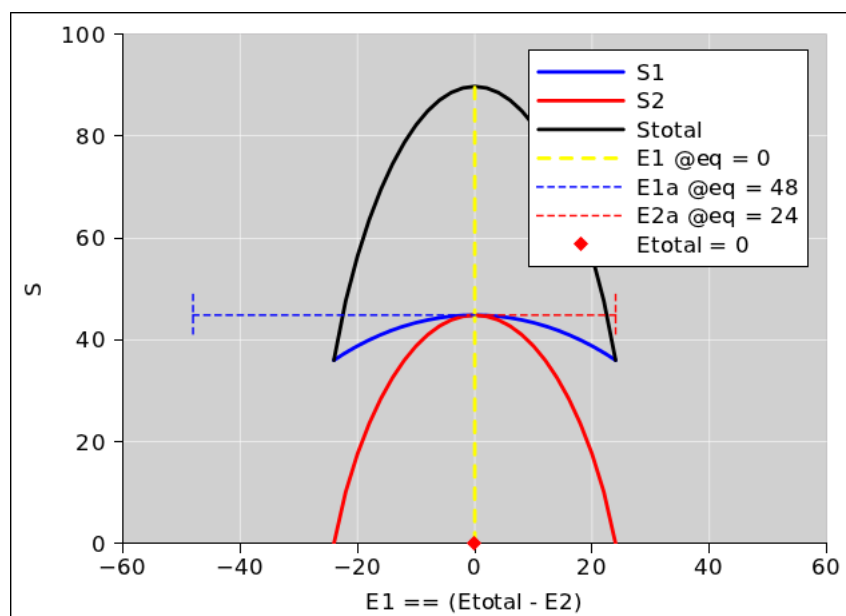
In the scenario considered in this section, the red system has more microstates, so it has more probability. As a consequence, it soaks up more energy, disproportionately more, as you can see by comparing the figures in this section with the corresponding figures in section 13.1. Be sure to notice the red and blue dashed horizontal lines.

In this scenario, equilibrium is isothermal ... as it must be, in any situation where subsystems reach equilibrium by exchanging energy. As a consequence, at equilibrium, the red slope and the blue slope are equal and opposite, as you can see in the diagrams.



Figure 13.9: Entropy versus Energy :  $E_{\text{total}} = -60$ Figure 13.10: Entropy versus Energy :  $E_{\text{total}} = -48$

Figure 13.11: Entropy versus Energy :  $E_{\text{total}} = -36$ Figure 13.12: Entropy versus Energy :  $E_{\text{total}} = -24$

Figure 13.13: Entropy versus Energy :  $E_{\text{total}} = -12$ Figure 13.14: Entropy versus Energy :  $E_{\text{total}} = 0$

## 13.4 Discussion: Constants Drop Out

It should be noted that temperature is not the same as energy. It's not even the same as energy per unit volume or energy per unit mass. Dimensionally, it's the same as energy per unit entropy, but even then it's not the ratio of gross energy over gross entropy. In fact, temperature is the *slope*, namely  $T_1 = \partial E_1 / \partial S_1 | N_1, V_1$ .

That means, among other things, that constants drop out. That is to say, if you shift  $E_1$  by a constant and/or shift  $S_1$  by a constant, the temperature is unchanged.

As a specific example, suppose you have a box of gas on a high shelf and a box of gas on a low shelf. You let them reach equilibrium by exchanging energy. They will have the same temperature, pressure, et cetera. The box on the high shelf will have a higher gravitational potential energy, but that will not affect the equilibrium temperature at all.

By the same token, a wound-up spring will be in thermal equilibrium with an unstressed spring at the same temperature, and a charged-up capacitor will be in thermal equilibrium with an uncharged capacitor at the same temperature. The same goes for kinetic energy: A rapidly spinning flywheel will be in thermal equilibrium with a stationary flywheel at the same temperature. It's not the energy that matters. It's the slope  $\partial E / \partial S$  that matters.

It is common knowledge that a parcel of air high in the atmosphere will be colder than a parcel of air at lower altitude. That tells us the atmosphere is not in thermal equilibrium. The temperature profile of the troposphere is more nearly adiabatic than isothermal, because it is vigorously stirred. Thunderstorms contribute quite a lot to the stirring, and it is no accident that the height of a typical thunderstorm is comparable to the altitude of the tropopause.

## 13.5 Calculations

We can easily compute the spin entropy as a function of energy, using the obvious combinatoric formula.

$$\begin{aligned} M_1 &= (E_1 + N_1)/2 &= \text{number of UP spins} \\ W_1 &= \binom{N_1}{M_1} &= \text{multiplicity} \\ S_1 &= \log(W_1) &= \text{entropy} \end{aligned} \tag{13.4}$$

Note that the binomial coefficient  $\binom{N}{m}$  is generally pronounced “ $N$  choose  $m$ ”.

It is implemented in typical spreadsheet programs by the `combin(N,m)` function.

One tricky task is calculating the starting point and ending point of each of the curves in the diagrams. This task is not trivial, and can be understood with the help of figure 13.15. The colored rectangles represent the feasible ways in which energy can be allocated to the two subsystems. Each black line is a contour of constant total energy  $E$ , where  $E = E_1 + E_2$ . As you can see, depending on  $E$  and on the amount of energy that each subsystem can hold, there are at least 13 different ways in which the available energy can cross the boundaries of the feasible region.

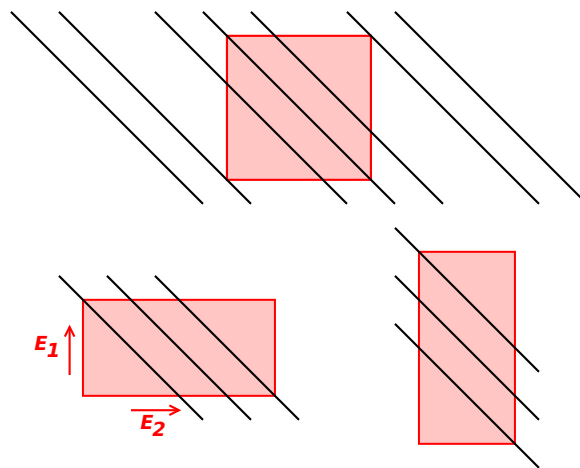


Figure 13.15: Starting Points and Ending Points

By diagramming the task in this way, we reduce it to a problem in computer graphics, for which well-known solutions exist. It pays to code this systematically; otherwise you'll spend unreasonable amounts of time debugging a bunch of special cases.

The spreadsheet used to produce the diagrams is available; see reference 40.

## 13.6 Chemical Potential

In this chapter we have demonstrated that

- Whenever the two parcels have reached equilibrium by exchanging energy, they will have the same temperature.

As we shall see in section 14.4, a closely parallel argument demonstrates that

- Whenever the two parcels have reached equilibrium by exchanging particles as well as energy, they will have the same chemical potential (and the same temperature).

## Chapter 14

# Spontaneity, Reversibility, Equilibrium, Stability, Solubility, etc.

### 14.1 Fundamental Notions

#### 14.1.1 Equilibrium

See section 10.1 for an introduction to the fundamental concept of equilibrium.

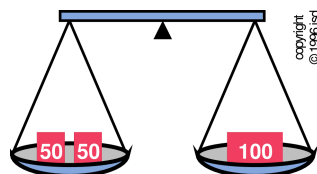


Figure 14.1: Equilibrium – Forces in Balance

#### 14.1.2 Stability

See section 10.5 for an introduction to the fundamental concept of stability.

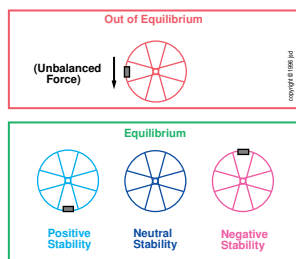


Figure 14.2: Equilibrium and Stability

### 14.1.3 A First Example: Heat Transfer

Suppose we have an object at a temperature  $T_2$  and we want to transfer some so-called “heat” to another object at some lesser temperature  $T_1$ . We assume both temperatures are positive. As usual, trying to quantify “heat” is a losing strategy; it is easier and better to formulate the analysis in terms of *energy* and *entropy*.

Under the given conditions we can write

$$\begin{aligned} dE_1 &= T_1 dS_1 \\ dE_2 &= T_2 dS_2 \end{aligned} \tag{14.1}$$

In more-general situations, there would be other terms on the RHS of such equations, but for present purposes we require all other terms to be negligible compared to the  $TdS$  term. This requirement essentially defines what we mean by *heat transfer* or equivalently *thermal transfer of energy*.

By conservation of energy we have

$$\begin{aligned} dE_1 + dE_2 &= 0 \\ dE_1 &= -dE_2 \end{aligned} \tag{14.2}$$

One line of algebra tells us that the total entropy of the world changed by an amount

$$dS_1 + dS_2 = (1/T_1 - 1/T_2)dE_1 \tag{14.3}$$

or equivalently

$$\begin{aligned} dS_1 + dS_2 &= \frac{T_2 - T_1}{T_1 T_2} dE_1 \\ &> 0 \quad (\text{since } T_2 > T_1) \end{aligned} \tag{14.4}$$

From the structure of equation 14.4 we can see that entropy is created by a thermal transfer *from* the hotter object *to* the cooler object. Therefore such a transfer can (and typically will) proceed spontaneously. In contrast, a transfer in the reverse direction cannot proceed, since it would violate the second law of thermodynamics.

We can also see that:

- If the two temperatures are very close together, a thermal transfer of energy can be very nearly reversible.
- Conversely, if the temperature difference is large, the transfer creates a lot of entropy, and is therefore strongly dissipative, strongly irreversible.

A useful rule of thumb says that any reversible thermal transfer will be rather slow. This can be understood in terms of the small temperature difference in conjunction with a finite thermal conductivity. In practice, people usually accept a goodly amount of inefficiency as part of the price for going fast. This involves engineering tradeoffs. We still need the deep principles of physics (to know what’s possible) ... but we need engineering on top of that (to know what’s practical).



### 14.1.4 Graphical Analysis – One Dimension

Questions about equilibrium, stability, and spontaneity are often most conveniently formulated as maximization or minimization problems. This is not the most general way to do things, but it is a convenient and intuitive starting point. In the one-dimensional case especially, the graphical approach makes it easy to see what is going on.

In figure 14.3, we are trying to minimize some objective function  $R$  (the “regret”). In an economics problem,  $R$  might represent the cost. In a physics or chemistry problem,  $R$  might represent something like the energy, or the Gibbs free enthalpy, or the negative of the entropy, or whatever. For now let’s just call it  $R$ .

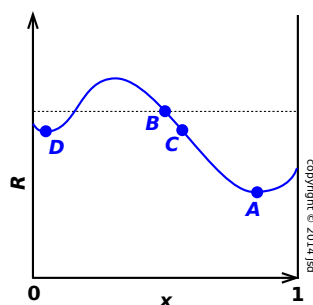


Figure 14.3: Minimizing Some Objective Function

In a physics problem, the abscissa  $x$  could be the position of a particle rolling in a potential well, subject to damping. In a chemistry problem,  $x$  could be the reaction coordinate, where  $x = 0$  corresponds to 100% reactants and  $x = 1$  corresponds to 100% products.

Global equilibrium corresponds to the lowest point on the curve. (If there are multiple points all at the minimum value, global equilibrium corresponds to this set of points.) This can be stated in mathematical terms as follows: For some fixed point  $A$ , if  $R(B) - R(A)$  is greater or equal to zero for all points  $B$ , then we know  $A$  is a global equilibrium point.

Note: Sometimes people try to state the equilibrium requirement in terms of  $\Delta R$ , where  $\Delta R := R(B) - R(A)$ , but this doesn't work very well. We are better off if we speak about point  $A$  and point  $B$  directly, rather than hiding them inside the  $\Delta$ .

Now suppose we restrict point  $B$  to be *near* point  $A$ . Then if  $R(B) - R(A)$  is greater or equal to zero for all nearby points  $B$ , then we say  $A$  is a *local* equilibrium point. (A local equilibrium point may or may not be a global equilibrium point also.) In figure 14.3, point  $A$  is a global minimum, while point  $D$  is only a local minimum.

Now let's consider the direction of spontaneous reaction. Given two points  $B$  and  $C$  that are near to each other, then if  $R(C)$  is less than  $R(B)$  then the reaction will proceed in the direction from  $B$  to  $C$  (unless it is forbidden by some other law). In other words, the reaction proceeds in the direction that produces a negative  $\Delta R$ .

Note that this rule applies only to small deltas. That is, it applies only to pairs of *nearby* points. In particular, starting from point  $B$ , the reaction will *not* proceed spontaneously toward point  $D$ , even though point  $D$  is lower. The *local* slope is the only thing that matters. If you try to formulate the rule in terms of  $\Delta R$  in mathematical terms, without thinking about what it means, you will get fooled when the points are not nearby.

### 14.1.5 Graphical Analysis – Multiple Dimensions

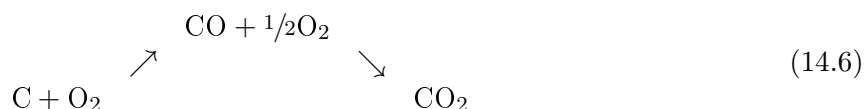
It is often interesting to ask whether a given situation is unstable, i.e. expected to change spontaneously – and if so, to ask in which direction it is expected to change. So, let's examine what it means to talk about “direction” in thermodynamic state-space.

For starters, consider reactions involving carbon, oxygen, and carbon dioxide.



Under some conditions we have a simple reaction that proceeds mostly in the left-to-right direction equation 14.5, combining carbon with oxygen to form carbon dioxide. Meanwhile, under other conditions the reverse proceeds mostly in the opposite direction, i.e. the decomposition of carbon dioxide to form carbon and oxygen.

More generally, however, we need to consider other possibilities, such as the possible presence of carbon monoxide.



This is now a multi-dimensional situation, as shown schematically in figure 14.4.

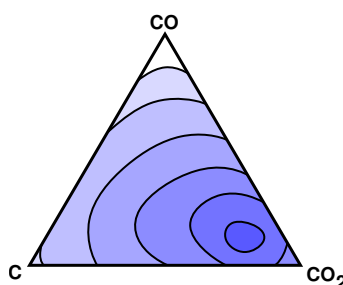


Figure 14.4: Reaction Space with Objective Function Contours

Consider the contrast:

<p>In one dimension, we can speak of a given transformation proceeding forward or backward. Forward means proceeding left-to-right as written in equation 14.5, while backwards means proceeding right-to-left.</p>	<p>In multiple dimensions, we cannot speak of forward, backward, right, or left. We need to specify in detail what sort of step is taken when the transformation proceeds.</p>
---	--

Remark on terminology: In this document, the term “transformation” is meant to be very general, including chemical reactions and phase transitions among other things. (It is not necessary to distinguish “chemical processes” from “physical processes”)

At any temperature other than absolute zero, a chemical reaction will never really go to completion. There will always be some leftover reactants and leftover intermediates, along with the nominal products. This is indicated by the contours in figure 14.4. The endpoint of the reaction is inside the smallest contour.

So far, in figure 14.3 and figure 14.4 we have been coy about what the objective function represents. Here’s a definite example: For an isolated system, the endpoint is the point of maximum entropy  $S$ , and we can interpret the contours in figure 14.4 as contours of constant  $S$ . As the reaction proceeds, it moves uphill in the direction of increasing  $S$ .

For a system that is not isolated – such as a system in contact with a heat bath – the system entropy  $S$  is not the whole story. We need to worry about  $S$ , but we also need to worry about how much entropy has been pushed across the boundary of the system.

### 14.1.6 Reduced Dimensionality

The ideas of section 14.1.4 and section 14.1.5 can be combined, as shown in figure 14.5. Sometimes a system is constrained to move along a lower-dimensional *reaction pathway* within a higher dimensional space.

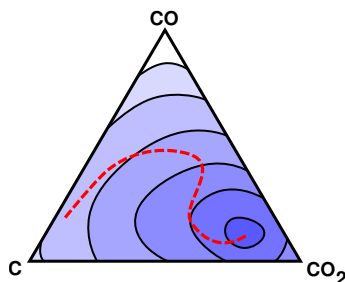


Figure 14.5: Hypothetical Reaction Pathway

Often it's hard to know exactly what the pathway is. Among other things, catalysis can drastically change the pathway.

### 14.1.7 General Analysis

In all cases, the analysis is based on the second law of thermodynamics, equation 2.1, which we restate here:

$$\begin{array}{ccc} \text{change in entropy} & \geq & \text{net flow of entropy} \\ \text{(inside boundary)} & & \text{(inward minus outward across boundary)} \end{array} \quad (14.7)$$

We can use this to clarify our thinking about equilibrium, stability, reversibility, et cetera:

- A transformation will not proceed in any direction that violates equation 14.7.
- For any reversible transformation, equation 14.7 is restricted to a strict equality, not an inequality:

$$\begin{array}{ccc} \text{change in entropy} & = & \text{net flow of entropy} \\ \text{(inside boundary)} & & \text{(inward minus outward across boundary)} \end{array} \quad (14.8)$$

In other words, entropy is conserved during a reversible transformation.

- For any irreversible transformation, equation 14.7 is restricted to a strict inequality, not an equality:

$$\begin{array}{ccc} \text{change in entropy} & > & \text{net flow of entropy} \\ \text{(inside boundary)} & & \text{(inward minus outward across boundary)} \end{array} \quad (14.9)$$

In other words, entropy is created during a reversible transformation. You know the transformation is irreversible, because the reverse would violate the second law, equation 14.7.

To repeat: in all cases, we can analyze the system by direct application of the second law of thermodynamics, in the form given by equation 14.7. This is always possible, but not necessarily convenient.

Be that as it may, people really like to have an objective function. They like this so much that they often engineer the system to have a well-behaved objective function (exactly or at least approximately). That is, they engineer it so that moving downhill in figure 14.4 corresponds to creating entropy, i.e. an irreversible transformation. Moving along an isopotential contour corresponds to a transformation that does not create entropy, i.e. a reversible transformation.

Beware: This is tricky, because “the amount of entropy created” is not (in general) a function of state. Sometimes we can engineer it so that there is some function of state that tells us what we need to know, but this is not guaranteed.

Having an objective function is useful for multiple reasons. For one thing, it provides a way to visualize and communicate what’s going on. Also, there are standard procedures for using variational principles as the basis for analytical and computational techniques. These techniques are often elegant and powerful. The details are beyond the scope of this document.

You don’t know what the objective function is until you see how the system is engineered. In all cases the point of the objective function (if any) is to express a corollary to the second law of thermodynamics.

- A microcanonical system cannot exchange energy, entropy or anything else with its surroundings. It is isolated. Maximizing the entropy  $S$  is the relevant objective function. See section 14.2.1.
- Sometimes one subsystem can do something that creates entropy somewhere else. Minimizing the subsystem energy  $E$  may be the relevant objective function. See section 14.2.2.
- A canonical system can exchange energy and entropy (but nothing else) by means of thermal contact with a heat bath at some definite temperature  $T$ . In this case, minimizing the system free energy  $E - TS$  is the relevant objective function. See section 14.2.3.
- Sometimes the system can do work by expanding against the ambient pressure  $P$ , and can also exchange energy and entropy with a heat bath at temperature  $T$ . Then the Gibbs free enthalpy  $E + PV - TS$  may be the relevant objective function. See section 14.2.4.
- Sometimes something else.
- Sometimes none of the above, in which case you have to rely directly on the second law directly. This requires considering two contributions: the change in entropy  $\Delta S$  inside the region of interest, plus whatever entropy (if any) flowed out of the region across the boundaries during the transformation. The entropy  $S$  is a function of state of the region. However, the flow term cannot be expressed in terms of a function of state, so we have to be careful.

This completes the analysis of the general principle. The second law is the central, fundamental idea. However, direct application of the second law is often inconvenient.

Therefore the rest of this chapter is mostly devoted to developing less-general but more-convenient techniques. In various special situations, subject to various provisos, we can find quantities that are convenient to measure that serve as *proxies* for the amount of entropy created.

### 14.1.8 What’s Fundamental and What’s Not

At this point we should discuss the oft-quoted words of David Goodstein. In reference 41, the section on “Variational Principles in Thermodynamics” begins by saying:

Fundamentally there is only one variational principle in thermodynamics. According to the Second Law, an isolated body in equilibrium has the maximum entropy that physics circumstances will allow.

The second sentence is true and important. For an isolated system, the maximum-entropy principle is an immediate corollary of the second law of thermodynamics, equation 14.7.

The first sentence in that quote seems a bit overstated. It only works if you consider a microcanonical (isolated) system to be more fundamental than, say, a canonical (constant-temperature) system. Note the contrast:

The maximum entropy principle is not true in general; it is only true for an isolated system.	The second law of thermodynamics is true in all generality.
---	---

The book goes on to say:

However, given in this form, it is often inconvenient to use.

It’s true that the maximum-entropy principle is often inconvenient, but it’s even worse than that. For a non-isolated system, maximizing the system entropy  $S$  is not even the correct variational principle. It’s not just inconvenient, it’s invalid. For a non-isolated system, in general there might not even be a valid variational principle of the type we are talking about.

## 14.2 Proxies for Predicting Spontaneity, Reversibility, Equilibrium, etc.

In this section, we apply the general law to some important special cases. We derive some simplified laws that are convenient to apply in such cases.

### 14.2.1 Isolated System; Proxy = Entropy

Consider a completely isolated system. No entropy is flowing across the boundary of the system. If we know the entropy of the system, we can use that to apply the second law directly.

It tells us that the system entropy cannot decrease. Any transformation that leaves the system entropy unchanged is reversible, and any change that increases the system entropy is irreversible.

It must be emphasized that these conclusions are very sensitive to the provisos and assumptions of this scenario. The conclusions apply only to a system that isolated from the rest of the universe.

### 14.2.2 External Damping; Proxy = Energy

In figure 14.6, we have divided the universe into three regions:

- the *interior region* – the mass and the spring
- the *neighborhood* – the damper, separate from the interior region but inside the black rectangle; and
- the rest of the universe – outside the black rectangle.

The combination of interior region + neighborhood will be called the *local region*. We assume the local region is thermally isolated from the rest of the universe.

We have engineered things so that the linkage from the internal region to the damper to be thermally insulating. That means the internal region can do mechanical work on the damper, but cannot exchange entropy with it.

The decision to consider the damper as not part of the interior was a somewhat arbitrary, but not unreasonable. There are plenty of real-world systems where this makes sense, such as a charged harmonic oscillator (where radiative damping is not considered interior to the system) or a marble oscillating in a bowl full of fluid (where the viscous damping is not considered interior to the marble).

Local conservation of energy tells us:

$$dE = -dE_n \quad (14.10)$$

We use unadorned symbols such as  $E$  and  $S$  etc. to denote the energy and entropy etc. inside the interior region. We use a subscript “n” as in  $E_n$  and  $S_n$  etc. to represent the energy and entropy etc. in the neighborhood. In this example, the neighborhood is the damper.

Now, suppose the oscillator starts out with a large amount of energy, large compared to  $kT$ . As the oscillator moves, energy will be dissipated in the damper. The entropy of the damper will increase. The entropy of the interior region is unknown and irrelevant, because it remains constant:

$$\begin{aligned} dS &= 0 \\ S &= \text{unknown} \end{aligned} \quad (14.11)$$

We require that the damper has some definite temperature. That allows us to relate its energy to its entropy in a simple way:

$$dE_n = T_n dS_n \quad (14.12)$$

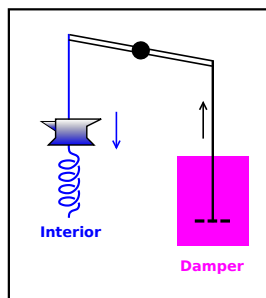


Figure 14.6: Oscillator with Damper



with no other terms on the RHS. There are undoubtedly other equally-correct ways of expanding  $dE_n$ , but we need not bother with them, because equation 14.12 is correct and sufficient for present purposes.

Physically, the simplicity of equation 14.12 depends on (among other things) the fact that the energy of the neighborhood does not depend on the position of the piston within the damper ( $x$ ), so we do not need an  $F \cdot x$  term in equation 14.12. The frictional force depends on the velocity ( $dx/dt$ ) but not on the position ( $x$ ).

We assume that whatever is going on in the system is statistically independent of whatever is going on in the damper, so the entropy is extensive. Physically, this is related to the fact that we engineered the linkage to be thermally non-conducting.

$$dS_{\text{local}} = dS + dS_n \quad (14.13)$$

Combining the previous equations, we find:

$$dS_{\text{local}} = -dE/T_n \quad (14.14)$$

This means that for any positive temperature  $T_n$ , we can use the energy of the system as a proxy for the entropy of the local region.

So, let's apply the second law to the local region. Recall that no entropy is flowing across the boundary between the local region and the rest of the universe. So the second law tells us that the system energy cannot increase. Any transformation that leaves the system energy unchanged is reversible, and any change that decreases the system energy is irreversible.

It must be emphasized that these conclusions are very sensitive to the provisos and assumptions of this scenario. The restrictions include: The system must be connected by a thermally-insulating linkage to a damper having some definite positive temperature, and the system must be isolated from the rest of the universe.

Another restriction is that the entropy within the system itself must be negligible or at least constant. If we implement the spring in figure 14.6 using a gas spring, i.e. a pneumatic cylinder, we would not be able to lower the system energy by condensing the gas, since that would require changing the system entropy.

Note: In elementary non-thermal mechanics, there is an unsophisticated rule that says "balls roll downhill" or something like that. That is *not* an accurate statement of the physics, because in the absence of dissipation, a ball that rolls down the hill will immediately roll up the hill on the other side of the valley.

If you want the ball to roll down *and stay down*, you need some dissipation, and we can understand this in terms of equation 14.14.

Note that the relevant temperature is the temperature of the damper,  $T_n$ . The system itself and the other non-dissipative components might not even have a well-defined temperature. The temperature (if any) of the non-dissipative components is irrelevant, because it doesn't enter into the calculation of the desired result (equation 14.14). This is related to the fact that  $dS$  is zero, so we know  $TdS$  is zero even if we don't know  $T$ .

### 14.2.3 Constant $V$ and $T$ ; Proxy = Helmholtz Free Energy

In this section, we consider a different system, and a different set of assumptions.

We turn our attention to a sample of gas, held under conditions of constant volume and constant positive temperature. We shall see that this allows us to answer questions about spontaneity using the system Helmholtz free energy  $F$  as a proxy for the amount of entropy created.

The situation is shown in figure 14.7. We have divided the universe into three regions:

- the *interior region* – inside the blue cylinder;
- the *neighborhood* – outside the cylinder but inside the black rectangle; and
- the rest of the universe – outside the black rectangle.

The combination of interior region + neighborhood will be called the *local region*.

Inside the blue cylinder is some gas. In the current scenario, the volume of the cylinder is constant. (Compare this to the constant-pressure scenario in section 14.2.4, where the volume is not constant.)

Inside the neighborhood is a heat bath, as represented by the magenta region in the figure. It is in thermal contact with the gas inside the cylinder. We assume the heat capacity of the heat bath is very large.

We assume the combined local system (interior + neighborhood) is isolated from the rest of the universe. Specifically, no energy or entropy can flow across the boundary of the local system (the black rectangle in figure 14.7).

We use unadorned symbols such as  $F$  and  $S$  etc. to denote the free energy and entropy etc. inside the interior region. We use a subscript “n” as in  $E_n$  and  $S_n$  etc. to represent the energy and entropy etc. in the neighborhood.

Here’s an outline of the usual calculation that shows why  $\mathbf{d}F$  is interesting. Note that  $F$  is the free energy inside the interior region. Similarly  $S$  is the entropy inside the interior region (in contrast to  $S_{\text{local}}$ , which includes all the local entropy,  $S_{\text{local}} = S + S_n$ ). This is important, because it is usually much more convenient to keep track of what’s going on in the interior region than to keep track of everything in the neighborhood and the rest of the universe.

We start by doing some math:

$$\begin{aligned}
 F &:= E - TS && \text{by definition of } F \\
 \mathbf{d}F &= \mathbf{d}E - T\mathbf{d}S - S\mathbf{d}T && \text{by differentiating} \\
 &= \mathbf{d}E - T\mathbf{d}S && \text{since } \mathbf{d}T = 0 \text{ by hypothesis} \\
 \mathbf{d}S - \mathbf{d}E/T &= -\mathbf{d}F/T && \text{by rearranging}
 \end{aligned} \tag{14.15}$$

Next we relate certain inside quantities to the corresponding outside quantities:

$$\begin{aligned}
 T &= T_n && \text{temperature the same everywhere} \\
 E + E_n &= \text{const} && \text{isolation + local conservation of energy} \\
 \mathbf{d}E &= -\mathbf{d}E_n && \text{by differentiating}
 \end{aligned} \tag{14.16}$$

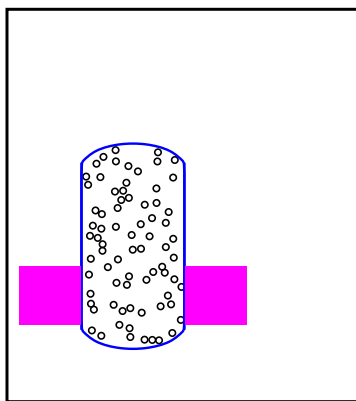


Figure 14.7: Constant Volume and Temperature; Cylinder + Neighborhood

Next we assume that the heat bath is internal equilibrium. This is a nontrivial assumption. We are making use of the fact that it is a heat bath, not a bubble bath. We are emphatically not assuming that the interior region is in equilibrium, because one of the major goals of the exercise is to see what happens when it is not in equilibrium. In particular we are emphatically not going to assume that  $E$  is a function of  $S$  and  $V$  alone. Therefore we cannot safely expand  $\mathbf{d}E = T\mathbf{d}S - P\mathbf{d}V$  in the interior region. We can, however, use the corresponding expansion in the neighborhood region, because it is in equilibrium:

$$\begin{aligned}
 \mathbf{d}E_n &= T\mathbf{d}S_n - P\mathbf{d}V_n && \text{bath in equilibrium} \\
 &= T\mathbf{d}S_n && \text{constant } V \text{ by hypothesis} \\
 \mathbf{d}S_n &= \mathbf{d}E_n/T && \text{by rearranging} \\
 &= -\mathbf{d}F/T && \text{by conservation of energy, equation 14.16}
 \end{aligned} \tag{14.17}$$

Next, we assume the entropy is an extensive quantity. That is tantamount to assuming that the probabilities are uncorrelated, specifically that the distribution that characterizes the interior is uncorrelated with the distribution that characterizes the neighborhood. This is usually a very reasonable assumption, especially for macroscopic systems.

We are now in a position to finish the calculation.

$$\begin{aligned}
 \mathbf{d}S_{\text{local}} &= \mathbf{d}S + \mathbf{d}S_n && \text{entropy is extensive} \\
 &= \mathbf{d}S - \mathbf{d}E/T && \text{by equation 14.17} \\
 &= -\mathbf{d}F/T && \text{by equation 14.15}
 \end{aligned} \tag{14.18}$$

This means that for any positive temperature  $T_n$ , we can use the Helmholtz free energy of the system as a proxy for the entropy of the local region.

So, let's apply the second law to the local region. Recall that no entropy is flowing across the boundary between the local region and the rest of the universe. So the second law tells us that the system free energy cannot increase. Any transformation that leaves the system free energy unchanged is reversible, and any change that decreases the system free energy is irreversible.

It must be emphasized that these conclusions are very sensitive to the provisos and assumptions of this scenario. The conclusions apply only to a constant-volume system that can exchange energy thermally with a heat bath at some positive temperature ... and is otherwise isolated from the rest of the universe.

#### 14.2.4 Constant $P$ and $T$ ; Proxy = Gibbs Free Enthalpy

We now turn our attention to conditions of constant pressure and constant positive temperature. This is closely analogous to section 14.2.3; the only difference is constant pressure instead of constant volume. We shall see that this allows us to answer questions about spontaneity using the system's Gibbs free enthalpy  $G$  as a proxy for the overall entropy  $S_{\text{total}}$ .

In figure 14.8, we have divided the universe into three regions:

- the *interior region* – inside the blue cylinder;
- the *neighborhood* – outside the cylinder but inside the black rectangle; and

- the rest of the universe – outside the black rectangle.

The combination of interior region + neighborhood will be called the *local region*.

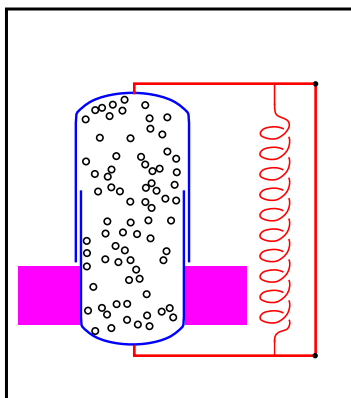


Figure 14.8: Constant Pressure and Temperature; Cylinder + Neighborhood

Inside the blue cylinder is some gas. The cylinder is made of two pieces that can slide up and down relative to each other, thereby changing the boundary between the interior region and the neighborhood. (Contrast this against the constant-volume scenario in section 14.2.3.)

Inside the neighborhood is a heat bath, as represented by the magenta region in the figure. It is in thermal contact with the gas inside the cylinder. We assume the heat capacity of the heat bath is very large.

Also in the neighborhood there is a complicated arrangement of levers and springs, which maintains a constant force (and therefore a constant force per unit area, i.e. pressure) on the cylinder.

We also assume that the kinetic energy of the levers and springs is negligible. This is a nontrivial assumption. It is tantamount to assuming that whatever changes are taking place are not too sudden, and that the springs and levers are somehow kept in thermal equilibrium with the heat bath.

We assume the combined local system (interior + neighborhood) is isolated from the rest of the universe. Specifically, no energy or entropy can flow across the boundary of the local system (the black rectangle in figure 14.8).

Note that  $G$  is the free enthalpy inside the interior region. Similarly  $S$  is the entropy inside the interior region (in contrast to  $S_{\text{local}}$ , which includes all the local entropy,  $S_{\text{local}} = S + S_n$ ). This is important, because it is usually much more convenient to keep track of what's going on in the interior region than to keep track of everything in the neighborhood and the rest of the universe.

We start by doing some math:

$$\begin{aligned}
 G &:= H - TS && \text{by definition of } G \\
 dG &= dH - TdS - SdT && \text{by differentiating} \\
 &= dH - TdS && \text{since } dT = 0 \text{ by hypothesis} \\
 dS - dH/T &= -dG/T && \text{by rearranging}
 \end{aligned} \tag{14.19}$$

Next we relate certain inside quantities to the corresponding outside quantities. We will make use of the fact that in this situation, enthalpy is conserved, as discussed in section 14.3.2.

$$\begin{aligned}
 T &= T_n && \text{temperature the same everywhere} \\
 H + H_n &= \text{const} && \text{local conservation of enthalpy} \\
 dH &= -dH_n && \text{by differentiating}
 \end{aligned} \tag{14.20}$$

Next we assume that the heat bath is internal equilibrium. This is a nontrivial assumption. We are making use of the fact that it is a heat bath, not a bubble bath. We are emphatically not assuming that the interior region is in equilibrium, because one of the major goals of the exercise is to see what happens when it is not in equilibrium. In particular we are emphatically not going to assume that  $H$  is a function of  $S$  and  $P$  alone. Therefore we cannot safely expand  $dH = TdS + VdP$  in the interior region. However, we don't need to do that. It suffices to rely on the assumption that the neighborhood is a well-behaved heat bath:

$$\begin{aligned}
 dH_n &= TdS_n + VdP_n && \text{bath in equilibrium} \\
 &= TdS_n && \text{constant } P \text{ by hypothesis} \\
 dS_n &= dH_n/T && \text{by rearranging} \\
 &= -dH/T && \text{using equation 14.20}
 \end{aligned} \tag{14.21}$$

Again we assume the entropy is extensive. We are now in a position to finish the calculation.

$$\begin{aligned}
 dS_{\text{local}} &= dS + dS_n && \text{entropy is extensive} \\
 &= dS - dH/T && \text{by equation 14.21} \\
 &= -dG/T && \text{by equation 14.19}
 \end{aligned} \tag{14.22}$$

So, let's apply the second law to the local region. Recall that no entropy is flowing across the boundary between the local region and the rest of the universe. So the second law tells us that the system free enthalpy cannot increase. Any transformation that leaves the system free enthalpy unchanged is reversible, and any change that decreases the system free enthalpy is irreversible.

It must be emphasized that these conclusions are very sensitive to the provisos and assumptions of this scenario. The conclusions apply only to a constant-pressure system that can exchange energy thermally with a heat bath at some positive temperature ... and is otherwise isolated from the rest of the universe.

## 14.3 Discussion: Some Fine Points

Let's take a moment to discuss a couple of tangentially-related points that sometimes come up.

### 14.3.1 Local Conservation

In ultra-simple situations, it is traditional to divide the universe into two regions: “the system” versus “the environment.” Sometimes other terminology is used, such as “interior” versus “exterior,” but the idea is the same.

In more complicated situations, such as fluid dynamics, we must divide the universe into a great many regions, aka parcels. We can ask about the energy, entropy, etc. internal to each parcel, and also ask about the transfer of energy, entropy, etc. to adjacent parcels.

This is important because, as discussed in section 1.2, a *local* conservation law is much more useful than a global conservation law. If some energy disappears from my system, it does me no good to have a global law that says the energy will eventually reappear “somewhere” in the universe. The local laws says that energy is conserved right here, right now.

For present purposes, we can get by with only three regions: the interior, the immediate neighborhood, and the rest of the universe. Examples of this can be seen in section 14.2.3, section 14.2.4, and section 14.2.2.

### 14.3.2 Lemma: Conservation of Enthalpy, Maybe

Energy is always strictly and locally conserved. It is conserved no matter whether the volume is changing or not, no matter whether the ambient pressure is changing or not, no matter whatever.

Enthalpy is sometimes conserved, subject to a few restrictions and provisos. The primary, crucial restriction requires us to work under conditions of constant pressure.

Another look at figure 14.8 will help us fill in the details.

As always, the enthalpy is:

$$\begin{aligned} H &= E + PV && \text{(interior region)} \\ H_n &= E_n + PV_n && \text{(neighborhood region)} \end{aligned} \tag{14.23}$$

By conservation of energy, we have

$$E + E_n = E_0 = \text{const} \tag{14.24}$$

Next we are going to argue for a “local conservation of volume” requirement. There are several ways to justify this. One way is to argue that it is corollary of the previous assumption that the local

system is isolated and not interacting with the rest of the universe. Another way is to just impose it as a requirement, explicitly requiring that the volume of the local region (the black rectangle in figure 14.8) is not changing. A third way is to arrange, as we have in this case, that there is no pressure acting on the outer boundary, so for purposes of the energy calculation we don't actually care where the boundary is; this would require extra terms in equation 14.26 but the extra terms would all turn out to be zero.

We quantify the idea of constant volume in the usual way:

$$V + V_n = V_0 = \text{const} \quad (14.25)$$

Now we do some algebra:

$$\begin{array}{llll}
 E + E_n & = & E_0 & \text{by equation 14.24} \\
 E + PV + E_n - PV & = & E_0 & \text{add and subtract } PV \\
 E + PV + E_n - P(V_0 - V_n) & = & E_0 & \text{by equation 14.25} \\
 E + PV + E_n + PV_n & = & E_0 + PV_0 & \text{by rearranging} \\
 H + H_n & = & \text{const} & \text{by equation 14.23}
 \end{array} \quad (14.26)$$

### 14.3.3 Energy and Entropy (as opposed to «Heat»)

Note that each of the calculations in section 14.2 was carried out by keeping track of the energy and entropy.

Thinking in terms of energy and entropy is good practice.	Thinking in terms of “heat” and “work” would be a fool’s errand.
---	--

Energy is conserved.	Neither heat nor work is separately conserved.
----------------------	--

We can easily understand that the linkage that connects the interior to the damper carries zero entropy and carries nonzero energy.	We see “work” leaving the interior in the form of $PdV$ or $F \cdot dx$ . We see no heat leaving the interior. Meanwhile, we see heat showing up in the damper, in the form of $TdS$ . This would be confusing, if we cared about heat and work, but we don’t care, so we escape unharmed.
---	--

Keeping track of the energy and the entropy is the easy and reliable way to solve the problem.

### 14.3.4 Spontaneity

Spontaneity is not quite the same as irreversibility, for the following reason: If a transformation can proceed in a direction that creates entropy, then as a rule of thumb, in many cases the transformation *will* spontaneously proceed in such a direction. However, this is only a rule of thumb, not a guarantee. As a counterexample, consider the transformation of diamond into graphite. Calculations show that under ordinary conditions this creates entropy, so it is definitely irreversible. However, the rate is so slow as to be utterly negligible, so we cannot say that the reaction occurs spontaneously, in any practical sense.



If a transformation does not occur, it might be because of the second law ... or because of any of the other laws of nature. It might be restricted by huge activation barriers, symmetries, selection rules, or other laws of physics such that it cannot create entropy at any non-negligible rate.

### 14.3.5 Conditionally Allowed and Unconditionally Disallowed

A reminder: The second law is only one law among many. Other laws include conservation of energy (aka the first law of thermodynamics), other conservation laws, various symmetries, spectroscopic selection rules, mathematical theorems, et cetera.

A process can proceed only if it complies with *all* the laws. Therefore if a process is forbidden by one of the laws, it is unconditionally forbidden. In contrast, if the process is allowed by one of the laws, it is only conditionally allowed, conditioned on compliance with all the other laws.

Based on a second-law analysis alone, we can determine that a process absolutely *will not* proceed spontaneously, in situations where doing so would violate the second law. In contrast, a second-law analysis does not allow us to say that a process “will” proceed spontaneously. Until we do a more complete analysis, the most we can say is that it *might* proceed spontaneously.

We are on much firmer ground when it comes to reversibility. In the context of ordinary chemical reactions, if we know the reaction can proceed in one direction, it is reversible if and only if it does not create entropy. That’s because of the reversibility of all the relevant<sup>1</sup> fundamental laws governing such reactions, except the second law. So if there is no barrier to the forward reaction, there should be no barrier to the reverse reaction, other than the second law. (Conversely, if the reaction cannot proceed in any direction, because of conservation laws or some such, it is pointless to ask whether it is reversible.)

### 14.3.6 Irreversible by State or by Rate

Consider the contrast:

In figure 14.9, there is no attempt to make the process reversible. The descent of the anvil is grossly dissipative. You can tell how much energy is dissipated during the process just by looking at the initial state and the final state.

We can call this “irreversible by state” and say that the amount of dissipation per operation is zeroth order in the rate (i.e. independent of the rate).

In figure 14.10, the process is very nearly reversible. There will probably always be “some” friction, but we may be able to engineer the bearing so that the friction is small, maybe even negligible. Typically, to a good approximation, the power dissipation will be second order in the *rate* of the process, and the total energy dissipated per cycle will be first order in the rate.

We can call this “irreversible by rate” and say that the amount of dissipation per operation is first order in the rate.

---

<sup>1</sup>We disregard the non-time-reversible behavior of weak nuclear processes, such as decay of neutral Kaons. Such processes are irrelevant to ordinary chemistry, biochemistry, and mechanics.

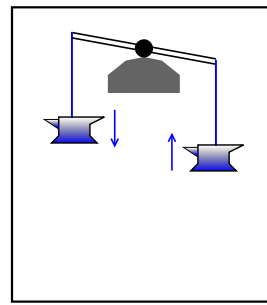
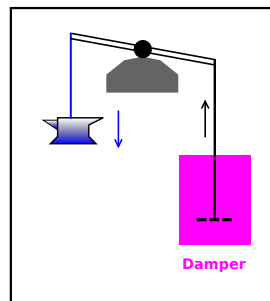


Figure 14.9: Irreversible by State

Figure 14.10: Irreversible by Rate

## 14.4 Temperature and Chemical Potential in the Equilibrium State

In this section, we derive a couple of interesting results. Consider a system that is isolated from the rest of the universe, and can be divided into two parcels. We imagine that parcel #1 serves as a heat bath for parcel #2, and vice versa. Then:

- Whenever the two parcels have reached equilibrium by exchanging energy, they will have the same temperature.

- Whenever the two parcels have reached equilibrium by exchanging particles as well as energy, they will have the same chemical potential (and the same temperature).

We begin with a review of basic ideas of temperature in the equilibrium state, as introduced in chapter 13. With great generality, we can say that at thermodynamic equilibrium, the state of the system may change, but only in directions that do not create any entropy.

We write the gradient as  $\mathbf{d}S$  rather than  $\nabla S$  for technical reasons, but either way, the gradient is a vector. It is a vector in the abstract thermodynamic state-space (or more precisely, the tangent space thereof).

As usual, subject to mild conditions, we can expand  $\mathbf{d}S$  using the chain rule:

$$\mathbf{d}S = \left. \frac{\partial S}{\partial N} \right|_E \mathbf{d}N + \left. \frac{\partial S}{\partial E} \right|_N \mathbf{d}E \quad (14.27)$$

We assume constant volume throughout this section. We also assume all the potentials are sufficiently differentiable.

We recognize the partial derivative in front of  $\mathbf{d}E$  as being the inverse temperature, as defined by equation 24.4, which we repeat here:

$$\beta := \left. \frac{\partial S}{\partial E} \right|_N \quad (14.28)$$

We can rewrite the other partial derivative by applying the celebrated cyclic triple partial derivative rule:

$$\left. \frac{\partial S}{\partial N} \right|_E \left. \frac{\partial N}{\partial E} \right|_S \left. \frac{\partial E}{\partial S} \right|_N = -1 \quad (14.29)$$

For an explanation of where this rule comes from, see section 14.10. We can re-arrange equation 14.29 to obtain:

$$\left. \frac{\partial S}{\partial N} \right|_E = -1 \left. \frac{\partial E}{\partial N} \right|_S \left. \frac{\partial S}{\partial E} \right|_N \quad (14.30)$$

Note that if you weren't fastidious about keeping track of the "constant  $E$ " "constant  $S$ " and "constant  $N$ " specifiers, it would be very easy to get equation 14.30 wrong by a factor of  $-1$ .

We recognize one of the factors on the RHS as the chemical potential, as defined by equation 7.32, which we repeat here:

$$\mu := \left. \frac{\partial E}{\partial N} \right|_S \quad (14.31)$$

Putting together all the ingredients we can write:

$$\begin{aligned} \mathbf{d}S &= 0 \\ &= -\mu\beta\mathbf{d}N + \beta\mathbf{d}E \end{aligned} \quad (14.32)$$

Since we can choose  $\mathbf{d}N$  and  $\mathbf{d}E$  independently, both terms on the RHS must vanish separately.

If we divide the system into two parcels, #1, and #2, then

$$\begin{aligned} \mathbf{d}S_1 &= -\mathbf{d}S_2 && \text{since } \mathbf{d}S = 0 \text{ at equilibrium} \\ \mathbf{d}N_1 &= -\mathbf{d}N_2 && \text{since } N \text{ is conserved} \\ \mathbf{d}E_1 &= -\mathbf{d}E_2 && \text{since } E \text{ is conserved} \end{aligned} \quad (14.33)$$

plugging that into the definitions of  $\beta$  and  $\mu$ , we conclude that at equilibrium:

$$\begin{aligned} \beta_1 &= \beta_2 && \text{if parcels exchange energy} \\ \beta_1\mu_1 &= \beta_2\mu_2 && \text{if parcels exchange energy and particles} \\ T_1 &= T_2 \\ \mu_1 &= \mu_2 \end{aligned} \quad (14.34)$$

The last two lines assume nonzero  $\beta$  i.e. non-infinite temperature.

So, we have accomplished the goal of this section. If/when the two parcels have reached equilibrium by exchanging energy, they will have the same inverse temperature. Assuming the inverse temperature is nonzero, then:

- If/when the two parcels have reached equilibrium by exchanging energy, they will have the same temperature. In other words, *equilibrium is isothermal*. See chapter 13 and reference 39 for more on this.
- If/when the two parcels have reached equilibrium by exchanging particles as well as energy, they will have the same chemical potential (and the same temperature).

One way to visualize this is in terms of the gradient vector  $\mathbf{d}S$ . The fact that  $\mathbf{d}S = 0$  implies the projection of  $\mathbf{d}S$  in *every* feasible direction must vanish, including the  $\mathbf{d}E$  direction and the  $\mathbf{d}N$  direction among others. Otherwise the system would be at non-equilibrium with respect to excursions in the direction(s) of non-vanishing  $\mathbf{d}S$ .

## 14.5 The Approach to Equilibrium

### 14.5.1 Non-Monotonic Case

Figure 14.11 shows the position and momentum of a damped harmonic oscillator. The system starts from rest at a position far from equilibrium, namely (position, momentum) = (1, 0). It then undergoes a series of damped oscillations before settling into the equilibrium state (0, 0). In this plot, time is an implicit parameter, increasing as we move clockwise along the curve.

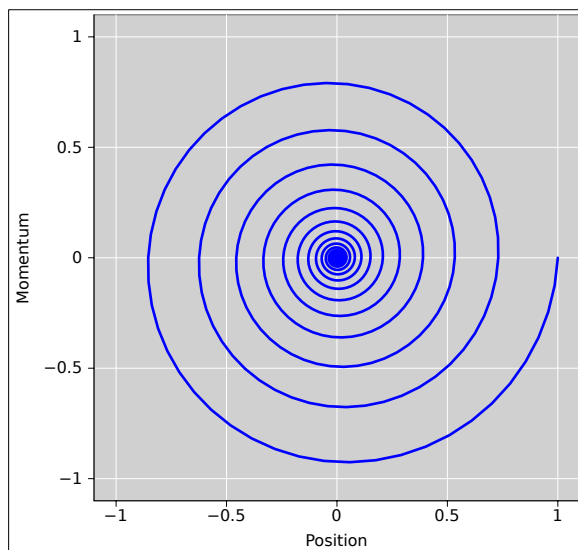


Figure 14.11: Phase Space : Under-Damped Harmonic Oscillator

You will note that neither variable moves directly toward the equilibrium position. If we divide the phase space into quadrants, and look at the sequence of events, ordered by time:

- In quadrant IV, the momentum is negative and becoming more so, i.e. moving away from equilibrium.
- In quadrant III, the position is negative and becoming more so, i.e. moving away from equilibrium.
- In quadrant II, the momentum is positive and becoming more so, i.e. moving away from equilibrium.
- In quadrant I, the position is positive and becoming more so, i.e. moving away from equilibrium.

This should convince you that the approach to equilibrium is *not* necessarily monotonic. Some variables approach equilibrium monotonically, but others do not.

When analyzing a complex system, it is sometimes very useful to identify a variable that changes monotonically as the system evolves. In the context of ordinary differential equations, such a variable is sometimes called a *Lyapunov function*.

In any physical system, the overall entropy (of the system plus environment) must be a monotone-increasing Lyapunov function, as we know by direct application of the second law of thermodynamics. For a system with external damping, such as the damped harmonic oscillator, decreasing system energy is a convenient proxy for increasing overall entropy, as discussed in section 14.2.2; see especially equation 14.14. Note that contours of constant energy are circles in figure 14.11, so you can see that energy decreases as the system evolves toward equilibrium.

### 14.5.2 Monotonic Case

For a critically damped or overdamped system, the approach to equilibrium is non-oscillatory. If the system is initially at rest, the position variable is monotonic, and the momentum variable is “almost” monotonic, in the sense that its absolute value increases to a maximum and thereafter decreases monotonically. More generally, each variable can cross through zero at most once. The critically damped system is shown in figure 14.12.

### 14.5.3 Approximations and Misconceptions

We know that if two or more regions have reached equilibrium by the exchange of energy, they have the same temperature, subject to mild restrictions, as discussed in section 14.4. This is based on fundamental notions such as the second law and the definition of temperature.

We now discuss some much less fundamental notions:

- Roughly speaking, ordinarily, hot things tend to cool off and cold things tend to warm up.

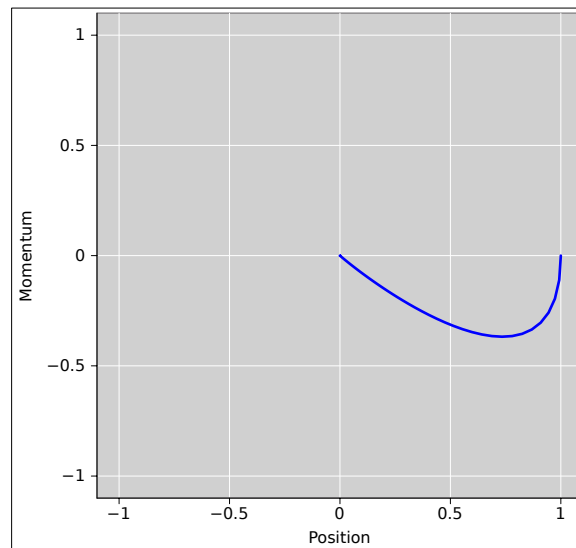


Figure 14.12: Phase Space : Critically-Damped Harmonic Oscillator

- Roughly speaking, ordinarily, hot things cool off monotonically and cold things warm up monotonically. In other words, in each region, temperature “ordinarily” behaves like a Lyapunov function.

Beware that some authorities go overboard and elevate these rules of thumb to the status of axioms. They assert that heat can never spontaneously flow from a lower-temperature region to a higher-temperature region. Sometimes this overstatement is even touted as “the” second law of thermodynamics.

This overstatement is not reliably true, as we can see from the following examples.

As a simple first example, consider a spin system. Region #1 is at a moderate positive temperature, while region #2 is at a moderate negative temperature. We allow the two regions to move toward equilibrium by exchanging energy. During this process, the temperature in region #1 will become more positive, while the temperature in region #2 will become more negative. The temperature difference between the two regions will initially increase, not decrease. Energy will flow from the negative-temperature region into the positive-temperature region.

This example tells us that the concept of *inverse temperature* is more fundamental than temperature itself. In this example, the difference in inverse temperature between the two regions tends to decrease. This is the general rule when two regions are coming into equilibrium by exchanging energy. It is a mistake to misstate this in terms of temperature rather than inverse temperature, but it is something of a technicality, and the mistake is easily corrected.

Here is another example that raises a much more fundamental issue: Consider the phenomenon of *second sound* in superfluid helium. The temperature oscillates like one of the variables in figure 14.11. The approach to equilibrium is non-monotonic.

Let’s be clear: For an ordinary material such as a hot potato, the equation of thermal conductivity is heavily overdamped, which guarantees that temperature approaches equilibrium monotonically

... but this is a property of the material, not a fundamental law of nature. It should not be taken as the definition of the second law of thermodynamics, or even a corollary thereof.

## 14.6 Natural Variables, or Not

### 14.6.1 The “Big Four” Thermodynamic Potentials

Suppose we assume, hypothetically and temporarily, that you are only interested in questions of stability, reversibility, and spontaneity. Then in this scenario, you might choose to be interested in one of the following thermodynamic potentials:

$$\begin{aligned}
 \Delta E & \text{ at constant } S \text{ and } V : \text{Energy} \\
 \Delta F & \text{ at constant } T \text{ and } V : \text{Helmholtz Free Energy} \\
 \Delta G & \text{ at constant } T \text{ and } P : \text{Gibbs Free Enthalpy} \\
 \Delta H & \text{ at constant } S \text{ and } P : \text{Enthalpy}
 \end{aligned}
 \tag{14.35}$$

For more about these potentials and the relationships between them, see chapter 15.

There is nothing fundamental about the choice of what you are interested in, or what you choose to hold constant. All that is a choice, not a law of nature. The only fundamental principle here is the non-decrease of overall entropy,  $S_{\text{total}}$ .

In particular, there is no natural or fundamental reason to think that there are any “natural variables” associated with the big four potentials. Do not believe any assertions such as the following:

$E$	is “naturally” $E(S, V)$	$\otimes$	(14.36)
$F$	is “naturally” $F(T, V)$	$\otimes$	
$G$	is “naturally” $G(T, P)$	$\otimes$	
$H$	is “naturally” $H(S, P)$	$\otimes$	

I typeset equation 14.36 on a red background with skull-and-crossbones symbols to emphasize my disapproval. I have never seen any credible evidence to support the idea of “natural variables.” Some evidence illustrating why it cannot be generally true is presented in section 14.6.2.

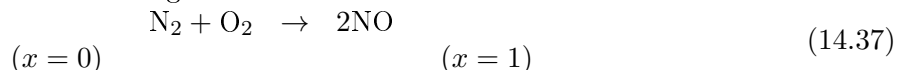
### 14.6.2 A Counterexample: Heat Capacity

Consider an ordinary heat capacity measurement that measures  $\Delta T$  as a function of  $\Delta E$ . This is perfectly well behaved operationally and conceptually.

The point is that  $E$  is perfectly well defined even when it is not treated as the dependent variable. Similarly,  $T$  is perfectly well defined, even when it is not treated as the independent variable. We are allowed to express  $T$  as  $T(V, E)$ . This doesn’t directly tell us much about stability, reversibility, or spontaneity, but it does tell us about the heat capacity, which is sometimes a perfectly reasonable thing to be interested in.

## 14.7 Going to Completion

Suppose we are interested in the following reaction:



which we carry out under conditions of constant pressure and constant temperature, so that when analyzing spontaneity, we can use  $G$  as a valid proxy for  $S_{\text{total}}$ , as discussed in section 14.2.4.

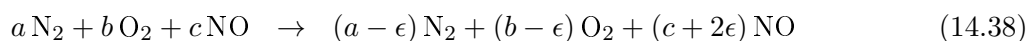
Equation 14.37 serves some purposes but not all.

- It defines what we mean by reactants and products.
- It defines a direction (a very specific direction) in parameter space.
- In some cases but *not* all, the reaction will essentially go to completion, so that the  $\Delta G$  of interest will be  $G(\text{RHS}) - G(\text{LHS})$  where RHS and LHS refer to this equation, namely equation 14.37.

Let  $x$  represent some notion of reaction coordinate, proceeding in the direction specified by equation 14.37, such that  $x = 0$  corresponds to 100% reactants, and  $x = 1$  corresponds to 100% products.

Equation 14.37 is often interpreted as representing the largest possible  $\Delta x$ , namely leaping from  $x = 0$  to  $x = 1$  in one step.

That's OK for some purposes, but when we are trying to figure out whether a reaction will go to completion, we usually need a more nuanced notion of what a reaction is. We need to consider *small* steps in reaction-coordinate space. One way of expressing this is in terms of the following equation:



where the parameters  $a$ ,  $b$ , and  $c$  specify the “current conditions” and  $\epsilon$  represents a small step in the  $x$ -direction. We see that the RHS of this equation has been displaced from the LHS by an amount  $\epsilon$  in the direction specified by equation 14.37.

To say the same thing another way, equation 14.37 is the *derivative* of equation 14.38 with respect to  $x$  (or, equivalently, with respect to  $\epsilon$ ). Equation 14.37 is obviously more compact, and is more convenient for most purposes, but you should not imagine that it describes everything that is going on; it only describes the local derivative of what's going on.

The amount of free enthalpy liberated by equation 14.38 will be denoted  $\delta G$ . It will be infinitesimal, in proportion to  $\epsilon$ . If divide  $\delta G$  by  $\epsilon$ , we get the directional derivative  $\nabla_x G$ .

Terminology note: In one dimension, the directional derivative  $\nabla_x G$  is synonymous with the ordinary derivative  $dG/dx$ .

This tells us what we need to know. If  $\nabla_x G$  is positive, the reaction is allowed proceed spontaneously by (at least) an infinitesimal amount in the  $+x$  direction. We allow it to do so, then re-evaluate  $\nabla_x G$  at the *new* “current” conditions. If  $\nabla_x G$  is still positive, we take another step. We iterate



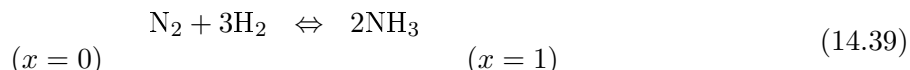
until we come to a set of conditions where  $\nabla_x G$  is no longer positive. At this point we have found the equilibrium conditions (subject of course to the initial conditions, and the constraint that the reaction equation 14.37 is the only allowed transformation).

Naturally, if we ever find that  $\nabla_x G$  is negative, we take a small step in the  $-x$  direction and iterate.

If the equilibrium conditions are near  $x = 1$ , we say that the reaction goes to completion as written. By the same token, if the equilibrium conditions are near  $x = 0$ , we say that the reaction goes to completion in the opposite direction, opposite to equation 14.37.

## 14.8 Example: Shift of Equilibrium

Let's consider the synthesis of ammonia:



We carry out this reaction under conditions of constant  $P$  and  $T$ . We let the reaction reach equilibrium. We arrange the conditions so that the equilibrium is nontrivial, i.e. the reaction does not go to completion in either direction.

The question for today is, what happens if we increase the pressure? Will the reaction remain at equilibrium, or will it now proceed to the left or right?

We can analyze this using the tools developed in the previous sections. At constant  $P$  and  $T$ , subject to mild restrictions, the reaction will proceed in whichever direction minimizes the free enthalpy:

$$\begin{array}{ll} dG/dx > 0 & \Rightarrow \text{proceed to the left} \\ dG/dx < 0 & \Rightarrow \text{proceed to the right} \\ dG/dx = 0 & \Rightarrow \text{equilibrium} \end{array} \quad (14.40)$$

where  $x$  is the reaction coordinate, i.e. the fraction of the mass that is in the form of  $\text{NH}_3$ , on the RHS of equation 14.39. See section 14.2.4 for an explanation of where equation 14.40 comes from.

Note that  $P$  and  $T$  do not need to be constant “for all time,” just constant while the  $d/dx$  equilibration is taking place.

As usual, the free enthalpy is defined to be:

$$G = E + PV - TS \quad (14.41)$$

so the gradient of the free enthalpy is:

$$dG = dE + PdV - TdS \quad (14.42)$$

There could have been terms involving  $VdP$  and  $SdT$ , but these are not interesting since we are working at constant  $P$  and  $T$ .

In more detail: If  $(P_1, T_1)$  corresponds to equilibrium, then we can combine equation 14.42 with equation 14.40 to obtain:

$$\begin{aligned} \left. \frac{dG}{dx} \right|_{(P_1, T_1)} &= \left[ \frac{dE}{dx} + P \frac{dV}{dx} - T \frac{dS}{dx} \right]_{(P_1, T_1)} \\ &= 0 \end{aligned} \quad (14.43)$$

To investigate the effect of changing the pressure, we need to compute

$$\left. \frac{dG}{dx} \right|_{(P_2, T_1)} = \left[ \frac{dE}{dx} + P \frac{dV}{dx} - T \frac{dS}{dx} \right]_{(P_2, T_1)} \quad (14.44)$$

where  $P_2$  is slightly different from the equilibrium pressure; that is:

$$P_2 = (1 + \delta)P_1 \quad (14.45)$$

We now argue that  $E$  and  $dE$  are insensitive to pressure. The potential energy of a given molecule depends on the bonds in the given molecule, independent of other molecules, hence independent of density (for an ideal gas). Similarly the kinetic energy of a given molecule depends on temperature, not on pressure or molar volume. Therefore:

$$\left. \frac{dE}{dx} \right|_{(P_2, T_1)} = \left. \frac{dE}{dx} \right|_{(P_1, T_1)} \quad (14.46)$$

Having examined the first term on the RHS of equation 14.44, we now examine the next term, namely the  $P dV/dx$  term. It turns out that this term is insensitive to pressure, just as the previous term was. We can understand this as follows: Let  $N$  denote the number of gas molecules on hand. Let's say there are  $N_1$  molecules when  $x = 1$ . Then for general  $x$  we have:

$$N = N_1 (2 - x) \quad (14.47)$$

That means  $dN/dx = -N_1$ , independent of pressure. Then the ideal gas law tells us that

$$\left. \frac{d}{dx}(PV) \right|_{(P, T)} = \left. \frac{d}{dx}(NkT) \right|_{(P, T)} \quad (14.48)$$

Since the RHS is independent of  $P$ , the LHS must also be independent of  $P$ , which in turn means that  $P dV/dx$  is independent of  $P$ . Note that the  $V dP/dx$  term is automatically zero.

Another way of reaching the same conclusion is to recall that  $PV$  is proportional to the kinetic energy of the gas molecules:  $PV = NkT = (\gamma - 1)E$ , as discussed in section 26.5. So when the reaction proceeds left to right, for each mole of gas that we get rid of, we have to account for  $RT/(\gamma - 1)$  of energy, independent of pressure.

This is an interesting result, because you might have thought that by applying pressure to the system, you could simply “push” the reaction to the right, since the RHS has a smaller volume. But it doesn't work that way. Pressurizing the system decreases the volume on both sides of the equation by the same factor. In the  $PdV$  term, the  $P$  is larger but the  $dV$  is smaller.

Also note that by combining the pressure-independence of the  $dE/dx$  term with the pressure-independence of the  $P dV/dx$  term, we find that  $dH/dx$  is pressure-independent also, where  $H$  is the enthalpy.

Now we come to the  $-TdS/dx$  term. Entropy depends on volume.

By way of background, let us temporarily consider a slightly different problem, namely one that has only  $N_1$  molecules on each side of the equation. Consider what would happen if we were to run this new reaction backwards, i.e. right to left, i.e. from  $x = 1$

to  $x = 0$ . The system volume would double, and we would pick up  $N_1 k \ln(2)$  units of entropy ... independent of pressure. It is the volume *ratio* that enters into this calculation, inside the logarithm. The ratio is independent of pressure, and therefore cannot contribute to explaining any pressure-related shift in equilibrium.

Returning to the main problem of interest, we have not  $N_1$  but  $2N_1$  molecules when  $x = 0$ . So when we run the real reaction backwards, in addition to simply letting  $N_1$  molecules expand, we have to create another  $N_1$  molecules from scratch.

For that we need the full-blown Sackur-Tetrode equation, equation 26.17, which we repeat here. For a pure monatomic nondegenerate ideal gas in three dimensions:

$$S/N = k \ln\left(\frac{V/N}{\Lambda^3}\right) + \frac{5}{2}k \quad (14.49)$$

which gives the entropy per particle in terms of the volume per particle, in an easy-to-remember form.

For the problem at hand, we can re-express this as:

$$S_i = N_i(x) k \ln\left(\frac{kT/P}{\Lambda_i^3}\right) + \frac{5}{2}k \quad (14.50)$$

where the index  $i$  runs over the three types of molecules present ( $N_2$ ,  $H_2$ , and  $NH_3$ ). We have also used the ideal gas law to eliminate the  $V$  dependence inside the logarithm in favor of our preferred variable  $P$ . We have (finally!) identified a contribution that depends on pressure and also depends on  $x$ .

We can understand the qualitative effect of this term as follows: The  $-TS$  term always contributes a drive to the left. According to equation 26.17, at higher pressure this drive will be less. So if we are in equilibrium at pressure  $P_1$  and move to a higher pressure  $P_2$ , there will be a net drive to the right.

We can quantify all this as follows: It might be tempting to just differentiate equation 14.50 with respect to  $x$  and examine the pressure-dependence of the result. However, it is easier if we start by subtracting equation 14.43 from equation 14.44, and then plug in equation 14.50 before differentiating. A lot of terms are unaffected by the change from  $P_1$  to  $P_2$ , and it is helpful if we can get such terms to drop out of the calculation sooner rather than later:

$$\begin{aligned} \begin{bmatrix} \frac{dG}{dx}(P_2) \\ -0 \end{bmatrix} &= \begin{bmatrix} \frac{dE}{dx}(P_2) + P \frac{dV}{dx}(P_2) - T \frac{dS}{dx}(P_2) \\ -\frac{dE}{dx}(P_1) - P \frac{dV}{dx}(P_1) + T \frac{dS}{dx}(P_1) \end{bmatrix} \\ &= 0 + 0 - T \frac{d}{dx} [S(P_2) - S(P_1)] \\ &= -kT \frac{d}{dx} \sum_i N_i(x) [\ln(1/P_2) - \ln(1/P_1)] \\ &= +kT \frac{d}{dx} \sum_i N_i(x) \ln(1 + \delta) \\ &= -kT \ln(1 + \delta) \end{aligned} \quad (14.51)$$

This quantitative result reinforces the previous qualitative analysis: If  $P_2$  is greater than  $P_1$ , the reaction will proceed in the  $+x$  direction, since that is what will minimize  $G$ .

The calculation involved many steps, but each step was reasonably easy.

Remark: The result is surprisingly simple. Whenever a complicated calculation produces a simple result, it is a sign that we don't really understand what's going on. I suspect there is an easier way

to obtain this result. In particular, since we have figured out that the entropy term is running the show, I conjecture that it may be possible to start from first principles and just keep track of the entropy.

**Remark:** In equation 14.51, by a first order expansion of the logarithm on the last line, you can verify that when the reaction is pushed toward equilibrium, the amount of push is proportional to  $\delta$ , which makes sense.

**Exercise:** Use a similar argument to show that increasing the temperature will shift the equilibrium of equation 14.39 to the left. Hint: In a gas-phase reaction such as this, the side with more moles of gas will have more entropy.

## 14.9 Le Châtelier’s «Principle», Or Not

One sometimes sees equilibrium (and or the shift of equilibrium) «explained» by reference to Le Châtelier’s «principle». This is highly problematic.

Le Châtelier in his lifetime gave two inconsistent statements of his so-called «principle». Restating them in modern terms:

- (1) The first version says, in effect, that all chemical equilibria are stable.
- (2) The second version says, in effect, that all *stable* chemical equilibria are stable.

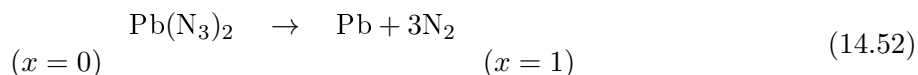
Version 2 is tautological. As such, it is not wrong ... but it is utterly uninformative.

Version 1 is just wrong, as we can see from the following examples:

**Example #1:** As a familiar, important situation, consider ice in equilibrium with liquid water, under the usual constant-pressure conditions. Let  $x$  represent the reaction coordinate, i.e. the fraction of the mass of the system that is in the form of ice. If you perturb the system by changing  $x$  – perhaps by adding water, adding ice, or adding energy – the system will exhibit zero tendency to return to its previous  $x$ -value. This system exhibits zero stability, aka neutral instability, aka neutral stability, as defined in figure 14.2.

**Example #2:** Consider an equilibrium mixture of helium and neon. Let the reaction coordinate ( $x$ ) be the fraction of the mass that is helium. If you perturb the system by increasing  $x$ , there will be no tendency for the system to react so as to decrease  $x$ . This is another example of zero stability aka neutral instability.

**Example #3:** Consider the decomposition of lead azide, as represented by the following reaction:



Initially we have  $x = 0$ . That is, we have a sample of plain lead azide. It is in thermodynamic equilibrium, in the sense that it has a definite pressure, definite temperature, et cetera.

If we perturb the system by increasing the temperature a small amount, the system will not react so as to counteract the change, not even a little bit. Indeed, if we increase the temperature enough, the system will explode, greatly *increasing* the temperature.

We say that this system is unstable. More specifically, it exhibits negative stability, as defined in figure 14.2.

**Example #4:** Suppose you raise the temperature in an RBMK nuclear reactor. Do you think the reaction will shift in such a way as to counteract the increase? It turns out that's not true, as some folks in Chernobyl found out the hard way in 1986.

**Example #5:** A simple mechanical example: Consider a perfectly balanced wheel, as in the middle example in the bottom row of figure 14.2. It is in equilibrium. If you rotate it to a new angle, it will not exhibit any tendency to return to its original state. This is neutral stability, i.e. neutrally-stable equilibrium.

**Example #6:** Consider a wheel with an off-center weight at top dead center, as in the lower-right example in figure 14.2. It is in equilibrium, but it is unstable.

**Suggestion:** If you want to talk about equilibrium and stability, use the standard terminology, namely equilibrium and stability, as defined in figure 14.1 and figure 14.2. There is no advantage to mentioning Le Châtelier's ideas in any part of this discussion, because the ideas are wrong. If you want to remark that "most" chemical equilibria encountered in the introductory chemistry course are stable equilibria, that's OK ... but such a remark must not be elevated to the status of a "principle," because there are many counterexamples.

Note that Lyapunov's detailed understanding of what stability means actually predates Le Châtelier's infamous «principle» by several years.

When first learning about equilibrium, stability, and damping, it is best to start with a one-dimensional system, such as the bicycle wheel depicted in figure 14.2. Then move on to multi-dimensional systems, such as an egg, which might be stable in one direction but unstable in another direction. Also note that in a multi-dimensional system, even if the system is stable, there is no reason to expect that the restoring force will be directly antiparallel to the perturbation. Very commonly, the system reacts by moving sideways, as discussed in section 14.5.

## 14.10 Appendix: The Cyclic Triple Derivative Rule

In this section, we derive a useful identity involving the partial derivatives of three variables. This goes by various names, including cyclic chain rule, cyclic partial derivative rule, cyclic identity, Euler's chain rule, et cetera. We derive it twice: once graphically in section 14.10.1 and once analytically in section 14.10.3. For an example where this rule is applied, see section 14.4.

### 14.10.1 Graphical Derivation

In figure 14.13 the contours of constant  $x$  are shown in blue, the contours of constant  $y$  are shown in black, and the contours of constant  $z$  are shown in red. Even though there are three variables, there are only two degrees of freedom, so the entire figure lies in the plane.

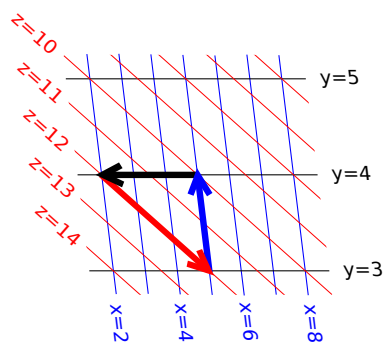


Figure 14.13: Cyclic Triple Derivative

The red arrow corresponds to:

$$\left. \frac{\partial x}{\partial y} \right|_z = \frac{3}{-1} = \frac{\# \text{ of blue contours}}{\# \text{ of black contours}} \text{ crossed by red arrow} \quad (14.53)$$

and can be interpreted as follows: The arrow runs along a contour of constant  $z$ , which is the  $z$  on the LHS of equation 14.53. The arrow crosses three contours of constant  $x$ , which is the numerator in equation 14.53. It crosses one contour of constant  $y$ , in the direction of decreasing  $y$ , which is the denominator.

Collecting results for all three vectors, we have:

$$\begin{aligned} \left. \frac{\partial x}{\partial y} \right|_z &= \frac{3}{-1} = \frac{\# \text{ of blue contours}}{\# \text{ of black contours}} \text{ crossed by red arrow} \\ \left. \frac{\partial y}{\partial z} \right|_x &= \frac{1}{-2} = \frac{\# \text{ of black contours}}{\# \text{ of red contours}} \text{ crossed by blue arrow} \\ \left. \frac{\partial z}{\partial x} \right|_y &= \frac{2}{-3} = \frac{\# \text{ of red contours}}{\# \text{ of blue contours}} \text{ crossed by black arrow} \end{aligned} \quad (14.54)$$

and if we multiply those together, we get:

$$\left. \frac{\partial x}{\partial y} \right|_z \left. \frac{\partial y}{\partial z} \right|_x \left. \frac{\partial z}{\partial x} \right|_y = -1 \quad (14.55)$$

### 14.10.2 Validity is Based on Topology

Note the contrast:

The cyclic triple derivative identity is a *topological* property. That is, you can rotate or stretch figure 14.13 however you like, and the result will be the same: the product of the three partial derivatives will always be  $-1$ . All we have to do is count contours, i.e. the number of contours crossed by each of the arrows.

The result does not depend on any *geometrical* properties of the situation. No metric is required. No dot products are required. No notion of length or angle is required. For example, as drawn in figure 14.13, the  $x$  contours are not vertical, the  $y$  contours are not horizontal, and the  $x$  and  $y$  contours are not mutually perpendicular ... but more generally, we don't even need to have a way of knowing whether the contours are horizontal, vertical, or perpendicular.

Similarly, if you rescale one set of contours, perhaps by making the contours twice as closely spaced, it has no effect on the result, because it just increases one of the numerators and one of the denominators in equation 14.55.

The validity of equation 14.55 depends on the following topological requirement: The three vectors must join up to form a triangle. This implies that the contours  $\{dx, dy, dz\}$  must *not* be linearly independent. In particular, you cannot apply equation 14.55 to the Cartesian  $X$ ,  $Y$ , and  $Z$  axes.

Validity also depends on another topological requirement: The contour lines must not begin or end within the triangle formed by the three vectors. We are guaranteed this will always be true, because of the fundamental theorem that says  $d(\text{anything})$  is exact ... or, equivalently,  $d(d(\text{anything})) = 0$ . In words, the theorem says “the boundary of a boundary is zero” or “a boundary cannot peter out.” This theorem is discussed in reference 4.

We apply this idea as follows: Every contour line that goes into the triangle has to go out again. From now on, let's only count *net* crossings, which means if a contour goes out across the same edge where it came in, that doesn't count at all. Then we can say that any blue line that goes in across the red arrow must go out across the black arrow. (It can't go out across the blue arrow, since the blue arrow lies along a blue contour, and contours can't cross.) To say the same thing more quantitatively, the number of net blue crossings inward across the red arrow equals the number of net blue crossings outward across the black arrow. This number of crossings shows up on the LHS of equation 14.55 twice, once as a numerator and once as a denominator. In one place or the other, it will show up with a minus sign. Assuming this number is nonzero, its appearance in a numerator cancels its appearance in a denominator, so all in all it contributes a factor of  $-1$  to the product. Taking all three variables into account, we get three factors of  $-1$ , which is the right answer.

Here is yet another way of saying the same thing. To simplify the language, let's interpret the  $x$ -value as "height." The blue arrow lies along a contour of constant height. The black arrow goes downhill a certain amount, while the red arrow goes uphill by the *same* amount. The amount must be the same, for the following two reasons: At one end, the red and black arrows meet at a point, and  $x$  must have some definite value at this point. At the other end, the red and black arrows terminate on the same contour of constant  $x$ . This change in height, this  $\Delta x$ , shows up on the LHS of equation 14.55 twice, once as a numerator and once as a denominator. In one place or the other, it shows up with a minus sign. This is guaranteed by the fact that when the arrows meet, they meet tip-to-tail, so if one of the pair is pointed downhill, the other must be pointed uphill.

### 14.10.3 Analytic Derivation

Let's start over, and derive the result again. Assuming  $z$  can be expressed as a function of  $x$  and  $y$ , and assuming everything is sufficiently differentiable, we can expand  $dz$  in terms of  $dx$  and  $dy$  using the chain rule:

$$dz = \left. \frac{\partial z}{\partial x} \right|_y dx + \left. \frac{\partial z}{\partial y} \right|_x dy \quad (14.56)$$

By the same token, we can expand  $dx$  in terms of  $dy$  and  $dz$ :

$$dx = \left. \frac{\partial x}{\partial y} \right|_z dy + \left. \frac{\partial x}{\partial z} \right|_y dz \quad (14.57)$$

Using equation 14.57 to eliminate  $dx$  from equation 14.56, we obtain:

$$dz = \frac{\partial z}{\partial x|y} \left( \frac{\partial x}{\partial y|z} dy + \frac{\partial x}{\partial z|y} dz \right) + \frac{\partial z}{\partial y|x} dy \quad (14.58)$$

hence

$$\left( 1 - \frac{\partial z}{\partial x|y} \frac{\partial x}{\partial z|y} \right) dz = \left( \frac{\partial z}{\partial x|y} \frac{\partial x}{\partial y|z} + \frac{\partial z}{\partial y|x} \right) dy \quad (14.59)$$

We are free to choose  $dz$  and  $dy$  arbitrarily and independently, so the only way that equation 14.59 can hold in general is if the parenthesized factors on each side are identically zero. From the LHS of this equation, we obtain the rule for the reciprocal of a partial derivative. This rule is more-or-less familiar from introductory calculus, but it is nice to know how to properly generalize it to partial derivatives:

$$\left. \frac{\partial z}{\partial x} \right|_y \left. \frac{\partial x}{\partial z} \right|_y = 1 \quad (14.60)$$



Meanwhile, from the parenthesized expression on the RHS of equation 14.59, with a little help from equation 14.60, we obtain the cyclic triple chain rule, the same as in section 14.10.1:

$$\left. \frac{\partial x}{\partial y} \right|_z \left. \frac{\partial y}{\partial z} \right|_x \left. \frac{\partial z}{\partial x} \right|_y = -1 \quad (14.61)$$

#### 14.10.4 Independent and Dependent Variables, or Not

In this situation, it is clearly not worth the trouble of deciding which are the “independent” variables and which are the “dependent” variables. If you decide based on equation 14.56 (which treats  $z$  as depending on  $x$  and  $y$ ) you will have to immediately change your mind based on equation 14.57 (which treats  $x$  as depending on  $y$  and  $z$ ).

Usually it is best to think primarily in terms of abstract points in thermodynamic state-space. You can put your finger on a point in figure 14.13 and thereby identify a point without reference to its  $x$ ,  $y$ , or  $z$  coordinates. The point doesn’t care which coordinate system (if any) you choose to use. Similarly, the vectors in the figure can be added graphically, tip-to-tail, without reference to any coordinate system or basis.

If and when we have established a coordinate system:

- Given a point, you can determine its  $x$ ,  $y$ , and  $z$  coordinates.
- Given  $x$  and  $y$ , you can locate the point and determine its properties, including its  $z$  coordinate.
- Equally well, given  $y$  and  $z$ , you can locate the point and determine its properties, including its  $x$  coordinate.

#### 14.10.5 Axes, or Not

Note that there are no axes in figure 14.13, strictly speaking. There are contours of constant  $x$ , constant  $y$ , and constant  $z$ , but no actual  $x$ -axis,  $y$ -axis, or  $z$ -axis.

In other, simpler situations, you can of course get away with a plain horizontal axis and a plain vertical axis, but you don’t want to become too attached to this approach. Even in cases where you can get away with plain axes, it is a good habit to plot the grid also (unless there is some peculiar compelling reason not to). Modern software makes it super-easy to include the grid.

Make it a habit to include the contours.

For more on this, see reference 42.

### 14.11 Entropy versus “Irreversibility” in Chemistry

In chemistry, the word “irreversible” is commonly used in connection with multiple inconsistent ideas, including:

- The reaction is spontaneous.

- The reaction strongly goes to completion.
- The reaction is thermodynamically irreversible.

Those ideas are not completely unrelated ... but they are not completely identical, and there is potential for serious confusion.

You cannot look at a chemical reaction (as written in standard form) and decide whether it is spontaneous, let alone whether it goes to completion. For example, consider the reaction



If you flow steam over hot iron, you produce iron oxide plus hydrogen. This reaction is used to produce hydrogen on an industrial scale. It goes to completion in the sense that the iron is used up. Conversely, if you flow hydrogen over hot iron oxide, you produce iron and  $\text{H}_2\text{O}$ . This is the reverse of equation 14.62, and it also goes to completion, in the sense that the iron oxide is used up.

What's more, none of that has much to do with whether the reaction was thermodynamically reversible or not.

In elementary chemistry classes, people tend to pick up wrong ideas about thermodynamics, because the vast preponderance of the reactions that they carry out are grossly irreversible, i.e. irreversible by state, as discussed in section 14.3.6. The reactions are nowhere near isentropic.

Meanwhile, there are plenty of chemical reactions that are very nearly reversible, i.e. irreversible by rate, as discussed in section 14.3.6. In everyday life, we see examples of this, such as electrochemical reactions, e.g. storage batteries and fuel cells. Another example is the  $\text{CO}_2$ /carbonate reaction discussed below. Alas, there is a tendency for people to forget about these reversible reactions and to unwisely assume that all reactions are grossly irreversible. This unwise assumption can be seen in the terminology itself: widely-used tables list the “standard heat of reaction” (rather than the standard *energy* of reaction), apparently under the unjustifiable assumption that the energy liberated by the reaction will always show up as heat. Similarly reactions are referred to as “exothermic” and “endothermic,” even though it would be much wiser to refer to them as exergonic and endergonic.

It is very difficult, perhaps impossible, to learn much about thermodynamics by studying bricks that fall freely and smash against the floor. Instead, thermodynamics is most understandable and most useful when applied to situations that have relatively little dissipation, i.e. that are nearly isentropic.

Lots of people get into the situation where they have studied tens or hundreds or thousands of reactions, all of which are irreversible by state. That's a trap for the unwary. It would be unwise to leap to the conclusion that all reactions are far from isentropic ... and it would be even more unwise to leap to the conclusion that “all” natural processes are far from isentropic.

Chemists are often called upon to teach thermodynamics, perhaps under the guise of a “P-Chem” course (i.e. physical chemistry). This leads some people to ask for purely *chemical* examples to illustrate entropy and other thermodynamic ideas. I will answer the question in a moment, but first let me register my strong objections to the question. Thermodynamics derives its great power and elegance from its wide generality. Specialists who cannot cope with examples outside their own narrow specialty ought not be teaching thermodynamics.

Here's a list of reasons why a proper understanding of entropy is directly or indirectly useful to chemistry students.

1. Consider electrochemical reactions. Under suitable conditions, some electrochemical reactions can be made very nearly reversible in the thermodynamic sense. (See reference 43 for some notes on how such cells work.) In these cases, the heat of reaction is very much less than the energy of reaction, and the entropy is very much less than the energy divided by  $T$ .
2. Consider the reaction that children commonly carry out, adding vinegar to baking soda, yielding sodium acetate and carbon dioxide gas. Let's carry out this reaction in a more grown-up apparatus, namely a sealed cylinder with a piston. By pushing on the piston with weights and springs, we can raise the pressure of the  $\text{CO}_2$  gas. If we raise the pressure high enough, we push  $\text{CO}_2$  back into solution. This in turn raises the activity of the carbonic acid, and at some point it becomes a strong enough acid to attack the sodium acetate and partially reverse the reaction, liberating acetic acid. So this is clearly and inescapably a *chemistry* situation.

Much of the significance of this story revolves around the fact that if we arrange the weights and springs just right, the whole process can be made thermodynamically reversible (nearly enough for practical purposes). Adding a tiny bit of weight will make the reaction go one way, just as removing a tiny bit of weight will make the reaction go the other way.

Now some interesting questions arise: Could we use this phenomenon to build an engine, in analogy to a steam engine, but using  $\text{CO}_2$  instead of steam, using the carbonate  $\leftrightarrow \text{CO}_2$  chemical reaction instead of the purely physical process of evaporation? How does the  $\text{CO}_2$  pressure in this system vary with temperature? How much useful work would this  $\text{CO}_2$  engine generate? How much waste heat? What is the best efficiency it could possibly have? Can we run the engine backwards so that it works as a refrigerator?

There are more questions of this kind, but you get the idea: once we have a reaction that is more-or-less thermodynamically reversible, we can bring to bear the entire machinery of thermodynamics.

3. Consider the colligative effects of a solute on the on freezing point, boiling point, and vapor pressure of a solvent. The fact that they're colligative – i.e. insensitive to the chemical properties of the solute – is strong evidence that entropy is what's driving these effects, not enthalpy, energy, or free energy.
4. Similarly: consider the Gibbs Gedankenexperiment (section 11.6). Starting with a sample of  $^4\text{He}$ , we get an increase in entropy if we mix it with  $^3\text{He}$ , or Ne, or Xe ... but we get no effect if we “mix” it with more of the same  $^4\text{He}$ .
5. People who take chemistry classes often go on to careers in other fields. For example, you might need knowledge of chemistry, physics, and engineering in order to design a rocket engine, or a jet engine, or a plain old piston engine. Such things commonly involve a chemical reaction followed by a more-or-less isentropic expansion. Even though the chemical reaction is grossly irreversible, understanding the rest of the system requires understanding thermodynamics.

To be really specific, suppose you are designing something with multiple heat engines in series. This case is considered as part of the standard “foundations of thermodynamics” argument, as illustrated figure 14.14. Entropy is conserved as it flows down the totem-pole of heat engines. The crucial conserved quantity that is the same for all the engines is *entropy* ... not energy,

free energy, or enthalpy. No entropy is lost during the process, because entropy cannot be destroyed, and no entropy (just work) flows out through the horizontal arrows. No entropy is created, because we are assuming the heat engines are 100% reversible. For more on this, see reference 8.

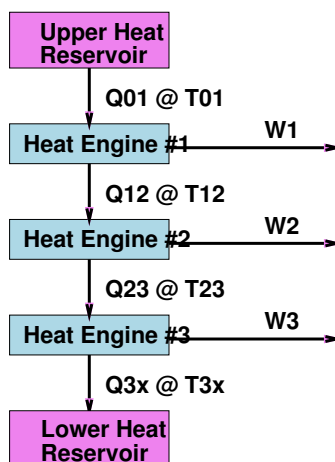


Figure 14.14: Heat Engines In Series

6. Consider “Design of Experiment,” as discussed in reference 13. In this case the entropy of interest is not the entropy of the reaction, but still it is entropy, calculated in accordance with equation 2.2, and it is something a chemist ought to know. Research chemists and especially chemical engineers are often in the situation where experiments are very expensive, and someone who doesn’t understand Design of Experiment will be in big trouble.

## Chapter 15

# The “Big Four” Energy-Like State Functions

### 15.1 Energy

The energy is one of the “big four” thermodynamic potentials.

The concept of energy has already been introduced; see chapter 1.

### 15.2 Enthalpy

#### 15.2.1 Integration by Parts; $PV$ and its Derivatives

Let’s establish a little bit of background. We will soon need to use the fact that

$$d(PV) = PdV + VdP \quad (15.1)$$

which is just the rule for differentiating a product. This rule applies to any two variables (not just  $P$  and  $V$ ), provided they were differentiable to begin with. Note that this rule is intimately related to the idea of integrating by parts, as you can see by writing it as

$$PdV = d(PV) - VdP \quad (15.2a)$$

$$\int PdV = \Delta(PV) - \int VdP \quad (15.2b)$$

Note that the RHS of equation 15.2b could be obtained two ways, either by direct integration of the RHS of equation 15.2a, or by integrating the LHS, integrating by parts.

The idea behind integration by parts can be visualized using the indicator diagram in figure 15.1. The black dividing line represents the equation of state, showing  $P$  as a function of  $V$  (and vice versa). The blue-shaded region is the area under the curve, namely  $\int PdV$ , where the  $P$  in the integrand is not any old  $P$ , but rather the  $P(V)$  you compute as a function of  $V$ , according to the equation of state. By the same logic, the red-shaded region has area  $\int VdP$ , by which we mean  $\int V(P)dP$ , computing  $V(P)$  as function of  $P$ , according to the equation of state.

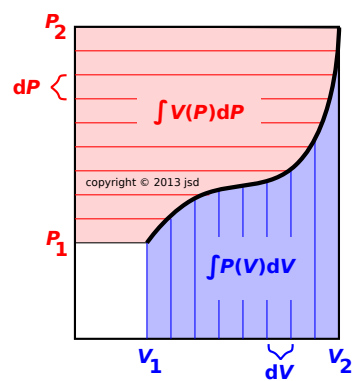


Figure 15.1: Integration by Parts

The large overall rectangle has height  $P_2$ , width  $V_2$ , and area  $P_2 V_2$ . The small unshaded rectangle has area  $P_1 V_1$ . So the entire shaded region has area  $\Delta(PV)$ . One way of looking at it is to say the area of any rectangle is  $\int \int dP dV$  where here  $P$  and  $V$  range over the whole rectangle, without regard to the equation of state.

If you think of  $\int P dV$  as the work, you can think of  $\int V dP$  as the *krow*, i.e. the reverse work, i.e. the shaded area minus the work. Equivalently we can say that it is the entire area of the indicator diagram, minus the small unshaded rectangle, minus the work.

We can rewrite equation 15.2b as

$$\int P dV + \int V dP = \Delta(PV) \quad (15.3)$$

which has the interesting property that the RHS depends only on the endpoints of the integration, independent of the details of the equation of state. This is in contrast to the LHS, where the work  $\int P dV$  depends on the equation of state and the *krow*  $\int V dP$  also depends on the equation of state. If we hold the endpoints fixed and wiggle the equation of state, any wiggle that increases the work decreases the *krow* and vice versa.

In the case where  $P$  happens to be a decreasing function of  $V$ , the picture looks different, but the bottom-line mathematical formula is the same and the meaning is the same.

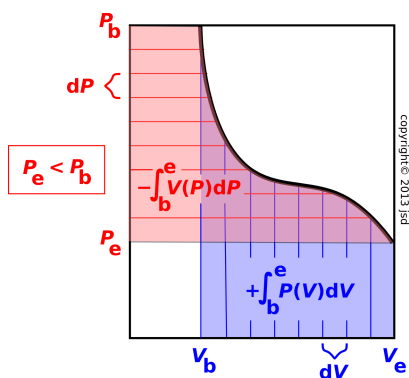


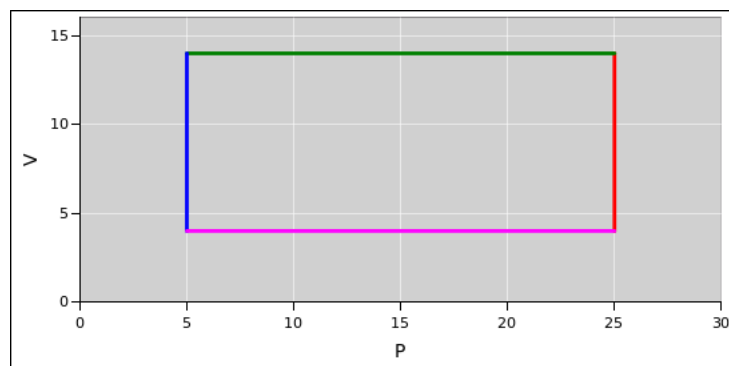
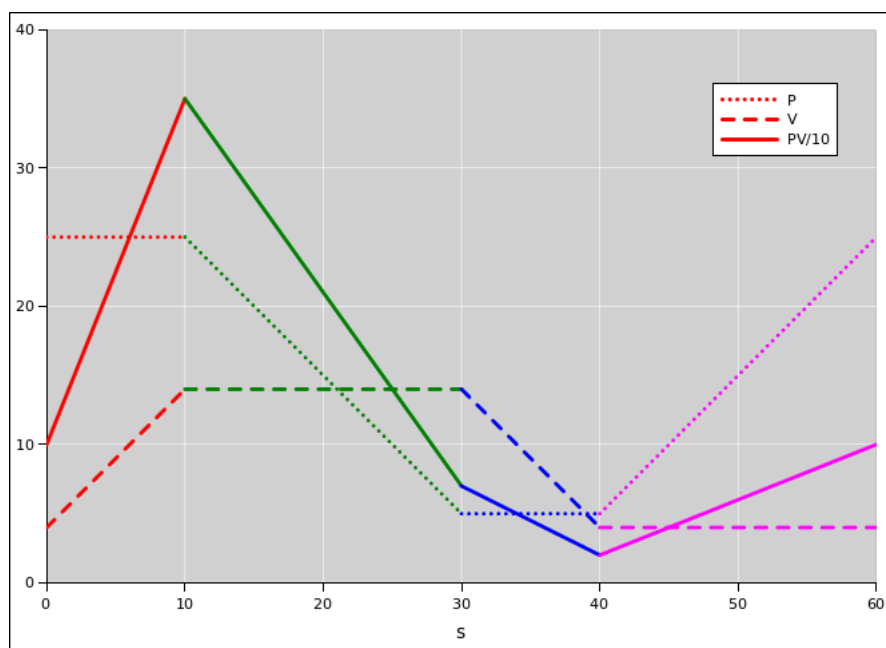
Figure 15.2: Integration by Parts, Decreasing Function

Here subscript “b” stands for beginning, while subscript “e” stands for ending. The trick is to notice that the final pressure  $P_e$  is less than the initial pressure  $P_b$ , so the integral of  $V dP$  is negative. If we keep the endpoints the same and wiggle the equation of state, anything that makes the work  $\int P dV$  more positive makes the *krow*  $\int V dP$  more negative.

### 15.2.2 More About $PdV$ versus $VdP$

Consider the thermodynamic cycle shown on the indicator diagram in figure 15.3.

In figure 15.4, we plot  $P$ ,  $V$ , and  $PV$  as functions of arc length  $s$ , as we go around the cycle, starting from the southeast corner. These are functions of state, so at the end of the cycle, they will certainly return to their initial values.

Figure 15.3: Indicator Diagram:  $(P, V)$  SpaceFigure 15.4: Functions of State:  $P$ ,  $V$ , and  $PV$



In figure 15.5, we plot some integrals. The integral of  $PdV$  is not a function of state; its value depends on the path whereby we reached the state. Ditto for the integral of  $VdP$ .

Interestingly, the sum of  $\int PdV$  plus  $\int VdP$  is a function of state, because it is the integral of a gradient, namely the integral of  $d(PV)$ . In contrast,  $PdV$  is not the gradient of anything, and  $VdP$  is not the gradient of anything. See section 8.2 for details on this.

You can verify that on a point-by-point basis, the  $\int d(PV)$  curve is the algebraic sum of the other two curves.

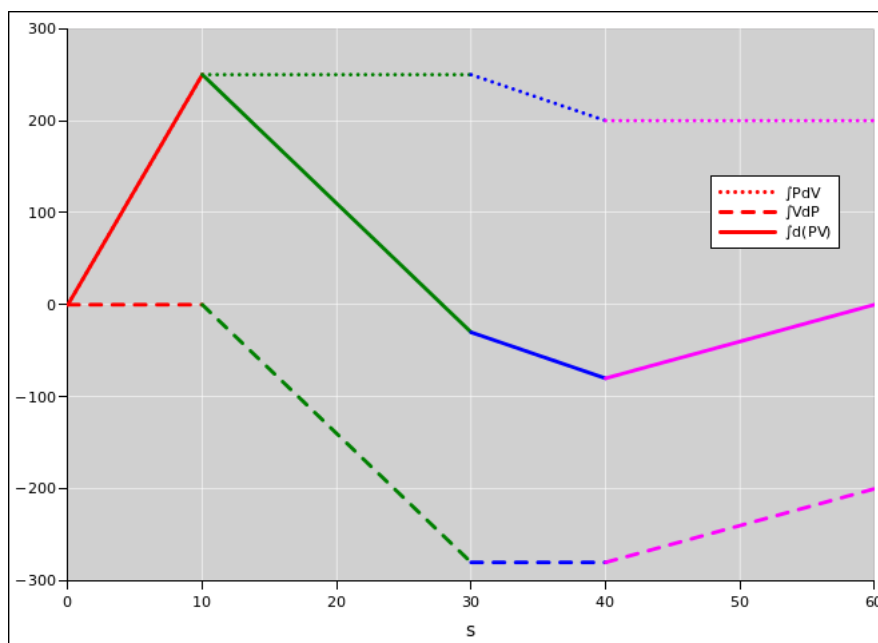


Figure 15.5: Integrals Along the Path

Figure 15.6 is another way of presenting the same basic idea. The shaded areas represent the derivative of  $PV$ .

In this section, in a departure from our usual practice, we are differentiating things with respect to the arc length  $s$  (rather than using the more sophisticated idea of gradient). This is not a good idea in general, but it is expedient in this special case. The whole notion of arc length is arbitrary and unphysical, because there is no natural notion of distance or angle in thermodynamic state space. If we rescaled the axes, it would have not the slightest effect on the real physics, it would change the arc length.

Because  $PV$  is a function of state, we know the area above the axis is equal to the area below the axis. When integrated over the whole cycle, the  $PdV$  contributions (red and blue) must be equal-and-opposite to the  $VdP$  contributions (green and magenta).

In other words, when we integrate over the whole cycle, we find that the total work is equal-and-opposite to the total krow. This applies to the whole cycle only; if we look at any particular leg

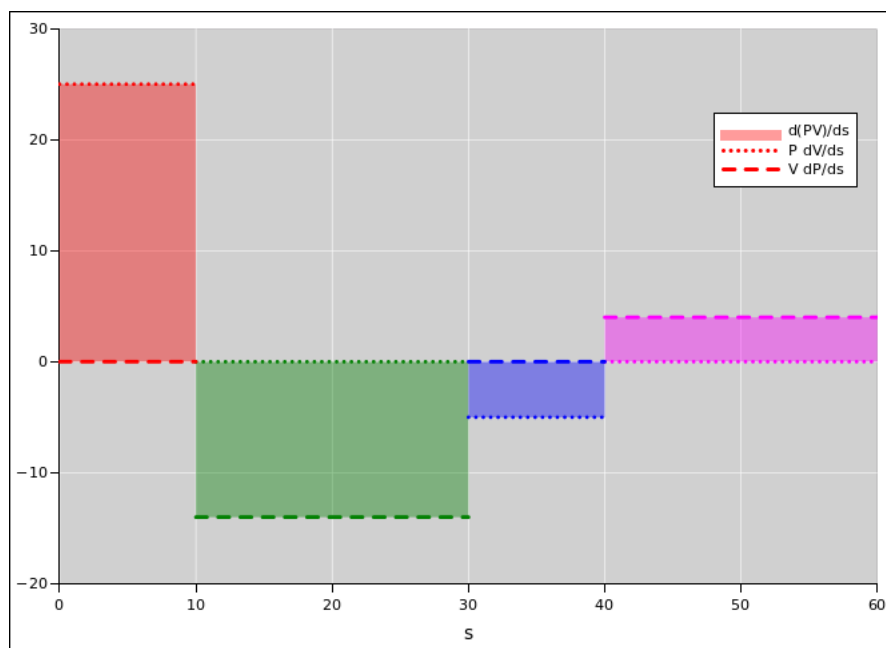


Figure 15.6: Areas Under the Curve

of the cycle, or any other subset of the cycle, there is no reason to expect any simple relationship between the total work and the total know. It is usually better to think in terms of the simple yet powerful local relationship between the derivatives:  $d(PV) = PdV + VdP$ .

### 15.2.3 Definition of Enthalpy

We hereby define the enthalpy as:

$$H := E + PV \quad (15.4)$$

where  $H$  is the near-universally conventional symbol for enthalpy,  $E$  is the energy,  $V$  is the volume of the system, and  $P$  is the pressure on the system. We will briefly explore some of the mathematical consequences of this definition, and then explain what enthalpy is good for.

Differentiating equation 15.4 and using equation 7.8 and equation 15.1, we find that

$$\begin{aligned} dH &= -PdV + TdS + PdV + VdP \\ &= VdP + TdS \end{aligned} \quad (15.5)$$

which runs nearly parallel to equation 7.8; on the RHS we have transformed  $-PdV$  into  $VdP$ , and of course the LHS is enthalpy instead of energy.

This trick of transforming  $x dy$  into  $-y dx$  (with a leftover  $d(xy)$  term) is an example of a *Legendre transformation*. In general, the Legendre transformation is a very powerful tool with many applications, not limited to thermodynamics. See reference 44.

In the chemistry lab, it is common to carry out reactions under conditions of constant pressure. If the reaction causes the system to expand or contract, it will do work against the surroundings. This work will change the energy ... but it will not change the enthalpy, because  $dH$  depends on  $VdP$ , and  $dP$  is zero, since we assumed constant pressure.

This means that under conditions of constant pressure, it is often easier to keep track of the enthalpy than to keep track of the energy.

For simple reactions that take place in aqueous solution, usually both  $dP$  and  $dV$  are zero to a good approximation, and there is no advantage to using enthalpy instead of energy. In contrast, gas-phase reactions commonly involve a huge change in volume. Consider for example the decay of ozone:  $2\text{O}_3 \rightarrow 3\text{O}_2$ .

Enthalpy is important in fluid dynamics. For example, Bernoulli’s principle, which is often a convenient way to calculate the pressure, can be interpreted as a statement about the enthalpy of a parcel of fluid. It is often claimed to be a statement about energy, but this claim is bogus. The claim is plausible at the level of dimensional analysis, but that’s not good enough; there is more to physics than dimensional analysis.

See section 15.8 for an example of how enthalpy can be put to good use.

### 15.2.4 Enthalpy is a Function of State

It is important to keep in mind that whenever we write something like  $H = E + PV$ , it is shorthand for

$$H_C = E_C + P_C V_C \quad (15.6)$$

where every variable in this equation is a function of state, i.e. a function of the state of the stuff in the cell  $C$ . Specifically

- $H_C$  refers to the enthalpy *within the cell*,
- $E_C$  refers to the energy *within the cell*,
- $P_C$  refers to the pressure *within the cell*,
- $V_C$  refers to the volume *within the cell*,

It is important that these things be functions of state. This is an important issue of principle. The principle is sometimes obscured by the fact that

- Sometimes it is possible to equate the pressure in the cell  $P_C$  to the ambient pressure  $P_A$ .
- Sometimes it is possible to related the entropy in the cell  $S_C$  to the entropy in the heat bath  $S_H$ .
- et cetera.

So, you might be tempted to write something like;

$$H_C = E_C + P_A V_C \quad \times \quad (15.7)$$

However, that is risky, and if you do something like that, it is up to you to demonstrate, on a case-by-case basis, that it is OK. Beware beware that there are lots of situations where it is not OK. For example:

- In an air gun, the pressure in the cell  $P_C$  is definitely not equal to the ambient atmospheric pressure  $P_A$ .
- In a dissipative system,  $dS_C$  (in the cell) is definitely not equal to  $-dS_H$  (in the heat bath).
- et cetera.

### 15.2.5 Derivatives of the Enthalpy

It is informative to differentiate  $H$  with respect to  $P$  and  $S$  directly, using the chain rule. This gives us:

$$dH = \left. \frac{\partial H}{\partial P} \right|_S dP + \left. \frac{\partial H}{\partial S} \right|_P dS \quad (15.8)$$

which is interesting because we can compare it, term by term, with equation 15.5. When we do that, we find that the following identities must hold:

$$V = \left. \frac{\partial H}{\partial P} \right|_S \quad (15.9)$$

and

$$T = \left. \frac{\partial H}{\partial S} \right|_P \quad (15.10)$$

which can be compared to our previous definition of temperature, equation 7.7, which says

$$T := \left. \frac{\partial E}{\partial S} \right|_V \quad (15.11)$$

(assuming the derivative exists). Equation 15.10 is not meant to redefine  $T$ ; it is merely a useful corollary, based on of our earlier definition of  $T$  and our definition of  $H$  (equation 15.4).

## 15.3 Free Energy

In many situations – for instance when dealing with heat engines – it is convenient to keep track of the *free energy* of a given parcel. This is also known as the *Helmholtz potential*, or the Helmholtz free energy. It is defined as:

$$F := E - T S \quad (15.12)$$

where  $F$  is the conventional symbol for free energy,  $E$  is (as always) the energy,  $S$  is the entropy, and  $T$  is the temperature of the parcel.

The free energy is extremely useful for analyzing the spontaneity and reversibility of transformations taking place at constant  $T$  and constant  $V$ . See chapter 14 for details.

See section 15.5 for a discussion of what is (or isn’t) “free” about the free energy.

The energy and the free energy are related to the partition function, as discussed in chapter 24.

## 15.4 Free Enthalpy

Combining the ideas of section 15.2 and section 15.3, there are many situations where it is convenient to keep track of the *free enthalpy*. This is also known as the *Gibbs potential* or the Gibbs free enthalpy. It is defined as:

$$\begin{aligned} G &= E + P V - T S \\ &= H - T S \end{aligned} \tag{15.13}$$

where  $G$  is the conventional symbol for free enthalpy. (Beware:  $G$  is all-too-commonly called the Gibbs free “energy” but that is a bit of a misnomer. Please call it the free *enthalpy*, to avoid confusion between  $F$  and  $G$ .)

The free enthalpy has many uses. For starters, it is extremely useful for analyzing the spontaneity and reversibility of transformations taking place at constant  $T$  and constant  $P$ , as discussed in chapter 14. (You should not however imagine that  $G$  is *restricted* to constant- $T$  and/or constant- $P$  situations, for reasons discussed in section 15.6.)

## 15.5 Thermodynamically Available Energy – Or Not

The notion of “available energy” content in a region is mostly a bad idea. It is an idea left over from cramped thermodynamics that does not generalize well to uncramped thermodynamics.

### 15.5.1 Overview

The notion of “free energy” is often misunderstood. Indeed the term “free energy” practically begs to be misunderstood.

It is superficially tempting to divide the energy  $E$  into two pieces, the “free” energy  $F$  and the “unfree” energy  $TS$ , but that’s just a pointless word-game as far as I can tell, with no connection to the ordinary meaning of “free” and with no connection to useful physics, except possibly in a few unusual situations.

- As discussed in section 1.7, the amount of energy available for doing work is not the same as the total energy  $E$ .

- As we shall see in this section, the available energy is also not equal to the free energy  $F = E - TS$ .

To repeat: You should not imagine that the free energy is the “thermodynamically available” part of the energy. Similarly you should not imagine that  $TS$  is the “unavailable” part of the energy.

The free energy of a given parcel is a function of state, and in particular is a function of the thermodynamic state *of that parcel*. That is, for parcel #1 we have  $F_1 = E_1 - T_1 S_1$  and for parcel #2 we have  $F_2 = E_2 - T_2 S_2$ .

Suppose we hook up a heat engine as shown in figure 15.7. This is virtually the same as figure 1.2, except that here we imagine that there are two heat-sinks on the cold side, namely region #2 and region #3. Initially heat-sink #2 is in use, and heat-sink #3 is disconnected. We imagine that the heat-reservoir on the high side (region #1) is has much less heat capacity than either of the heat-sinks on the low side. Also, we have added an anvil so that the heat engine can do work against the gravitational field.

Assume the heat engine is maximally efficient. That is to say, it is reversible. Therefore its efficiency is the Carnot efficiency,  $(T_1 - T_2)/T_2$ . We see that the amount of “thermodynamically available” energy depends on  $T_2$ , whereas the free energy of parcel #1 does not. In particular, if  $T_2$  is cold enough, the work done by the heat engine will exceed the free energy of parcel #1. Indeed, in the limit that parcel #2 is very large and very cold (approaching absolute zero), the work done by the heat engine will converge to the entire energy  $E_1$ , not the free energy  $F_1$ .

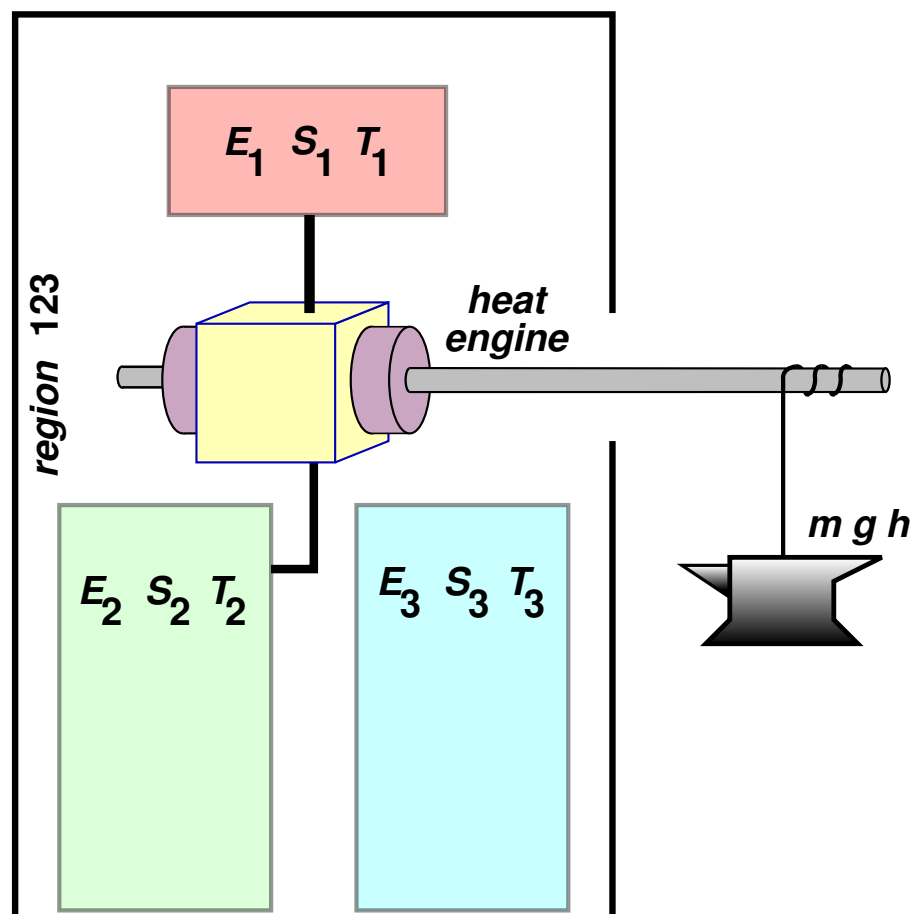
We can underline this point by switching the cold-side connection from region #2 to region #3. This changes the amount of energy that we can get out of region #1, even though there has been no change in the state of region #1. This should prove beyond all doubt that “available energy” is not equal to the Helmholtz free energy  $F = E - TS$ . Similarly it is not equal to the Gibbs free enthalpy  $G = H - TS$ . It’s not even a function of state.

- It is possible to construct scenarios in which there is a well-defined notion of available energy – available within that scenario. See section 15.5.2 for an example.
- On the other hand, all-too-often people throw around the term “available energy” in situations where it cannot possibly have any well-defined meaning. For example, there is – in general – no way to attribute a definite amount of “available energy” to a gallon of gasoline, for reasons discussed in section 15.5.2.

You may *wish* there were a definite state-function that would quantify the “available energy” of the gasoline, but wishing does not make it so.

We can reconcile the two previous itemized points by making the distinction between a scenario-function and a state-function. Something that is well defined in a careful scenario (involving two reservoirs and numerous restrictions) is not well defined for a single reservoir (all by itself with no restrictions).

Every minute spent learning about “available energy” is two minutes wasted, because everything you learn about it will have to be unlearned before you can do real, uncramped thermodynamics.



Constructive suggestion: In any case, if you find yourself trying to quantify the so-called “thermal energy” content of something, it is likely that you are asking the wrong question. Sometimes you can salvage part of the question by considering microscopic versus macroscopic forms of energy, but even this is risky. In most cases you will be much better off quantifying something else instead. As mentioned in section 17.1:

- Sometimes “heat” means  $TdS$ . In such a case, quantify  $TdS$ .
- Don’t bother trying to quantify the “heat content” or “thermal energy content.” Instead quantify the energy and entropy. The latter are unambiguous and unproblematic.
- Very often, when people speak about “heat energy” you can just cross out the word “heat” and assume they meant “energy.” Ditto for “thermal energy.” This is consistent with the previous item, in the sense that the “heat energy” is one component of the “heat.”

See chapter 19 for more on this.

In general, you should never assume you can figure out the nature of a thing merely by looking at the name of a thing. As discussed in reference 45, a titmouse is not a kind of mouse, and buckwheat is not a kind of wheat. As Voltaire remarked, the Holy Roman Empire was neither holy, nor Roman, nor an empire. By the same token, free energy is not the “free” part of the energy.

### 15.5.2 A Calculation of “Available” Energy

You are encouraged to skip this section. It exists primarily to dispel some misconceptions. However, experience indicates that discussing a misconception is almost as likely to consolidate the misconception as to dispel it. Your best option is to accept the idea that energy and entropy are primary and fundamental, whereas heat, work, and “available energy” are not. Accept the fact that any notion of “useful energy” belongs more to microeconomics than to physics.

Alas, some people stubbornly *wish* for there to be some state-function that tells us the “available energy”; and the purpose of this section is to disabuse them of that notion.

The plan is to analyze in more detail the system shown in figure 15.7. This provides a more-detailed proof of some of the assertions made in section 15.5.1.

The idea behind the calculation is that we start out with region 1 hotter than region 2. We operate the heat engine so as to raise the weight, doing work against gravity. This extracts energy and entropy from region 1. When  $T_1$  becomes equal to  $T_2$  we have extracted all of the energy that was available in this scenario, and the height of the weight tells us how much “available” energy we started with. We can then operate the heat engine in reverse, using it as a heat pump driven by the falling weight. We can arrange the direction of pumping so that energy and entropy are extracted from region 1, so that region 1 ends up colder than region 2, even though it started out warmer.

A major purpose here is to do everything in sufficient detail so that there is no doubt what we mean by the amount of energy “available” for doing work in this scenario. We measure the “available



energy” quite simply and directly, by doing the work and measuring how much work is done. At any moment, the “available energy” is the amount of energy not (yet) transferred to the load.

This calculation is tantamount to rederiving the Carnot efficiency formula.

By conservation of energy we have:

$$\begin{aligned} E_1 + E_2 + mgh &= \text{const} \\ dE_1 + dE_2 + mgdh &= 0 \end{aligned} \quad (15.14)$$

Since the heat engine is reversible, we have:

$$dS_1 + dS_2 = 0 \quad (15.15)$$

Differentiating the energy, and using the fact that the reservoirs are held at constant volume, we have:

$$\begin{aligned} dE_1 &= T_1 dS_1 - P_1 dV_1 \\ &= T_1 dS_1 \\ dE_2 &= T_2 dS_2 - P_2 dV_2 \\ &= T_2 dS_2 \end{aligned} \quad (15.16)$$

We assume that region 1 has a constant heat capacity over the range of interest. We assume that region 2 is an enormous heat bath, with an enormous heat capacity, and therefore a constant temperature.

$$\begin{aligned} E_1 &= C v_1 T_1 \\ dT_1 &= (1/C v_1) dE_1 \\ dT_2 &= 0 \end{aligned} \quad (15.17)$$

Plugging equation 15.17 into equation 15.16 we obtain:

$$\begin{aligned} dE_1 &= (E_1/C v_1) dS_1 \\ d \ln(E_1) &= (1/C v_1) dS_1 \end{aligned} \quad (15.18)$$

Doing some integrals gives us:

$$\begin{aligned} \ln(E_1) &= (1/C v_1) S_1 + \text{const}_1 \\ E_1 &= \alpha \exp(S_1/C v_1) \\ E_2 &= T_2 S_2 + \text{const}_2 \\ &= -T_2 S_1 + \text{const}_3 \end{aligned} \quad (15.19)$$

Remember: At any moment, the available energy is the energy that has not (yet) been transferred to the external load. Using conservation of energy (equation 15.14), we obtain:

$$\begin{aligned} -mgh &\equiv \text{available energy} \\ &= E_1 + E_2 \\ &= E_1 - T_2 S_1 + \text{const}_5 \end{aligned} \quad (15.20)$$

Here’s the punchline: the RHS here is not the free energy  $F_1 := E_1 - T_1 S_1$ . It is “almost” of the right form, but it involves  $T_2$  instead of  $T_1$ . It is provably not possible to write the available energy as  $F_1$  or  $F_2$  or  $F_1 + F_2$  (except in the trivial case where  $T_1 \equiv T_2$ , in which case nobody is interested, because the available work is zero).

Equation 15.20 is a thinly-disguised variation of the Carnot efficiency formula. You should have known all along that the available energy could not be written as  $F_1$  or  $F_2$  or any linear combination of the two, because the Carnot efficiency formula depends on *both*  $T_1$  and  $T_2$ , and it is nonlinear.

Here's another way you should have known that the available energy cannot possibly correspond to the free energy – without doing more than one line of calculation: Look at the definition  $F_1 := E_1 - T_1 S_1$  and consider the asymptotic behavior. In the case of an ideal gas or anything else with a constant heat capacity, at high temperatures  $F$  must be negative. It must have a downward slope. Indeed it must be concave downward. Surely nobody thinks that a hot fluid has a negative available energy, or that the hotter it gets the less useful it is for driving a heat engine.

Thirdly and perhaps most importantly, we note again that we can change the amount of “available energy” by switching the lower side of the heat engine from heat-sink 2 to heat-sink 3. This changes the amount of work that the engine can do, without change the state of any of the reservoirs.

We now introduce  $\Delta S$  to represent the amount of entropy transferred to region 1 from region 2 (via the heat engine), and we choose the constants such that  $\Delta S$  is zero when the two regions are in equilibrium. That allows us to write the “available energy” as a simple function of  $\Delta S$ :

$$-mgh = Cv_1 T_2 [\exp(\Delta S/Cv_1) - (\Delta S/Cv_1)] \quad (15.21)$$

This function is plotted in figure 15.8. You can see that the available energy is positive whenever region 1 is hotter *or colder* than region 2. It is usually more practical to store available energy in a region that is hotter than the heat sink, and you can understand why this is, because of the exponential in equation 15.21. However, at least in principle, anything that is colder than the heat sink also constitutes a source of available energy.

## 15.6 Relationships among $E$ , $F$ , $G$ , and $H$

We have now encountered four quantities  $\{E, F, G, H\}$  all of which have dimensions of energy. The relationships among these quantities can be nicely summarized in two-dimensional charts, as in figure 15.9.



Figure 15.9: Energy, Enthalpy, Free Energy, and Free Enthalpy

Figure 15.10: Some Derivatives of  $E$ ,  $F$ ,  $G$ , and  $H$

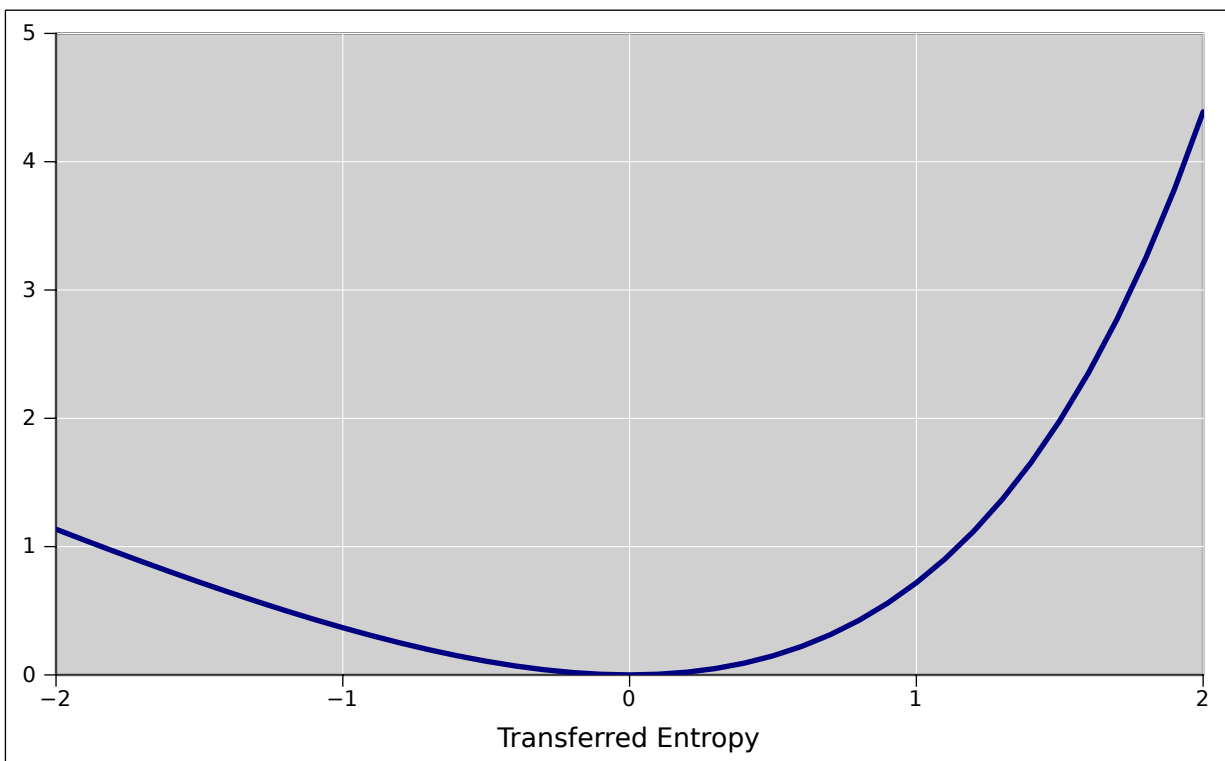


Figure 15.8: Available Energy vs Transferred Entropy

The four expressions in figure 15.10 constitute all of the expressions that can be generated by starting with equation 7.8 and applying Legendre transformations, without introducing any new variables. They are emphatically *not* the only valid ways of differentiating  $E$ ,  $F$ ,  $G$ , and  $H$ . Equation 7.12 is a very practical example – namely heat capacity – that does not show up in figure 15.10. It involves expressing  $dE$  in terms of  $dV$  and  $dT$  (rather than  $dV$  and  $dS$ ). As another example, equation 26.10 naturally expresses the energy as a function of temperature, not as a function of entropy.

Beware: There is a widespread misconception that  $E$  is “naturally” (or necessarily) expressed in terms of  $V$  and  $S$ , while  $H$  is “naturally” (or necessarily) expressed in terms of  $P$  and  $S$ , and so on for  $F(V, T)$  and  $G(P, S)$ . To get an idea of how widespread this misconception is, see reference 46 and references therein. Alas, there are no good reasons for such restrictions on the choice of variables.

These restrictions may be a crude attempt to solve the problems caused by taking shortcuts with the notation for partial derivatives. However, the restrictions are neither necessary nor sufficient to solve the problems. One key requirement for staying out of trouble is to always specify the direction when writing a partial derivative. That is, do not leave off the “at constant  $X$ ” when writing the partial derivative at constant  $X$ . See section 7.5 and reference 3 for more on this.

Subject to some significant restrictions, you can derive a notion of conservation of enthalpy. Specifically, this is restricted to conditions of constant pressure, plus some additional technical restrictions. See chapter 14. (This stands in contrast to energy, which obeys a strict local conservation law without restrictions.) If the pressure is changing, the safest procedure is to keep track of the pressure and volume, apply the energy conservation law, and then calculate the enthalpy from the definition (equation 15.4) if desired.

## 15.7 Yet More Transformations

Starting from equation 7.33 there is another whole family of Legendre transformations involving  $\mu N$ .

## 15.8 Example: Hydrogen/Oxygen Fuel Cell

### 15.8.1 Basic Scenario

Consider the simplified fuel cell shown in figure 15.11.

Ideally, the fuel cell carries out the following reaction:



That equation is an OK place to start, but there’s a lot it isn’t telling us. For one thing, we have not yet begun to account for the energy and entropy.

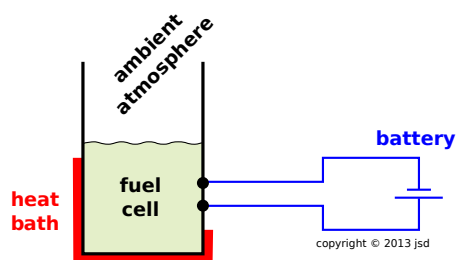


Figure 15.11: Fuel Cell Et Cetera

There are four main energy reservoirs we need to keep track of:

$$\begin{aligned} E_A &= \text{energy in the ambient atmosphere} \\ E_B &= \text{energy in the battery} \\ E_C &= \text{energy in the chemicals in the fuel cell} \\ E_H &= \text{energy in the heat bath} \end{aligned} \tag{15.23}$$

The product on the RHS of equation 15.22 is a liquid, so, on a mole-for-mole basis, it takes up a lot less volume than the corresponding reactants. This means that as the reaction proceeds left-to-right, the collection of chemicals shrinks. In turn, that means that the ambient atmosphere does  $PdV$  work on the chemicals.

Also, although it's not accounted for in equation 15.22 the way we have written it, the reaction must be balanced with respect to electrons (not just with respect to atoms). To make a long story short, this means that as the reaction proceeds, the chemicals do work against the battery.

Notation: The conventional symbols for electrical quantities include the charge  $Q$  and the voltage  $V$ . However, those letters mean something else in thermodynamics, so we are going to write the electrical work as  $F \cdot dx$ . You can think of it as the force on the electrons times the distance the electrons move in the wire. Or you can imagine replacing the battery with a highly-efficient reversible motor, and measuring the genuine force and distance on the output shaft of the motor.

At this point we have enough variables to write down a useful expansion for the derivative of the energy:

$$dE_C = T_C dS_C - F_C \cdot dx_C - P_C dV_C \tag{15.24}$$

## 15.8.2 Enthalpy

For reasons that will be explained in a moment, it is useful to introduce the enthalpy  $H$ , and its gradient  $dH$ :

$$\begin{aligned} H &:= E + PV \\ dH &:= dE + PdV + VdP \end{aligned} \tag{15.25}$$

Plugging this into equation 15.24 we obtain

$$dH_C = T_C dS_C - F_C \cdot dx_C + V_C dP_C \tag{15.26a}$$

$$dH_C = T_C dS_C - F_C \cdot dx_C \tag{15.26b}$$

Equation 15.26b is predicated on the assumption that the pressure in our fuel cell ( $P_C$ ) is constant all during the reaction, so that the derivative  $dP_C$  vanishes.

Here's one reason why it is worthwhile to analyze the system in terms of enthalpy: We can look at tabulated data to find the enthalpy of reaction for equation 15.22 (subject to mild restrictions). If the reaction is carried out under “standard conditions”, tables tell us the enthalpy of the chemicals  $H_C$  is 286 kJ/mol less on the RHS compared to the LHS. The enthalpy includes the energy of the reaction  $\Delta E$  plus the  $PdV$  work (specifically  $\Delta E_C$  plus  $P_C dV_C$ ) ... neither of which we have evaluated separately. We could evaluate them, but there is no need to.

Tangential remark: Before we go on, let’s take another look at equation 15.24. From the point of view of mathematics, as well as fundamental physics, the  $PdV$  term and the  $F \cdot dx$  term play exactly the same role. So the question arises, why do we include  $PdV$  in the definition of enthalpy ... but not  $F \cdot dx$ ???

Partly it has to do with convention, but the convention is based on convenience. It is fairly common for the pressure to be constant, which makes it not very interesting, so we use the enthalpy formalism to make the uninteresting stuff drop out of the equation.

On the other hand, if you ever find yourself in a situation where there is an uninteresting  $F \cdot dx$  term, you are encouraged to define your own enthalpy-like state function that deals with  $F \cdot x$  the same way enthalpy deals with  $PV$ .

To make progress, we need more information. We need to know the entropy. It turns out that the entropy is known for oxygen, hydrogen, and water under standard conditions. This can be found, to a sufficient approximation, by starting at some very low temperature and integrating the heat capacity. (This procedure overlooks a certain amount of spectator entropy, but that is not important for present purposes.) Tabulated data tells us that the entropy (under standard conditions) of the RHS of equation 15.22 is 164 J/K/mol lower compared to the LHS.

At this point we can calculate the maximum amount of energy that can be delivered to the battery. Re-arranging equation 15.26b we have

$$F_C \cdot dx_C = T_C dS_C - dH_C \quad (15.27)$$

The large negative  $\Delta H$  means that a large amount of  $F \cdot dx$  work can be done on the battery, although the rather large negative  $\Delta S$  interferes with this somewhat. At 20° C, the  $TdS$  term contributes about 47.96 J/mol in the unhelpful direction, so the best we can expect from our fuel cell is about 238 kJ/mol. That represents about 83% efficiency. The efficiency is less than 100% because a certain amount of “waste heat” is being dumped into the heat bath.

Note that the Carnot efficiency would be essentially zero under these conditions, because the highest temperature in the device is the same as the lowest temperature. So the fuel cell exceeds the Carnot efficiency, by an enormous margin. This does not violate any laws of physics, because the Carnot formula applies to heat engines, and the fuel cell is not a heat engine.

If we were to burn the hydrogen and oxygen in an ordinary flame, and then use the flame to operate a heat engine, it would hardly matter how efficient the engine was, how reversible the engine was, because the chemistry of combustion is grossly irreversible, and this would be the dominant inefficiency of the overall process. See section 15.8.5 for more discussion of combustion.

### 15.8.3 Gibbs Free Enthalpy

For reasons that will be explained in a moment, it is useful to introduce the Gibbs free enthalpy:

$$\begin{aligned} G &= E + PV - TS \\ &= H - TS \\ dG &= dH - TdS - SdT \end{aligned} \quad (15.28)$$

Plugging into equation 15.26b we obtain

$$dG_C = -F_C \cdot dx_C - S_C dT_C \quad (15.29a)$$

$$dG_C = -F_C \cdot dx_C \quad (15.29b)$$

where the last line is predicated on the assumption of constant temperature.

You should recognize the tactic. If our fuel cell is operating at constant temperature, we can declare the  $TdS$  term “uninteresting” and make it go away. This is the same tactic we used to make the  $PdV$  term go away in section 15.8.2.

In favorable situations, this means we have an even easier time calculating the  $F \cdot dx$  work done by the fuel cell. All we need to do is look up the tabulated number for the Gibbs free enthalpy of reaction, for the reaction of interest. (This number will of course depend on temperature and pressure, but if the reaction takes place under “standard” or well-known conditions, it is likely that tabulated data will be available.)

#### 15.8.4 Discussion: Assumptions

In this section, we have made a great many simplifying assumptions. Let’s try to be explicit about some of the most-important and least-obvious assumptions. Here is a partial list:

- We assume no entropy crosses the C/A boundary (from the cell to the atmosphere or vice versa).
- We assume no entropy flows down the wires, crossing the C/B boundary.
- We assume no entropy or energy flows across the A/H boundary, the A/B boundary, or the B/H boundary.
- There is certainly entropy flowing across the C/H boundary.

However, unless otherwise stated, we assume that the overall process is reversible. In particular, there are no contributions to  $dS_C$  other than a reversible flows of entropy across the C/H boundary. In mathematical terms:  $dS_H = -dS_C$ .

- We assume  $xP_C = P_A$ . This is a pretty safe assumption, for this apparatus, based on previous assumptions plus the third law of motion. However, if you change the apparatus, all bets are off.
- We assume  $F_C dx_C = -F_B dx_B$ . This should be pretty safe, based on previous assumptions plus conservation of energy.
- etc. etc. etc.



### 15.8.5 Plain Combustion $\Rightarrow$ Dissipation

Back in the bad old days, before fuel cells were invented, there was no way to carry out the reaction in such a way that equation 15.26b applied. Approximately the only thing you could do was plain old combustion. If we remove the electrical connections from the fuel cell, all that remains is combustion, and the relevant enthalpy equation reduces to:

$$dH_C = T_C dS_C \quad (15.30a)$$

$$\Delta H_C = \int T_C dS_C \quad \text{along some path } \Gamma \quad (15.30b)$$

Previously, when analyzing the fuel cell, we had essentially two equations in two unknowns: We used entropy considerations (the second law) to determine how much energy was flowing from the cell into the heat bath (in the form of  $TdS$  heat), and then used energy conservation (the first law) to determine how much energy was flowing from the cell into the battery (in the form of  $F \cdot dx$  work).

Now, with the electrical connections removed, we have a problem. We now have two equations in only one unknown, and the two equations are inconsistent. Something has to give.

To make a long story short, the only way we can restore consistency is to stop assuming that the process is reversible. The laws of thermodynamics are telling us that if our heat bath is at some ordinary reasonable temperature, any attempt to use combustion of  $H_2$  in  $O_2$  to deliver energy to the heat bath will be grossly irreversible.

We now have a two-step process: We use the first law to determine how much energy will be delivered to the heat bath. We then create *from scratch* enough entropy so that this energy can be delivered in the form of  $TdS$  heat. This newly-created entropy is a new variable. We have returned to having two equations in two unknowns. This allows us to satisfy the second law while still satisfying the first law.

Let’s analyze this situation in more detail, using pictures. We start with figure 15.12, which is a pictorial representation of equation 15.24. This remains a valid equation.

Neither equation 15.24 nor figure 15.12 expresses conservation of energy. Indeed, they don’t express very much at all, beyond the assumption that  $E$  can be expressed as a differentiable function of three state-variables, namely  $S$ ,  $x$ , and  $V$ .

It must be emphasized that each and every term in equation 15.24 is a function of state. Ditto for each and every term depicted in figure 15.12. The equation and the figure apply equally well to thermally-isolated and non-isolated systems. They apply equally well to reversible transformations and irreversible transformations.

Having said that, we can take two further steps: We assume that the heat flow across the  $A/H$  boundary is reversible. (This implies, among other things, that the cell and the heat bath are at the same temperature.) We then apply local conservation laws to each of the boundaries of the cell:

- We have reversible flow of  $PdV$  work across the  $A/C$  boundary. Energy leaves the ambient atmosphere and flows into the cell
- We have reversible flow of  $F \cdot dx$  work across the  $C/B$  boundary. Energy flows out of the cell into the battery.

The diagram illustrates the decomposition of the differential of internal energy,  $dE_C$ , into its constituent parts. On the left, a light green rectangle is labeled  $dE_C$ . To its right, the word "equals" is written. On the right, a right-angled triangle is shown, divided into three colored regions: a red region at the bottom-left labeled  $T_C dS_C$ , a green region at the top-left labeled  $-P_C dV_C$ , and a blue region at the top-right labeled  $-\vec{F}_C \cdot d\vec{x}_C$ . Below the triangle, the text "copyright © 2013 jsd" is written.

$$dE_C = T_C dS_C - P_C dV_C - \vec{F}_C \cdot d\vec{x}_C$$

copyright © 2013 jsd

Figure 15.12:  $dE$  – Derivatives in Place

- We have reversible heat flow across the  $C/H$  boundary. Energy flows out of the cell into the heat bath.

These three flows are diagrammed in figure 15.13.

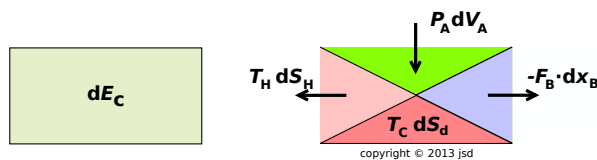
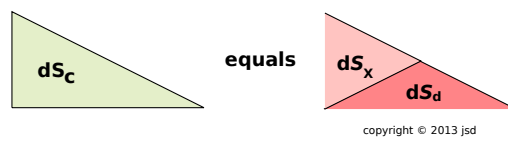


Figure 15.13:  $dE$  – Flowing, or Not

For a dissipative process, there is another term that must be included, in order to account for the possibility of entropy being created inside the cell. This term has got nothing to do with any flow. This is the  $S_d$  term in figure 15.15, and the corresponding  $T_C dS_d$  term in figure 15.13.

We have divided  $dS$  into two pieces:

- $dS_X$ , where the subscript  $x$  indicates “xfer” (as in transfer). This represents the reversible

Figure 15.14:  $dS$  – Derivatives in Place

transfer of entropy across the boundary, as shown in figure 15.15. Any entropy lost from the cell via this mechanism will be gained by the heat bath.

- $dS_d$ , where the subscript d indicates “dissipation.” This does not represent the flow energy across any boundary. This represents endogenous entropy, created in place. For a reversible process, this contribution is zero.



Figure 15.15:  $dS$  – Flowing, or Not

To summarize:

1. It is always OK to interpret equation 15.24 in terms of derivatives in place. Each and every term is a function of state. Each and every term is independent of what, if anything, is going

on in adjacent regions. This works for thermally-isolated and non-isolated systems. This works for reversible transformations and irreversible transformations. Ditto for figure 15.12.

2. You can to a *limited* degree add a layer of interpretation that interprets each term on the RHS of equation 15.24 in terms of flow.
  - You can re-interpret some of the terms all of the time.
  - You can re-interpret all of the terms some of the time.
  - You cannot re-interpret all of the terms all of the time.

In particular:

For a reversible system, entropy is locally conserved, and of course energy is locally conserved. In such a system, each term on the RHS of equation 15.24 corresponds both to an *internal change* within the system and a *flow* across the boundary.

For an irreversible system, the dissipative part of  $dS$  is purely internal. It is not associated with any flow. It will not show up on any flow diagram.

### 15.8.6 Underdetermined

When we disconnected the fuel cell from the electrical load, the problem became overdetermined, and we had to add a new variable – the amount of dissipation – in order to make sense of the situation.

Now imagine that we keep the new variable and re-connect the electrical load. That makes the problem *underdetermined*. The amount of dissipation cannot be determined from the original statement of the problem, and without that, we cannot determine how much energy goes into the heat bath, or how much energy goes into the battery.

This is a correct representation of the actual physics.

As always, a good way to deal with an underdetermined problem is to obtain more information.

For example, if you think  $I^2R$  losses are the dominant contribution to the dissipation, you can measure the current  $I$  and measure the resistivity of the electrolyte. The current of course depends on the *rate* of the reaction. This is not something you can determine by looking at the reaction equation (equation 15.22), but it is easy enough to measure.

If you slow down the reaction so that  $I^2R$  losses are no longer dominant, you still have to worry about overvoltage issues, side reactions, and who-knows-what all else.

### 15.8.7 $H$ Stands For Enthalpy – Not «Heat»

In this context, the RHS of equation 15.30b is commonly called «the heat». We see that in this special case, the enthalpy is equal to «the heat». For this not-very-good reason, the documents that tabulate  $\Delta H$  all-too-often call it the «heat» of reaction, even though it really should be called the

*enthalpy* of reaction. The enthalpy is always the enthalpy, but it is not always equal to the heat. Our fuel cell is the poster child for illustrating that  $H$  is not heat, as you can see in equation 15.26b for example.

Beware that in other contexts, lots of other things are sometimes called «the heat» – including, for example, the differential form  $TdS$  on the RHS of equation 15.30a. Other possibilities are discussed in section 17.1.





## Chapter 16

# Adiabatic Processes

### 16.1 Various Definitions of “Adiabatic”

The word *adiabatic* is another term that suffers from multiple inconsistent meanings. The situation is summarized in figure 16.1. The underlined terms capture some of the possible meanings of adiabatic.

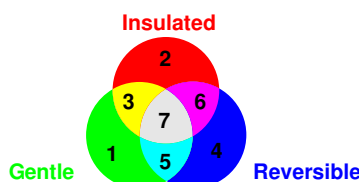


Figure 16.1: Some Concepts Related to Adiabatic

- a) a) Some thoughtful experts use the term *adiabatic* to denote a process where no entropy is transferred across the boundary of the region of interest. In other words, the system is thermally insulated. This was probably the original meaning, according to several lines of evidence, including the Greek etymology:  $a + dia + batos =$  not passing across. This implies that the process must be *fast enough* so that the inevitable small heat leaks are insignificant. As a corollary, we conclude the entropy of the region does not decrease. This is represented by regions 2, 3, 6, and 7 in the diagram.
- b) b) Other thoughtful experts use the term *adiabatic approximation* (in contrast to the sudden approximation) to describe a transformation carried out sufficiently gently that each initial state can be identified with a corresponding final state, and the probabilities (aka occupation numbers) are not changed – at least not changed by the transformation process itself.
- c) c) Dictionaries and textbooks commonly define “adiabatic” to mean no flow of entropy across the boundary *and* no creation of entropy. This is represented by regions 6 and 7 in the diagram. As a corollary, this implies that the system is isentropic. That is to say, its entropy is not changing.

Here are some illustrations of the relationships between the various definitions:

- As an example where definition (a) applies, but definition (c) does not, see reference 47. It speaks of an irreversible adiabatic process. That means the system is thermally insulated, but entropy is being created deep within the interior of the system. This is represented by regions 2 and 3 in the figure.

To say the same thing the other way: In the dream-world where only reversible processes need be considered, definitions (a) and (c) would be equivalent, but that's not much help to us in the real world.

(In contrast, if we were discussing energy, the ideas of “no flow across the boundary” and “no change” would be equivalent, since energy can never be created or destroyed.)

- As an example where definition (b) applies but definition (a) might or might not, consider the refrigeration technique known as *adiabatic demagnetization*. The demagnetization is carried out gently, so that the notion of corresponding states applies to it – or at least it would, if the system were isolated. It would cause the temperature of the spin system to decrease.

The interesting thing is that people still call it “adiabatic” demagnetization, to indicate gentleness, even when the system is not insulated, i.e. when definition (a) does not apply. Specifically, consider the subcase where there is a steady flow of heat inward across the boundary of the system, balanced by a steady demagnetization, so as to maintain constant temperature. Lots of entropy is flowing across the boundary, violating the first definition, but it is still called adiabatic demagnetization in accordance with the second definition. This subcase is represented by regions 1 and 5 in the diagram.

- As an example where all three definitions would apply, but only definition (b) is intended, consider the magnetic resonance technique known as “adiabatic fast passage” or “rapid adiabatic passage”. The word “adiabatic” in this context means the process is slow and gentle enough that there will be corresponding states, and occupation numbers will be preserved. Meanwhile, in this context the notion of no entropy flow across the boundary is not implied by the word “adiabatic”; so the word “fast” is used to indicate that not much entropy flows across the boundary in the available time. To repeat: adiabatic fast passage involves both ideas: it must be both “fast enough” and “slow enough”. This example sits in region 7 in the diagram.
- It is possible to have a transformation that is isentropic but does not produce states with corresponding probabilities. This is illustrated by figure 16.2. The red probability distribution and the blue probability distribution have the same entropy, namely 2 bits. However, there is no way to match red states with blue states of the same probability. This should be clear from the fact that two of the red states have exactly the same probability, whereas no two blue states do. Also, you can see that the lowest two red states cannot be matched with any blue state of the same probability, not even close.

We can arbitrarily identify certain states as “corresponding” according to their identification number, 1 through 11, but the entropy doesn't care about that. It only cares about the probabilities.

For another example where there are corresponding state ID numbers, but the probability of each state might *or might not* be preserved, see section 16.2.

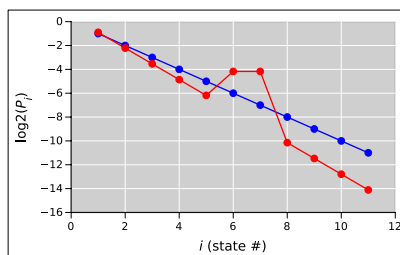


Figure 16.2: Non-Corresponding Probabilities

- Although a reversible insulated system must be isentropic, the converse does not hold. Consider for example a metal wire with one end connected to a hot heat bath, and the other end connected to a cold heat bath. In the steady state, the entropy of the wire is not changing. The entropy just flows through without accumulating.

My recommendation is to avoid using the term adiabatic whenever possible. Some constructive suggestions include:

- If you mean thermally insulated, say thermally insulated.
- If you mean a non-sudden perturbation, say non-sudden or gentle.
- If you mean isentropic, say isentropic.
- Instead of the nouns “adiabat” or “adiabatic line,” say “contour of constant entropy”.

## 16.2 Adiabatic versus Isothermal Expansion

Suppose we have some gas in a cylinder with a piston, and we gradually move the piston so that the gas expands to twice its original volume.

Further suppose that we do this fast enough that there is no thermal energy transport through the walls of the cylinder ... yet slow enough that there is a 1-to-1 correspondence between the states before and after. So this is adiabatic in every sense of the word. The process is diagrammed in figure 16.3.

We consider 20 states, as shown by the dots in the diagram. These are, for practical purposes, the only accessible states. That is to say, all the higher-energy states are unoccupied, to a good approximation. In the diagram, there are three panels. The left panel shows the situation before expansion, and the right panel shows the situation after. The middle panel is a “energy level diagram” that shows how the energy of each mode changes during the expansion.

You can see that within each pair of corresponding dots, the probability is the same before and after. Therefore the entropy of the gas is exactly the same. The energy of the gas has gone down, and the temperature has gone down in equal measure. The slope of the red line is an indication of temperature, in accordance with the Boltzmann factor.

We can contrast this with the isothermal expansion shown in figure 16.4. The gas in this case is in contact with a heat bath, so that during the expansion the temperature does not decrease. The

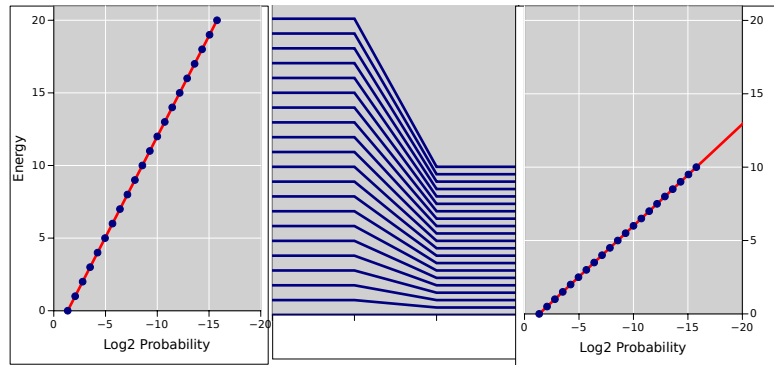


Figure 16.3: Isentropic Expansion

energy of each mode goes down as before, but the occupation numbers do not stay the same. The lion's share of the probability is now distributed over twice as many states. Therefore the entropy of the gas goes up. Within each pair of corresponding states, each state is very nearly half as likely after, as compared to before.

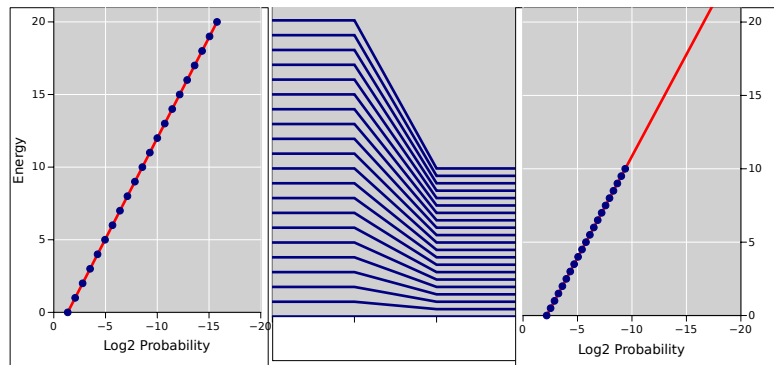


Figure 16.4: Isothermal Expansion

The red line is shifted to the right by one unit, reflecting the  $2\times$  lower probability of each state, but it keeps the same slope, representing the constant temperature.

# Chapter 17

## Heat

### 17.1 Definitions

The term “heat” is a confusing chimera. It is partly energy, partly entropy, partly temperature, and partly who-knows-what. It shows up in a variety of idiomatic expressions, some of which are useful, as discussed in section 17.2. However “heat” by itself is not particularly useful.

By itself, the word “heat” has at least six reasonable and widely-used but mutually-inconsistent technical meanings, as enumerated below. In addition there are innumerable nontechnical, metaphorical, jocular, and/or not-so-reasonable meanings. It is not worth arguing about the relative merits of these meanings, except to say that each has some merit. I observe that a typical thoughtful expert will use each of these meanings, depending on context. It would be nice to have a single, universally-accepted meaning, but I doubt that will happen anytime soon.

In some cases you can figure out the meaning from context. In other cases it is better to avoid the term “heat” altogether, and quantify something else instead (perhaps temperature, energy, and/or entropy).

Here are some of the relatively-sensible technical definitions of “heat.” Studying this list in detail is not worthwhile, but you should skim it, just to get some appreciation for how much trouble you get into when you try to define “heat”:

- (1) Sometimes “heat” simply means hotness, i.e. relatively high temperature. Example: if we’re having a heat wave, it means a spell of hot weather. The corresponding verb, heating, simply means making something hotter. This type of heat is an intensive scalar quantity, and can be measured in degrees.
- (2) Sometimes the word “heat” is used to refer to the  $T dS$  term in equation 7.8. This type of heat is a vector quantity, not a scalar. In particular it is a one-form. In uncramped thermodynamics, it is an ungrady one-form. The corresponding verb, heating, happens if and only if there is a change in the entropy of the region.

This type of heat is extensive, in the sense that if you build twice as many heat engines and operate them all in the same way, you get twice as much  $TdS$ .

- (3) Sometimes “heat” is defined to be the integral of  $T dS$ . For the moment, let’s consider the definite integral  $\int_{\Gamma} T dS$ , integrated along some pre-specified path  $\Gamma$ .

For any given path  $\Gamma$ , this heat is a scalar with dimensions of energy. However, it is very sensitive to how you choose the path, so it might be better to think of it as a *functional* of the path, as discussed in item 4. The overall functional  $f[\Gamma]$  is distinct from the specific scalar  $f[\Gamma]$ .

- (4) Sometimes it pays to work at a higher level, and define “heat” as the indefinite integral  $\int T dS$ , ready to be integrated along some as-yet-unspecified path. Equivalently, you can think of it as a lambda-expression:  $(\int_{\Gamma} T dS \text{ for all } \Gamma)$ .

When we consider the various paths  $\Gamma$  from state  $A$  to state  $B$ , the heat is not a function of state ( $h(B)$  for all  $B$ ) or even a function of the two states ( $h(A, B)$  for all  $A$  and  $B$ ), but rather a *functional* depending on every detail of the path ( $h[\Gamma]$  for all  $\Gamma$ ). See chapter 7 for a discussion of functions of state. See reference 48 for the definition of functional. We write  $h[\dots]$  with square brackets to emphasize that it is a functional, not a mere function.

This heat is a mapping from paths to scalars. Thinking of it as a mapping highlights the connection with item 2, insofar as a one-form is a mapping from vectors to scalars.

- (5) Sometimes “heat” is defined as “energy that is transferred from one body to another as the result of a difference in temperature.” This implies a transfer of entropy across the boundary of the region. This definition is quite prevalent in encyclopedias, dictionaries, and grade-school textbooks. Some people learn this by rote, and rely on it as if it were the 11th commandment, and fail to appreciate its limitations. It works OK within a modest range of “textbook” situations, but it can be hard to quantify and can lead to nasty inconsistencies when applied to other situations, notably when dissipation is occurring, as discussed in section 11.5.6.
- (6) Sometimes people use the terms “heat energy” or “heat content” or “thermal energy” (in contrast to “mechanical energy”) to identify the parts of a compound cramped system, such as the Slinktato™, as discussed in section 19.1. This makes sense in a compound cramped system and not otherwise. In contrast, in a non-compound cramped system such as an ideal potato by itself, there is no advantage to talking about the “heat” energy as opposed to the plain old energy; it’s harmless but pointless. A vastly worse problem arises in uncramped systems, where there is provably no such thing as “heat content” or “thermal energy content.” This problem cannot be fixed by changing the terminology, because when a thing does not exist, it doesn’t matter what you call it. This type of heat (if it exists at all) is an extensive scalar, and can be measured in joules.

In the narrow set of situations where the concept makes sense, it goes by various names including heat content, aka thermal energy, aka caloric, aka  $Q$ . An example of this is discussed in section 11.5.3. This is an extensive scalar, and can be measured in joules. Beware that this notion cannot be extended to uncramped thermodynamics. It cannot even be safely extended from one cramped situation to another, as you can see from the fact that  $\Delta Q = C_V \Delta T$  is different from  $\Delta Q = C_P \Delta T$  ... yet each is called “heat” within its own cramped subspace (constant  $V$  or constant  $P$  respectively).

In addition, one sometimes encounters some less-than-reasonable definitions, including:

4. Chemists commonly use “heat” as an all-purpose synonym for enthalpy, for instance in expressions such as “heat of reaction” or “heat of formation.” This includes cases where there the “heat” (i.e. enthalpy) is not flowing across a boundary. Even more remarkably, it includes cases where the enthalpy is predominantly nonthermal, for instance in an electrochemical fuel cell. This usage is quite common, but I consider it a very unhelpful misnomer. I recommend crossing out terms like “heat of formation” and replacing them with terms like “enthalpy of formation” at every opportunity. Similarly the terms “exothermic” and “endothermic” in most cases should be crossed out and replaced with “exergonic” and “endergonic” respectively ... or perhaps “exenthalpic” and “endenthalpic.”

You have to be careful, though, because even though heat *usually* is connected to enthalpy (as in the heat capacity  $C_P$ ), sometimes it is connected to energy instead. For example, as discussed in section 7.5, the heat capacity  $C_P$  measures enthalpy per unit temperature, whereas the heat capacity  $C_V$  measures energy per unit temperature. The fact that both are called the “heat” capacity makes it impossible to assign an unambiguous meaning to “heat.”

5. Some non-experts, when asked to define “heat,” describe something that is, in effect, the infrared portion of the electromagnetic spectrum. This notion is the basis of the phrase “heat rays,” and of the cliché “it gives off more heat than light.” Alas, this cliché makes no sense from a scientific point of view: It’s true that a black body that gives off primarily infrared radiation is hot ... but a black body that gives off primarily visible light is *hotter*. To consider IR as more “heat-like” than visible light is completely backwards, from a basic-physics point of view.
6. There are endless attempts to extend the vague idea of “heat content” from cramped to uncramped thermodynamics, i.e. to situations where thermodynamic *cycles* are possible. This is highly pernicious. Such attempts never succeed, for reasons discussed in chapter 19.

As an example where definition #1 and definition #2 apply, but definition #5 does not, consider the notion that a microwave oven heats a potato. Clearly (1) the food gets hotter. Clearly (2) the entropy of the food changes. However, (5) no entropy was transferred across the *boundary* of the food. Energy was transferred, but the entropy was created from scratch, within the food. According to any reasonable definition of temperature, the microwave-generating components inside the oven aren’t very hot, so you can’t say the energy was transferred “as the result of a difference in temperature.” This is one of the examples shown in the Venn diagram in figure 17.1.

As an even more spectacular example, consider a system consisting of a closed loop of wire. At time  $t = 0$  there is a current flowing in the wire. This current is decaying, on a timescale given by  $L/R$ . The entropy and temperature of the wire are increasing, even though no energy is being transferred (thermally or otherwise) across the boundary of the system. NMR  $\tau_1$  processes are in the same category.

The distinction between (2) and (5) is an instance of the boundary/interior issue, as discussed in section 8.6.

As an example where definition #2 and definition #5 apply, but definition #1 does not, consider a glass of ice water sitting on the table. We say that heat leaks into the system and melts the ice. The temperature does not change during the process.

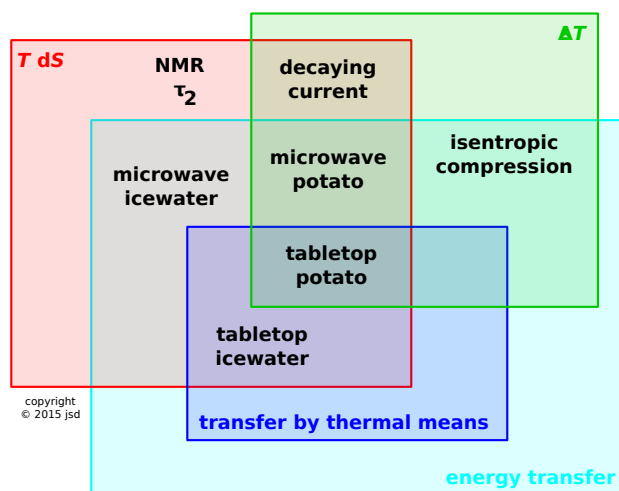


Figure 17.1: Heat: Multiple Definitions, Multiple Examples



As an example where definition #2 applies, but definition #5 and definition #1 do not, consider a glass of ice water in the microwave oven. The ice melts, but the temperature does not change, and no energy was transferred in thermal form across the boundary.

As an even more spectacular example, consider the NMR  $\tau_2$  process. Entropy is increasing, even though no energy whatsoever is being transferred (thermally or otherwise).

As an example where definition #1 applies but definition #2 and definition #5 do not, consider the reversible thermally-insulated compression of a parcel of gas. We say the gas heats up, and arguably one could say there is an increase in the amount of “thermal energy” within the region. On the other hand, clearly no heat or entropy was transferred across the boundary, and there was no change in the entropy within the region. I have seen experts refer to this as *adiabatic heating*; this makes perfect sense in context, but comes as a shock to anyone who thinks definition #5 is the only definition.

We now discuss in more detail the advantages and disadvantages of definition #5:

Definition #5 is the most prevalent, perhaps in part because it is easily expressed in non-mathematical words. Many students have been forced to learn this definition by rote.	Rote learning is a poor substitute for understanding.
---	---

Definition #5 makes sense in some situations, such as a simple non-moving heat exchanger in a non-dissipative system.	Such situations are not representative of the general case.
---	---

Definition #5 focuses attention on flow across a boundary. This is good, because we believe all the laws of physics should be stated in local form, and flows across a boundary are crucial for this.	It focuses on temperature and heat. It would be better to focus on energy and entropy. Certainly energy and entropy can flow between systems that don’t even have a well-defined temperature (let alone a difference in temperature). Also remember that heat is not a conserved quantity, and it is hard to know what “flow” means when applied to non-conserved quantities. Whenever you talk about heat flow, you run the risk that non-experts will visualize heat as some sort of conserved fluid.
---	---

Heat is non-conserved twice over. First of all, even in reversible processes, heat is non-conserved because non-Locrian energy can be converted to Locrian energy and (within limits) vice versa. As mentioned in section 11.5.7 energy is conserved, but heat (by itself) is not conserved. Secondly, in irreversible processes heat is not conserved because entropy is not conserved.

## 17.2 Idiomatic Expressions

The word “heat” occurs in a great number of idiomatic expressions.

As is always the case with idiomatic expressions, the meaning of the expression as a whole cannot be inferred from the meaning of the individual words. In particular, most of the following heat-related idiomatic expressions are reasonably well defined, even though the word “heat” by itself is not:

- heat engine
- heat pump
- heat exchanger
- heat bath
- heat sink
- heat source
- heat capacity (ambiguous and sometimes misleading)
- heat leak and/or heat flow (problematic)
- heat of reaction (very problematic)
- et cetera.

At the next level of detail, here are some notes on the last three items:

- The term “heat capacity” is ambiguous for multiple reasons, some of which have nothing to do with heat. One problem is that there are multiple different heat capacities, such as the heat capacity at constant temperature, heat capacity at constant volume, et cetera. Also note that in all cases, “heat capacity” is something of a misnomer, since it would be simpler to think of it as the entropy capacity, as discussed in section 7.5.
- In the context of “heat flow” the term “heat” almost certainly means energy, because energy is a conserved quantity, and the concept of “flow” is hard to define for non-conserved quantities. See reference 6. We conclude that “heat flow” must refer to energy that is flowing via some thermal process (perhaps conduction or radiation or something like that).
- As discussed in section 14.11, whenever you see the phrase “heat of reaction” you should cross it out and replace it with “enthalpy of reaction” or something similar. Also beware that Hess’s law is often taught in such a way that it seems to express conservation of heat, as discussed in connection with figure 7.2. That’s terrible! Heat is not conserved!

As a related point of terminology: Keep in mind that the symbol  $H$  conventionally stands for enthalpy; it does not stand for heat. Alas, many texts don’t distinguish between heat and enthalpy. That’s a problem because sometimes the enthalpy of reaction ( $\delta H$ ) shows up as heat, and sometimes as something else (such as electrical energy).

### 17.3 Resolving or Avoiding the Ambiguities

If you mean hotness, as in definition #1 above, it is better to speak of temperature rather than heat. This avoids an ambiguous use of the term “heat.”

When experts talk about the  $T\mathbf{d}S$  vector (definition #2) they commonly call it literally  $T\mathbf{d}S$  (pronounced literally “tee dee ess”). This is nicely unambiguous. The term “heat vector” is a

slightly more elegant way of talking about the same idea. The point is that saying “heat vector” rather than merely “heat” makes it clear we are talking about  $T \mathbf{d}S$ , thereby removing a great deal of ambiguity. Remember that this vector is a one-form (as opposed to a pointy vector), and lives in abstract thermodynamic state-space (unlike everyday position vectors). The RHS of figure 8.4 shows you how to visualize the  $T \mathbf{d}S$  vector. For an introduction to one-forms and how they apply to thermodynamics, see reference 4.

In almost all cases where the “transfer across a boundary” idea is used (definition #5), the  $T \mathbf{d}S$  vector idea (definition #2) would be a more precise and more reliable way of describing what is going on. This removes the inconsistencies associated with the “transfer across a boundary” idea. Also, whether or not energy is being transferred across a boundary, visualizing  $T \mathbf{d}S$  as a vector resolves a goodly number of conceptual problems.

Here is a helpful analogy:

The problematic concept of phlogiston was replaced by two precise concepts (namely oxygen and energy).	The problematic concept of heat has been replaced by two precise concepts (namely energy and entropy).
--	--

As another analogy, consider the comparison between “heat” and “blue,” another common four-letter word.

Nobody in his right mind would try to quantify what “blue” means. Instead of quantifying the blueness, you should quantify something else, perhaps power versus wavelength.	Instead of quantifying heat, you should quantify the energy and entropy.
---	--

Actually “heat” is far more problematic than “blue,” because there’s something even worse than imprecision, namely holy wars between the Big-Endians and the Little-Endians, each of whom think they know “the one true definition” of what “heat” is.



# Chapter 18

## Work

### 18.1 Definitions

The definition of *work* suffers from one major problem plus several minor nuisances.

The major problem is that there are two perfectly good but inconsistent notions:

1. Mechanical transfer of energy across a boundary. Here mechanical means non-thermal and non-advective.
2. Force times distance.

These two notions are closely related but certainly not identical. This is an instance of the boundary/interior issue, as discussed in section 8.6. This is a recipe for maximal confusion. (Wildly different ideas are easily distinguished, and identical ideas need not be distinguished.)

Within the force-times-distance family, there are the following nuisance factors, which will be discussed below:

- Done “on” versus done “by.”
- Differential versus integral formulation.
- Microscopic versus coarse-grained on some length-scale  $\lambda$ .
- Local versus overall.

We start by considering the case where the energy is a nice differentiable function of state, and is known as a function of two variables  $V$  and  $S$  alone. Then we can write

$$\begin{aligned}\mathbf{d}E &= \left. \frac{\partial E}{\partial V} \right|_S \mathbf{d}V + \left. \frac{\partial E}{\partial S} \right|_V \mathbf{d}S \\ &= -P\mathbf{d}V + T\mathbf{d}S\end{aligned}\tag{18.1}$$

which is just a repeat of equation 7.5 and equation 7.8. This gives us the differential formulation of work, as follows:

The first term on the RHS, namely  $-P\mathbf{d}V$ , is commonly called the work done on the system. Positive work done on the system increases the energy of the system.

The negative thereof, namely  $P\mathbf{d}V$ , is the work done by the system. Positive work done by the system decreases the energy of the system.

As an elaboration, consider the common case where  $V$  itself is known as a differentiable function of some other variables (say)  $A$ ,  $B$ , and  $C$ .

**Example #1:** Suppose the system is the parallelepiped spanned by the vectors  $A$ ,  $B$ , and  $C$ . Then the volume is  $V = A \wedge B \wedge C$ .

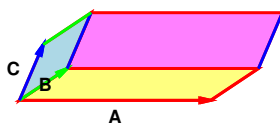


Figure 18.1: Parallelepiped

**Example #2:** Suppose the system is a spring as shown in figure 18.2. It has one end attached to point  $A$  and the other end attached to point  $B$ , where both  $A$  and  $B$  are points on a long one-dimensional track. Then  $V$  is just the length of the spring,  $V = B - A$ .

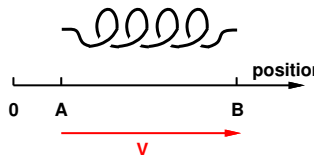


Figure 18.2: Spring

We can differentiate  $V$  to obtain

$$\mathbf{d}V = \left. \frac{\partial V}{\partial A} \right|_{B,C} \mathbf{d}A + \left. \frac{\partial V}{\partial B} \right|_{C,A} \mathbf{d}B + \left. \frac{\partial V}{\partial C} \right|_{A,B} \mathbf{d}C \quad (18.2)$$

and plug that into equation 18.1 to obtain

$$\mathbf{d}E = \left. \frac{\partial E}{\partial V} \right|_S \left. \frac{\partial V}{\partial A} \right|_{B,C} \mathbf{d}A + \left. \frac{\partial E}{\partial V} \right|_S \left. \frac{\partial V}{\partial B} \right|_{C,A} \mathbf{d}B + \left. \frac{\partial E}{\partial V} \right|_S \left. \frac{\partial V}{\partial C} \right|_{A,B} \mathbf{d}C + \left. \frac{\partial E}{\partial S} \right|_V \mathbf{d}S \quad (18.3)$$

We can write this more compactly as:

$$\mathbf{d}E = -F_{A|B,C} \mathbf{d}A - F_{B|C,A} \mathbf{d}B - F_{C|A,B} \mathbf{d}C + T \mathbf{d}S \quad (18.4)$$

where we have defined the notion of force in a given direction according to:

$$F_{A|B,C} := - \left. \frac{\partial E}{\partial A} \right|_{B,C} \quad (18.5)$$

and similarly for the other directions. Compare equation 7.30.

It is conventional but very risky to write  $F_A$  (meaning force “in the  $A$  direction”) as shorthand for  $F_{A|B,C}$ . This is risky because the notion of “the  $A$  direction” is not well defined. It is OK to speak of the direction of constant  $B$  and  $C$ , but not the direction of changing  $A$ . Specifically, in example #2, when we evaluate  $\partial E/\partial A$ , we get very different results depending on whether we evaluate it at constant  $B$  or at constant  $V$ .

There is no reliable, general way to disambiguate this by assuming that  $B$  and  $C$  are the directions “perpendicular” to  $A$ . As an aside, note that in the two examples above, if  $A$  and  $B$  are interpreted as position-vectors in real space, they are definitely not perpendicular. More to the point, when  $A$  and  $B$  are interpreted as part of the abstract thermodynamic state-space, we cannot even define a notion of perpendicular.

In the present context,  $F_A$  is unambiguous because  $F_{A|B,C}$  is by far the strongest candidate for what it might mean. But in another context, the symbol  $F_A$  might be highly ambiguous.

### 18.1.1 Integral versus Differential

We can convert to the integral formulation of work by integrating the differential representation along some path  $\Gamma$ . The work done by the system is:

$$\text{workby}[\Gamma] = \int_{\Gamma} P \cdot dV \quad (18.6)$$

Consider the contrast:

<p>The differential formulation of work (<math>PdV</math>) is a vector, specifically a one-form. A one-form can be considered as a mapping from pointy vectors to scalars.</p>	<p>The integral formulation of work (<math>\text{workby}[\cdot \cdot \cdot]</math>) is a functional. It is a mapping from paths to scalars.</p>
--	---

In particular, if  $\Gamma$  is a path from point  $X$  to point  $Y$ , you should not imagine that the work is a function of  $X$  and/or  $Y$ ; rather it is a functional of the entire path. If  $PdV$  were a grady one-form, you could express the work as a function of the endpoints alone, but it isn't so you can't.

### 18.1.2 Coarse Graining

For each length scale  $\lambda$ , we get a different notion of work; these include microscopic work, mesoscopic work, and holoscopic work (aka macroscopic work, aka pseudowork, as discussed in section 18.5). These are all similar in spirit, but the differences are hugely important. To illustrate this point, consider a flywheel in a box:

- The holoscopic KE is zero, because the CM of the box is not moving.
- If we look inside the box, we see that the flywheel has mesoscopic KE, because it is spinning.
- If we look more closely, we find additional KE in the thermally-excited phonon modes, because the flywheel has nonzero temperature.
- If we look yet more closely, we find yet more KE, including the KE of electrons whizzing around inside atoms.

More generally, there are innumerable gray areas, depending on the length scale  $\lambda$ .

In thermodynamics, it is usually – but not necessarily – appropriate to assume that “work” refers to either mesoscopic or holoscopic work.

### 18.1.3 Local versus Overall

Sometimes it is useful to consider the force and displacement acting locally on *part* of the boundary, and sometimes it is useful to consider the overall force and overall displacement.

To say the same thing in mathematical terms, let’s multiply both sides of equation 18.2 by  $P$  to obtain:

$$PdV = F_{A|B,C}dA + F_{B|C,A}dB + F_{C|A,B}dC \quad (18.7)$$

In some contexts, it would make sense to speak of just *one* of the terms on the RHS as “the” work.

## 18.2 Energy Flow versus Work

Let’s consider systems that have some internal structure.

Our first example is shown in figure 18.3, namely a spring with a massive bob at one end. The other end is anchored. The mass of the spring itself is negligible compared to the mass of the bob. Dissipation is negligible. I am pushing on the bob, making it move at a steady speed  $v \equiv dA/dt$ . This requires adjusting the applied force  $F$  so that it always just balances the force of the spring.

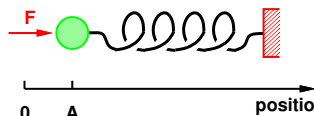


Figure 18.3: Spring with Bob

When we ask how much “work” is involved, we have a bit of a dilemma.

<p>It certainly <i>feels</i> to me like I am doing work on the spring+bob system. Energy is flowing across the boundary from me into the bob.</p>	<p>The <i>overall</i> work on the spring+bob system is zero. The force of my push on one end is exactly balanced by the force of constraint on the other end. Zero total force implies zero macroscopic work (aka pseudowork). Having zero macroscopic work is consistent with the work/KE theorem, since the KE of the system is not changing.</p>
---	---

This dilemma does not go away if we break the system into sub-systems. The applied force on the bob is just balanced by the force of the spring, so there is no net force (hence no overall work) on the bob considered as a subsystem. The same goes for each small subsection of the spring: No net force, no acceleration, no work, and no change in KE.



The “local work” at the moving end is  $F \cdot dx$ .

The “local work” at the fixed end is zero, since it is  $F \cdot 0$ .

It is OK to think of energy pouring into the spring as a whole at the rate  $dE/dt = F \cdot v$ . It is OK to think of energy as being like an abstract fluid flowing across the boundary.

It seems highly problematic to treat work as if it were a fluid flowing across the boundary. In particular, a naive attempt to apply the work/KE theorem is a disaster, because the energy inside the spring is virtually all potential energy; the KE inside the spring is negligible. The alleged work-fluid is flowing into the spring from the bob, and not flowing out anywhere, yet no work or KE is accumulating inside the spring.

As a second example, consider the oil bearing in section 11.5.5. Again we have a boundary/interior issue. Again we have a dilemma, due to conflicting definitions of work:

<p>I am doing work in the sense of force (at a given point) times distance (moved by that point). I am doing work in the sense of pouring net energy across the boundary of the system.</p>	<p>There is no overall force, no overall work, no acceleration, and no change in KE.</p>
---	--

Part of the lesson here is that you need to think carefully about the conditions for validity of the work/KE theorem. A non-exhaustive list is:

- It suffices to have a rigid body, i.e. no motion of one part relative to another, i.e. no internal forces except forces of constraint. This implies no change in the internal potential energy.
- It suffices to have a dismembered body, i.e. no internal forces between the parts, i.e. parts free to move independently of each other. Again this implies no change in the internal potential energy.
- It suffices to carry out a full accounting for the internal forces, not just the external forces. This implies accounting for the changing internal potential energy.

There are some interesting parallels between the oil bearing and the spring:

- In both cases, momentum flows into the system on one side and simultaneously flows out the other side, so there is no net accumulation of momentum within the system.
- Meanwhile, energy flows into the system on one side and does *not* flow out the other side, so that energy accumulates within the system.
- In one case the accumulated energy is in microscopic form, raising the temperature of the oil, while in the other case it takes the form of non-Locrian potential energy in the spring.

If you want a third parallel system, consider a force applied to a free body, such as the bob in figure 18.3 without the spring and without the anchor. Energy and momentum flow into the system and accumulate. The accumulated energy takes the form of non-Locrian kinetic energy.

From this we see that the work/KE theorem is intimately connected to the accumulation of *momentum* within the system, not the accumulation of energy per se.

A related thought is that momentum is conserved and energy is conserved, while work (by itself) is not conserved. KE (by itself) is not conserved.

### 18.3 Remarks

Keep in mind that “work” is ambiguous. If you decide to speak in terms of work, you need to spell out exactly what you mean.

Also keep in mind that dissipative processes commonly convert mesoscopic KE into microscopic KE as well as non-kinetic forms of energy. Energy is conserved; mesoscopic KE is not (by itself) conserved.

### 18.4 Hidden Energy

You can’t hide momentum; if an object has momentum its center-of-mass will be moving, and this will be easy to notice. In contrast, you can easily hide energy in an object’s internal degrees of freedom, perhaps in the form of spinning flywheels, taut springs, random microscopic energy, or other things having nothing to do with center-of-mass motion.

Here is an example of hidden energy: Consider a cart with two flywheels on board. Initially everything is at rest. Apply a pair of forces (equal and opposite) to the front flywheel, causing it to spin up, clockwise. Apply a similar pair of forces to the back flywheel, causing it to spin up, counterclockwise. The net force on the cart is zero. The motion of the cart’s center of mass is zero. The net force dot the overall motion is zero squared. The cart’s overall angular momentum is also zero. Yet the cart has gained kinetic energy: internal, mesoscopic kinetic energy.

Examples like this are a dime a dozen. In some sense what we are seeing here is the difference between holoscopic and mesoscopic kinetic energy. If you don’t recognize the difference, and recklessly talk about “the” kinetic energy, you’re going to have trouble.

### 18.5 Pseudowork

Sometimes it is appropriate to focus attention on the coarsest level of course-graining. We continue using the ideas and notation introduced in section 18.1.2.

In the large- $\lambda$  limit we have:

$$d(P^2/(2M)) = F_{\text{tot}} \cdot dx_{\text{cm}} \quad (18.8)$$

where  $P = \sum p_i$  is the total momentum of the system,  $M := \sum m_i$  is the total mass,  $F_{\text{tot}} := \sum F_i$  is total force applied to the system, and  $x_{\text{cm}}$  is the distance travel led by the center of mass. This is similar in form to equation 18.6, but the meaning is different. See reference 18 for a derivation and discussion.

The RHS of equation 18.8 is called the *pseudowork*. The LHS represents the change in something we can call the pseudokinetic energy. This is just a synonym for the holoscopic kinetic energy.

There is an easy-to-prove theorem that says that for any length scale  $\lambda$ , an object’s total  $\text{KE}[\lambda]$  measured in the lab frame is equal to the  $\text{KE}[\lambda]$  of the *relative* motion of the components of the object

(i.e. the  $\text{KE}[\lambda]$  measured in a frame comoving with the CM of the object) ... plus the holoscopic KE associated with the motion of the CM relative to the lab frame (as given by equation 18.8).

Mesoscopic work and holoscopic work (aka pseudowork) are consistent with the spirit of thermodynamics, because they don't require knowing the microscopic forces and motions.

However, the pseudowork is *not* equal to the “thermodynamic”  $w$  that appears in the oft-abused equation 7.5. Here's a counterexample: Suppose you apply a combination of forces to a system and its center of mass doesn't move. Then there are at least three possibilities:

- Maybe there is no energy transfer at all, e.g. static equilibrium;
- Maybe there is a completely nonthermal transfer of energy, e.g. spinning up a flywheel; or
- Maybe the energy is completely thermalized, as in boring a cannon with a completely dull tool (section 11.5.3).

According to the meaning of  $w$  usually associated with equation 7.5,  $w$  is zero in the first case, nonzero in the second case, and who-knows-what in the third case. It is a common mistake to confuse  $w$  with work or pseudowork. Don't do it.



## Chapter 19

# Cramped versus Uncramped Thermodynamics

### 19.1 Overview

In this chapter we consider some basic questions about what can be done with thermodynamics. These include:

- Is it possible to define a nontrivial quantitative notion of “heat content” or “thermal energy content” of a system?
- Is it possible to use the system as the working fluid in a heat engine?

We shall see that at most one of these things is possible for any given system. Sometimes neither is possible, but never both. There are actually four scenarios that are worth considering. An overview of the situation can be found in figure 19.1.

Inside the playpen we have three toys: An ideal spring, an ideal heat container (represented by a potato), and a combination of the two (the legendary Slinktato™). Outside the playpen we have a heat engine, and for present purposes we are interested in the thermodynamics of the working fluid within the heat engine.

In more detail:

1. You can perform  $F \cdot d\mathbf{x}$  work on the spring. In contrast, the spring is thermally isolated from the rest of the world, so you cannot perform appreciable  $TdS$  heating on the spring. Furthermore, and more importantly, even if you did change the temperature, it would have no effect on the length or spring-constant of the spring. (Real springs are somewhat sensitive to temperature, but this is an ideal spring, engineered to minimize the temperature dependence.)

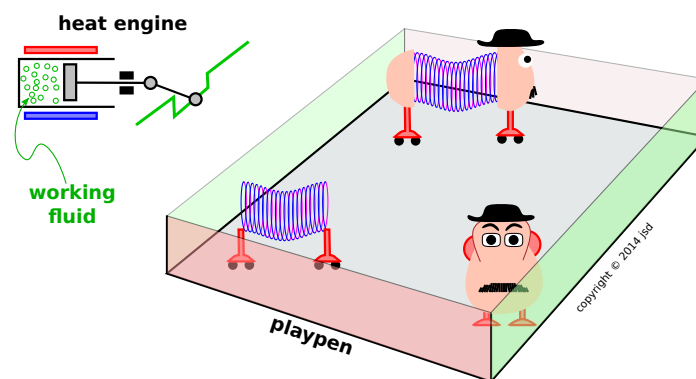


Figure 19.1: Heat Engine versus Slinktato™

2. You can perform  $TdS$  heating on the potato. In contrast, the potato is rigid and well protected from outside forces, so you cannot perform appreciable  $F \cdot d\mathbf{x}$  work on it. Furthermore, and more importantly, even if you did apply a force, it would have no effect on the temperature, heat capacity, size, or mechanical properties of the potato.
3. You can write an equation for the Slinktato™ that contains both  $TdS$  and  $F \cdot d\mathbf{x}$  terms ... but there is no coupling between the two terms. Changing the length of the spring does not change the temperature of the potato, and changing the entropy of the potato does not change the force on the spring.

It is not possible to build a heat engine using a Slinktato™ or anything else you find inside the playpen.

4. Outside the playpen, things are much more interesting. The working fluid in the heat engine can take in  $TdS$  heat and put out  $F \cdot d\mathbf{x}$  work.

Life would be much simpler if we could just get rid of the notion of “thermal energy content” altogether. This is “almost” but not quite possible. Let’s check the various cases one by one:

1. For the ideal spring, we can get rid of the notion of “thermal energy” because there isn’t any.
2. For the ideal potato, we have the option of getting rid of the notion of “thermal energy.” We can replace it with the simpler notion of “energy” because 100% of the relevant energy (or energy-change) is thermal.
3. A system where it is possible to have a cycle, but the cycle cannot be used to build a heat engine is called a *compound* cramped system. The Slinktato™ is an example; see section 19.3 for other examples. This is the only case where it is both possible and worthwhile to define a notion of “thermal energy content.” The energy in the potato part of the Slinktato™ is thermal energy, while the energy in the spring part is mechanical energy. It is useful to have terminology to describe these two contributions.
4. For a heat engine, it is essential to get rid of the notion of “thermal energy content.” Any such notion makes it completely impossible to understand what’s going on.

The situation is summarized in the following table:

<u>system:</u>	<u>heat content?</u>	<u>variables?</u>	<u>heat engine?</u>
spring:	trivial (zero)	too few	no
potato:	trivial (total)	too few	no
Slinktato™:	<b>nontrivial</b>	decoupled	no
heat engine working fluid:	impossible	<b>coupled</b>	<b>yes</b>

To make it easier to discuss things, we define an *uncramped* system to be one that has sufficiently many variables and sufficiently few restrictions, so that it could be used as the working fluid in a heat engine. Anything else is considered *cramped*. In figure 19.1, cramped systems are inside the playpen, while uncramped systems are outside.

There are three main ways of obtaining a cramped system, if that’s what you want:

- You might be so lucky that the first system you stumble upon is a cramped system.
- You might sort through a large number of systems and *select* one where the variables are decoupled in the desired way.
- You might work hard and *engineer* a cramped system.

However, if you are not sufficiently lucky, selective, and/or industrious, it is very likely that a typical system will *not* be cramped. The relevant variables will *not* be decoupled.

Furthermore, cramped is not always desirable. If you’re trying to build a heat engine, you want a working fluid that is not cramped.

Often there are timescales that must be considered. For example, if you have a huge current of water in the ocean, the thermal variables are more-or-less decoupled from the mechanical variables on a timescale of days or weeks, but not forever.

Bottom line: Sometimes it makes sense to talk about “heat energy” as something distinct from “mechanical energy” – and sometimes it doesn’t. There is nothing in the laws of physics that makes it certain – or even likely – that such a distinction makes sense. In particular, if you want to use such concepts, it is your responsibility to prove that they make sense, in any given situation.

## 19.2 A Closer Look

Let’s look into the situation more carefully, more quantitatively.

1. Let’s start with a plain-vanilla uncramped system. For simplicity, we assume the energy is known as a function of  $V$  and  $S$ . Indeed, we assume this function is differentiable, so we can write:

$$dE = -F \cdot dx + TdS \quad (19.1)$$

In such a system, we can have nontrivial *thermodynamic cycles*.

2. The easiest way (but not the only way) to create a cramped system is to reduce the number of variables. For instance, starting from equation 19.1 we can consider cases where the  $PdV$  term is negligible, and all that remains on the RHS is the  $TdS$  term. This corresponds to the potato in figure 19.1.

By the same token, we can consider systems where the  $TdS$  term is negligible, and all that remains on the RHS of equation 19.1 is the  $F \cdot dx$  term. This corresponds to the spring in figure 19.1.

When there are so few nontrivial variables, you couldn’t draw a thermodynamic cycle if you wanted to. The system is effectively one-dimensional. The notion of “area” on the indicator diagram is undefined and undefinable.

3. The other type of cramped system is a compound cramped system. The Slinktato™ is an example. Such a system has enough variables to make it possible to draw a thermodynamic cycle, but the cycle is trivial in the sense that  $\oint F \cdot dx = 0$  and  $\oint TdS = 0$  when we integrate around any cycle.



The distinction between cramped and uncramped is indicated in figure 19.2. This is a copy of figure 2 as seen in section 0.3.

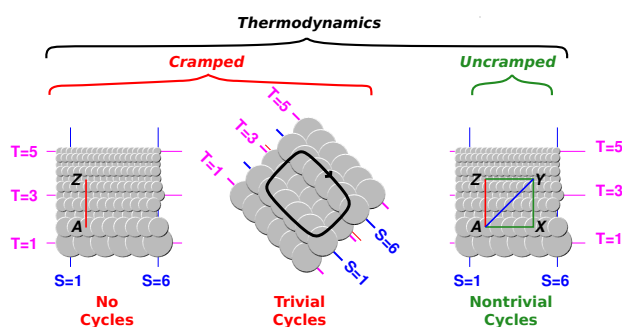


Figure 19.2: Cramped versus Uncramped

Remember: The defining property of a cramped system is that it is so constrained that it is not possible to build a heat engine. That means either

- No cycles at all, as in the left part of figure 19.2. The ideal potato by itself is an example.
- If there is a cycle it is trivial, in the sense that the integral around the cycle is always zero, as in the middle of figure 19.2. The Slinktato™ is an example.

As will be explained in a moment, in any uncramped system, there cannot be any notion of “heat content” or “thermal energy content.” To say the same thing the other way, if the system has a well-defined heat-function  $Q$ , it cannot be used as the working fluid in a heat engine.

(The converse does not hold; we are not saying that every cramped system has a well-defined  $Q$ -function.)

We can begin to understand this as follows: In an uncramped system, you can put energy into the system via  $TdS$  heat and take energy out via  $F \cdot dx$  work. That’s pretty much the defining property of a heat engine. It guarantees that you cannot define the “thermal energy content” of the system by keeping track of how much heat you put in.

We can formalize this as follows: Suppose there is a  $Q$  function, representing thermal energy content, such that  $Q$  is a function of state. Consistency with other notions of heat requires that  $dQ = TdS$ .

Now  $\oint \mathbf{d}Q = 0$  since  $\oint \mathbf{d}(\text{anything}) = 0$  automatically. That means  $\oint T \mathbf{d}S = 0$ , which makes it impossible to build a heat engine.

We are not saying that the notion of “thermal energy content” is always wrong; we are just saying that it cannot exist in any situation where it is possible to build a heat engine.

Here is another calculation we can do that sheds additional light on the situation. Again suppose that we have

$$\mathbf{d}Q = T \mathbf{d}S \quad (19.2)$$

Taking the exterior derivative of both sides, we obtain

$$\begin{aligned} \mathbf{d}\mathbf{d}Q &= \mathbf{d}(T \mathbf{d}S) \\ &= T \mathbf{d}\mathbf{d}S + \mathbf{d}T \wedge \mathbf{d}S \end{aligned} \quad (19.3)$$

hence

$$\mathbf{d}T \wedge \mathbf{d}S = 0 \quad (19.4)$$

since  $\mathbf{d}(\mathbf{d}(\text{anything})) = 0$ . That is equivalent to saying “the boundary of a boundary is zero” as discussed in reference 4. Equation 19.4 can be satisfied if  $T$  is constant. It can also be satisfied if the contours of constant  $T$  run parallel to the contours of constant  $S$ . There are lots of ways this could happen; you could have  $T \propto S$  or  $T \propto S^2$  or whatever. An example is shown in the middle part of figure 19.2.

We can gain additional insight if we expand  $\mathbf{d}T$ . In accordance with basic principles, we can expand any exterior derivative along the following lines:

$$\mathbf{d}T = \frac{\partial T}{\partial x|S} \mathbf{d}x + \frac{\partial T}{\partial S|x} \mathbf{d}S \quad (19.5)$$

and plugging into equation 19.4 we obtain

$$\mathbf{d}T \wedge \mathbf{d}S = \frac{\partial T}{\partial x|S} \mathbf{d}x \wedge \mathbf{d}S + \frac{\partial T}{\partial S|x} \mathbf{d}S \wedge \mathbf{d}S \quad (19.6)$$

The second term on the RHS is obviously zero, so equation 19.4 holds if and only if

$$\frac{\partial T}{\partial x|S} \mathbf{d}x \wedge \mathbf{d}S = 0 \quad (19.7)$$

This equation could hold if  $\mathbf{d}x \wedge \mathbf{d}S = 0$ , which is the easy way to create a cramped system. However, things are more interesting if we assume this factor is not zero. In that case equation 19.7

$$\frac{\partial T}{\partial x|S} = 0 \quad (19.8)$$

This equation says in mathematical terms what we said in words back in section 19.1 when we defined the Slinktato™. This is how we arranged for it to have a well-defined heat content.

Now we refer back to the definition of temperature, i.e. equation 7.7, which can be rewritten (subject to mild restrictions) as

$$T = \frac{\partial E}{\partial S|x} \quad (19.9)$$

which means (subject to the assumptions made above) that the system is cramped if and only if

$$\frac{\partial^2 E}{\partial x \partial S} = 0 \quad (19.10)$$

Note that if you re-did the calculation to find the conditions under which the integral of work (around a cycle) was zero, you would get

$$\frac{\partial^2 E}{\partial S \partial x} = 0 \quad (19.11)$$

which is the same thing, because of the equivalence of mixed partials. We knew in advance that it had to work out this way, because of equation 19.1. The integral of the LHS is zero, so on the RHS, if one term is zero the other must be zero also.

The dichotomy between cramped and uncramped is an endless source of misconceptions.

People tend to have a great deal of hands-on experience with cramped systems such as the heating and cooling of a potato, or the heating and cooling of a baby bottle. The concepts of energy, temperature, entropy, and heat-content are well defined for such a system.

People have relatively little experience with heat engines. If you generalize from cramped to uncramped thermodynamics, you get to keep the ideas of energy, temperature, and entropy ... but you do *not* get to keep any notion of heat content or thermal energy content.

### 19.3 Real-World Compound Cramped Systems

The Slinktato™ is an ultra-simplified example of a compound cramped system.

The ocean is a real-world example. It makes sense to speak of the “thermal energy content” of the ocean. There is also mechanical energy in the ocean, in the form of tides and waves and huge currents such as the Gulf Stream. The mechanical variables are decoupled from the thermal variables to a good approximation, albeit not exactly. There is a thermodynamic process that drives the Gulf Stream, but virtually none of this energy leaves the ocean in mechanical form, so this doesn’t count as much of a heat engine. There exist tidal stream generators that extract energy from the ocean, but this is a negligible part of the overall energy budget of the ocean. Most of the budget involves purely thermal processes, such as sunlight in and evaporation out.

A lot of people have a knee-jerk response that says “there is no such thing as heat content” but we have to make an exception for compound cramped systems. There is nothing wrong with talking about the “heat content” of the ocean (on an appropriate timescale).

### 19.4 Heat Content, or Not

Here is yet another way of looking at the distinction between cramped and uncramped thermodynamics. Consider the elementary example of so-called “heat content” or “thermal energy content” that might arise in connection with a measurement of the heat capacity of a cylinder of compressed gas. We have a problem already, because there are two heat capacities: the heat capacity at constant pressure, and the heat capacity at constant volume. So it is unclear whether the heat content should be  $C_P T$  or  $C_V T$ . Now we get to play whack-a-mole: You can remove the ambiguity by rigorously restricting attention to either constant volume or constant pressure ... but that restriction makes it impossible to analyze a Carnot-type heat engine.

To repeat: It may at first be tempting to think that the gas cylinder has a so-called “thermal energy” related to  $T$  and  $S$ , plus a “nonthermal energy” related to  $P$  and  $V$ , but if you try to build a theory of thermodynamics on this basis you are guaranteed to fail. The sooner you give up, the happier you will be.

Cramped thermodynamics is a legitimate topic. It is only a small subset of thermodynamics, but it's not crazy. Almost everyone learns about cramped thermodynamics before they learn about uncramped thermodynamics. Consider for example warming the milk in a baby-bottle. This is almost always carried out under conditions of constant pressure. You're not trying to build a steam engine (or any other kind of engine) out of the thing. In this case, for this narrow purpose, there is a valid notion of the “heat content” of the system.

Within limits, the choice is yours: If you want to do cramped thermodynamics, you can do cramped thermodynamics. Just please don't imagine your results apply to thermodynamics in general. Cramped thermodynamics by definition is restricted to situations where the state-space is so low-dimensional that there is no hope of building a heat engine or a refrigerator or anything like that. There are no Carnot cycles, nor indeed any other kind of nontrivial cycles.

Long ago, there was a fairly elaborate theory of caloric. This was superseded by thermodynamics during the 19th century.

To repeat, it is OK to talk about “heat content” in the context of warming up a baby bottle. It is OK to talk about “caloric” in connection with a swimming pool as it warms up in the spring and cools down in the fall. It is OK to talk about “thermal energy” in connection with the heat capacity of a chunk of copper in a high-school lab experiment.

However, just because it works in cramped situations doesn't mean it works in uncramped situations.

It is not OK to talk about “heat content” or “thermal versus nonthermal energy” or “caloric” in the context of uncramped thermodynamics, i.e. in any situation where it is possible to build a heat engine.

Energy is energy. Energy doesn't recognize the distinction between thermal and nonthermal, and thermodynamics allows us to convert between the two (in any situation where it is possible to have a nontrivial thermodynamic *cycle*, i.e. in any situation where it is possible in principle to build a heat engine).

The problem is that the  $Q$  that appears in equation 19.2 simply cannot exist in the context of uncramped thermodynamics.

- Writing  $dQ$  doesn't make the problem go away.
- Writing  $\delta Q$  with a slash through the derivative operator doesn't make the problem go away.
- Writing  $DQ$  doesn't make the problem go away.
- Writing  $\Delta Q$  doesn't make the problem go away.
- Writing it inside an integral such as  $\int \cdots dQ$  doesn't make the problem go away.
- Writing  $dQ_{\text{rev}}$  and restricting attention to reversible paths doesn't make the problem go away; see section 19.5.
- Restricting attention to “chemical” situations doesn't make the problem go away.
- Permutations and combinations of the above don't make the problem go away.

The problem *still* is that  $Q$  exists as a state function only within cramped thermodynamics, not more generally, not in any situation where a heat engine is possible. In uncramped thermodynamics,  $Q$  may exist as a functional of some path, but not as a function of state.

For a list of constructive suggestions about things that actually do exist as functions of state, see section 8.2.

You can visualize the situation by reference to figure 19.2.

<p>On the LHS, if we restrict attention to the subspace defined by the red line, there is only one path from <math>A</math> to <math>Z</math>.</p>	<p>On the RHS, there are many ways of getting from <math>A</math> to <math>Z</math>, including <math>A \rightarrow Z</math>, or <math>A \rightarrow Y \rightarrow Z</math>, or even paths that include cycles, such as <math>A \rightarrow X \rightarrow Y \rightarrow A \rightarrow X \rightarrow Y \rightarrow A \rightarrow Z</math>, and so forth.</p>
--	--

<p>Within the subspace defined by the red line in figure 19.2, you can represent <math>Q</math> as height, and this <math>Q</math> is well defined everywhere in this small, cramped subspace.</p>	<p>You cannot define a <math>Q</math> value as a function of position in a way that is consistent throughout the <math>(T, S)</math> space. The peculiar thing is that you can take almost any simple one-dimensional <i>subspace</i> in the plane and define a consistent <math>Q</math> function there, but you cannot extend this to cover the entire space. You can't pin down the location of the problem, because it is nowhere in particular, yet it is everywhere: you cannot assign a consistent height to points in this space.</p>
--	---

**Pedagogical remarks:** Virtually everyone begins the study of thermodynamics by considering cramped situations. This is traditional ... but it is a pedagogical disaster for anyone trying to learn uncramped thermodynamics. Cramped thermodynamics is not a good foundation for learning uncramped thermodynamics; it is aggressively deceptive.

Virtually every newcomer to thermodynamics *tries* to extend the “heat content” idea from cramped thermodynamics to uncramped thermodynamics. It always *almost* works ... but it never *really* works.

The next time you feel the need for a measure of “heat content” in the context of uncramped thermodynamics, lie down until the feeling goes away.

## 19.5 No Unique Reversible Path

Sometimes you hear people talking about “the” reversible from  $A$  to  $B$ . This makes no sense whatsoever.

The existence of the Carnot cycle  $A \rightarrow X \rightarrow Y \rightarrow Z \rightarrow A$  implies that there are (at least!) two inequivalent paths from  $A$  to  $Z$ , including the simple path  $A \rightarrow Z$  along a contour of constant entropy, and the more complex path  $A \rightarrow X \rightarrow Y \rightarrow Z$ .

Furthermore, not all cycles are Carnot cycles. The path  $A \rightarrow Y \rightarrow Z \rightarrow A$  is another perfectly legitimate thermodynamic cycle. Compared to a Carnot-cycle engine, a reversible heat engine that uses the  $A \rightarrow Y \rightarrow Z \rightarrow A$  cycle is more complex and harder to analyze, but only slightly so.

Within uncramped thermodynamics, you are allowed to build things that aren't heat engines. That is, you can have cycles that don't convert any heat-bath energy into useful work. The various possibilities are summarized in the Venn diagram in figure 19.3, which can be considered a less-detailed version of figure 19.2.

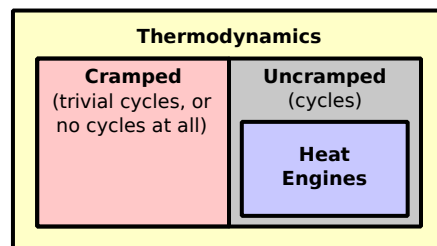


Figure 19.3: Cramped versus Uncramped Thermodynamics : Venn Diagram

## 19.6 Vectors: Direction and Magnitude

We now focus attention on the immediate neighborhood of point  $A$  in figure 19.2. It must be emphasized that paths can depart from point  $A$  in innumerable many directions. The Carnot cycle uses only two of these directions (namely the contour of constant  $T$  during one part of the cycle, and the contour of constant  $S$  during another part of the cycle). However, there are infinitely many non-Carnot cycles, and infinitely many ways in which a reversible path can depart from point  $A$  such that neither  $T$  nor  $S$  is constant. The blue line in figure 19.2 is just one of many such paths.

In the immediate neighborhood of point  $A$ , we can distinguish these paths by their *direction*. The red line in figure 19.2 represents a change in  $T$  in the direction of constant  $S$ , while the blue line represents a change in  $T$  along some other direction.

Therefore, uncramped thermodynamics requires us to treat  $dT$  as a vector. If you think of  $dT$  as representing some kind of “change in  $T$ ” you need to specify the direction of the change (as well as the magnitude). Whenever something has a direction and a magnitude, you should suspect that it is a vector.

For large excursions, we would need to specify the entire path, but in the immediate neighborhood of a given point, it suffices to know the magnitude and direction. Therefore a vector such as  $dT$  can be considered a function of state. It depends on the local state, not on the entire path. It is a vector-valued function of state.

## 19.7 Reversibility

The existence of *reversible* heat engines is sufficient to guarantee the existence of innumerable many inequivalent paths from  $A$  to  $Z$ , and also to guarantee the existence of innumerable many directions for the vectors located at point  $A$ .

In this chapter, for simplicity, we have not mentioned any irreversible processes.

Irreversibility is not the cause of the multiple paths and multiple directions. Conversely, multiple paths and multiple directions are not the cause of irreversibility.





## Chapter 20

# Ambiguous Terminology

### 20.1 Background

As always, ideas are primary and fundamental. Terminology is important only insofar as it helps us formulate and communicate the ideas.

By way of analogy, consider the relationship between organic vegetables and organic chemistry. There is nothing wrong with either of those two ideas. Neither one – by itself – is a misconception. However, we have a problem with the terminology. The word “organic” is being used with two different definitions. You can *create* a misconception by using one definition in a situation where the other is appropriate, but that is a secondary, artificial, and needless problem. The primary problem is the terminology.

Sometimes a word has two definitions that are so different that no confusion arises, for example “dove” (the past-tense verb) and “dove” (the bird). In contrast, you can easily get into trouble if a word has two meanings are similar enough to be deceptive, yet not so similar that the difference can be ignored.

Many of the most-important terms in thermodynamic can be highly deceptive if you’re not careful.

In such a situation, good learning and good critical thinking demand that you learn each concept in its proper context, and then learn to *reconcile* them all. That requires learning how to distinguish one context from the other, so that each concept can be used appropriately. The goal should be to understand the *whole* situation. To say the same thing the other way: You’ll never really understand the subject if you learn one concept and ignore the other. Studying one part of the elephant in great detail will not make the rest of the elephant go away.

I don’t mind struggling with something if it is intrinsically hard and complicated. In contrast, it’s painful to see people struggling with things that are complicated for no good reason, for instance because the terminology is messed up. That’s a wanton waste of resources. We should be able to fix the terminology, so that nobody has to struggle with it!

## 20.2 Overview

In physics, there is only meaning of “energy,” but the physics meaning conflicts with the plebeian meaning, as discussed in section 1.8.1 and section 20.3.

In physics, there is almost only one meaning of “conservation,” but the physics meaning conflicts with the plebeian meaning, as discussed in section 1.8.2 and section 20.4.

There are multiple inconsistent technical meanings for “heat,” not to mention innumerable nontechnical meanings, as discussed in chapter 17.

There are multiple inconsistent technical meanings for “work” as discussed in chapter 18.

There are multiple inconsistent technical meanings for “adiabatic” as discussed in chapter 16.

In the literature, the term “state” is used inconsistently. It can either mean microstate or macrostate, as discussed in section 2.7 and section 12.1.

Similarly, “phase space” is ambiguous:

<p>Phase-space means one thing in classical canonical mechanics; it corresponds to what we have been calling state-space, as discussed in section 12.3.</p>	<p>Phase space means something else in classical thermodynamics; it has to do with macroscopic phases such as the liquid phase and the solid phase.</p>
---	---

(Ironically, Gibbs has his name associated with both of these notions.)

I’m not even talking about quantum mechanical phase  $\phi$ , as in  $\exp(i\phi)$ ; that’s a third notion, which is not terribly troublesome because you can usually figure out the meaning based on context.

Given how messed-up our language is, it’s a miracle anybody ever communicates anything.

## 20.3 Energy

As mentioned in section 1.8.1, the plebeian notion of energy corresponds roughly to “available” energy or “useful” energy. This is important, but very hard to quantify.

The “available energy” is not equal to the physics energy  $E$ , and also not equal to the Helmholtz free energy  $F$  or the Gibbs free enthalpy  $G$  (as defined in chapter 15). The simple way to understand this is to realize that  $E$ ,  $F$ , and  $G$  are functions of state, whereas the plebeian notion of “useful energy” is not. The physics energy of a particular parcel depends only on the properties of the parcel itself, whereas the usefulness of that energy depends on properties of the parcel *and* properties of the surrounding world.

By way of analogy: A grocery store in Iowa sells a lot more bags of ice in mid-summer than it does in mid-winter. The thermodynamic *state* of the ice is the same in either case, but its usefulness is wildly different.

In relative terms,  $F$  and  $G$  are “closer” to capturing the idea of “available” energy, but in absolute terms they are not close enough. They are not viable contenders for quantifying the “useful” or “available” energy.

- In some cases, the difference between “useful” and “non-useful” energy has almost nothing to do with entropy. For example, sometimes people spend several billion dollars and incur enormous risk to build a nuclear power plant. The plant does not increase the amount of energy in the uranium. Indeed the plant wastes most of the theoretically-available energy of the uranium. In some sense, the purpose of the plant is to *concentrate* the energy. The fraction of the energy that is not wasted is made vastly more concentrated in space and time, and this makes it more useful.
- In other cases, the difference between “useful” and “non-useful” energy depends much more directly on entropy. For example, if you have a hot potato and a cold potato, you can use them to run a heat engine and do useful work. In contrast, if you have two tepid potatoes with the same total energy, the energy is not useful, in the sense that you cannot run a heat engine and do useful work. This can be understood in terms of entropy, namely the entropy and energy of the two potatoes together. (The available energy cannot be written as a state-function pertaining to either potato separately.)

Any proper theory of “useful energy” would involve a great deal of microeconomics, not just physics. There is an elaborate theory of microeconomic utility.

## 20.4 Conservation

1. In physics, the main meaning of conservation refers to continuity of flow, as expressed in equation 1.1.
2. Unfortunately, even within physics the word “conservative” has been given a second meaning, as we now discuss. Sometimes a vector field is the gradient of some scalar potential, in which case we say the vector field is *grady*. Mathematicians would call it an exact differential, or an exact one-form. This math terminology is not recommended, because it conflicts too strongly with the common-sense concept of “exact” as in strictly true, not an approximation. Also, to my way of thinking, there is no such thing as an inexact differential; if it’s not exact, it’s not a differential at all. You could call it an inexact non-differential, or an inexact one-form, or (preferably) a non-grady one-form.

Sometimes in physics, a grady force-field is called a “conservative” force, and by the same token a non-grady force-field is called a “non-conservative” force. This terminology is emphatically not recommended, because it conflicts with the definition of conservative flow as expressed in equation 1.1.

Whenever you want to describe a force field that is not the gradient of any potential, I strongly recommend calling it a non-grady force field. Similarly, if you ever see the term “non-conservative force”, cross it out and substitute “non-grady force”. A non-grady force, such as you find in a transformer or in a betatron, does *not* violate any conservation laws.

3. We now consider a third definition of “conservation”, namely the plebeian definition, i.e. the non-physics definition. The key idea here has to do with saving, preserving, not wasting, not dissipating. For example, we speak of conservation of endangered wildlife.

This is an important concept, but very difficult to quantify.

Wasting something, almost by definition, is irreversible, which suggests that the plebeian notion of conservation is loosely related to the idea of entropy – but only loosely.

## 20.5 Other Ambiguities

Numerous other conflicts are discussed in reference 45. This includes a great many basic terms used in math, physics, and thermodynamics ... including force, acceleration, gravity, closed system, et cetera.

## Chapter 21

# Thermodynamics, Restricted or Not

There are various ways of restricting the applicability of thermodynamics, including

- microcanonical only (i.e. constant energy)
- equilibrium only
- reversible only
- ideal gases only
- vertical size small compared to  $kT/mg$ .
- et cetera.

Indeed, there are some people who seem to think that thermodynamics applies only to microcanonical reversible processes in a fully-equilibrated ideal gas.

To make progress, we need to carefully distinguish two ideas:

- a) Simplifying assumptions made in the context of a particular scenario. Depending on details, these may be entirely appropriate. Sometimes the gases involved are ideal, to an excellent approximation ... but not always. Sometimes a process is reversible, to an excellent approximation ... but not always.
- b) Restrictions applied to the foundations of thermodynamics. We must be very careful with this. There must not be too many restrictions, nor too few. Some restrictions are necessary, while other restrictions are worse than useless.

Some thermodynamic concepts and/or formulas necessarily have restricted validity.

- As discussed in section 11.4, there are situations where it is impossible to define a temperature.
- The Boltzmann distribution law (equation 9.1 and figure 9.1) is valid only in equilibrium.
- The notion of equiprobable states (equation 9.8) applies exactly only in microcanonical equilibrium, although it may be a worthwhile approximation in other situations.
- Deciding how many macroscopic variables are needed to describe the macrostate requires some judgment, and depends on knowing the context. For example, equation 7.8 and similarly equation 15.13 are restricted to cases where advection of energy is insignificant, changes in the number of particles are insignificant, changes in magnetic fields or other applied fields have no significant effect, et cetera. If you want to lift these restrictions, you have to add additional terms to the equations.

In contrast, very importantly, the law of conservation of energy applies *without restriction*. Similarly, the law of paraconservation of entropy applies *without restriction*. You must not think of  $E$  and/or  $S$  as being undefined in regions where “non-ideal” processes are occurring. Otherwise, it would be possible for some energy and/or entropy to flow into the “non-ideal” region, become undefined, and never come out again, thereby undermining the entire notion of conservation.

The ideas in the previous paragraph should not be overstated, because an approximate conservation law is not necessarily useless. For example, ordinary chemistry is based on the assumption that each of the chemical elements is separately conserved. But we know that’s only approximately true; if we wait long enough uranium will decay into thorium. Still, on the timescale of ordinary chemical reactions, we can say that uranium is conserved, to an excellent approximation.

When a law has small exceptions, you shouldn’t give up on the law entirely. You shouldn’t think that just because a process is slightly non-ideal, it becomes a free-for-all, where all the important quantities are undefined and none of the laws apply.

If you want to make simplifying assumptions in the context of a specific scenario, go ahead ... but don’t confuse that with restrictions on the fundamental laws.

Also, in an elementary course, it might be necessary, for pedagogical reasons, to use simplified versions of the fundamental laws ... but you need to be careful with this, lest it create misconceptions.

- As an example: an imperfect notion of entropy in terms of multiplicity (equation 9.8) is better than no notion of entropy at all. However sooner or later (preferably sooner) you need to understand that entropy is really defined in terms of statistics (equation 2.2 or equation 27.6), not multiplicity.
- As another example: In an elementary course, it might be appropriate to start by applying thermo to ideal gases. However, sooner or later (preferably sooner) it is very important to consider other systems; otherwise you risk horrific misconceptions, as discussed in section 9.3.3.

Finally, it must be emphasized that one should not ask whether thermodynamics “is” or “is not” applicable to a particular situation, as if it were an all-or-nothing proposition. Some concepts (such as energy and entropy) are always valid, while other concepts (such as equilibrium and temperature) might or might not be valid, depending on the situation.

## Chapter 22

# The Relevance of Entropy

The concept of entropy is important in the following areas, among others:

- 1) cryptography and cryptanalysis, i.e. secret codes
- 2) communications, data storage, and information theory, including error-correcting codes and data-compression codes
- 3) computer science, including machine learning, etc.
- 4) pattern recognition, including speech recognition, optical character recognition, face recognition, etc.
- 5) librarianship
- 6) the design of experiments (reference 13)
- 7) physics in general, including the physics of computation, cosmology, and astrophysics
- 8) the design of thermal machinery, including refrigerators, heat pumps, piston engines, turbine engines, rocket engines, etc.
- 9) nuclear engineering, including reactors and weapons
- 10) fluid dynamics
- 11) chemistry and chemical engineering

Very roughly speaking, the items higher on the list can be assigned to the “information theory” camp, while the items lower on the list can be assigned to the “thermodynamics” camp. However, there is tremendous overlap between the two camps. The approach of understanding the microscopic quantum states and using that to explain macroscopic observables such as energy, entropy, temperature, etc. is called *statistical mechanics*; see e.g. reference 28 and reference 49. Examples of this include

- a) The physics of computation is squarely in both camps; see reference 36, reference 37, and reference 50.
- b) Things like Maxwell demons and Szilárd engines are squarely in both camps; see reference 51 and reference 52.
- c) Demagnetization refrigerators (as described in section 11.10) are in both camps, because you can quantify the entropy either by microscopic state-counting or by macroscopic thermal measurements. The macroscopic molar heat capacity of copper nuclei is  $R \ln 4$ , where the 4 is the number of microscopic states per nucleon.

- d)** When you have a three-dimensional gas phase in equilibrium with a two-dimensional gas phase (i.e. adsorbed particles freely moving along a surface), the equilibrium between the two phases depends partly on a Boltzmann factor (involving the surface binding energy), but also depends on a state-counting factor. There are more states in the bulk than there are on the surface. This involves a Boltzmann factor (which gives the probability per microstate) as well as state-counting factors that scale like volume/ $\Lambda^3$  and area/ $\Lambda^2$ . This allows us to calculate the entropy via  $\sum_i P_i \log(1/P_i) \dots$  and the result agrees with the entropy associated with macroscopic measurements of energy and temperature. In particular it gives us a very simple and useful way of calculating the surface density as a function of bulk density and temperature.
- e)** Similar microstate-counting factors appear in the Saha equation, which uses statistical mechanics to make a quantitative prediction of ionization levels as a function of density. See reference 53 for a discussion of how such things scale as a function of the temperature and volume of the system.
- f)** Similar microstate-counting factors permit an understanding of the scaling (with respect to density or concentration) of the equilibrium quotient and equilibrium “constant” in chemical reactions. The state-counting and the scaling are definitely nontrivial in the case of reactions such as  $F_2 \leftrightarrow 2F$ , where the number of reactant particles is different from the number of product particles. This is discussed in some detail in reference 53. Once again, microscopic state-counting is key to understanding macroscopic observations such as densities and reaction rates.

So: we have multiple lines of evidence telling us that entropy is entropy. It is the same entropy, no matter whether you measure it in bits or in joules per kelvin (section 9.5).

As mentioned in chapter 2, you can’t do thermodynamics without entropy.

Also: entropy is one of the great elegant ideas of all time. C.P. Snow compared not knowing about the second law to never having read a work by Shakespeare.



## Chapter 23

# Equilibrium, Equiprobability, Boltzmann Factors, and Temperature

### 23.1 Background and Preview

In this chapter we investigate the following propositions. They are often assumed to be true, and sometimes even “proved” to be true, but we shall see that there are exceptions.

Understanding these propositions, and their limitations, is central to any real understanding of thermodynamics.

For any system with a constant number of particles, in thermal equilibrium:

- 1) For any two accessible microstates that have the same energy, their probabilities are equal. That is,

$$P_i = P_j \quad \text{if} \quad E_i = E_j \quad (23.1)$$

for any two microstates  $i$  and  $j$ .

- 2) Supposedly, for any two accessible microstates that differ in energy, their probabilities are related by a Boltzmann factor. Specifically:

$$\frac{P_i}{P_j} = \exp\left(\frac{E_j - E_i}{kT}\right) \quad (23.2)$$

for some value of  $T$ . This  $T$  is called the *temperature*. In the numerator on the RHS, the sign is such that the microstate with the greater energy has the lesser probability (assuming the temperature is positive and finite).

- 3) Supposedly, equilibrium is symmetric: If  $A$  is in equilibrium with  $B$ , then  $B$  is in equilibrium with  $A$ , and has the same temperature. Also, equilibrium is transitive: if  $A$  is in equilibrium with  $B$ , and  $B$  is in equilibrium with  $C$ , then  $A$  is in equilibrium with  $C$ .

As a corollary, equilibrium is weakly reflexive. That is, if a system is in equilibrium with anything else, it is in equilibrium with itself.

We postpone until section 23.5 any discussion of proposition (3).

It is widely believed and often “proved” that proposition (1) is equivalent to proposition (2), i.e. that each one follows from the other. We shall see that in fact, the two propositions are almost *but not exactly* equivalent. The discussion will shed light on some quite fundamental issues, such as what we mean by “thermal equilibrium” and “temperature”.

It is trivial to show that proposition (1) follows from proposition (2), since the former is just a special case of the latter, namely the case where  $E_i = E_j$ .

The converse is quite a bit more of a challenge. The rest of this section is devoted to figuring out under what conditions we might be able to derive equation 23.2 from equation 23.1. The derivation requires many steps, each of which is simple enough, but the cumulative effect is rather complicated, so it is sometimes hard to see the whole picture at once. Complicating factors include:

- We will be considering the heat bath  $B$ , the special subsystem  $S$ , and the overall system  $B + S$ .
- We will be considering individual sites, microstates (each a collection of  $N$  sites) and macrostates (each a set of many, many microstates).
- We will be considering conditional probabilities as well as unconditional probabilities.
- When the energy of  $S$  goes up, the energy of  $B$  goes down, so it is misleading to talk of a “high energy” state or a “low energy” state.
- We need to carefully define what we mean by “heat bath”, because otherwise what we are trying to prove is not necessarily true.
- We will use what we know about the constant-energy case to tell us what happens in the constant-temperature case.

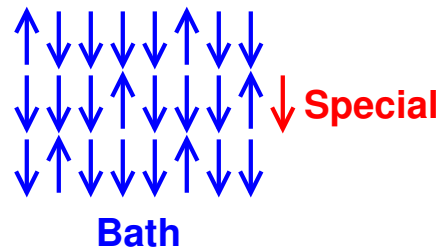
We begin by considering some numerical examples.

## 23.2 Example: $N = 1001$

Our first example consists of the system shown in figure 23.1. The system is divided into two subsystems: Subsystem “ $B$ ” is shown in blue, and will sometimes serve as “bath” (i.e. heat bath). Subsystem “ $S$ ” is shown in scarlet, and will sometimes be referred to as the “small” or “special” subsystem.

In general there are  $N_S$  sites in the scarlet subsystem,  $N_B$  sites in the blue subsystem, for a total of  $N = N_S + N_B$  sites overall. We start by considering the case where  $N_S = 1$ ,  $N_B = 1000$ , and  $N = 1001$ .

For clarity, there are only  $N_B = 24$  blue sites shown in the figure, so you will have to use your imagination to extrapolate to  $N_B = 1000$ .

Figure 23.1: Heat Bath + Special Subsystem ( $x = 0$ )

The overall system  $B + S$  is isolated so that its energy is constant. The various sites within the overall system are weakly interacting, so that they can exchange energy with each other. In our first example, all  $N$  sites in the overall system are equivalent. That is, we have arbitrarily designated one of the sites as “special” but this designation has no effect on the physics of the overall system.

Each of the  $N$  sites can be in one of two states, either UP or DOWN. The energy of the UP state is higher than the energy of the DOWN state by one unit.

We have arranged that  $m$  of the  $N$  sites are in the UP state. We choose the zero of energy such that  $E = m$ . We shall be particularly interested in the case where  $m = 250$  and  $N = 1001$ .

The overall system has only one macrostate, namely the set of all microstates consistent with the given (constant) values of total  $N$  and total energy  $E$ . There are  $W$  microstates in the given macrostate, where  $W$  is called the multiplicity.

Figure 23.1 is a snapshot, showing only one microstate of the overall system. By conservation of energy we have constant  $m$ , so we can find all the other microstates by simply finding all permutations, i.e. all ways of assigning  $m$  UP labels to  $N$  sites. That means the multiplicity can be computed in terms of the binomial coefficient:

$$\begin{aligned} W &= \binom{N}{m} \\ &= \frac{N!}{(N-m)!m!} \end{aligned} \quad (23.3)$$

Note that the binomial coefficient  $\binom{N}{m}$  is generally pronounced “ $N$  choose  $m$ ”. It is implemented in typical spreadsheet programs by the `combin(N,m)` function.

For the present example, the numerical values are:

$$\begin{aligned}
 N &= 1001 \\
 m &= 250 \\
 W &= 6.43 \times 10^{242} \\
 \log_2(W) &= 806.6\text{bits}
 \end{aligned}
 \tag{23.4}$$

The microstates are all equivalent, so the probability of the  $i$ th microstate is  $P_i = 1/W$  for all  $i$ .

Let's think about the symmetry of the situation. All  $N$  sites are equivalent, so we expect that anything that happens at one site is equally likely to happen at any other site.

Because (by construction)  $m$  is very nearly one fourth of  $N$ , if we pick any site at random, it is very nearly three times as likely to be in the DOWN state as in the UP state. Since we imagine that the sites are freely exchanging energy, we can replace the average over sites by a time average at a single site, whereupon we see that the scarlet site (or any other site) is three times as likely to be found in the DOWN state as in the UP state. In symbols:

$$\begin{aligned}
 \frac{P_S(\text{UP})}{P_S(\text{DOWN})} &= \frac{m}{N-m} \\
 &\approx \frac{1}{3}
 \end{aligned}
 \tag{23.5}$$

We can define two categories of microstates: one where the special site is in the DOWN state (which we call  $x = 0$ , as in figure 23.1), and another where the special site is in the UP state (which we call  $x = 1$ , as in figure 23.2).

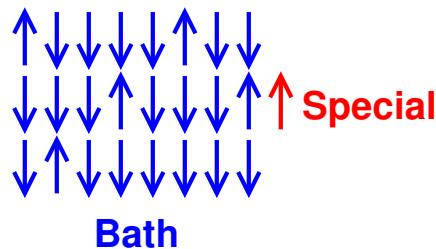


Figure 23.2: Heat Bath + Special Site ( $x = 1$ )

These categories are in some ways just like macrostates, in the sense that they are sets of microstates. However, for clarity we choose to call them categories not macrostates. We can calculate the multiplicity of each category separately. As before, all we need to do is count permutations. Within

each category, however, we only consider permutations *of the blue sites* because the state of the scarlet site is fixed.

We can streamline the discussion by borrowing some notation that is commonly applied to chemical reactions. Here  $x$  is the *reaction coordinate*. The reaction of interest involves the transfer of one unit of energy to the special site from the heat bath. That is:

$$\begin{array}{rclcl}
 x & = & 0 & \rightarrow & 1 & \text{reaction coordinate} \\
 W_B & = & \binom{N-1}{m} & \rightarrow & \binom{N-1}{m-1} & \text{multiplicity of the bath} \\
 & = & \binom{1000}{250} & \rightarrow & \binom{1000}{249} & \\
 & = & 4.82 \times 10^{242} & \rightarrow & 1.61 \times 10^{242} & \\
 \text{ratio} & = & & & 3 : 1 & \\
 \log_2(W_B) & = & 806.18\text{bits} & \rightarrow & 804.59\text{bits} & \\
 \ln(W_B) & = & 558.80\text{nats} & \rightarrow & 557.70\text{nats} & 
 \end{array} \tag{23.6}$$

The multiplicity of the  $x = 1$  category is less, because when we do the permutations, there is one fewer “up” state to play with.

Whether or not we assign these microstates to categories, they are still microstates of the overall system. Therefore they all have the same energy, since the system is isolated. Therefore the microstates are all equally probable, in accordance with proposition (1) as set forth at the beginning of section 23.1.

If you look at the numbers in equation 23.6, you see that the  $x = 0$  microstates are very nearly threefold more numerous than the  $x = 1$  microstates. We can calculate this exactly in terms of  $m$  and  $N$ :

$$\begin{aligned}
 \frac{P_S(\text{UP})}{P_S(\text{DOWN})} &= \frac{\binom{N-1}{m-1}}{\binom{N-1}{m}} \\
 &= \frac{m}{N-m} \\
 &\approx \frac{1}{3}
 \end{aligned} \tag{23.7}$$

You can verify algebraically that the ratio of multiplicities is exactly equal to  $m/(N - m)$ . This factor shows up in both equation 23.5 and equation 23.7, which means the probability we get by counting microstates is provably identical to the probability we get from symmetry arguments.

Consistency is always nice, but in this case it doesn’t tell us much beyond what we already knew. (Things will get much more exciting in a moment.)

Feel free to skip the following tangential remark. It is just another consistency check. The rest of the development does not depend on it.

Let’s check that the multiplicity values for the categories are consistent with the multiplicity of the overall system.

Each category has a certain multiplicity. If we add these two numbers together, the sum should equal the multiplicity of the overall system.

We know this “should” be true, because we have exhausted all the possibilities.

We can verify that it is in fact true by using the mathematical properties of the binomial coefficients, especially the fact that each entry in Pascal’s triangle is the sum of the two

entries above it on the previous row. To say the same thing more formally, you can easily verify the following algebraic identity:

$$\binom{N}{m} = \binom{N-1}{m} + \binom{N-1}{m-1} \quad (23.8)$$

### 23.3 Example: $N = 1002$

To obtain a clearer picture of what is going on, and to obtain a much stricter check on the correctness of what we have done, we now increase the number of scarlet sites to  $N_S = 2$ . To keep things simple we increase the total  $N$  to 1002 and increase  $m$  to 251. The reaction coordinate can now take on the values  $x = 0$ ,  $x = 1$ , and  $x = 2$ . I'm not going to bother redrawing the pictures.

The trick of calculating the scarlet-subsystem probabilities by appeal to symmetry still works (although it isn't guaranteed to work for more complicated systems). More importantly, we can always calculate the probability by looking at the microstates; that always works. Indeed, since the microstates of the overall system are equiprobable, all we need to do is count them.

$x$	$=$	0	$\rightarrow$	1	$\rightarrow$	2	rxn coord
$W_B$	$=$	$\binom{N-2}{m}$	$\rightarrow$	$\binom{N-2}{m-1}$	$\rightarrow$	$\binom{N-2}{m-2}$	multiplicity
	$=$	$\binom{1000}{251}$	$\rightarrow$	$\binom{1000}{250}$	$\rightarrow$	$\binom{1000}{249}$	
	$=$	$1.44 \times 10^{243}$	$\rightarrow$	$4.82 \times 10^{242}$	$\rightarrow$	$1.61 \times 10^{242}$	
ratio	$=$	$3 : 1$		$3 : 1$			
$\log_2(W_B)$	$=$	807.76bits	$\rightarrow$	806.18bits	$\rightarrow$	804.59bits	
$\ln(W_B)$	$=$	559.89nats	$\rightarrow$	558.80nats	$\rightarrow$	557.70nats	

(23.9)

The situation is shown in figure 23.3. We see that every time the scarlet subsystem energy goes up (additively), the bath energy goes down (additively), the multiplicity goes down (multiplicatively), and therefore the log multiplicity goes down (additively). Specifically, the log multiplicity is very nearly linear in the energy, as you can see from the fact that (to an excellent approximation) the points fall on a straight line in figure 23.3.

If we define temperature to be the negative reciprocal of the slope of this line, then this example upholds proposition (2). This definition is consistent with the previous definition, equation 7.7.

Our example is imperfect in the sense that the three points in figure 23.3 do not fall *exactly* on a straight line. Therefore our example does not *exactly* uphold proposition (2). On the other hand, it is quite a good approximation. The points fall so nearly on a straight line that you probably can't see any discrepancy by looking at the figure. We shall demonstrate in a moment that there is some nonzero discrepancy. This is not tragic; we can rationalize it by saying that a bath consisting of 1000 sites is a slightly imperfect heat bath. In the limit as  $N$  and  $m$  go to infinity, the bath becomes perfect.

We can quantify the imperfection as follows: The probability ratio between the upper two points is:

$$\frac{W(x=1)}{W(x=0)} = \frac{m}{N-1-m} \quad (23.10)$$

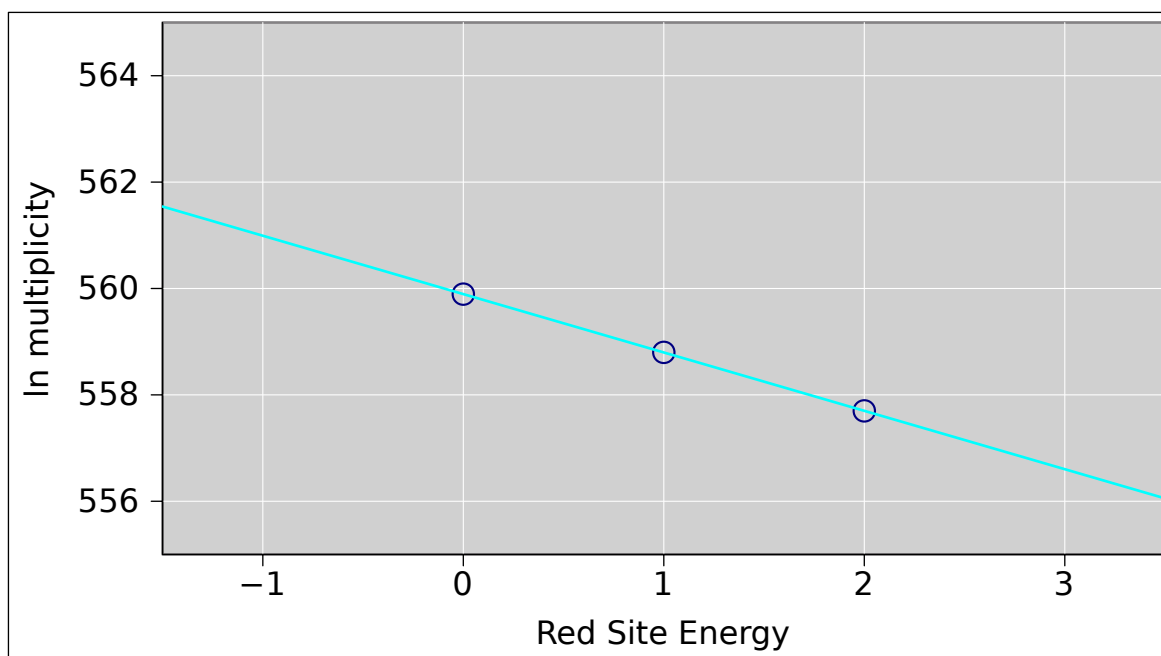


Figure 23.3: Log Multiplicity versus Energy

Meanwhile, the ratio between the lower two points is:

$$\frac{W(x=2)}{W(x=1)} = \frac{m-1}{N-m} \quad (23.11)$$

which is obviously not the same number. On the other hand, if you pass to the limit of large  $N$  and large  $m$ , these two ratios converge as closely as you like. (Also note that these two ratios bracket the ratio given in equation 23.7.)

## 23.4 Example: $N = 4$

We have just seen that the advantages of having a heat bath with a large number of sites.

To emphasize this point, let's see what happens when  $N_B$  is small. In particular, consider the case where  $N_B = 2$ ,  $N_S = 2$ , and  $m = 2$ . Then the ratios in equation 23.10 and equation 23.11 are 2:1 and 1:2 respectively ... which are spectacularly different from each other. The situation is shown in figure 23.4.

Obviously these points do not lie on a straight line. The probabilities do not follow a Boltzmann distribution, not even approximately. A major part of the problem is that the blue subsystem, consisting of  $N_B = 2$  sites, is not a good heat bath, not even approximately.

In this situation, temperature is undefined and undefinable, even though the system satisfies Feynman's definition of thermal equilibrium, i.e. when all the fast things have happened and the slow things have not. This is the maximum entropy macrostate, the most entropy the system can have

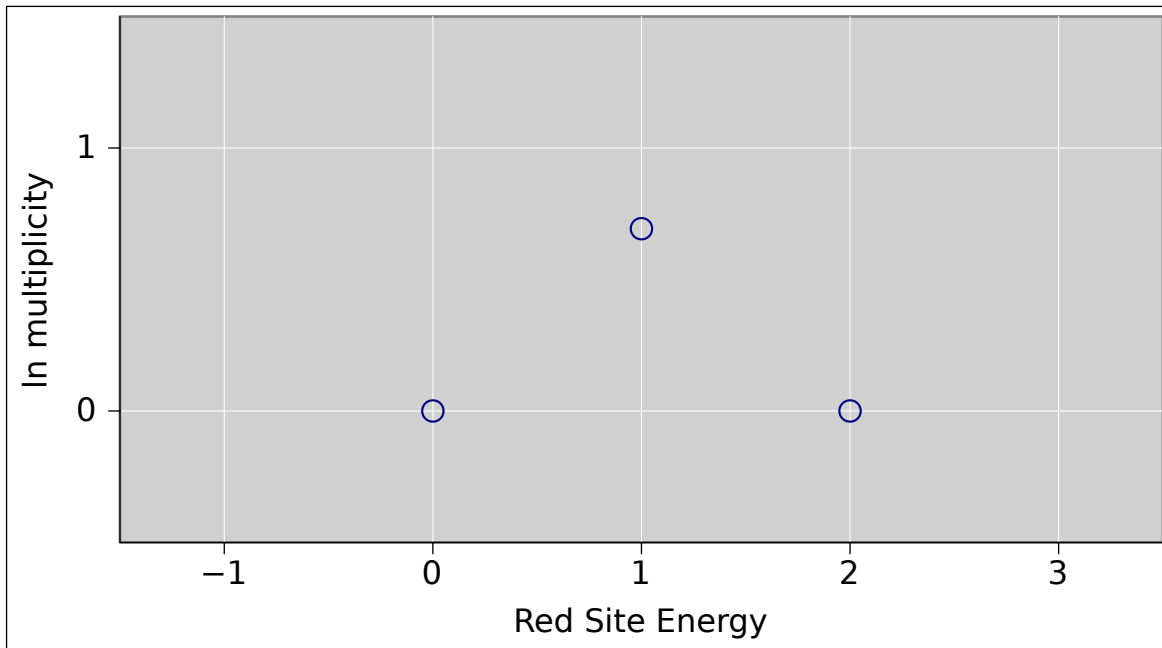


Figure 23.4: Log Multiplicity versus Energy – No Temperature

subject to the stipulated constraints (constant  $N$  and constant  $m$ ). This is a very peculiar state, but as far as I can tell it deserves to be labeled the equilibrium state. Certainly there is no other state with a stronger claim to the label.

Note that the 1:2:1 ratio we are discussing, as shown in figure 23.4, gives the probability per microstate for each of the four microstates. If you are interested in the probability of the three energy levels, the answer is 1:4:1, because the  $x = 1$  energy level has twice the multiplicity of the others. Always remember that the probability of a macrostate depends on the number of microstates as well as the probability per microstate.

### 23.5 Role Reversal: $N = 1002$ ; $T_M$ versus $T_\mu$

The situation shown in figure 23.4 may seem slightly contrived, since it applies to thermal equilibrium in the absence of any well-behaved heat bath. However, the same considerations sometimes come into play even when there is a heat bath involved, if we use it bath-backwards. In particular, we now return to the case where  $N = 1002$  and  $N_S = 2$ . We saw in section 23.3 that in this situation, the scarlet subsystem exhibits a Boltzmann distribution, in accordance with proposition (2). But what about the blue subsystem?

It turns out that each and every microstate of the blue subsystem in the  $x = 0$  and  $x = 2$  categories has the same probability, even though they do not all have the same energy. This means that the blue microstate probabilities do not follow a Boltzmann distribution.

Furthermore, each blue microstate in the  $x = 1$  category has twice as much probability as any  $x = 0$  or  $x = 2$  microstate, because there are two ways it can happen, based on the multiplicity



of the corresponding microstates of the overall subsystem. That is, when  $x = 1$ , there are two microstates of the overall system for each microstate of the blue subsystem (due to the multiplicity of the scarlet subsystem), and the microstates of the overall system are equiprobable. The result is closely analogous to the situation shown in figure 23.4.

The way to understand this is to recognize that when  $N_S = 2$ , the scarlet subsystem is too small to serve as a proper heat bath for the blue subsystem.

At this point, things are rather complicated. To help clarify the ideas, we rearrange the Boltzmann distribution law (equation 23.2) as follows:

$$T_\mu := \frac{-1}{k} \frac{E_i - E_j}{\ln(P_i) - \ln(P_j)} \quad (23.12)$$

for any two microstates  $i$  and  $j$ . We take this as the definition of  $T_\mu$ , where  $\mu$  refers to *microstate*.

We contrast this with the conventional definition of temperature

$$T_M := \frac{-1}{k} \frac{E_i - E_j}{\ln(W_i) - \ln(W_j)} \quad (23.13)$$

We take this as the definition of  $T_M$ , where  $M$  refers to *macrostate*. As far as I can tell, this  $T_M$  is what most people mean by “the” temperature  $T$ . It more-or-less agrees with the classical definition given in equation 7.7.

It must be emphasized that when two subsystems are in contact, the Boltzmann property of one system depends on the bath-like behavior of the other. The  $T_\mu$  of one subsystem is equal to the  $T_M$  of the other. That is,  $S$  is Boltzmann-distributed if  $B$  is a well-behaved heat bath; meanwhile  $S$  is Boltzmann-distributed if  $B$  is a well-behaved heat bath.

To say the same thing the other way, you cannot think of a subsystem as serving as a bath for itself. In the present example, for the blue subsystem  $T_M$  is well defined but  $T_\mu$  is undefined and undefinable, while for the scarlet subsystem the reverse is true:  $T_\mu$  is well defined but  $T_M$  is undefined and undefinable.

Among other things, we have just disproved proposition (3).

If you think that is confusing, you can for homework consider the following situation, which is in some ways even more confusing. It serves to even more dramatically discredit the idea that two subsystems in equilibrium must have the same temperature.

We have just considered the case where the scarlet subsystem consisted of two spin-1/2 particles, so that it had four microstates and three energy levels. Now replace that by a single spin-1 particle, so that it has only three microstates (and three energy levels).

In this scenario, there are three macrostates of the blue subsystem, corresponding to three different energies. The odd thing is that each and every microstate of the blue subsystem has exactly the same probability, even though they do not all have the same energy.

In some perverse sense these blue microstates can be considered consistent with a Boltzmann distribution, if you take the inverse temperature  $\beta$  to be zero (i.e. infinite temperature).

This situation arises because each energy level of the scarlet system has the same multiplicity,  $W_S = 1$ . Therefore the log multiplicity is zero, and  $\beta = (d/dE)\ln(W) = 0$ .

This situation is mighty peculiar, because we have two subsystems in equilibrium with each other both of which are Boltzmann-distributed, but which do not have the same temperature. We are attributing an infinite temperature to one subsystem and a non-infinite temperature to the other. This can't be good.

Note that in all previous scenarios we were able to calculate the probability in two different ways, by symmetry and by counting the microstates. However, in the present scenario, where we have a spin-1 particle in equilibrium with a bunch of spin-1/2 particles, we cannot use the symmetry argument. We can still count the microstates; that always works.

## 23.6 Example: Light Blue

We can deepen our understanding by considering yet another example.

At each of the blue sites, we replace what was there with something where the energy splitting is only half a unit. Call these “light blue” sites if you want. Meanwhile, the scarlet sites are the same as before; their energy splitting remains one full unit.

In this situation,  $m$  is no longer a conserved quantity. Whenever the reaction coordinate  $x$  increases by one, it annihilates two units of  $m_B$  and creates one unit of  $m_S$ . Energy is of course still conserved:  $E = m_S + m_B/2$ .

We wish the scarlet subsystem to remain at the same temperature as in previous examples, which means we want its UP/DOWN ratio to remain at 1/3. To do this, we must drastically change the UP/DOWN ratio of the blue subsystem. Previously it was 1/3 but we shall see that now it must be  $\sqrt{1/3}$ .

In our numerical model, we represent this by  $N_B = 1000$  and  $m_B = 368 - 2x$ .

Now, whenever we increase  $x$  by one, we now have *two* fewer UP states to play with in the blue subsystem, so the multiplicity changes by two factors, each of which is very nearly  $m_B/(N_B - m_B)$  i.e. very nearly  $\sqrt{3}$ . The two factors together mean that the multiplicity changes by a factor of 3, which means the probability of the  $S$  microstates changes by a factor of 3, as desired.

## 23.7 Discussion

One reason for working through the “light blue” scenario is to emphasize that the RHS of equation 23.13 is indeed properly written in terms of  $W$  and  $E$  ... in contrast to various other quantities that you might have thought would be more directly important.

There is a long list of things that might seem directly important, but are not. When two systems are in equilibrium:

- That does not mean they have the same degree of polarization or the same UP/DOWN ratio per site.
- That does not mean they have the same energy, or the same energy per site.
- That does not mean they have the same entropy or the same multiplicity.
- That does not mean that they have the same heat capacity or the same heat capacity per site.

The list of things that do matter is much shorter: When two subsystems are in thermal equilibrium:

- The microstate probability of one subsystem is determined by the macrostate multiplicity of the other subsystem.
- For a well-behaved heat bath, the change in log multiplicity is proportional to the change in energy. The constant of proportionality is  $\beta$ , the inverse temperature.

These points are particularly clear in the “light blue” scenario (section 23.6). When setting up the problem, we needed to supply “just enough” energy to achieve the desired temperature, i.e. the desired  $T_M$ , i.e. the desired  $\Delta \ln(W)/\Delta E$ . The amount of energy required to do this, 183 units, might not have been obvious *a priori*.

Suppose you have one heat bath in contact with another.

If they start out at different temperatures, energy will flow from one to the other. This will continue until the  $T_M$  of one lines up with the  $\Delta \ln(W)/\Delta E$  of the other.

This depends on  $\ln(W)$  being a convex function of  $E$ . This is not an entirely trivial assumption. For one thing, it means that in two dimensions, a single particle in a box would not be a good heat bath, since its density of states is independent of  $E$ . Multiple particles in a box works fine, even in two dimensions, because the combinatorial factors come to the rescue.

## 23.8 Relevance

Sometimes it is suggested that the discrepancies and limitations discussed in this chapter are irrelevant, because they go away in the large- $N$  limit, and thermodynamics only applies in the large- $N$  limit.

Well, they do go away in the large- $N$  limit, but that does not make them irrelevant. Vast parts of thermodynamics do make sense even for small- $N$  systems. It is therefore important to know which parts we can rely on and which parts break down when stressed. Important small- $N$  applications include reversible computing and quantum computing. Also, the standard textbook derivation of the Boltzmann factor uses a small- $N_S$  argument. If we are going to make such an argument, we ought to do it correctly.



## Chapter 24

# Partition Function

### 24.1 Basic Properties

In thermal equilibrium, the probability of each microstate is proportional to its Boltzmann factor:

$$P_i \propto \exp(-\hat{E}_i/kT) \quad (24.1)$$

where  $P_i$  is the probability of the  $i$ th microstate, and  $\hat{E}_i$  is the energy of the  $i$ th microstate. You can think of the Boltzmann factor  $\exp(-\hat{E}_i/kT)$  as an *unnormalized probability*. In some cases an unnormalized probability is satisfactory, or even desirable, but in other cases you really want the normalized probability, normalized so that  $\sum P_i = 1$ . That is easily arranged:

$$P_i = \frac{\exp(-\hat{E}_i/kT)}{\sum_j \exp(-\hat{E}_j/kT)} \quad (24.2)$$

The normalization denominator in equation 24.2 is something we are going to encounter again and again, so we might as well give it a name. It is called the *partition function* and is denoted  $Z$ . That is:

$$Z := \sum_j \exp(-\hat{E}_j/kT) \quad (24.3)$$

Actually there is more to the story; we shall see that  $Z$  serves in many roles, not just as a normalization denominator. However, that is more than we need to know at the moment. For the time being, it suffices to think of  $Z$  as the normalization denominator. Additional motivation for caring about  $Z$  will accumulate in the next few sections.

Before continuing, we need to take care of some housekeeping details.

We will find it convenient to express some things in terms of inverse temperature. Following convention, we define

$$\beta := \frac{1}{kT} \quad (24.4)$$

The factor of  $k$  means that  $1/\beta$  is measured in units of energy (per particle). This means we don't need to bother with units of temperature; all we need are units of energy.

In this section, we assume constant  $N$ , i.e. constant number of particles. We also assume that the system is fully in equilibrium. That is, this analysis applies only to the Locrian modes, and any non-Locrian modes will have to be handled by other means.

Remark: The partition function is almost universally denoted  $Z$ , which is traceable to the German word *Zustandsumme*, meaning literally “sum over states.” This etymological remark seems somewhat glib because although equation 24.3 truly is a sum over all microstates, there are innumerable other expressions that also take the form of a sum over states. Still, the fact remains that  $Z$  is so important that whenever anybody talks about “the” sum over states, you can assume they mean equation 24.3 or equivalently equation 24.6.

Here are some basic facts about probabilities and Boltzmann factors:

The probability of the  $i$ th state is  $P_i$ .

The Boltzmann factor for state  $i$  is  $\exp(-\beta \hat{E}_i)$ , where  $\hat{E}_i$  is the energy of the state.

The probabilities are normalized such that

$$\sum P_i = 1 \quad (24.5)$$

The sum of the Boltzmann factors is called the partition function:

$$Z := \sum_i e^{-\beta \hat{E}_i} \quad (24.6)$$

Knowing the probability  $P_i$  for every state somewhat useful, but as we shall see, it is not nearly as useful as knowing the Boltzmann factors  $\exp(-\beta \hat{E}_i)$ .

If you know the Boltzmann factors, you can calculate all the probabilities in accordance with equation 24.7, but the converse does not hold: knowing all the probabilities does not suffice to calculate the Boltzmann factors.

In fact, we shall see that if you know the partition function, you can calculate everything there is to know about Locrian thermodynamics.

Among its many uses, the partition function can be used to write:

$$P_i = \frac{\exp(-\beta \hat{E}_i)}{Z} \quad (24.7)$$

## 24.2 Calculations Using the Partition Function

A direct application of basic probability ideas is:

$$\begin{aligned} \langle X \rangle &= \sum_i x_i P_i && \text{for any probability distribution} \\ &= \frac{1}{Z} \sum_i x_i e^{-\beta \hat{E}_i} && \text{for a Boltzmann distribution} \end{aligned} \quad (24.8)$$

where  $\langle \dots \rangle$  denotes the *expectation value* of some property. The idea of expectation value applies to the macrostate. Here  $x_i$  is the value of the  $X$ -property in the  $i$ th microstate. So we see that

equation 24.8 is a weighted average, such that each  $x_i$  is weighted by the probability of state  $i$ . This averaging process relates a macroscopic property  $X$  to the corresponding microscopic property  $x_i$ .

As a sanity check, you should verify that  $\langle 1 \rangle = 1$  by plugging into equation 24.8.

We now begin to explore the real power of the partition function, using it for much more than just a normalization factor.

We can start from the observation that  $Z$ , as defined by equation 24.6, is a perfectly good state function, just as  $P$ ,  $V$ ,  $T$ ,  $S$ , et cetera are state functions. We will soon have more to say about the physical significance of this state function.

We now illustrate what  $Z$  is good for. Here is a justly-famous calculation that starts with  $\ln(Z)$  and differentiates with respect to  $\beta$ :

$$\begin{aligned} \frac{\partial \ln(Z)}{\partial \beta | \{\hat{E}_i\}} &= \frac{1}{Z} \sum_i (-\hat{E}_i) e^{-\beta \hat{E}_i} \\ &= -\langle \hat{E} \rangle \\ &= -E \end{aligned} \tag{24.9}$$

Recall that  $\hat{E}_i$  is the energy of the  $i$ th microstate, while  $E$  is the energy of the macrostate.

Equation 24.9 tells us that one of the directional derivatives of the partition function is related to the energy. For a particle in a box, or for an ideal gas, all the energy levels are determined by the volume of the box, in which case we can write  $E = -\partial \ln(Z) / \partial \beta$  at constant volume.

You have to pay attention to understand what is happening here. How can the macroscopic energy  $\langle E \rangle$  be changing when we require all the  $\hat{E}_i$  to be constant? The answer is that the expectation value  $\langle \dots \rangle$  is a *weighted* average, weighted according to the probability of finding the system in the  $i$ th microstate, and by changing the inverse temperature  $\beta$  we change the weighting.

As another example calculation using the partition function, it is amusing to express the entropy in terms of the partition function. We start with the workhorse expression for entropy, equation 2.2 or equation 9.3, and substitute the probability from equation 24.7.

$$S[P] = -k \sum_i P_i \ln(P_i) \tag{24.10a}$$

$$= -k \sum_i \frac{e^{-\beta \hat{E}_i}}{Z} \ln\left(\frac{e^{-\beta \hat{E}_i}}{Z}\right) \tag{24.10b}$$

$$= -k \sum_i \frac{e^{-\beta \hat{E}_i}}{Z} [-\beta \hat{E}_i - \ln(Z)] \tag{24.10c}$$

$$= k\beta \sum_i \frac{\hat{E}_i e^{-\beta \hat{E}_i}}{Z} + k \ln(Z) \sum_i \frac{e^{-\beta \hat{E}_i}}{Z} \tag{24.10d}$$

$$= k\beta \langle \hat{E} \rangle + k \ln(Z) \langle 1 \rangle \tag{24.10e}$$

$$= k\beta E + k \ln(Z) \tag{24.10f}$$

$$= -k \frac{\partial \ln(Z)}{\partial \ln(\beta) | \{\hat{E}_i\}} + k \ln(Z) \tag{24.10g}$$

We obtained the last line by plugging in the value of  $E$  obtained from equation 24.9. This gives us a handy formula for calculating the entropy directly from the partition function.

Here we have used the fact that  $\langle \ln(Z) \rangle \equiv \ln(Z)$ , as it must be since  $Z$  is not a function of the dummy index  $i$ . Also, in the last line we have used equation 24.9.

The next-to-last line of equation 24.10 tells us that  $E - TS = -kT \ln(Z)$  ... and equation 15.12 tells us that the free energy is  $F := E - TS$ . Combining these expressions yields a surprisingly simple expression for the free energy:

$$F = -kT \ln(Z) \tag{24.11}$$

As an exercise in algebra, you find the entropy in terms of the free energy, namely

$$S[P] = -\frac{\partial F}{\partial T|\{\hat{E}_i\}} \quad (24.12)$$

by carrying out the derivative in equation 24.12 and comparing with equation 24.10.

We have just established a connection between the free energy  $F$ , the temperature  $T$ , and the partition function  $Z$ . If at any point you know two of the three, you can immediately calculate the third.

As another example, consider the case where the microstate energy depends linearly on some parameter  $B$ :

$$\hat{E}_i(B) = \hat{E}_i(0) + B M_i \quad \text{for all } i \quad (24.13)$$

From there, it is straightforward to show that

$$\langle M \rangle = -\frac{1}{\beta} \frac{\partial \ln(Z)}{\partial B} \Big|_{\beta, \{\hat{E}_i(0)\}} \quad (24.14)$$

The notation was chosen to suggest that  $B$  might be an overall applied magnetic field, and  $M_i$  might be the magnetization of the  $i$ th state ... but this interpretation is not mandatory. The idea applies for any parameter that affects the energy linearly as in equation 24.13. Remember Feynman's proverb: the same equations have the same solutions.

### 24.3 Example: Harmonic Oscillator

The partition function  $Z$  is defined in terms of a series, but sometimes it is possible to sum the series analytically to obtain a closed-form expression for  $Z$ . The partition function of a quantum harmonic oscillator is a simple example of this. As discussed in reference 54, it involves a summing a geometric series, which is about as easy as anything could be. The result is

$$Z = 1/2 \operatorname{csch}(1/2\beta\hbar\omega) \quad (24.15)$$

where  $\operatorname{csch}$  is the hyperbolic cosecant, i.e. the reciprocal of the hyperbolic sine.

Using methods described in section 24.2 we can easily find the energy of the harmonic oscillator in thermal equilibrium. The result is given by equation 24.16 and diagrammed in figure 24.1.

$$E = 1/2\hbar\omega \coth(1/2\beta\hbar\omega) \quad (24.16)$$

The entropy of a harmonic oscillator is:

$$\begin{aligned} S &= k\beta E + k \ln(Z) \\ S/k &= 1/2\beta\hbar\omega \coth(1/2\beta\hbar\omega) + \ln[1/2 \operatorname{csch}(1/2\beta\hbar\omega)] \\ &= \beta\hbar\omega \frac{e^{-\beta\hbar\omega}}{1-e^{-\beta\hbar\omega}} - \ln(1 - e^{-\beta\hbar\omega}) \end{aligned} \quad (24.17)$$

In the high temperature limit ( $\beta \rightarrow 0$ ) this reduces to:

$$\begin{aligned} S &= 1 - \ln(\beta\hbar\omega) \\ &= 1 + \ln\left(\frac{kT}{\hbar\omega}\right) \end{aligned} \quad (24.18)$$



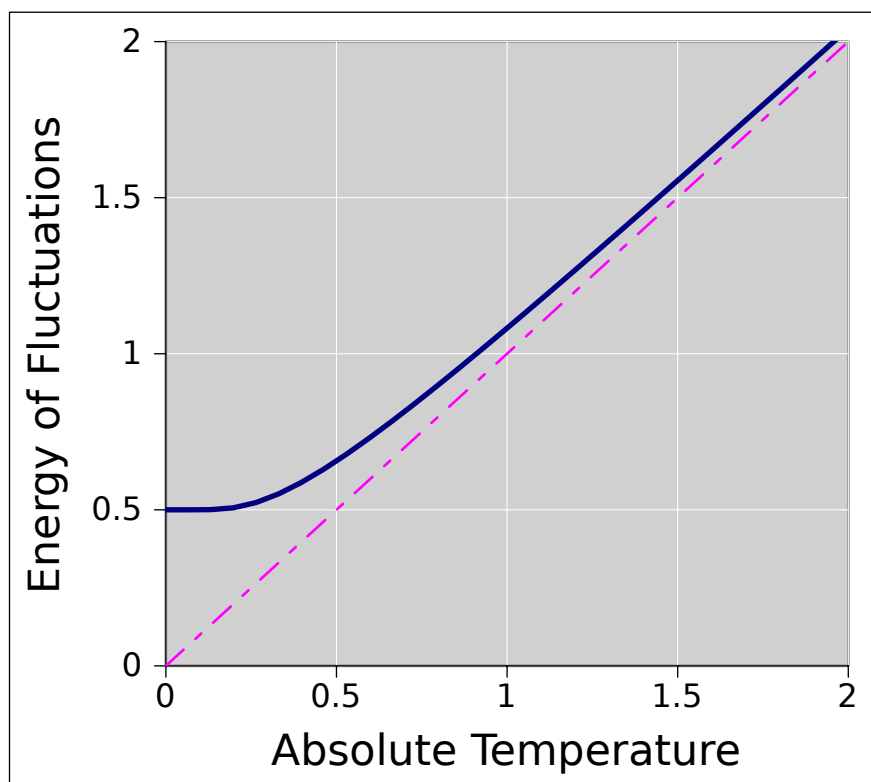


Figure 24.1: Energy vs Temperature for a Harmonic Oscillator

The microstates of a harmonic oscillator are definitely not equally populated, but we remark that the entropy in equation 24.18 is the same as what we would get for a system with  $e kT/\hbar\omega$  equally-populated microstates. In particular it does *not* correspond to a picture where every microstate with energy  $\hat{E} < kT$  is occupied and others are not; the probability is spread out over approximately  $e$  times that many states.

In the low-temperature limit, when  $kT$  is small, the entropy is very very small:

$$S/k = \frac{\hbar\omega}{kT} \exp(-\frac{\hbar\omega}{kT}) \quad (24.19)$$

This is most easily understood by reference to the definition of entropy, as expressed by e.g. equation 2.3. At low temperature, all of the probability is in the ground state, except for a very very small bit of probability in the first excited state.

For details on all this, see reference 54.

## 24.4 Example: Two-State System

Suppose we have a two-state system. Specifically, consider a particle such as an electron or proton, which has two spin states, UP and DOWN, or equivalently  $|\uparrow\rangle$  and  $|\downarrow\rangle$ . Let's apply a magnetic field  $B$ , so that the two states have energy

$$\begin{aligned} \hat{E}(\text{UP}) &= +\mu B \\ \hat{E}(\text{DOWN}) &= -\mu B \end{aligned} \quad (24.20)$$

where  $\mu$  is called the magnetic moment. For a single particle, the partition function is simply:

$$\begin{aligned} Z_1 &= \sum_i e^{-\beta \hat{E}(i)} \\ &= e^{-\beta \mu B} + e^{+\beta \mu B} \\ &= 2 \cosh(\beta \mu B) \end{aligned} \quad (24.21)$$

Next let us consider  $N$  such particles, and assume that they are very weakly interacting, so that when we calculate the energy we can pretend they are non-interacting. Then the overall partition function is

$$Z = Z_1^N \quad (24.22)$$

Using equation 24.9 we find that the energy of this system is

$$\begin{aligned} E &= -\frac{\partial \ln(Z)}{\partial \beta} \\ &= -N\mu B \tanh(\beta \mu B) \end{aligned} \quad (24.23)$$

We can calculate the entropy directly from the workhorse equation, equation 2.2, or from equation 24.10, or from equation 24.12. The latter is perhaps easiest:

$$\begin{aligned} S &= k\beta E + k \ln(Z) \\ &= -Nk\beta \mu B \tanh(\beta \mu B) + Nk \ln(2 \cosh(\beta \mu B)) \end{aligned} \quad (24.24)$$

You can easily verify that at high temperature ( $\beta = 0$ ), this reduces to  $S/N = k \ln(2)$  i.e. one bit per spin, as it should. Meanwhile, at low temperatures ( $\beta \rightarrow \infty$ ), it reduces to  $S = 0$ .

It is interesting to plot the entropy as a function of entropy, as in figure 24.2.

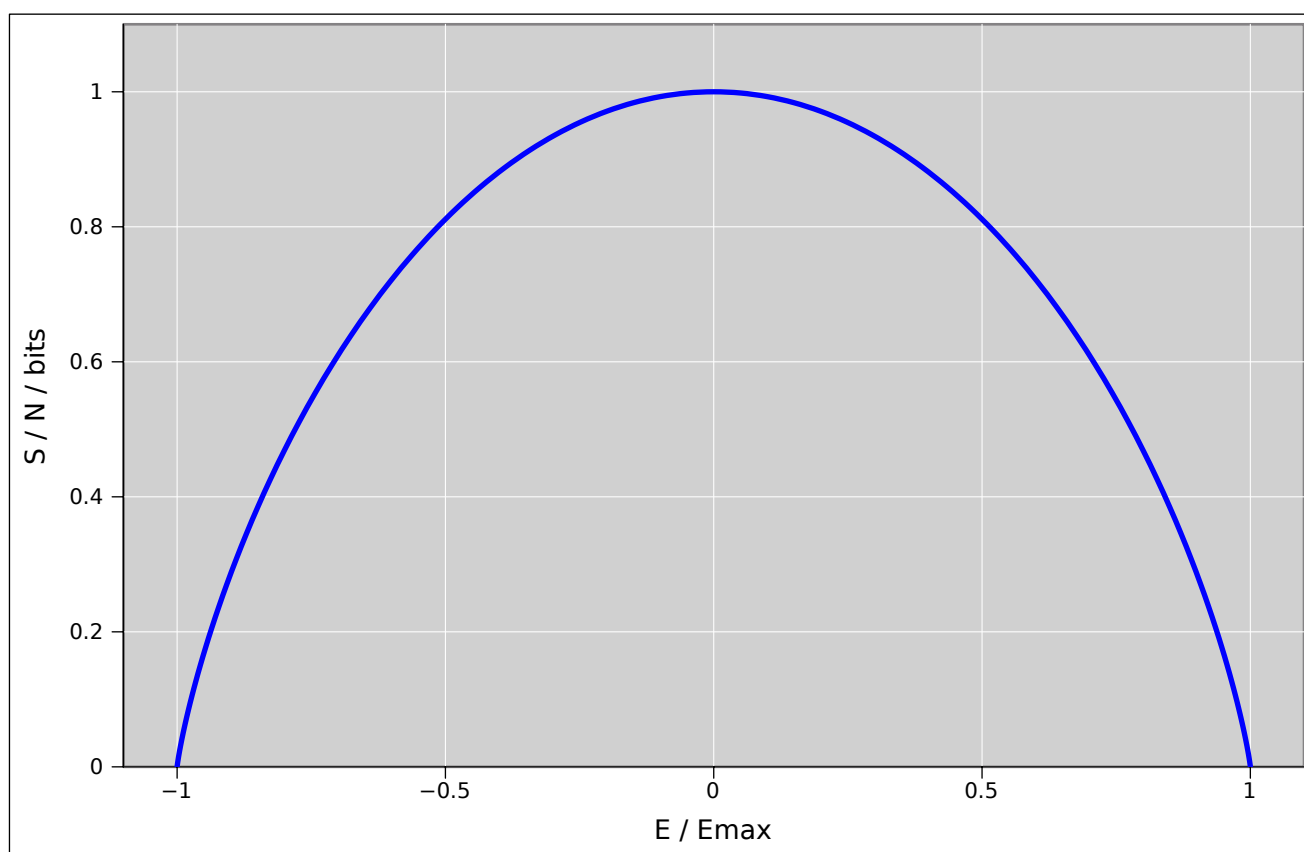


Figure 24.2: Entropy versus Energy – Two State System

In this figure, the slope of the curve is  $\beta$ , i.e. the inverse temperature. It may not be obvious from the figure, but the slope of the curve is infinite at both ends. That is, at the low-energy end the temperature is positive but only slightly above zero, whereas at the high-energy end the temperature is negative but only slightly below zero. Meanwhile, the peak of the curve corresponds to infinite temperature, i.e.  $\beta = 0$ . The temperature is shown in figure 24.3.

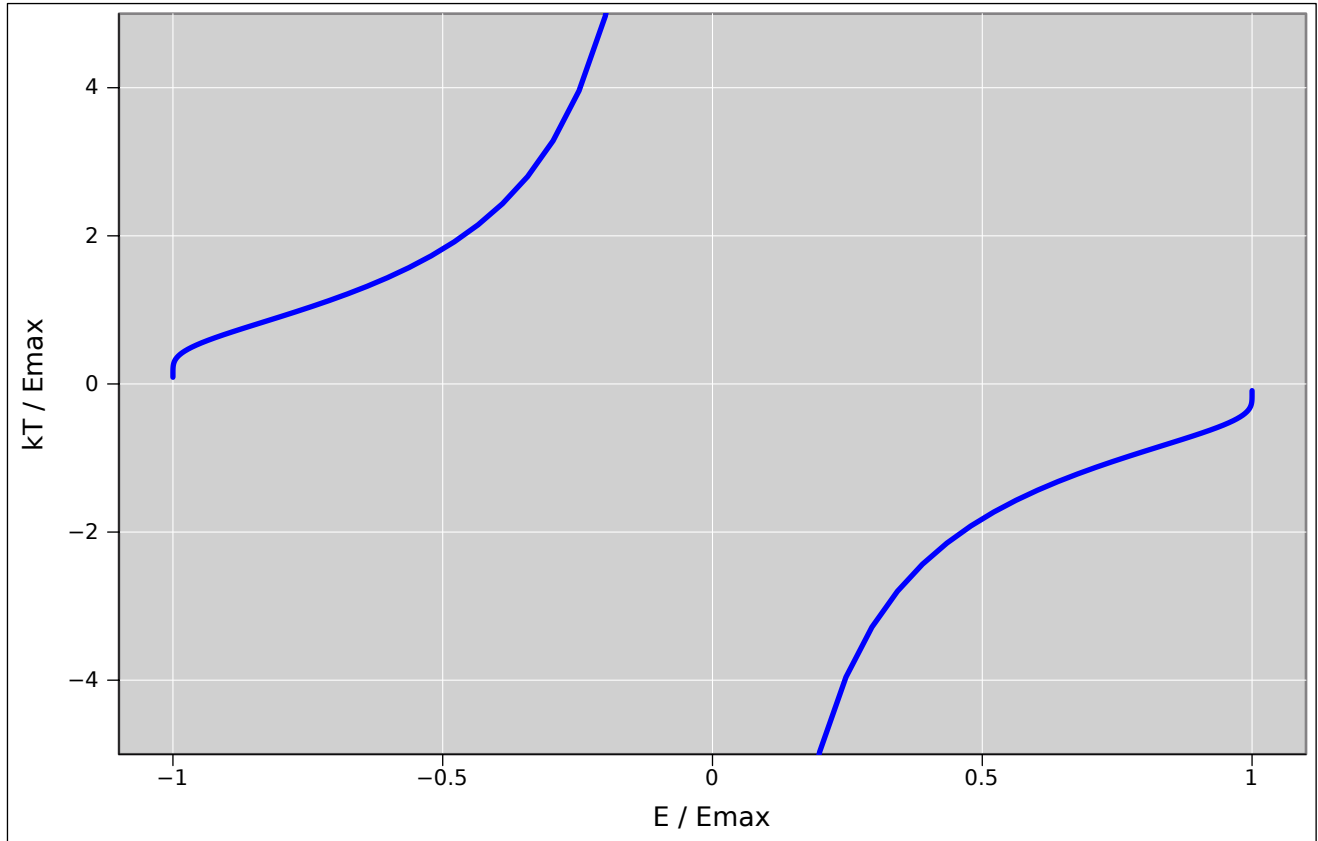


Figure 24.3: Temperature versus Energy – Two State System

In this system, the curve of  $T$  as a function of  $E$  has infinite slope when  $E = E_{\min}$ . You can prove that by considering the inverse function,  $E$  as a function of  $T$ , and expanding to first order in  $T$ . To get a fuller understanding of what is happening in the neighborhood of this point, we can define a new variable  $b := \exp(-\mu B/kT)$  and develop a Taylor series as a function of  $b$ . That gives us

$$\begin{aligned} \frac{E - E_{\min}}{N} &= 2\mu B e^{-2\mu B/kT} && \text{for } T \text{ near zero} \\ kT &= \frac{2\mu B}{\ln(2N\mu B) - \ln(E - E_{\min})} \end{aligned} \quad (24.25)$$

which is what we would expect from basic principles: The energy of the excited state is  $2\mu B$  above the ground state, and the probability of the excited state is given by a Boltzmann factor.

Let us briefly mention the pedestrian notion of “equipartition” (i.e.  $1/2kT$  of energy per degree of freedom, as suggested by equation 25.7). This notion makes absolutely no sense for our spin system. We can understand this as follows: The pedestrian result calls for  $1/2kT$  of energy per *quadratic*

degree of freedom in the *classical* limit, whereas (a) this system is not classical, and (b) it doesn't have any quadratic degrees of freedom.

For more about the advantages and limitations of the idea of equipartition, see chapter 25.

Indeed, one could well ask the opposite question: Given that we are defining temperature via equation 7.7, how could «equipartition» ever work at all? Partly the answer has to do with “the art of the possible.” That is, people learned to apply classical thermodynamics to problems where it worked, and learned to stay away from systems where it didn't work. If you hunt around, you can find systems that are both harmonic and non-quantized, such as the classical ideal gas, the phonon gas in a solid (well below the melting point), and the rigid rotor (in the high temperature limit). Such systems will have  $1/2kT$  of energy in each quadratic degree of freedom. On the other hand, if you get the solid too hot, it becomes anharmonic, and if you get the rotor too cold, it becomes quantized. Furthermore, the two-state system is always anharmonic and always quantized. Bottom line: Sometimes equipartition works, and sometimes it doesn't.

## 24.5 Rescaling the Partition Function

This section is a bit of a digression. Feel free to skip it if you're in a hurry.

We started out by saying that the probability  $P_i$  is “proportional” to the Boltzmann factor  $\exp(-\beta\hat{E}_i)$ .

If  $P_i$  is proportional to one thing, it is proportional to lots of other things. So the question arises, what reason do we have to prefer  $\exp(-\beta\hat{E}_i)$  over other expressions, such as the pseudo-Boltzmann factor  $\alpha \exp(-\beta\hat{E}_i)$ .

We assume the fudge factor  $\alpha$  is the same for every microstate, i.e. for every term in the partition function. That means that the probability  $P_i^\dagger$  we calculate based on the pseudo-Boltzmann factor is the same as what we would calculate based on the regular Boltzmann factor:

$$\begin{aligned} P_i^\dagger &= \frac{\alpha \exp(-\beta\hat{E}_i)}{\sum_j \alpha \exp(-\beta\hat{E}_j)} \\ &= P_i \end{aligned} \tag{24.26}$$

All the microstate probabilities are the same, so anything – such as entropy – that depends directly on microstate probabilities will be the same, whether or not we rescale the Boltzmann factors.

Our next steps depend on whether  $\alpha$  depends on  $\beta$  or not. If  $\alpha$  is a constant, independent of  $\beta$ , then rescaling the Boltzmann factors by a factor of  $\alpha$  has no effect on the entropy, energy, or anything else. You should verify that any factor of  $\alpha$  would drop out of equation 24.9 on the first line.

We now consider the case where  $\alpha$  depends on  $\beta$ . (We are still assuming that  $\alpha$  is the same for every microstate, i.e. independent of  $i$ , but it can depend on  $\beta$ .)

If we were only using  $Z$  as a normalization denominator, having a fudge factor that depends on  $\beta$  would not matter. We could just pull the factor out front in the numerator and denominator of equation 24.26 whereupon it would drop out.

In contrast, if we are interested in derivatives, the derivatives of  $Z' := \beta Z$  are different from the derivatives of plain  $Z$ . You can easily verify this by plugging  $Z'$  into equation 24.9. The  $\beta$ -dependence matters in equation 24.9 even though it doesn't matter in equation 24.10. We summarize this by saying that  $Z$  is not just a normalization factor.

A particularly interesting type of fudge factor is  $\exp(-\beta\phi)$  for some constant  $\phi$ . You can easily verify that this corresponds to shifting all the energies in the problem by  $\phi$ . This can be considered a type of gauge invariance. In situations where relativity is not involved, such as the present situation, you can shift all the energies in the problem by some constant without changing the observable physics. The numerical value of the energy is changed, but this has no observable consequences. In particular, shifting the energy does not shift the entropy.

# Chapter 25

## Equipartition

### 25.1 Generalized Equipartition Theorem

In this chapter we temporarily lower our standards and derive some results that apply only in the classical limit, specifically in the “energy continuum” limit. That is, we assume that the temperature is high compared to the spacing between energy levels, so that when evaluating the partition function we can approximate the sum by an integral. We further assume that the system occupies a bounded region in phase space. That is, we assume there is zero probability that any of the position variables or momentum variables will ever take on super-large values.

Subject to these provisos, we<sup>1</sup> can write the partition function as:

$$\begin{aligned} Z &\equiv \sum_i \exp[-\beta E_i] \\ &\approx \int \int \exp[-\beta E(x, v)] dx dv \end{aligned} \quad (25.1)$$

Here we intend  $x$  and  $v$  to represent, somewhat abstractly, whatever variables contribute to the energy. (We imagine that  $x$  represents the variable or variables that contribute to the potential energy, while  $v$  represents the variable or variables that contribute to the kinetic energy, but this distinction is not important.)

Using 20/20 hindsight, we anticipate that it will be interesting to evaluate the expectation value of  $\partial E / \partial \ln(x)|v$ . We can evaluate this in the usual way, in terms of the partition function:

$$\begin{aligned} \left\langle \frac{\partial E}{\partial \ln(x)|v} \right\rangle &= \sum_i P_i \frac{\partial E_i}{\partial \ln(x)|v} \\ &= \frac{1}{Z} \sum_i \frac{\partial E_i}{\partial \ln(x)|v} \exp[-\beta E_i] \\ &= \frac{1}{Z} \int \int \frac{\partial E}{\partial \ln(x)|v} \exp[-\beta E(x, v)] dx dv \\ &= \frac{-1}{\beta Z} \int \int \frac{\partial \exp[-\beta E(x, v)]}{\partial \ln(x)|v} dx dv \end{aligned} \quad (25.2)$$

We now integrate by parts. The boundary terms vanish, because we have assumed the system occupies a bounded region in phase space.

$$\begin{aligned} \left\langle \frac{\partial E}{\partial \ln(x)|v} \right\rangle &= \sum_i P_i \frac{\partial E_i}{\partial \ln(x)|v} \\ &= \frac{1}{\beta Z} \int \int \exp[-\beta E(x, v)] dx dv \\ &= \frac{1}{\beta} \\ &= kT \end{aligned} \quad (25.3)$$

---

<sup>1</sup>This calculation in this section parallels Appendix G in M.D. Sturge, **Statistical and Thermal Physics** (2003)

We can of course write a corresponding result for the  $v$ -dependence at constant  $x$ :

$$\left\langle \frac{\partial E}{\partial \ln(v)|x} \right\rangle = kT \quad (25.4)$$

and if there are multiple variables  $\{x_i\}$  and  $\{v_i\}$  we can write a corresponding result for each of them. This is called the *generalized equipartition theorem*.

## 25.2 Corollaries: Power-Law Equipartition

An interesting corollary is obtained in the case where the energy contains a power-law term of the form  $E_j = |x|^M$ :

$$\begin{aligned} \left\langle \frac{\partial E_j}{\partial \ln(x)|v} \right\rangle &= kT \\ \langle M E_j \rangle &= kT \\ \langle E_j \rangle &= kT/M \end{aligned} \quad (25.5)$$

In the very common case where  $M = 2$ , this reduces to

$$\langle E_j \rangle = 1/2 kT \quad (25.6)$$

and if the total energy consists of a sum of  $D$  such terms, the total energy is

$$\langle E \rangle = 1/2 D_{\square} kT \quad (25.7)$$

This result is the *quadratic corollary to the equipartition theorem*. The symbol  $D_{\square}$  (pronounced “ $D$  quad”) is the number of quadratic degrees of freedom of the system. Here we are assuming every degree of freedom appears in the Hamiltonian either quadratically or not at all.

All too often, equation 25.7 is called simply “the” equipartition theorem. However, we should beware that it is only a pedestrian corollary. It only applies when every degree of freedom contributes to the energy either quadratically or not at all. This includes a remarkably large number of situations, including the harmonic oscillator, the particle in a box, and the rigid rotor ... but certainly not all situations, as discussed in section 25.3.

Also keep in mind that all the results in this chapter are based on the assumption that the system is classical, so that in the definition of the partition function we can approximate the sum by an integral.

## 25.3 Harmonic Oscillator, Particle in a Box, and Other Potentials

Let us consider a particle moving in one dimension in a power-law potential well. The energy is therefore

$$E = v^2 + |x|^M \quad (25.8)$$

where the first term represents the usual kinetic energy (in some units) and the second term represents the potential energy. The case  $M = 2$  corresponds to a harmonic oscillator, as shown in figure 25.2.

As  $M$  becomes larger and larger, the situation begins to more and more closely resemble a square-well potential, i.e. a particle in a box, as you can see in figure 25.3 and figure 25.4.



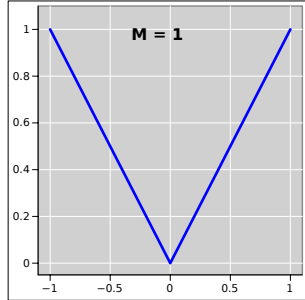


Figure 25.1:  $M = 1$  i.e. Linear Power-Law Potential Well

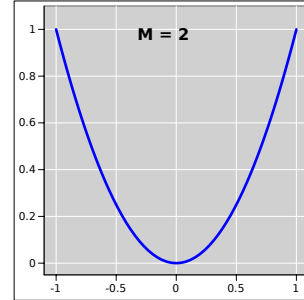


Figure 25.2:  $M = 2$  i.e. Quadratic Power-Law Potential Well

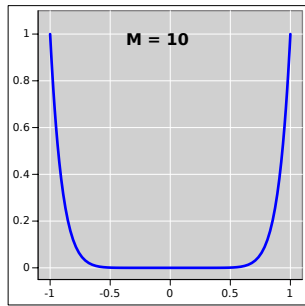


Figure 25.3:  $M = 10$  Power-Law Potential Well

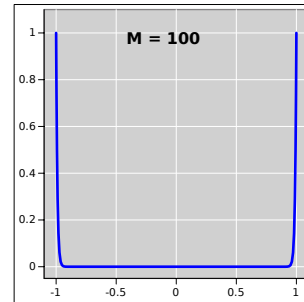


Figure 25.4:  $M = 100$  Power-Law Potential Well

Let us apply the generalized equipartition theorem, namely equation 25.3 and equation 25.4, to each of these situations.

$M$	$: \langle E \rangle / kT$		
1	$: 0.5 + 1.0$	$= 1.5$	gravitational field in 1D
1	$: 1.5 + 1.0$	$= 2.5$	gravitational field in 3D
2	$: 0.5 + 0.5$	$= 1$	harmonic oscillator
10	$: 0.5 + 0.1$	$= 0.6$	
100	$: 0.5 + 0.01$	$= 0.51$	
$\infty$	$: 0.5 + 0$	$= 0.5$	square well

(25.9)

## 25.4 Remarks

It must be emphasized that every degree of freedom counts as a degree of freedom, whether or not it contributes to the heat capacity. For example, consider a free particle in three dimensions:

There are six degrees of freedom.

There are only three quadratic degrees of freedom.

All six of the degrees of freedom contribute to the entropy. This should be obvious from the fact that the phase space of such a particle is six dimensional.

Only three of the degrees of freedom contribute to the heat capacity.

- For a one-dimensional harmonic oscillator, there are two quadratic degrees of freedom. Both contribute to the entropy, and both contribute to the energy.
- For the square well, there is only one quadratic degree of freedom, namely the one that gives rise to the kinetic energy. In contrast, the position is a non-quadratic degree of freedom, and contributes nothing to the average energy; there is no relevant potential energy.
- In the general case, you are not allowed to simply ignore non-quadratic degrees of freedom. In the example defined by equation 25.8, varying the exponent  $M$  from  $M = 2$  on up allows us to interpolate between the harmonic oscillator and the square well. At one extreme ( $M = 2$ ) the potential energy is quadratic, and at the other extreme ( $M = \infty$ ) the potential energy is irrelevant ... but in between, it is both relevant and non-quadratic.
- The  $M = 1$  case is easy to analyze, interesting, and reasonably common. For example, it applies to the  $m|g|h$  term in a tall column of gas in a gravitational field, as shown in figure 9.6.
- For any degree of freedom where the spacing between energy levels is *large* compared to  $kT$ , that degree of freedom contributes nothing very interesting to the average energy. (This is the polar opposite of the “energy continuum” limit.) We say that such a degree of freedom is “frozen out” by the laws of quantum mechanics.
- The idea of equipartition fails miserably for the spin system discussed in section 24.4, because the system does not have any quadratic degrees of freedom, and more importantly because the “energy continuum” approximation does not apply.
- Pedestrian ideas of equipartition can be applied to the ratio of specific heats of common gases, but only very imperfectly, as shown in table 26.1. This is because the “energy continuum” approximation applies to some of the degrees of freedom but not others. That is, some of the degrees of freedom have a moderately large spacing, so that they are quantum-mechanically “frozen out” at some temperatures but not others. For additional discussion of this point, see section 26.7.

## Chapter 26

# Partition Function: Some Examples

In this chapter, we consider the partition function for various interesting systems. We start with a single particle in a box (section 26.1). We then consider an ideal gas of  $N$  particles in a box (section 26.2), including a pure monatomic gas and mixtures of monatomic gases. We then consider the rigid rotor, as a model of a lightweight diatomic model (section 26.3). We use the partition function to derive some classic macroscopic thermodynamic formulas.

### 26.1 Preview: Single Particle in a Box

As discussed in section 26.9, the canonical partition function for a single high-temperature nonrelativistic pointlike particle in a box is:

$$Z_{\text{ppb}} = \frac{V}{\Lambda^3} \quad (26.1)$$

where  $V$  is the volume of the container. The subscript “ppb” stands for “point particle in a box”. The RHS is temperature-dependent because  $\Lambda$  scales like  $\sqrt{\beta}$ . Here  $\Lambda$  is the thermal de Broglie length

$$\Lambda := \sqrt{\left(\frac{2\pi\hbar^2}{mkT}\right)} \quad (26.2)$$

which is the same as equation 12.2.

In general, the partition function is defined in terms of an infinite series, but in many cases it is possible to sum the series. In this case the result is a compact, closed-form expression, namely equation 26.1.

*Using* partition functions is more fun than deriving them, so let’s start by doing some examples using equation 26.1, and postpone the derivation to section 26.9.

There’s a lot more we could say about this, but it’s easier to do the more-general case of the ideal gas (section 26.2) and treat the single particle in a box as a special case thereof, i.e. the  $N = 1$  case.

## 26.2 Ideal Gas of Point Particles

In this section, we generalize the single-particle partition function to a gas of multiple particles. We continue to assume all the particles are pointlike. At not-to-high temperatures, this is a good model for a monatomic gas.

### 26.2.1 Distinguishable Particles

We start by considering a single gas particle, and model it as a six-sided die. Since there are six possible states, the single-particle partition function has six terms:

$$Q = Q_1 + Q_2 + Q_3 + Q_4 + Q_5 + Q_6 \quad (26.3)$$

We do not assume the die is fair, so the terms are not necessarily equal.

We now proceed to the case of  $N = 2$  dice. The partition function will have  $6^2 = 36$  terms. We can calculate the probability of each two-particle state in terms of the corresponding one-particle states. In fact, since the gas is ideal, each particle is independent of the other, so the single-particle probabilities are statistically independent, so each two-particle probability is just a simple product, as shown in the following equation:

$$\begin{aligned} Z = & Q_1 R_1 + Q_1 R_2 + Q_1 R_3 + Q_1 R_4 + Q_1 R_5 + Q_1 R_6 \\ & + Q_2 R_1 + Q_2 R_2 + Q_2 R_3 + Q_2 R_4 + Q_2 R_5 + Q_2 R_6 \\ & + Q_3 R_1 + Q_3 R_2 + Q_3 R_3 + Q_3 R_4 + Q_3 R_5 + Q_3 R_6 \\ & + Q_4 R_1 + Q_4 R_2 + Q_4 R_3 + Q_4 R_4 + Q_4 R_5 + Q_4 R_6 \\ & + Q_5 R_1 + Q_5 R_2 + Q_5 R_3 + Q_5 R_4 + Q_5 R_5 + Q_5 R_6 \\ & + Q_6 R_1 + Q_6 R_2 + Q_6 R_3 + Q_6 R_4 + Q_6 R_5 + Q_6 R_6 \end{aligned} \quad (26.4)$$

where  $Q$  is the partition function for the first die,  $R$  is the partition function for the second die, and subscripts on  $Q$  and  $R$  identify terms within each single-particle partition function.

Using the distributive rule, we can simplify equation 26.4 quite a bit: we find simply  $Z = QR$ . If we now assume that the two dice are statistically the same (to a good approximation), then we can further simplify this to  $Z = Q^2$  (to the same good approximation). In the case of  $N$  dice, the general result is:

$$Z_{\text{dice}|D} = \left(\frac{V}{\Lambda^3}\right)^N \quad (26.5)$$

This is the correct result for the case where each particle has (almost) the same single-particle partition function, provided the particles remain enough different to be distinguishable in the quantum-mechanical sense ... and provided we know in advance which particles are going to appear in the mixture, so that there is no entropy of the deal (as defined in section 12.10).

## 26.2.2 Indistinguishable Particles; Delabeling

We now consider a situation that is the same as above, except that the particles are all identical in the quantum-mechanical sense. (We will consider mixtures in section 26.2.3.) If we apply equation 26.4 to this situation, every term below the diagonal has a mirror-image term above the diagonal that describes exactly the same state. For example,  $Q_2 R_6$  is the exact same state as  $R_6 Q_2$ . Therefore if we include all the terms in equation 26.4 in our sum over states, we have overcounted these off-diagonal states by a factor of two. So to a first approximation, neglecting the diagonal terms, we can correct for this by writing  $Z = QR/2$ . Generalizing from  $N = 2$  to general  $N$ , we have:

$$Z = \frac{1}{N!} \left( \frac{V}{\Lambda^3} \right)^N \quad (\text{nondegenerate pure gas}) \quad (26.6)$$

At the next level of detail, we should think a little more carefully about the diagonal terms in equation 26.4. There are two possibilities:

- For Fermi particles, such states do not exist, and we should not count the diagonal terms at all. (This is one manifestation of the Pauli exclusion principle.) Alas equation 26.6 overcounts these diagonal states; it counts them even though it shouldn't.
- For Bose particles, we should count all the diagonal terms. They should *not* be divided by the delabeling factor ( $N!$ ) because each state is correctly represented by exactly one term in equation 26.4. Alas equation 26.6 undercounts these diagonal states; it discounts them by a factor of  $N!$  even though it shouldn't.

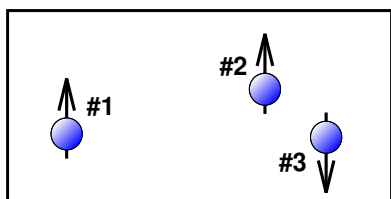
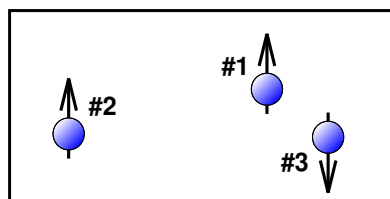
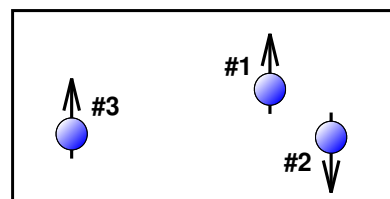
On the other hand, by hypothesis we are restricting attention to nondegenerate gases; therefore the chance of any particular slot being occupied is small, and the chance of any particular slot being occupied more than once is small squared, or smaller. That means there must be many, many terms in the sum over states, and the diagonal terms must be a small fraction of the total. Therefore we don't much care what we do with the diagonal terms. We could keep them all, discard them all, or whatever; it doesn't much matter. As  $N$  becomes larger, the diagonal terms become even less of a problem. The simplest thing is to arbitrarily use equation 26.6, which will be slightly too high for fermions and slightly too low for bosons, but close enough for most applications, provided the system really is nondegenerate.

## 26.2.3 Mixtures

Things get even more interesting when we consider mixtures.

Figure 26.1 shows three particles that are very very nearly identical. In particular, this could represent three atoms of  $^3\text{He}$ , where arrows in the diagram represent the nuclear spins. There are six ways of labeling these three particles, of which three ways are shown in the figure.

In fact diagram *A* and diagram *B* depict exactly the same state. Particle #1 and particle #2 are quantum-mechanically identical, so the physics doesn't care which is which. That is, these two labels are different only in our imagination.

**A****B****C**

Tangential technical remark: In a snapshot such as we see in figure 26.1, it could be argued that particle #1 and particle #2 are distinguishable by their positions, even though they are “identical” particles. Note that we are using some subtle terminology: identical is not the same as indistinguishable. The idea of “distinguishable by position” is valid at not-too-low temperatures, where the thermal de Broglie length is small compared to the spacing between particles. However, even in the best case it is only valid temporarily, because of the uncertainty principle: The more precisely we know the position at the present time, the less precisely we will know it at future times. For a minimum-uncertainty wave packet, the formula is

$$\Delta t = \frac{2m\Delta x^2}{\hbar} \quad (26.7)$$

which for atoms at ordinary atmospheric density is on the order of a nanosecond. This is much shorter than the timescale for carrying out a typical Gibbs-type mixing experiment, so under these conditions we should not think of the atoms as being distinguishable by position. (For a very dilute gas of massive particles on very fast timescales, the story might be different.)

The notion of “distinguishable by position” is sometimes important. Without it, classical physics would not exist. In particular, we would never be able to talk about an individual electron; we would be required to antisymmetrize the wavefunction with respect to every electron in the universe.

When we evaluate the partition function, each state needs to appear once and only once in the sum over states. In figure 26.1, we need to include  $A$  or  $B$  but not both, since they are two ways of describing the exact same state.

As is so often the case, we may find it convenient to count the states *as if* the particles were labeled, and then apply the appropriate delabeling factors.

For the gas shown in figure 26.1, we have  $N = 3$ . The delabeling factor will be  $2!$  for the spin-up component of the mixture, because there are 2 particles in this component. The delabeling factor will be  $1!$  for the spin-down component, because there is only one particle in this component. For general  $N$ , if we assume that all the particles have the same single-particle partition function – which is what happens when the Hamiltonian is spin-independent – then the partition function for the gas as a whole is

$$\begin{aligned} Z_{\text{mix}|D} &= \prod_j \frac{1}{N_j!} \left(\frac{V}{\Lambda^3}\right)^{N_j} \\ &= \left(\frac{V}{\Lambda^3}\right)^N \prod_j \frac{1}{N_j!} \end{aligned} \quad (26.8)$$

where the index  $j$  runs over all components in the mixture. To derive the second line we have used the obvious sum rule for the total number of particles:

$$\sum_j N_j = N \quad (26.9)$$

The last line of equation 26.8 is very similar to equation 26.6, in the sense that both contain a factor of  $(V/\Lambda^3)$  to the  $N$ th power. Only the delabeling factors are different.

### 26.2.4 Energy, Heat Capacity, and Entropy for a Pure Gas

We can use equation 26.6 to find the energy for the pure gas, with the help of equation 24.9. Plugging in, we find

$$\begin{aligned} E &= -\frac{\partial \ln(Z)}{\partial \beta | \{\bar{E}_i\}} \\ &= \frac{3}{2} \frac{N}{\beta} \\ &= \frac{3}{2} N k T \end{aligned} \tag{26.10}$$

as expected for the ideal monatomic nondegenerate nonrelativistic pure gas in three dimensions. (See equation 26.42 for a more general expression, applicable to a polytropic gas. See section 26.2.3 for a discussion of mixtures, including mixtures of isotopes, mixtures of spin-states, et cetera.)

If you're clever, you can do this calculation in your head, because the RHS of equation 26.6 depends on  $\beta$  to the  $-3N/2$  power, and all the other factors drop out when you take the logarithmic derivative.

Note that equation 26.10 naturally expresses the energy as a function of temperature, in contrast to (say) equation 7.8 which treats the energy as a function of entropy. There is nothing wrong with either way of doing it. Indeed, it is best to think in topological terms, i.e. to think of energy at each point in thermodynamic state-space. We can describe this point in terms of its temperature, or in terms of its entropy, or in innumerable other ways.

This expression for the energy is independent of  $V$ . On the other hand, we are free to treat it a function of  $V$  (as well as  $T$ ). We could multiply the RHS by  $V^0$ , which is, formally speaking, a function of  $V$ , even though it isn't a very interesting function. We mention this because we want to take the partial derivative along a contour of constant  $V$ , to find the heat capacity in accordance with equation 7.13.

$$\begin{aligned} C_V &:= \left. \frac{\partial E}{\partial T} \right|_V && \text{always} \\ &= \frac{3}{2} N k && \text{for our gas} \\ c_V &= \frac{3}{2} R \end{aligned} \tag{26.11}$$

where  $C_V$  with a capital  $C$  denotes the extensive heat capacity, while  $c_V$  with a small  $c$  denotes the molar heat capacity. Here  $R$  is the universal gas constant,  $R = N_A k$ , where  $N_A$  is Avogadro's number (aka Loschmidt's number).

Recall that our gas is a monatomic tabletop nondegenerate nonrelativistic ideal gas in three dimensions.

It is also worthwhile to calculate the heat capacity at constant pressure. Using the definition of enthalpy (equation 15.4) and the ideal gas law (equation 26.40) we can write  $H = E + PV =$



$E + N k T$  and plug that into the definition of  $C_P$ :

$$\begin{aligned}
 C_P &:= \left. \frac{\partial H}{\partial T} \right|_P && \text{always} \\
 &= \left. \frac{\partial E}{\partial T} \right|_P + N k && \text{for our gas} \\
 &= \frac{3}{2} N k + N k \\
 &= \frac{5}{2} N k \\
 c_P &= \frac{5}{2} R
 \end{aligned} \tag{26.12}$$

See the discussion leading up to equation 26.48 and equation 26.49 for a more general expression.

Let's calculate the entropy. We start with equation 24.10f, plug in equation 26.10 for the energy, and then plug in equation 26.6 for the partition function. That gives us

$$\begin{aligned}
 S_{\text{pure}} &= k \beta E + k \ln(Z) \\
 &= \frac{3}{2} k N + k \ln\left(\frac{1}{N!} \frac{V^N}{\Lambda^{3N}}\right)
 \end{aligned} \tag{26.13}$$

As discussed in reference 55 and references therein, the first Stirling approximation for factorials is:

$$\ln(N!) \approx (N + 1/2) \ln(N) - N + 1/2 \ln(2\pi) + (\text{terms of order } 1/N) \tag{26.14}$$

Plugging that into equation 26.13 gives us:

$$\begin{aligned}
 S &= \frac{3}{2} k N + k \ln\left(\frac{1}{N!} \frac{V^N}{\Lambda^{3N}}\right) \\
 &= N k \frac{3}{2} - k[N \ln(N) - N + 1/2 \ln(2\pi)] + N k \ln\left(\frac{V}{\Lambda^3}\right)
 \end{aligned} \tag{26.15}$$

We can make this easier to understand if we write the molar entropy  $S/N$  in terms of the molar volume  $V/N$ , which is logical since we expect both  $S$  and  $V$  to more-or-less scale like  $N$ . That gives us:

$$\frac{S/N}{k} = \ln\left(\frac{V/N}{\Lambda^3}\right) + \frac{5}{2} - \frac{1/2 \ln(N)}{N} - \frac{1/2 \ln(2\pi)}{N} \tag{26.16}$$

For large enough  $N$  we can ignore the last two terms on the RHS, which gives us the celebrated Sackur-Tetrode formula:

$$\frac{S/N}{k} = \ln\left(\frac{V/N}{\Lambda^3}\right) + \frac{5}{2} \tag{26.17}$$

This expresses the molar entropy  $S/N$  in terms of the molar volume  $V/N$  and the thermal de Broglie length  $\Lambda$ . Note that the RHS depends on temperature via  $\Lambda$ , in accordance with equation 26.2. The temperature dependence is shown explicitly in equation 26.18:

$$\frac{S/N}{k} = \ln\left(\frac{V}{N}\right) + \frac{3}{2} \ln(T) + \text{constants} \tag{26.18}$$

Note that validity of all these results is restricted to monatomic gases. It is also restricted to nondegenerate gases, which requires the molar volume to be large. Specifically,  $(V/N)/\Lambda^3$  must be large compared to 1. As an additional restriction, equation 26.16 requires  $N$  to be somewhat large compared to 1, so that the first Stirling approximation can be used. For systems with more than a

few particles this is not much of a restriction, since the first Stirling approximation is good to better than 1% when  $N = 4$  and gets better from there. (The second Stirling approximation is good to 144 parts per million even at  $N = 2$ , and gets rapidly better from there, so you can use that if you ever care about small- $N$  systems. And for that matter, equation 26.13 is valid for all  $N$  whatsoever, from  $N = 0$  on up.) In contrast, equation 26.17 requires that  $N$  be very large compared to 1, since that is the only way we can justify throwing away the last two terms in equation 26.16.

For a more general formula, see equation 26.58.

Before we go on, it is worth noting that equation 26.16 is more accurate than equation 26.17. In many thermodynamic situations, it is safe to assume that  $N$  is very large ... but we are not *required* to assume that if we don't want to. The basic laws of thermodynamics apply just fine to systems with only a few particles.

This is interesting because equation 26.16 tells us that the entropy  $S$  is not really an extensive quantity. If you increase  $V$  in proportion to  $N$  while keeping the temperature constant,  $S$  does not increase in equal proportion. This is because of the last two terms in equation 26.16. These terms have  $N$  in the denominator without any corresponding extensive quantity in the numerator.

When we have  $10^{23}$  particles, these non-extensive terms are utterly negligible, but when we have only a few particles, that's a different story.

This should not come as any big surprise. The energy of a liquid or solid is not really extensive either, because of things like surface tension and surface reconstruction. For more about non-extensive entropy, see section 12.8 and especially section 12.11.

### 26.2.5 Entropy of a Mixture

Equation 26.13 assumed a pure gas. We rewrite it here for convenience:

$$\begin{aligned} S_{\text{pure}} &= k\beta E + k \ln(Z) \\ &= \frac{3}{2} k N + k \ln\left(\frac{1}{N!} \frac{V^N}{\Lambda^{3N}}\right) \\ &= \frac{3}{2} k N + k \ln\left(\frac{V^N}{\Lambda^{3N}}\right) + k \ln\left(\frac{1}{N!}\right) \end{aligned} \quad (26.19)$$

We can easily find the corresponding formula for a mixture, using the same methods as section 26.2.4, except that we start from equation 26.8 (instead of equation 26.6). That gives us:

$$\begin{aligned} S_{\text{mix}|D} &= k\beta E + k \ln(Z) \\ &= \frac{3}{2} k N + k \ln\left(\frac{V^N}{\Lambda^{3N}}\right) + k \ln\left(\frac{1}{\prod_j N_j!}\right) \\ &= S_{\text{pure}} + k \ln\left(\frac{N!}{\prod_j N_j!}\right) \end{aligned} \quad (26.20)$$

The subscript “ $|D$ ” is a reminder that this is the conditional entropy, conditioned on the deal, i.e. conditioned on knowing in advance WHICH particles are in the mixture, i.e. not including the entropy of the deal.

The last term on the RHS of equation 26.20 is commonly called the *spin entropy*. More generally, the particles could be labeled by lots of things, not just spin. Commonly we find a natural mixture of isotopes, and artificial mixtures are also useful. Air is a mixture of different chemical species. Therefore the last term on the RHS should be called the *label entropy* or something like that. In

all cases, you can think of this term as representing the entropy of mixing the various components in the mixture. The argument to the logarithm – namely  $N!/\prod_j N_j!$  – is just the number of ways you could add physically-meaningful labels to a previously-unlabeled gas.

When we speak of physically-meaningful labels, the laws of physics take an expansive, inclusive view of what is a meaningful difference. Atomic number  $Z$  is relevant: oxygen is different from nitrogen. Baryon number  $A$  is relevant:  $^{14}\text{C}$  is different from  $^{12}\text{C}$ . Electron spin is relevant. Nuclear spin is relevant. Molecular rotational and vibrational excitations are relevant. Nuclear excited states are relevant (while they last). All these things contribute to the entropy. *Sometimes* they are just spectator entropy and can be ignored for *some* purposes, but sometimes not.

It is remarkable that if we carry out any process that does not change the labels, the entropy of mixing is just an additive constant. For example, in a heat-capacity experiment, the spin entropy would be considered “spectator entropy” and would not affect the result at all, since heat capacity depends only on derivatives of the entropy (in accordance with equation 7.24).

To say the same thing another way: Very often people use the expression for a pure one-component gas, equation 26.13, even when they shouldn’t, but they get away with it (for *some* purposes) because it is only off by an additive constant.

Beware: In some books, including all-too-many chemistry books, it is fashionable to ignore the spin entropy. Sometimes they go so far as to redefine “the” entropy so as to not include this contribution ... which is a very bad idea. It’s true that you can ignore the spin entropy under *some* conditions ... but not all conditions. For example, in the case of spin-aligned hydrogen, if you want it to form a superfluid, the superfluid phase necessarily contains no entropy whatsoever, including nuclear spin entropy. Spin entropy and other types of label entropy are entirely real, even if some people choose not to pay attention to them.

As a point of terminology: On the RHS of equation 26.20 some people choose to associate  $S_{\text{pure}}$  with the “external” coordinates of the gas particles, and associate the spin-entropy with the “internal” coordinates (i.e. spin state).

If we use the first Stirling approximation, we find that the molar entropy is given by

$$\frac{S/N}{k} = \ln\left(\frac{V/N}{\Lambda^3}\right) + \frac{5}{2} + \sum_j x_j \ln(1/x_j) \quad (26.21)$$

which can be compared with the Sackur-Tetrode formula for the pure gas, equation 26.17. We see that once again the entropy is “almost” extensive. There is however an extra constant term, representing the entropy of mixing. Here  $x_j$  is the mole fraction of the  $j$ th component of the mixture, i.e.

$$x_j := N_j/N \quad (26.22)$$

### 26.2.6 Extreme Mixtures

We now consider the extreme case where all of the gas particles are different. I call this “snow”; based on the proverbial notion that no two snowflakes are alike.

$$\begin{aligned} S_{\text{snow}|D} &= k\beta E + k \ln(Z) \\ &= \frac{3}{2} k N + k \ln\left(\frac{V^N}{\Lambda^{3N}}\right) \end{aligned} \quad (26.23)$$

If we rearrange this equation to put molar entropy on the LHS, we get:

$$\frac{S_{\text{snow}|D}/N}{k} = \frac{3}{2} + \ln\left(\frac{V}{\Lambda^3}\right) \quad (26.24)$$

where the RHS is blatantly not intensive (because it depends directly on  $V$ ) ... in contrast to the Sackur-Tetrode formula (equation 26.17) which is “almost” intensive in the sense that the RHS depends on the intensive molar volume  $V/N$  (rather than the extensive volume  $V$ ).

To make sense of this result, consider the following experiment, which can be considered a version of the Gibbs experiment. Start with a box with a partition. Place some snow to the left of the partition, and some more snow to the right of the partition. We now make the dubious assumption that we can tell the difference. If you want, imagine blue snow on one side and red snow on the other. This is not the usual case, and we would not expect random dealing to deal all the blue snow to one side and all the red snow to the other. On the other hand, we could engineer such an arrangement if we wanted to.

In other words, equation 26.24 may be slightly misleading, insofar as we are neglecting the entropy of the deal. This stands in contrast to equation 26.28, which may be more relevant to the usual situation.

When we pull out the partition, the snow on one side mixes with the snow on the other side, increasing the entropy. It’s all snow, but the combined sample of snow is more mixed than the two original samples, and has greater entropy. For more about non-extensive entropy, see section 12.8 and especially section 12.11.

Just to be clear: When talking about whether “the” entropy is extensive, I am assuming we measure “the” entropy long after the partition has been pulled out, after things have settled down ... and compare it to the situation before the partition was pulled out. Without this assumption, the whole question would be hopelessly ill-defined, because the entropy is changing over time.

### 26.2.7 Entropy of the Deal

Let’s now consider the entropy of the deal. As a simple example, consider two almost-identical scenarios involving a deck of cards.

Start with a deck in a known state. The entropy is zero.

Start with a deck in a known state. The entropy is zero.

Shuffle the deck. The entropy is now 226 bits.

Divide the deck in half, forming two hands of 26 cards apiece. The entropy is still zero.	Divide the deck in half, forming two hands of 26 cards apiece. The entropy is still 226 bits. The entropy of each hand separately is about $88\frac{1}{2}$ bits, so we have a total 177 bits for the entropy of permutation <i>within each hand</i> , plus another 49 bits for the entropy of the deal, i.e. not knowing which cards got dealt into which hand.
Shuffle each hand separately. The entropy goes up to $88\frac{1}{2}$ bits per hand, giving a total of 177 bits.	Shuffle each hand separately. The entropy remains 226 bits.
Put the two hands together and shuffle them together. The entropy goes up from 177 to 226 bits.	Put the two hands together and shuffle them together. The entropy remains 226 bits.

Note: This is analogous to pulling out the partition in a Gibbs experiment, allowing the gases to mix.

Divide the deck into hands again, with 26 cards per hand. The entropy is still 226 bits.	Divide the deck into hands again, with 26 cards per hand. The entropy is still 226 bits.
--	--

Note: This is analogous to re-inserting the partition in a Gibbs experiment. Re-insertion leaves the entropy unchanged. For distinguishable particles, the entropy includes a large contribution from the entropy of the deal.

Peek at the cards. This zeros the entropy, including the entropy of the deal.	Don't peek. The entropy remains 226 bits.
Shuffle each hand separately. The entropy goes back up to 177 bits.	Shuffle each hand separately. The entropy is unchanged, namely 226 bits.

We can apply the same idea – the entropy of the deal – to other systems. Let's do it more formally, with algebraic symbols rather than raw numbers. To keep the number of terms manageable, let's use a two-state system (coins) rather than dice, playing cards, or snowflakes.

Let's extend equation 26.4 to cover the case where there is a large set (the universal set) of  $M$  distinguishable particles, and our sample of  $N$  particles is drawn from that. We choose the case of  $M = 3$  (penny, nickel, and dime) with single-particle partition functions  $Q$ ,  $R$ , and  $S$  respectively.

We choose  $N = 2$  of them and toss them randomly.

$$\begin{aligned}
 Z &= Q_1 R_1 + Q_1 R_2 \\
 &+ Q_2 R_1 + Q_2 R_2 \\
 &+ Q_1 S_1 + Q_1 S_2 \\
 &+ Q_2 S_1 + Q_2 S_2 \\
 &+ R_1 S_1 + R_1 S_2 \\
 &+ R_2 S_1 + R_2 S_2 \\
 &= QR + QS + RS
 \end{aligned} \tag{26.25}$$

We interpret this as three copies of the “generic”  $N$ -particle partition function that we would get if we knew in advance WHICH  $N$  of the  $M$  possible particles would appear in our sample.

Assuming all the particles behave very nearly the same, and generalizing from  $M = 3, N = 2$  to general  $M$  and  $N$ , we obtain:

$$\begin{aligned}
 Z_{\text{snow}} &= \binom{M}{N} \left(\frac{V}{\Lambda^3}\right)^N \\
 &= \frac{M!}{(M-N)!N!} \left(\frac{V}{\Lambda^3}\right)^N
 \end{aligned} \tag{26.26}$$

where the first factor on the RHS is the binomial coefficient, pronounced “ $M$  choose  $N$ ”

Equation 26.26 can be seen as an extension of equation 26.5 and/or equation 26.8, extended to account for the entropy of the deal. For large  $M$ , large compared to  $N$ ,

$$Z_{\text{snow}} \approx \frac{M^N}{N!} \left(\frac{V}{\Lambda^3}\right)^N \quad \text{for } M \gg N \tag{26.27}$$

Tangential remark: The factor of  $N!$  in the denominator is necessary to uphold the rule that each state is represented once and only once in the sum over states. This factor may “look like” a delabeling factor, but it is not. All the particles here are distinguishable, which means essentially that they have built-in physically-significant labels. This factor of  $N!$  is related to the fact that in equation 26.25 the letters  $P$ ,  $Q$ , and  $R$  are in alphabetical order in every term where they appear. This “normal ordering” accounts for the fact that once we have considered the state where the dime is heads-up and the penny is heads-down, we must not consider the state where the penny is heads-down and the penny is heads-up, because that is just another way of saying the same thing. This is a purely classical counting job, having nothing to do with the delabeling of identical particles.

$$\begin{aligned}
 S_{\text{snow}} &= k\beta E + k \ln(Z) \\
 &= \frac{3}{2} k N + k \ln\left(\frac{V^N}{\Lambda^{3N}}\right) + k \ln\left(\frac{M^N}{N!}\right) \\
 &= S_{\text{pure}} + N k \ln(M)
 \end{aligned} \tag{26.28}$$

where  $S_{\text{pure}}$  is defined in equation 26.13. The last term on the RHS of equation 26.28 is a remarkably simple expression for the entropy of the deal, assuming  $M \gg N$ . The equation as a whole can be considered a modification of equation 26.24, modified to account for the entropy of the deal.

Equation 26.28 tells us that the entropy for an extreme mixture is the same as the entropy for a pure, monatomic gas ... plus an additive term that is extensive (i.e. proportional to  $N$ ) and possibly very large. At constant  $N$ , it is independent of temperature, pressure, and volume.

In particular, if we perform a Gibbs-type mixing experiment involving extreme mixtures, starting with  $N_1$  particles on one side and  $N_2$  particles on the other side, there will be no entropy of mixing. The entropy of the deal will be simply additive, namely  $(N_1 + N_2)k \ln(M)$ , before and after mixing.

It is remarkable how this result fits in with other things we know about Gibbs-type mixing experiments:

- Helium on one side and helium on the other  $\rightarrow$  no entropy of mixing, because the two samples are the same.
- Helium on one side and neon on the other  $\rightarrow$  considerable entropy of mixing, because the two samples are different.
- Snow on one side and some different snow on the other side  $\rightarrow$  no entropy of mixing, even though the two samples are as different as they possibly could be.

The total entropy  $S_{\text{snow}}$  (including the entropy of the deal) behaves wildly differently from the conditional entropy  $S_{\text{snow}|D}$  (not including the entropy of the deal).

Using the same approximations that led to the Sackur-Tetrode equation (equation 26.17), we obtain:

$$\frac{S_{\text{snow}}/N}{k} = \ln\left(\frac{V/N}{\Lambda^3}\right) + \frac{5}{2} + \ln(M) \quad (26.29)$$

## 26.3 Rigid Rotor

We now turn to the case of a particles that have some structure. That is, they are not pointlike. Let's start with a heteronuclear diatomic molecule such as CO. There is a range of temperatures that are high enough to excite rotations of such a molecule, yet not high enough to excite vibrations. This is particularly relevant because room temperature sits in the middle of this “rigid rotor regime” for CO.

The energy of a rigid rotor is:

$$\epsilon(J, m_J) = \frac{\hbar^2 J(J+1)}{2I^\circ} \quad (26.30)$$

where  $J$  and  $m_J$  are the quantum numbers for the angular momentum associated with the ordinary rotation (not including the “internal” degrees of freedom, i.e. nuclear spin, if any). The moment of inertia is (as always) given by

$$I^\circ = \int r^2 dm^\circ \quad (26.31)$$

where  $r$  is the radius (measured from the center of mass),  $m^\circ$  is the mass. We write  $m^\circ$  and  $I^\circ$  – with a little circle suggesting rotation – to avoid a collision with the magnetic quantum number ( $m$ ) and the nuclear spin ( $I$ ), which we will encounter in a moment.

The partition function is, as always, a sum over states:

$$Z = \sum_{J,m_J} \exp[-\beta \epsilon(J, m_J)] \quad (26.32a)$$

$$= \sum_J (2J+1) \exp[-\beta J(J+1)(\hbar^2/2I^\circ)] \quad (26.32b)$$

in equation 26.32b, we have replaced the sum over states with a sum over *energy levels*. This makes use of the fact that the RHS of equation 26.30 is independent of  $m_J$ , and the fact that for each  $J$ , there are  $(2J+1)$  allowed  $m_J$  values. In other words, each energy level has a multiplicity of

$$\omega = 2J+1 \quad (\text{states per energy level}) \quad (26.33)$$

There is no known way to get a closed-form expression for rigid-rotor partition function, equation 26.32. (This stands in contrast to the harmonic oscillator and the particle in a box, which do have nice closed-form expressions.)

At low temperatures, we can approximate equation 26.32 by grinding out the first few terms of the sum. That gives us a power series:

$$\begin{aligned} Z &= 1 + 3x^2 + 5x^6 + 7x^{12} + \dots \\ x &:= \exp(-\theta_r/T) \\ \theta_r &:= \frac{\hbar^2}{2kI^\circ} \end{aligned} \quad (\text{rotational temperature-scale}) \quad (26.34)$$

At high temperatures, we need a different approximation:

$$\begin{aligned} Z_{\text{warm\_rotor}} &= \frac{T}{\theta_r} \left[ 1 + \frac{1}{3}b + \frac{1}{15}b^2 + \frac{4}{315}b^3 + \dots \right] \\ b &:= \beta k \theta_r \end{aligned} \quad (26.35)$$

At high enough temperatures, the partition function converges to  $T/\theta_r$ . This asymptote is shown by the dashed line in figure 26.2.

The low-temperature approximation is shown by the blue curve in figure 26.2. It is good for all  $T/\theta_r \leq 2$ .

The high-temperature approximation is shown by the red curve in figure 26.2. It is good for all  $T/\theta_r \geq 2$ .

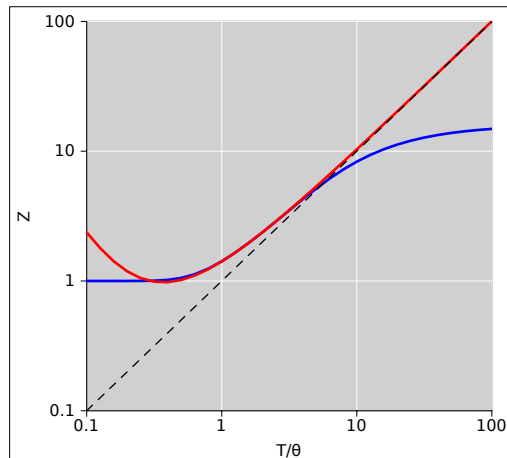


Figure 26.2: Rigid Rotor : Approximations to the Partition Function



We can rewrite the high-temperature limit in an interesting form:

$$\begin{aligned} Z_{\text{hot\_rotor}} &= \frac{T}{\theta_r} \\ &= 4\pi \frac{r_g^2}{\lambda^2} \end{aligned} \quad (26.36)$$

where  $r_g$  is the radius of gyration, defined as  $r_g^2 := I^{\text{circ}}/m^\circ$ . This is a well-known quantity in mechanics. It is a measure of the “effective” size of the rotor. The thermal de Broglie length  $\lambda$  is defined in equation 26.2.

It is interesting to contrast equation 26.1 with equation 26.35. Both involve the thermal de Broglie length,  $\lambda$ . However, the former compares  $\lambda$  to the size of the box, while the latter compares  $\lambda$  to the size of the molecule – quite independent of the size of the box.

## 26.4 Isentropic Processes

Scaling arguments are always fun. Let’s see what happens when we scale a box containing an ideal gas. We restrict attention to a tabletop nonrelativistic monatomic nondegenerate ideal gas in three dimensions except where otherwise stated). In particular, in this section we do not consider the rotational degrees of freedom mentioned in section 26.3.

Consider the case where our gas starts out in a three-dimensional box of volume  $V$ . Then we increase each of the linear dimensions by a factor of  $\alpha$ . Then the volume increases by a factor of  $\alpha^3$ . The energy of each microstate decreases by a factor of  $\alpha^2$  in accordance with the usual nonrelativistic kinetic energy formula  $p^2/(2m)$  where  $p = \hbar k$ . (Because the gas is monatomic and ideal, this kinetic energy is the total energy.)

This is interesting because if we also scale  $\beta$  by a factor of  $\alpha^2$ , then every term in equation 26.69 is left unchanged, i.e. every term scales like the zeroth power of  $\alpha$ . That implies that the partition function itself is unchanged, which in turn implies that the entropy is unchanged. We can summarize this as:

$$\begin{aligned} \beta V^{-2/3} &= f(S) && \text{for some function } f() \\ \beta V^{-2/3} &= \text{const} && \text{for any isentropic process} \end{aligned} \quad (26.37)$$

where the RHS of this equation is some as-yet-unknown function of entropy, but is not a function of  $\beta$  or  $V$ . (We continue to assume constant  $N$ , i.e. constant number of particles.)

Equation 26.37 is useful in a number of ways. For starters, we can use it to eliminate temperature in favor of entropy in equation 26.10. Plugging in, we get

$$E = \frac{3}{2} \frac{N V^{-2/3}}{f(S)} \quad (26.38)$$

That’s useful because pressure is defined as a derivative of the energy *at constant entropy* in accor-

dance with equation 7.6. Applying the definition to the present case, we get

$$\begin{aligned}
 P &:= - \left. \frac{\partial E}{\partial V} \right|_S \\
 &= \frac{3}{2} \frac{N (2/3) V^{-5/3}}{f(S)} \\
 &= \frac{2}{3} \frac{E}{V}
 \end{aligned} \tag{26.39}$$

Plugging the last line of equation 26.39 into equation 26.10, we find

$$\begin{aligned}
 P V &= N k T \\
 P V &= \frac{N}{\beta}
 \end{aligned} \tag{26.40}$$

which is the celebrated *ideal gas law*. It is quite useful. However, it is not, by itself, a complete description of the ideal gas; we need another equation (such as equation 26.37) to get a reasonably complete picture. All this can be derived from the partition function, subject to suitable restrictions.

It is worthwhile to use equation 26.40 to eliminate the  $\beta$  dependence from equation 26.37. That gives us, after some rearranging,

$$\begin{aligned}
 P V^{5/3} &= N/f(S) && \text{for some function } f() \\
 P V^{5/3} &= \text{const} && \text{for any canonical isentropic process}
 \end{aligned} \tag{26.41}$$

See equation 26.45 for a more general expression.

## 26.5 Polytropic Processes ... Gamma etc.

In this section we generalize the results of section 26.4 to cover polyatomic gases. We continue to restrict attention to a tabletop nonrelativistic nondegenerate ideal gas in three dimensions ... except where otherwise stated.

We need to be careful, because the energy-versus-temperature relationship will no longer be given by equation 26.10. That equation only accounts for the kinetic energy of the gas particles, whereas the polyatomic gas will have rotational and vibrational modes that make additional contributions to the energy.

We now hypothesize that the energy in these additional modes will scale in proportion to the kinetic energy, at least approximately. This hypothesis seems somewhat plausible, since we have seen that the total energy of a particle in a box is proportional to temperature, and the total energy of a harmonic oscillator is proportional to temperature except at the very lowest temperatures. So if it turns out that other things are also proportional to temperature, we won't be too surprised. On the other hand, a plausible hypothesis is not at all the same as a proof, and we shall see that the total energy is *not* always proportional to temperature.

To make progress, we say that any gas that upholds equation 26.42, where the RHS is constant, or at worst a slowly-varying function of temperature, is (by definition) a *polytropic gas*.

$$\begin{aligned}
 \frac{E}{N k T} &= \text{more-or-less constant} \\
 &= \frac{1}{\gamma-1}
 \end{aligned} \tag{26.42}$$

We write the RHS as a peculiar function of  $\gamma$  in accordance with tradition, and to simplify results such as equation 26.50. There are lots of physical systems that more-or-less fit this pattern. In particular, given a system of  $N$  particles, each with  $D_{\square}$  quadratic degrees of freedom, equipartition tells us that

$$\frac{E}{NkT} = \frac{D_{\square}}{2} \quad (26.43)$$

as discussed in section 25.2. Of course, not everything is quadratic, so equation 26.42 is more general than equation 26.43.

When non-quadratic degrees of freedom can be ignored, we can write:

$$\gamma = \frac{D+2}{D} \quad (26.44)$$

You can see that  $\gamma = 5/3$  for a monatomic gas in three dimensions. See table 26.1 for other examples. Because of its role in equation 26.50,  $\gamma$  is conventionally called the ratio of specific heats. This same quantity  $\gamma$  is also called the *adiabatic exponent*, because of its role in equation 26.45. It is also very commonly called simply the “gamma” of the gas, since it is almost universally denoted by the symbol  $\gamma$ .

Using the same sort of arguments used in section 26.4, we find that equation 26.37 still holds, since it the main requirement is a total energy that scales like  $\alpha^{-2}$ .

Continuing down the same road, we find:

$$\begin{aligned} PV^{\gamma} &= N/f(S) && \text{for some function } f() \\ PV^{\gamma} &= \text{const} && \text{for any canonical isentropic process} \\ kTV^{\gamma-1} &= 1/f(S) \\ kTV^{\gamma-1} &= \text{const} && \text{for any isentropic process} \end{aligned} \quad (26.45)$$

Some typical values for  $\gamma$  are given in table 26.1. As we shall see, theory predicts  $\gamma = 5/3$  for a monatomic nonrelativistic nondegenerate ideal gas in three dimensions. For polyatomic gases, the gamma will be less. This is related to the number of “quadratic degrees of freedom” as discussed in section 26.7.

**Terminology:** We define a *polytropic process* (not to be confused with polytropic gas) as any process that follows a law of the form  $PV^n = c$ , This includes but is not limited to the case where the exponent  $n$  is the adiabatic exponent  $\gamma$ . Interesting cases include

- $n = 0$  implies constant pressure,
- $n = 1$  for an ideal gas implies constant temperature,
- $n = \gamma$  implies constant entropy, and
- $n = \infty$  implies constant volume.

Let’s calculate the energy content of a polytropic gas. Specifically, we calculate the amount of energy you could extract by letting the gas push against a piston as it expands isentropically from volume  $V$  to infinity, as you can confirm by doing the integral of  $PdV$ :

$$\begin{aligned} E &= \int_V^{\infty} P dV \\ &= \frac{PV}{\gamma - 1} \end{aligned} \quad (26.46)$$

Gas	$\theta_r/\text{K}$	$T/\text{K}$	$\gamma$	$2/(\gamma - 1)$
He		293.15	1.66	3
H <sub>2</sub>	85.3	92.15	1.597	3.35
..	..	293.15	1.41	4.87
N <sub>2</sub>	2.88	293.15	1.4	5
O <sub>2</sub>	2.07	293.15	1.4	5
Dry air		273.15	1.403	4.96
..		293.15	1.402	4.98
..		373.15	1.401	4.99
..		473.15	1.398	5.03
..		673.15	1.393	5.09
..		1273.15	1.365	5.48
..		2273.15	1.088	22.7
CO	2.77	293.15	1.4	5
Cl <sub>2</sub>	.351	293.15	1.34	5.88
H <sub>2</sub> O		293.15	1.33	6
CO <sub>2</sub>		293.15	1.30	6.66

Table 26.1: Values of  $\theta_r$  and  $\gamma$  for common gases

This means the ideal gas law (equation 26.40) can be extended to say:

$$PV = NkT = (\gamma - 1)E \quad (26.47)$$

This is interesting because  $PV$  has the dimensions of energy, and it is a common mistake to think of it as “the” energy of the gas. However we see from equation 26.47 and table 26.1 that  $PV$  is only 66% of the energy for helium, and only 40% of the energy for air.

You shouldn’t ask where the “missing” energy went. There is no missing energy here. There was never a valid reason to think that  $PV$  was “the” energy. The integral of  $PdV$  has the same dimensions as  $PV$ , but is not equal to  $PV$ . There’s more to physics than dimensional analysis.

Let’s calculate the heat capacities for a polytropic ideal gas. We retrace the steps used in section 26.2. Rather than starting from equation 26.10 to derive equation 26.11 and equation 26.11, we now start from equation 26.47 to derive the following for constant volume:

$$\begin{aligned} C_v &= Nk \frac{1}{\gamma-1} \\ c_v &= R \frac{1}{\gamma-1} \end{aligned} \quad (26.48)$$

And similarly, for constant pressure:

$$\begin{aligned} C_p &= Nk \frac{\gamma}{\gamma-1} \\ c_p &= R \frac{\gamma}{\gamma-1} \end{aligned} \quad (26.49)$$

The ratio of specific heats is

$$\begin{aligned} \frac{C_p}{C_v} &= \gamma \\ \frac{c_p}{c_v} &= \gamma \end{aligned} \quad (26.50)$$

This is why  $\gamma$  deserves the name “ratio of specific heats” or “specific heat ratio”.

We can use equation 26.48 and equation 26.49 to get a useful expression for the entropy of a polytropic gas. We invoke the general definition of heat capacity – aka the entropy capacity, loosely speaking – namely equation 7.24.

$$\begin{aligned} C_v &= \left. \frac{\partial S}{\partial \ln(T)} \right|_V \\ &= Nk \frac{1}{\gamma-1} \end{aligned} \quad (26.51)$$

We can integrate that along a contour of constant  $V$  to obtain:

$$S = \frac{Nk \ln(T)}{\gamma-1} + f(V) \quad (26.52)$$

where  $f()$  is some as-yet-unspecified function. As a check, note that for the ideal monatomic gas,  $\gamma = 5/3$ , so equation 26.52 has the same temperature-dependence as equation 26.18, as it should.

Similarly:

$$\begin{aligned} C_p &= \left. \frac{\partial S}{\partial \ln(T)} \right|_P \\ &= Nk \frac{\gamma}{\gamma-1} \end{aligned} \quad (26.53)$$

$$S = \frac{Nk\gamma \ln(T)}{\gamma-1} + f(P) \quad (26.54)$$

Let's try to derive the entropy of a polytropic gas. We start by rewriting the partition function for a particle in a box (equation 26.1) as:

$$Z_{\text{ppb}} = \frac{V}{\Lambda^3} \quad (26.55)$$

We then replace the 3 by

$$b := \frac{2}{\gamma-1} \quad (26.56)$$

and turn the crank. As a generalization of equation 26.6 we have:

$$Z_N = \frac{1}{N!} \left( \frac{V L^{b-3}}{\Lambda^b} \right)^N \quad (26.57)$$

where  $L$  is some length I threw in to make the dimensions come out right.

Then in analogy to equation 26.13 we have

$$\begin{aligned} S_{\text{polytropic}} &= k\beta E + k \ln(Z) \\ &= \frac{kN}{\gamma-1} + k \ln\left(\frac{1}{N!}\right) + (b-3)Nk \ln(L) + Nk \ln(V) - \frac{2Nk}{\gamma-1} \ln(\Lambda) \\ &= f(N) + Nk \ln(V) + \frac{Nk}{\gamma-1} \ln(T) \end{aligned} \quad (26.58)$$

for some unspecified  $f(N)$ . All of the temperature dependence is in the last term. You can check that this term is plausible, insofar as it agrees with equation 26.52. Similarly, all the volume dependence is in the next-to-last term. You can check that this is plausible, by considering an adiabatic process such that  $PV^\gamma$  is constant, and  $PV = NkT$ . For such a process, equation 26.58 predicts zero change in entropy, as it should.

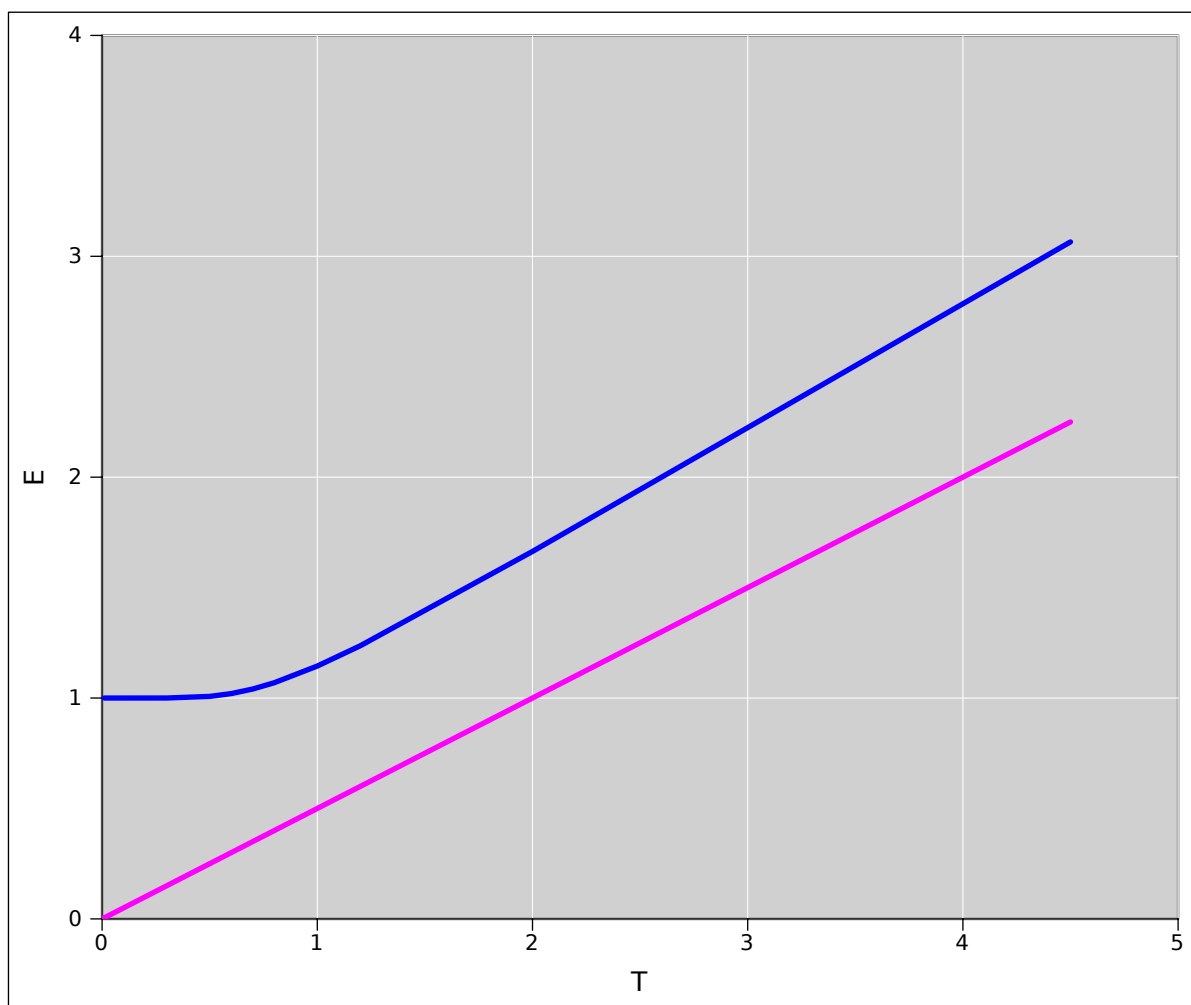


Figure 26.3: Particle in a Box : Energy at Low Temperature

## 26.6 Low Temperature

In this section we consider low temperatures, not just the high-temperature limit. For a single particle in a one-dimensional box, the partition function is given by equation 26.63. We calculate the energy from the partition function in the usual way, via equation 24.9.

Here the energy and temperature are measured in units of the ground-state energy (which depends on the size of the box). The blue curve shows the actual energy of the system; the magenta curve shows the high-temperature asymptote, namely  $E = 0.5T$ .

The famous zero-point energy is clearly visible in this plot.

As you can see in the diagram, the slope of the  $E$ -versus- $T$  curve starts out at zero and then increases. It actually becomes larger than 0.5. At higher temperatures (not shown in this diagram) it turns over, converging to 0.5 from above.

## 26.7 Degrees of Freedom, or Not

In this section we will temporarily lower our standards. We will do some things in the manner of “classical thermodynamics” i.e. the way they were done in the 19th century, before people knew about quantum mechanics.

Also in this section, we restrict attention to ideal gases, so that  $PV = NkT$ . This is quite a good approximation for typical gases under ordinary table-top conditions. We further assume that the gas is non-relativistic.

We now attempt to apply the pedestrian notion of equipartition, as expressed by equation 25.7. It tells us that for a classical system at temperature  $T$ , there is  $1/2kT$  of energy (on average) for each quadratic degree of freedom. In particular, if there are  $N$  particles in the system and  $D_{\square}$  classical quadratic degrees of freedom per particle, the energy of the system is:

$$\begin{aligned} E &= \frac{D_{\square}}{2} N kT \\ &= \frac{D_{\square}}{2} \frac{N}{\beta} \end{aligned} \tag{26.59}$$

We assert that a box of monatomic gas has  $D_{\square} = 3$  quadratic degrees of freedom per atom. That is, each atom is free to move in the  $X$ ,  $Y$ , and  $Z$  directions, but has no other degrees of freedom that contribute to the average energy. (To understand why the potential energy does not contribute, see section 25.3.) This means that equation 26.59 is consistent with equation 26.10. However, remember that equation 26.10 was carefully calculated, based on little more than the energy-versus-momentum relationship for a free particle ... whereas equation 26.59 is based on a number of bold assumptions.

Things get more interesting when we assert that for a small linear molecule such as  $N_2$  or  $CO$ , there are  $D_{\square} = 5$  degrees of freedom. The story here is that in addition to the aforementioned freedom to move in the  $X$ ,  $Y$ , and  $Z$  directions, the molecule is also free to rotate in two directions. We assert that the molecule is not able to rotate around its axis of symmetry, because that degree of freedom is frozen out ... but it is free to tumble around two independent axes perpendicular to the axis of symmetry.

Going back to equation 26.59 and comparing it to equation 26.47, we find the two expressions are equivalent if and only if

$$\begin{aligned} D_{\square} &= 2/(\gamma - 1) \\ \gamma &= (2 + D_{\square})/D_{\square} \\ &= 1 + 2/D_{\square} \end{aligned} \tag{26.60}$$

You can now appreciate why the rightmost column of table 26.1 tabulates the quantity  $2/(\gamma - 1)$ . The hope is that an experimental measurement of  $\gamma$  for some gas might tell us how many classical quadratic degrees of freedom there are for each particle in the gas, by means of the formula  $D_{\square} = 2/(\gamma - 1)$ . This hope is obviously unfulfilled in cases where formula gives a non-integer result. However, there are quite a few cases where we do get an integer result. This is understandable, because some of the degrees of freedom are not classical. In particular the “continuum energy” approximation is not valid. When the spacing between energy levels is comparable to  $kT$ , that degree of freedom is partially frozen out and partially not. For details on this, see chapter 25.

You have to be a little bit careful even when  $2/(\gamma - 1)$  is an integer. For instance, as you might guess from table 26.1, there is a point near  $T = 160K$  where the  $\gamma$  of molecular hydrogen passes through the value  $\gamma = 1.5$ , corresponding to  $D_{\square} = 4$ , but this is absolutely not because hydrogen has four degrees of freedom. There are more than four degrees of freedom, but some of them are partially frozen out, and it is merely fortuitous if/when  $\gamma$  comes out to be an integer.

The  $\gamma$  values for  $\text{Cl}_2$  and  $\text{CO}_2$  are lower than you would expect for small linear molecules. This is because vibrational degrees of freedom are starting to come into play.

For an even more spectacular example of where classical ideas break down, including the idea of “degrees of freedom,” and the idea of “equipartition of energy” (i.e.  $1/2kT$  of energy per degree of freedom), look at the two-state system discussed in section 24.4.

## 26.8 Discussion

Except for section 26.7, we derived everything we needed more-or-less from first principles: We used quantum mechanics to enumerate the microstates (figure 26.4), we calculated the microstate energy as  $p^2/(2m)$ , then constructed the partition function. The rest was just turning the crank, since there are well-known formulas for calculating the thermodynamic observables (energy, entropy, pressure, et cetera) in terms of the partition function.

## 26.9 Derivation: Particle in a Box

This section shows how to derive the canonical partition function for a single particle in a box.

The three lowest-lying energy eigenstates for a one-dimensional particle in a box are illustrated in figure 26.4.



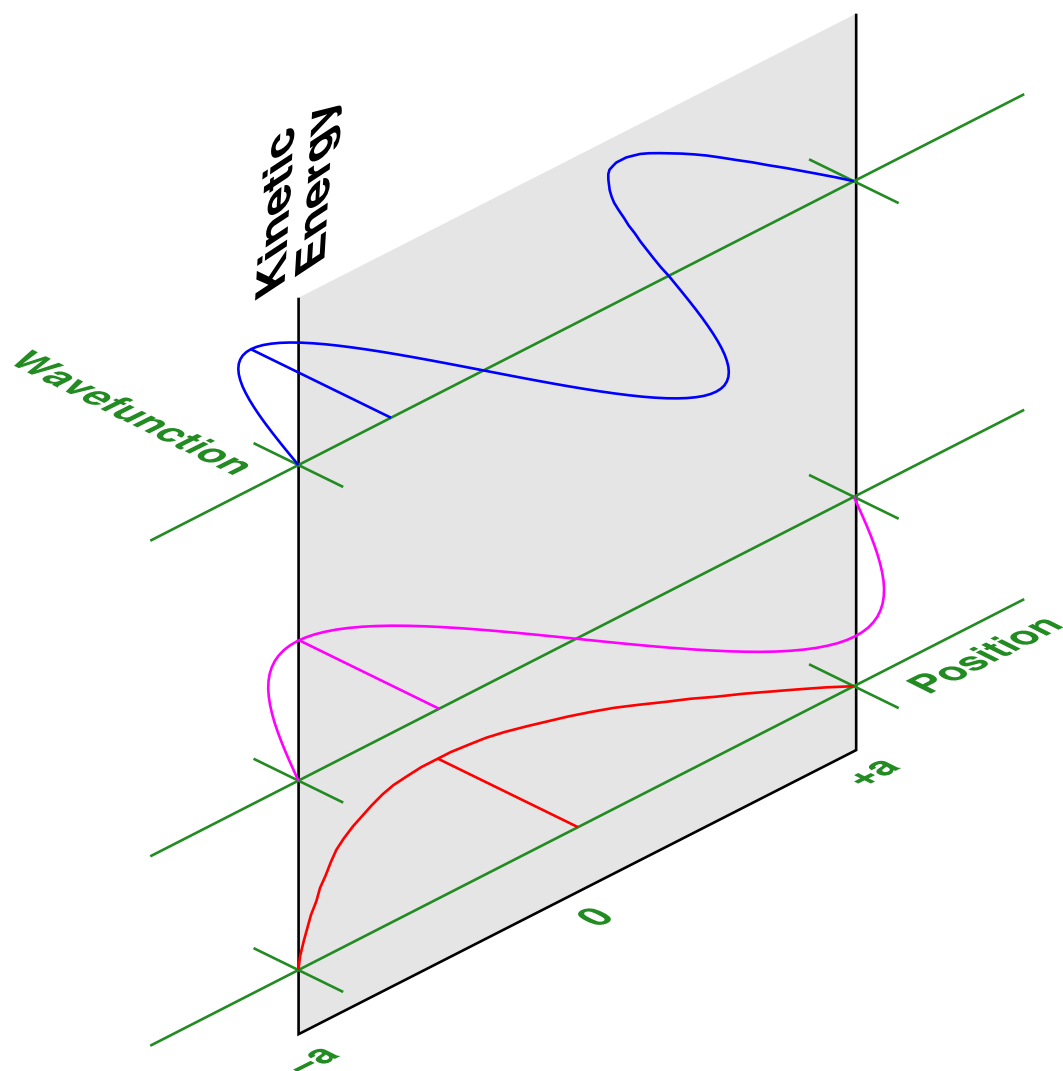


Figure 26.4: Eigenfunctions for a Particle in a Box

The wavevector of the  $n$ th state is denoted  $k_n$ , and can be determined as follows: Observe that the ground-state wavefunction ( $n = 1$ ) picks up  $\pi$  (not  $2\pi$ ) units of phase from one end of the box to the other, and the  $n$ th state has  $n$  times as many wiggles as the ground state. That is,

$$k_n L = n \pi \quad (26.61)$$

where  $L$  is the length of the box.

As always, the momentum is  $p = \hbar k$ , so for a non-relativistic particle the energy of the  $n$ th state is

$$\begin{aligned} E_n &= \frac{p^2}{2m} \\ &= \frac{n^2 \pi^2 \hbar^2}{2m L^2} \end{aligned} \quad (26.62)$$

and the partition function is therefore

$$\begin{aligned} Z &= \sum_{n=1}^{\infty} \exp\left(-\beta \frac{n^2 \pi^2 \hbar^2}{2m L^2}\right) \\ &= \sum_{n=1}^{\infty} \exp\left(-\frac{\pi}{4} n^2 \frac{\Lambda^2}{L^2}\right) \\ &= \sum_{n=1}^{\infty} X^{n^2} \end{aligned} \quad (26.63)$$

where (as always)  $\Lambda$  denotes the thermal de Broglie length (equation 26.2), and where the first term in the partition function is:

$$\begin{aligned} X &:= \exp\left(-\frac{\pi}{4} \frac{\Lambda^2}{L^2}\right) \\ &= \exp\left(\frac{-E_1}{kT}\right) \\ &= \exp\left(\frac{-p_1^2/2m}{kT}\right) \end{aligned} \quad (26.64)$$

and conversely

$$\frac{\Lambda}{L} = \sqrt{\left(\frac{4}{\pi} \frac{E_1}{kT}\right)} \quad (26.65)$$

where  $E_1$  and  $p_1$  are the kinetic energy and the momentum of the  $n = 1$  state of the particle in a box.

In the low temperature limit,  $X$  is small, so only the first few terms are important in the sum on the RHS of equation 26.63. When  $X$  is less than 0.5, we can approximate  $Z$  to better than 1ppm accuracy as:

$$Z = X + X^4 + X^9 + X^{16} \quad (26.66)$$

The probability of occupation of the two lowest-lying states are then:

$$\begin{aligned} P_1 &= 1 - X^3 \pm O(X^6) \\ P_2 &= X^3 \pm O(X^6) \end{aligned} \quad (26.67)$$

We now move away from the low temperature limit consider moderate and high temperatures. In this case, the sum in equation 26.63 can be approximated by an integral, to high accuracy.<sup>1</sup>

The integral is in fact a *Gaussian* integral, which makes things easy for us, since Gaussian integrals show up quite often in physics, and there are routine procedures for handling them. (See reference 56 for a review.) In fact, you can almost do this integral in your head, by making a scaling argument. The summand in equation 26.63 (which is also our integrand) is a Gaussian with a peak height essentially equal to unity, and with a width (along the  $n$  axis) that scales like  $L/\Lambda$ . So the area under the curve scales like  $L/\Lambda$ . If you do the math, the factors of 2 and factors of  $\pi$  drop out and you find that the area of a half-Gaussian is just  $L/\Lambda$ .

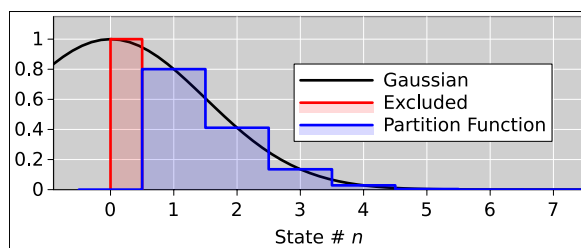


Figure 26.5: Gaussian Approximation to the Partition Function for a Particle in a Box

Figure 26.5 shows why we care about only half of the Gaussian. Also, if we want the area under the Gaussian curve to be a good approximation, it pays to account for the red shaded region that does not contribute to the partition function. Most textbooks overlook this, but it makes a dramatic improvement at the low end of the temperature range. Equation 26.68 gives much better than 1ppm accuracy all the way down to  $X = .5$ , where it mates up with equation 26.66 as shown in figure 26.6.

$$Z = \frac{L}{\Lambda} - 1/2 \quad (\text{Moderate or High Temperature}) \quad (26.68)$$

At sufficiently high temperature and/or low density, you can drop the  $1/2$ .

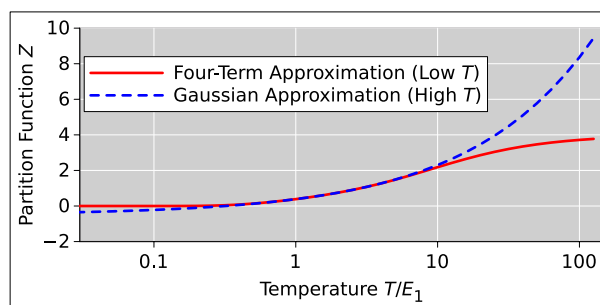


Figure 26.6: Gaussian versus Four-Term Approximation

We can use the partition function to derive anything we need to know about the thermodynamics of a particle in a box.

<sup>1</sup>This is similar to (but not exactly equivalent to) the Thomas-Fermi approximation. Also beware that there is something else called Thomas-Fermi *theory*, which is considerably more elaborate than the Thomas-Fermi approximation.

Let us now pass from one dimension to three dimensions. The partition function for a particle in a three-dimensional box can be derived using the same methods that led to equation 26.68. We won't bother to display all the steps here. The exact expression for  $Z$  can be written in various ways, including:

$$Z = \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} \exp\left(-\beta \frac{n_x^2 \pi^2 \hbar^2}{2m L_x^2}\right) \exp\left(-\beta \frac{n_y^2 \pi^2 \hbar^2}{2m L_y^2}\right) \exp\left(-\beta \frac{n_z^2 \pi^2 \hbar^2}{2m L_z^2}\right) \quad (26.69)$$

In the high-temperature limit this reduces to:

$$Z = \frac{V}{\Lambda^3} \quad (26.70)$$

where  $V$  is the volume of the box. The relationship between equation 26.68 and equation 26.70 is well-nigh unforgettable, based on dimensional analysis.

## 26.10 Area per State in Phase Space

It turns out that Planck used  $h$  in connection with thermodynamics many years before anything resembling modern quantum mechanics was invented. Thermodynamics did not inherit  $h$  from quantum mechanics; it was actually the other way around. More importantly, you shouldn't imagine that there is any dividing line between thermodynamics and quantum mechanics anyway. All the branches of physics are highly interconnected.

### 26.10.1 Particle in a Box

If we (temporarily) confine attention to the positive  $k$  axis, for a particle in a box, equation 26.61 the wavenumber of the  $n$ th basis state is  $k_n = n\pi/L$ . The momentum is therefore  $p_n = \hbar k_n = n\pi\hbar/L$ . Therefore, the spacing between states (along the the positive momentum axis) is  $\pi\hbar/L$ . Meanwhile, there is no spacing along the position axis; the particle is within the box, but cannot be localized any more precisely than that. Therefore each state is (temporarily) associated with an area in phase space of  $\pi\hbar$  or equivalently  $h/2$ . The states themselves do not have any extent in the  $p$  direction; area in phase space is the area *between* two states, the area bounded by the states.

For a particle in a box, running-wave states are not a solution to the equation of motion. Therefore, when we consider the  $k$  axis as a whole, we find that the area between one state and the next consists of *two* patches of area, one at positive  $k$  and another at negative  $k$ . Both values of  $k$ , positive and negative, correspond to the same physical state. Taking these two contributions together, the actual area per state is simply  $h$ .

Figure 26.7 shows the phase space for a particle in a box, with the lowest three areas color-coded. The area “associated” with a state is the area *between* that state and the preceding one.

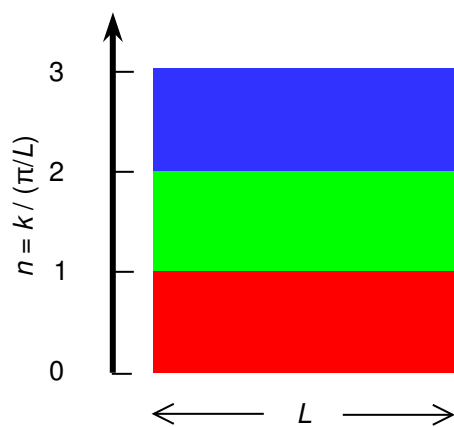


Figure 26.7: Phase Space : Particle in a Box

### 26.10.2 Periodic Boundary Conditions

We now consider a particle subject to periodic boundary conditions with period  $L$ . This is analogous to water in a circular trough with circumference  $L$ . Running-wave states are allowed. The calculation is nearly the same as for a particle in a box, except that the wavenumber of the  $n$ th basis state is  $k_n = 2n\pi/L$ , which differs by a factor of 2 from the particle-in-a-box expression. Also, there is a  $n = 0$  state that we didn't have in the box.

In this case, positive  $k$  corresponds to a rightward running wave, while negative  $k$  corresponds to a leftward running wave. These states are physically distinct, so each state has only one patch of area in phase space. The area is  $2\pi\hbar$  or simply  $h$ .

Figure 26.8 shows the phase space for this case, with the three lowest-energy basis states color coded. This is much simpler than the particle-in-a-box case (figure 26.7).

### 26.10.3 Harmonic Oscillator

Figure 26.9 shows the analogous situation for a harmonic oscillator. Once again, the states themselves occupy zero area in phase space. When we talk about area in phase space, we talk about the area bounded *between* two states. In the figure, the states are represented by the boundary between one color and the next. The boundary has zero thickness.

For the harmonic oscillator (unlike a particle in a box) each state has nontrivial extent in the position-direction, not just the momentum-direction.

### 26.10.4 Non-Basis States

Any state of the system can be expressed as a linear combination of basis states. For example, if you want to create a state that is spatially localized somewhere within the box, this can be expressed as a linear combination of basis states.

Now it turns out that the process of taking linear combinations always preserve area in phase space. So each and every state, including any spatially-localized state, will occupy an area  $h$  in phase space. This fact is used in section 12.3.

Actually it has been known since the 1800s that any physically-realizable transformation preserves area in phase space; this is known as Liouville's theorem. Any violation of this theorem would immediately violate many laws of physics, including the second law of thermodynamics, the Heisenberg uncertainty principle, the optical brightness theorem, the fluctuation/dissipation theorem, et cetera.

Phase space is conserved.

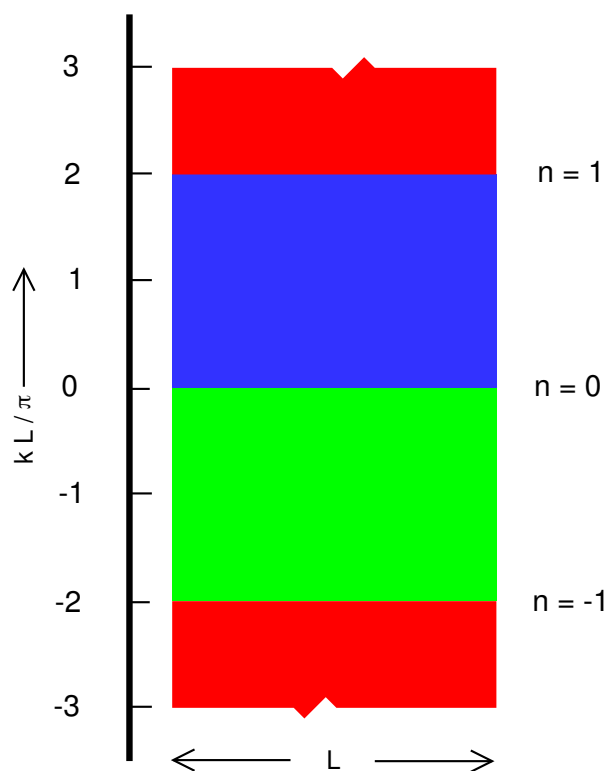


Figure 26.8: Phase Space : Particle with Periodic Boundary Conditions

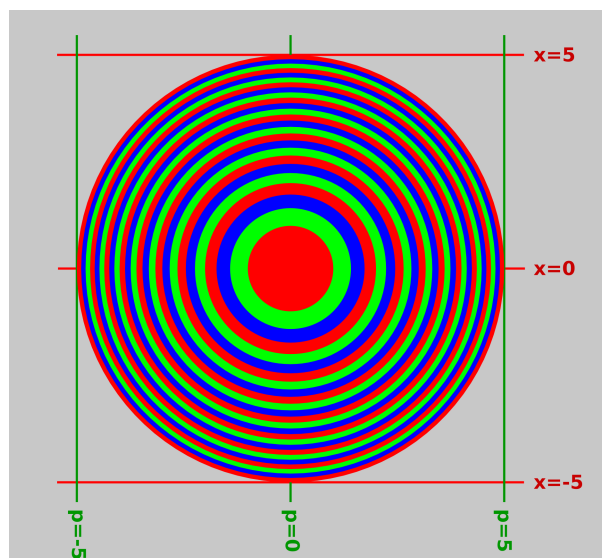


Figure 26.9: States of Definite Energy : Harmonic Oscillator



## Chapter 27

# Density Matrices

For the sake of completeness, this section makes explicit the limits of validity for equation 2.2, and explains how to handle the unusual situations where it is not valid.

Equation 2.2 is *almost* the most general formulation of the entropy idea.

If you are using an ordinary computer and an ordinary communication channel, measuring bits and the probability of bits, equation 2.2 serves just fine.

If you are dealing with a nondegenerate macroscopic system, equation 2.2 should serve just fine.

Equation 27.6 is truly the most general formulation.

If you are using a quantum computer and/or a quantum communication channel, measuring qubits and the amplitude of qubits, you presumably didn't get into that situation by accident, and you will be well aware of the need to use equation 27.6.

If you are dealing with (say) the heat capacity of a superfluid, superconductor, or other system that exhibits macroscopic quantum coherence, you will presumably be aware of the need to use equation 27.6.

Most commonly, equation 27.6 is used in conjunction with ultramicroscopic systems. As an example, consider the  $sp^3$  atomic orbitals, which are a coherent superposition of the  $\{s, p_x, p_y, p_z\}$  orbitals.

It is more-or-less impossible to formulate a theory of thermodynamics without a concept of microstate. Microstates are best understood in terms of quantum states, which is the approach taken throughout this document.

There is, of course, more to quantum mechanics than the notion of state. There is also the notion of *probability amplitude* (or simply *amplitude*); the absolute square of the amplitude is the probability.

For many purposes, the probability tells us everything we need to know, so we don't directly care about the amplitude.

However there are situations where the system exhibits *coherence* between the quantum states. We sometimes say the system is in an *entangled* state. Schrödinger's Cat is a well-known example, perhaps an unnecessarily-complicated example.

When a system exhibits both coherent superposition and incoherent (thermal) mixing, the best way to represent it is in terms of a density matrix.

Any pure quantum state  $|\psi\rangle$  is represented by a density matrix  $\rho$  which can be expressed as an outer product:

$$\rho = |\psi\rangle\langle\psi| \quad (27.1)$$

That means that for an  $N$ -dimensional state space, the density matrix will be an  $N \times N$  matrix.

Let's look at some examples. Suppose the state-space of the system is spanned by two basis states,  $|1\rangle$  and  $|2\rangle$ . Each of these states can be represented by a state vector, or by the corresponding density matrix.

$$\begin{aligned} \rho_1 &= |1\rangle\langle 1| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ \rho_2 &= |2\rangle\langle 2| = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \quad (27.2)$$

Things get more interesting when we consider a state that is a coherent superposition of the two basis states:

$$\begin{aligned} |a\rangle &= \frac{|1\rangle - |2\rangle}{\sqrt{2}} \\ \rho_a &= |a\rangle\langle a| = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \end{aligned} \quad (27.3)$$

Note that the diagonal elements of the density matrix can be interpreted as the probability of the basis states, and they necessarily sum to unity, as you can see in each of the examples. The off-diagonal elements represent correlations between the the basis states.

Things get even more interesting if we allow an arbitrary phase, as follows:

$$\begin{aligned} |b\rangle &= \frac{|1\rangle + e^{i\phi}|2\rangle}{\sqrt{2}} \\ \rho_b &= |b\rangle\langle b| = 1/2 \begin{bmatrix} 1 & e^{i\phi} \\ e^{-i\phi} & 1 \end{bmatrix} \end{aligned} \quad (27.4)$$

It turns out that in many situations, especially macroscopic situations, there are physical processes that perturb the phase of a superposition such as this. If we take the average over all phases, we get:

$$\langle \rho_b \rangle = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \quad (27.5)$$

which for the first time shows us the power of the density-matrix formalism. The object in equation 27.5 does not represent a pure quantum state, but rather the incoherent (thermal) mixing of states. This stands in contrast to previous equations such as equation 27.4 which did represent pure quantum states.

Note that equation 27.5 could have been obtained by taking a 50/50 mixture of  $\rho_1$  and  $\rho_2$  as given in equation 27.2. This is an example of the general rule that thermal mixtures can be computed by averaging the density matrices of the ingredients.

Notice the great power of the density matrix: Whereas a quantum state vector  $|a\rangle$  represents a microstate, a suitable density matrix  $\rho$  can fully represent a macrostate.

Reference 35 contains many more examples of density matrices.

There is a wonderfully simple test to detect pure states, by looking at the square of the density matrix. If and only if  $\rho^2 = \rho$ , the density matrix represents a pure state; otherwise it represents a mixture. Pure states have zero entropy; mixtures have entropy greater than zero, as we shall see in connection with equation 27.6.

Note that equation 27.4 represents a pure state while equation 27.5 does not – even though they have the same diagonal elements, i.e. the same state-by-state probabilities for the basis states. The off-diagonal terms, i.e. the correlations, make a significant contribution to the entropy.

In all generality, for a system characterized by a density matrix  $\rho$ , the entropy is given by

$$S := -\text{Tr } \rho \log \rho \quad (27.6)$$

This is the most robust definition of entropy. This is the gold standard. For many cases, i.e. when we can ignore quantum entanglement, it reduces to equation 2.2. Other expressions may be useful in more-restricted cases (as in section 9.6 for example) but you can never go wrong using equation 27.6.

Since the expectation value of any observable operator  $O$  is given by  $\text{Tr } \rho O$ , equation 27.6 can be interpreted as the expectation value of the surprisal, as discussed in section 2.7.1, where we define the operator form of the surprisal to be:

$$\$ := -\log \rho \quad (27.7)$$

In case you are wondering how to take the logarithm of a matrix, here's one way to do it: Expand  $\log(x)$  in a Taylor series. (It is smarter to expand about  $x = 1$  than about  $x = 0$ .) Then you can evaluate  $\log(x)$  in terms of powers of  $x$ , which requires nothing beyond matrix multiplication, scalar multiplication, addition, and other well-understood operations. Specifically,

$$\log(\rho) = -\sum_{N=1}^{\infty} \frac{(1-\rho)^N}{N} \quad (27.8)$$

Furthermore, in any basis where the density matrix is diagonal – i.e. where the off-diagonal elements vanish – there is an even easier way to evaluate the logarithm: just take the log of each diagonal element separately, element by element.

Also: In any basis where the density matrix is diagonal, equation 27.6 is manifestly identical to equation 2.2. Note that any matrix can be made diagonal by a suitable change of basis. Also

note that the value of the trace operator is unaffected by a change of basis; this can be seen as an immediate consequence of the “cyclic property” of the trace.

In a practical sense, what this section is saying is that if your density matrix  $\rho$  is not diagonal, it might be a good idea to perform a change of basis so that  $\rho$  becomes diagonal, and then evaluate equation 27.6 (or equivalently equation 2.2) in that basis. Equation 27.6 is just a compact way of saying this.

## Chapter 28

### Summary

- Thermodynamics inherits many results from nonthermal mechanics. Energy, momentum, and electrical charge are always well defined. Each obeys a strict local conservation law.
- Entropy is defined in terms of probability. It is always well defined. It obeys a strict local paraconservation law. Entropy is what sets thermodynamics apart from nonthermal mechanics.
- Entropy is not defined in terms of energy, nor vice versa. Energy and entropy are well defined even in situations where the temperature is unknown, undefinable, irrelevant, or zero.
- Entropy is not defined in terms of position. It involves probability spread out in state-space, not necessarily particles spread out in position-space.
- Entropy is not defined in terms of multiplicity. It is *equal* to the log of the multiplicity in the special case where all accessible states are equiprobable ... but not in the general case.
- Work suffers from two inconsistent definitions. Heat suffers from at least three inconsistent definitions. Adiabatic suffers from two inconsistent definitions. At the very least, we need to coin new words or phrases, so we can talk about the underlying reality with some semblance of clarity. (This is loosely analogous to the way phlogiston was replaced by two more-modern, more-precise concepts, namely energy and oxygen.)
- Heat and work are at best merely means for keeping track of certain contributions to the energy budget and entropy budget. In some situations, your best strategy is to forget about heat and work and account for energy and entropy directly.
- When properly stated, the first law of thermodynamics expresses conservation of energy ... nothing more, nothing less. There are several equally-correct ways to state this. There are also innumerable many ways of misstating it, some of which are appallingly widespread.
- When properly stated, the second law of thermodynamics expresses paraconservation of entropy ... nothing more, nothing less. There are several equally-correct ways to state this. There are also innumerable many ways of misstating it, some of which are appallingly widespread.

- Some systems (not all) are in internal equilibrium. They are described by a thermal distribution. They have a temperature.
- Even more importantly, some systems (not all) are in internal equilibrium *with exceptions*. They are described by a thermal distribution *with exceptions*. They have a temperature.
- Two systems that are each in internal equilibrium may or may not be in equilibrium with each other. Any attempted theory of thermodynamics based on the assumption that everything is in equilibrium would be trivial and worthless.
- The idea of distinguishing thermal versus nonthermal energy transfer across a boundary makes sense in selected situations, but has serious limitations.
  - Heat exchangers exist, and provide 100% thermal energy transfer.
  - Thermally-insulated pushrods exist, and (if properly used) provide nearly 100% nonthermal energy transfer.
  - The idea of distinguishing thermal from nonthermal on the basis of transfers across a boundary goes to pot in dissipative situations such as friction in an oil bearing.
- There is a well-founded way to split the energy-change  $dE$  into a thermal part  $T dS$  and a mechanical part  $P dV$  (subject to mild restrictions).
- There is a well-founded way to split the overall energy  $E$  into a Locrian (thermal) part and a non-Locrian (nonthermal) part (subject to mild restrictions).
- Not all kinetic energy contributes to the heat capacity. Not all of the heat capacity comes from kinetic energy. Not even close.

More generally, none of the following splits is the same as another:

- $T dS$  versus  $P dV$ .
- Locrian versus non-Locrian.
- Cramped versus uncramped.
- Kinetic versus potential energy.
- Overall motion of the center-of-mass versus internal motion relative to the center-of-mass.
- There is a simple relationship between force and momentum, for any system, macroscopic or microscopic.
- For pointlike systems (no internal degrees of freedom), there is a simple relationship between overall force and total kinetic energy ... but for more complex systems, the relationship is much more complicated. There are multiple inequivalent work-like quantities, depending on what length scale  $\lambda$  you look at.

## Chapter 29

# About the Book

I have made essentially all of this book available on the World Wide Web. The HTML version is at [www.av8n.com/physics/thermo/](http://www.av8n.com/physics/thermo/). The the PDF version is at [www.av8n.com/physics/thermo-laws.pdf](http://www.av8n.com/physics/thermo-laws.pdf).

The HTML was prepared from the  $\text{\LaTeX}$  sources using HeVeA, plus some custom post-processing.

Many readers have provided valuable feedback about the parts they liked and the parts that needed fixing. Many sections were written in response to readers' requests for more information on certain topics.





## Chapter 30

# References

Thanks to Carl Mungan for many helpful discussions.

1. John Denker,  
“Introduction to Probability”  
[www.av8n.com/physics/probability-intro.htm](http://www.av8n.com/physics/probability-intro.htm)
2. John Denker “Gas Laws”  
[www.av8n.com/physics/gas-laws.htm](http://www.av8n.com/physics/gas-laws.htm)
3. John Denker,  
“Partial Derivatives – Pictorial Representation”  
[www.av8n.com/physics/partial-derivative.htm](http://www.av8n.com/physics/partial-derivative.htm).
4. John Denker, “Thermodynamics and Differential Forms”  
[www.av8n.com/physics/thermo-forms.htm](http://www.av8n.com/physics/thermo-forms.htm)
5. C.W. Misner, K.S. Thorne, and J.A. Wheeler,  
**Gravitation** (W.H. Freeman, San Francisco, CA, 1973).
6. John Denker,  
“Conservative Flow and the Continuity of World-Lines”  
[www.av8n.com/physics/conservation-continuity.htm](http://www.av8n.com/physics/conservation-continuity.htm)
7. John Denker,  
“Welcome to Spacetime”  
[www.av8n.com/physics/spacetime-welcome.htm](http://www.av8n.com/physics/spacetime-welcome.htm)
8. Feynman, Leighton, and Sands,  
**The Feynman Lectures on Physics**.  
Volume I Chapter 4 deals with fundamental notions of conservation of energy, including the celebrated parable of Dennis and the blocks. Figure 44-8 on page 44-8 of Volume I illustrates heat engines in series. Chapter 46 (“Ratchet and Pawl”) discusses fluctuations.

9. John Denker,  
 “The Black-Box Approach”  
[www.av8n.com/physics/black-box.htm](http://www.av8n.com/physics/black-box.htm)
10. John Denker,  
 “Position, Displacement, Velocity, Momentum, Energy, and Mass in Spacetime”  
[www.av8n.com/physics/spacetime-momentum.htm](http://www.av8n.com/physics/spacetime-momentum.htm)
11. James Prescott Joule,  
 “On the Heat Evolved by Metallic Conductors of Electricity”  
 Philosophical Magazine, **19**, 260 (October, 1841).  
 Also in **Scientific Papers** 65.
12. Nina Byers,  
 “E. Noether’s Discovery of the Deep Connection Between Symmetries and Conservation Laws”  
<https://arxiv.org/pdf/physics/9807044>
13. John Denker, “The Twelve-Coins Puzzle”  
[www.av8n.com/physics/twelve-coins.htm](http://www.av8n.com/physics/twelve-coins.htm)
14. Ludwig Boltzmann,  
 “Über die Mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie”  
 Wiener Berichte **53** 195–220 (1866).
15. Ludwig Boltzmann,  
**Vorlesungen über Gastheorie** 2 vol.  
 J.A. Barth, Leipzig (1896 and 1898).
16. J W Gibbs,  
**Elementary Principles in Statistical Mechanics**  
 Ox Bow Press, 1902.
17. Claude Shannon, “A Mathematical Theory of Communication”  
 Bell System Technical Journal (1948).
18. John Denker,  
 “Kinetic Energy, Work, Momentum, Force times Time, and Force dot Distance”  
[www.av8n.com/physics/kinetic-energy.htm](http://www.av8n.com/physics/kinetic-energy.htm)
19. John Denker,  
 “Heat Capacity at a Phase Transition”  
[www.av8n.com/physics/phase-transition-heat.htm](http://www.av8n.com/physics/phase-transition-heat.htm)
20. Jonathan Swift, **Gulliver’s Travels**  
<http://www.gutenberg.org/dirs/etext97/gltrv10h.htm>
21. Feynman, **The Character of Physical Law**.

22. John Denker,  
“spreadsheet to demonstrate state functions and out-of-state functions” [out-of-state.xls](#)
23. John Denker,  
“Visualizing Non-Conservative Fields and Ungrady One-Forms”  
[www.av8n.com/physics/non-grady.htm](http://www.av8n.com/physics/non-grady.htm)
24. John Denker,  
“spreadsheet for plotting Carnot cycles” [../carnot-plot.xls](#)
25. John Denker,  
“Entropy – Increased by Stirring, Decreased by Observation”  
[www.av8n.com/physics/entropy-sim.htm](http://www.av8n.com/physics/entropy-sim.htm)
26. John Denker, “Dimensionless Units”  
[www.av8n.com/physics/dimensionless-units.htm](http://www.av8n.com/physics/dimensionless-units.htm)
27. John Denker, “Negative Temperatures”  
[www.av8n.com/physics/neg-temp.htm](http://www.av8n.com/physics/neg-temp.htm)
28. Feynman, **Statistical Mechanics: A Set of Lectures**  
ISBN 0201360764
29. Sadi Carnot, **Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance** (1824).
30. Benjamin Thompson  
“Heat is a Form of Motion: An Experiment in Boring Cannon”  
Philosophical Transactions **88** (1798).  
<http://www.chemteam.info/Chem-History/Rumford-1798.html>
31. B. A. Sherwood and W. H. Bernard,  
“Work and heat transfer in the presence of sliding friction”  
Am. J. Phys **52**(11) (1984).
32. John Denker,  
“Energy Flow – Principle of Virtual Work”  
[www.av8n.com/physics/isothermal-pressure.htm](http://www.av8n.com/physics/isothermal-pressure.htm)
33. E.T. Jayne, “The Gibbs Paradox,”  
in Smith, Erickson & Neudorfer (eds),  
**Maximum Entropy and Bayesian Methods** (1992).
34. D. Liu, M. Yue, J. Zhang, T.M. McQueen, J.W. Lynn, X. Wang, Y. Chen, J. Li, R.J. Cava,  
X. Liu, Z. Altounian and Q. Huang,  
“Origin and tuning of the magnetocaloric effect for the magnetic refrigerant MnFe(P1-xGex)”  
Physical Review **B79**, 014435 (2009).  
[http://www.nist.gov/public\\_affairs/techbeat/tb2009\\_0127.htm](http://www.nist.gov/public_affairs/techbeat/tb2009_0127.htm)

35. John Denker,  
“Wave Packets, Thermal or Otherwise”  
[www.av8n.com/physics/thermal-wave-packet.htm](http://www.av8n.com/physics/thermal-wave-packet.htm)
36. Rolf Landauer,  
“Irreversibility and Heat Generation in the Computing Process”  
IBM J. Res. Dev. 5, 183 (1961).  
<http://www.research.ibm.com/journal/rd/441/landauerii.pdf>
37. Charles H. Bennett,  
“The thermodynamics of computation – a review”  
International Journal of Theoretical Physics, **21** (12), 905–940 (December 1982).
38. Michael Baer,  
“A simple countable infinite-entropy distribution”  
<http://www.math.utah.edu/~erin/32103/HW25Solutions.pdf>
39. Thomas A. Moore and Daniel V. Schroeder,  
“A different approach to introducing statistical mechanics”  
[http://ajp.aapt.org/resource/1/ajpias/v65/i1/p26\\_s1](http://ajp.aapt.org/resource/1/ajpias/v65/i1/p26_s1)  
<http://physics.weber.edu/schroeder/statmech.pdf>
40. John Denker,  
“Spreadsheet to calculate energy, entropy, and temperature”  
./ds-de-temperature.gnumeric (my version)  
./ds-de-temperature.xls (should work, but may require fiddling, depending on your system details)
41. David L. Goodstein,  
**States of Matter** (1975).  
Republished by Dover (2002).
42. John Denker, “The Evil of Axes”  
[www.av8n.com/physics/axes.htm](http://www.av8n.com/physics/axes.htm)
43. John Denker, “How a Battery Works”  
[www.av8n.com/physics/battery.htm](http://www.av8n.com/physics/battery.htm)
44. R. K. P. Zia, Edward F. Redish, and Susan R. McKay,  
“Making Sense of the Legendre Transform”  
American Journal of Physics **77** (7) 614 (2009).  
<http://dx.doi.org/10.1119/1.3119512>  
<http://arxiv.org/abs/0806.1147>
45. John Denker, “Weird Terminology”  
[www.av8n.com/physics/weird-terminology.htm](http://www.av8n.com/physics/weird-terminology.htm)
46. Robert A. Alberty,  
“Use of Legendre Transforms in Chemical Thermodynamics”

- (IUPAC Technical Report) Pure Appl. Chem., **73** 8, pp. 1349 – 1380 (2001).  
<http://www.iupac.org/publications/pac/2001/pdf/7308x1349.pdf>
47. Carl E. Mungan,  
“Irreversible adiabatic compression of an ideal gas”  
<http://usna.edu/Users/physics/mungan/Publications/TPT4.pdf>
48. Wikipedia entry: “Functional (mathematics)”  
[https://en.wikipedia.org/wiki/Functional\\_\(mathematics\)](https://en.wikipedia.org/wiki/Functional_(mathematics))
49. Kittel and Kroemer,  
**Thermal Physics** (1980).  
This is far and away the most sensible thermo book I’ve seen.
50. Seth Lloyd,  
“Use of mutual information to decrease entropy: Implications for the second law of thermodynamics”  
Phys. Rev. A **39**, 5378–5386 (1989).
51. Wojciech Hubert Zurek,  
“Maxwell’s Demon, Szilard’s Engine and Quantum Measurements”  
<http://arxiv.org/pdf/quant-ph/0301076>
52. Koji Maruyama, Franco Nori, and Vlatko Vedral,  
“The Physics of Maxwell’s demon and information”  
Rev. Mod. Phys. **81**, 1 (2009).  
<http://arxiv.org/abs/0707.3400>
53. John Denker, “Introduction to Scaling Laws”  
[www.av8n.com/physics/scaling.htm](http://www.av8n.com/physics/scaling.htm)
54. John Denker, “Quantum Harmonic Oscillator”  
[www.av8n.com/physics/oscillator.htm](http://www.av8n.com/physics/oscillator.htm)
55. John Denker,  
“Stirling’s Approximation(s) for Factorials”  
[www.av8n.com/physics/stirling-factorial.htm](http://www.av8n.com/physics/stirling-factorial.htm)
56. W.R. Salzman,  
“Integrals Useful in the Kinetic Molecular Theory of Gases”  
<http://www.chem.arizona.edu/~salzmanr/480a/480ants/integral/integral.html>