Energy and Human Ambitions on a Finite Planet

Assessing and Adapting to Planetary Limits

Tom Murphy

Title
Energy and Human Ambitions on a Finite Planet

ISBN
978-0-578-86717-5

Author
Murphy, Thomas W, Jr

Publication Date
2021-03-11

DOI
10.21221/S2978-0-578-86717-5

Energy and Human Ambitions on a Finite Planet

Assessing and Adapting to Planetary Limits

Thomas W. Murphy, Jr.

University of California, San Diego

Published by eScholarship, University of California

ISBN: 978-0-578-86717-5 (electronic; DOI 10.21221/S2978-0-578-86717-5)
978-0-578-86696-3 (paperback)

Publisher: Initially released in electronic form in December 2020. Published by eScholarship, University of California. First published on eScholarship March 11, 2021. This version was produced on February 22, 2022. An electronic version of this book is freely available at https://escholarship.org/uc/item/9js5291m (this edition) or https://escholarship.org/uc/energy_ambitions for this and future editions. A printed book is available at-cost at https://www.lulu.com. To provide feedback and corrections, see https://tmurphy.physics.ucsd.edu/energy-text/.

Cover Photos: Space Shuttle photo courtesy of NASA/Jerry Cannon & George Roberts; Alpine lake on the Olympic Peninsula photo by Tom Murphy.

Colophon: This document was typeset using LaTeX and the kaobook class, which is built on a KOMA-Script foundation. The prevailing font is Palladio, 10 point.

This book is dedicated to Earth,
whose value is beyond measure.
May we learn to live within its bounds,
to the enduring benefit of all life.

Preface: Before Taking the Plunge

This book somewhat mirrors a personal journey that transformed my life, altered the way I look at the great human endeavor, and redefined my relationship to this planet. The transition that took a couple of decades for me is unlikely to be replicated for the reader in the short span of time it takes to absorb the content of this book. Nonetheless, the framework can be laid down so that readers might begin their own journeys and perhaps arrive at some profound realizations. This preface explains the approach and some overarching principles of the text.

We live in a physical world governed by physical law. Unlike the case for civil or criminal law, we are not even afforded the opportunity to break the laws of physics, except in fiction or entertainment. We do not need to create a physics police force or build physics jails or plead cases in front of some physics court. Nature provides perfect, automatic enforcement for free.

The domains of energy, the environment, economics, etc. are no exceptions, and can be put on a physical footing. It is worth exploring the emergent framework: reflecting on scale, efficiency, and thermodynamic limits of the human enterprise. By understanding the boundaries, we can begin to think about viable long-term plans in a way that too few are doing today. Thus far, heeding physical boundaries has not been necessary for the most part, as the scale of human endeavors has only recently become significant in a planetary context. We are now entering into a new reality: one in which our ambitions are on a collision course with natural limits on a finite planet. It is a slow-motion trajectory that has been apparent to some for an embarrassingly long time [1], but not yet acute enough to have grabbed the lasting attention of the majority.

The delirious ascent in energy and resource use witnessed over the past few centuries has been accomplished via the rapid, accelerating expenditure of a one-time inheritance of natural resources—a brief and singularly remarkable era in the long saga of human history. It has produced a dangerously distorted impression of what "normal" looks like on this planet. The fireworks show on display today is spectacular, fun, and inspirational, but also exceptionally unusual. Just as a meteorologist somehow born and trained within a 15-minute fireworks display likely cannot make useful predictions about weather and sky conditions over the next week, we are ill-equipped to intuitively understand what comes after the present phase. Luckily, science offers tools by which to transcend our narrow, warped perspectives, and can assist in discerning likely from wishful visions. The aim of this textbook is to set quantitative bounds on the present era as a way to better prepare for the possibility of a much different future. Our eventual success depends on serious attention to planetary limits.

This book is written to support a general education college course on energy and the environment. It was formulated as a physics course, but is written in the hope that it may also be accessible beyond this narrow setting. Physics is built on a mathematical foundation, and the domain of energy demands quantitative assessment. As a consequence, the book does not shy away from numbers. The math that is covered is presented in a way that aims to integrate intuition and the formality of equations. While math and quantitative elements are present throughout the book, Chapters 1, 3, and 6 are perhaps the most math-intense, featuring exponential functions, logarithms, and the lightest exposure to differential equations. But students need not master math beyond simple arithmetic operations, being able to rearrange equations, compute logarithms and exponentials, and raise a number to a power. Appendix A may serve as a useful math refresher.

An attempt is made to prevent students from equation-hunting, promoting instead development of a core understanding and intuition. This can require an adjustment on the part of students, who often treat equations as algorithmic tools to file away for use later when solving problems rather than as the embodiment of concepts to be internalized. Students often want a clear recipe so that when presented with a problem for homework, they can mimic a parallel example clearly laid out in the book. Doing so may be convenient and time-efficient, but short-circuits actual learning—bypassing the neural development that would accompany *mastering* the mental processes that are involved in solving a problem. Only the student can form these neural connections, and only through some struggle and effort. In this sense, learning is like climbing a hill: the only way to get to the top is by investing the effort to gain elevation—no shortcuts can bypass the inevitable climb.

Problems in this book are formulated to emphasize understanding the underlying concepts, rather than execution of a mathematical recipe. When students say they have math difficulties, it is usually not a problem carrying out the operations ($+$, $-$, $\times$, $\div$), but in *formulating* an approach. Therefore, the main difficulty is a conceptual one, but blamed on math because casting a problem in a mathematical framework forces a mastery of the conceptual underpinning: nowhere to hide. Given two numbers, should one divide or multiply them to get the answer sought? Resolving such questions requires a deeper understanding of the meaning *behind* the numbers in the problem (and associated units, often). By focusing on what the numbers represent and how they relate to each other, problems aim to build a more meaningful and permanent understanding of the content.

In soliciting feedback from students about problems, comments frequently pointed out that "Problem X used exactly the same math approach as Problem Y, so was redundant." This exposes a glaring difference in how students and instructors might view a problem. To the student providing such feedback, the problem seems to merely mirror an algorithm, devoid of contextual meaning. To the instructor, it is a window into a richer world: insight and personal ownership of the material is at stake. Problems are an opportunity to *learn*, as students are perhaps most actively engaged, mentally, when attempting to solve them. Instructors are trying to recreate their own learning experiences for students, through the imperfect mechanism of assigned work.

A similar revelation stems from comments that express the sentiment: "this problem has unnecessary information that is not required to solve the problem." Is the point to churn out a number, or to embed the result into a deeper context (i.e., *learn*)? It's a matter of **context** over **algorithm**. Context is where the *real* learning happens. It's where deep and lasting connections are made to the real world. The point is not to exercise a student's ability to perform mathematical operations, but to absorb a greater insight into the issue through its quantitative analysis. Math is like the airplane that delivers a skydiver to the jump. The jump/dive is the whole point, but the airplane is a necessary conveyance. When it comes time to jump, clinging to the familiar safety of the plane won't accomplish the goal. A student who bypasses the context for just the math operation has not embraced the intended experience and attendant mental growth.

The book's format sometimes weaves math and numbers into the text, which is unfamiliar to some students who are accustomed to clear delineation between math and text. Students are advised to approach sections containing mathematical developments by treating equations as statements of truth (within the appropriate context and assumptions) that help define and complete logical arguments. Or, think of equations as short-hand sentences that *encapsulate a concept*. Experts work to understand the concepts, by *reading* rather than memorizing equations. What is the equation trying to say? What truth does it impart? What relationships does it elucidate? Equations in the text are surrounded by sentences to help bring the equations to life as guides to intuition. Students who just want a step-by-step recipe to utilizing equations in an algorithmic autopilot mode are missing an opportunity to internalize ("own") the complete argument and concept. Once the **concept** is mastered, the equation is a natural consequence, and can be generated at need *from the concept* when solving a problem.

Most textbooks on energy and the environment for a general education audience stick to dry analyses of energy resources, their implementations, and the advantages and disadvantages of each. This textbook also does so, but is less reserved about providing contextual interpretations, like saying that resources such as waves, geothermal, tidal, or ocean currents are probably not worth serious attention, due to their small scale. In this sense, the book bears some resemblance to David MacKay's fabulous and inspirational *Sustainable Energy: Without the Hot Air* [2]. In fact, the decision to make this text fully available for free in electronic form as a PDF (available at https://escholarship.org/uc/energy_ambitions) was completely inspired by MacKay's first doing the same. The topic is too important to allow financial interests of a publisher to limit access. The price of the print version of this book—available at https://www.lulu.com/—is intended to cover production costs only. If using as a textbook for a course, consider its adaptation at https://www.kudu.com/physics_of_energy, allowing customization and unlimited editing of the course copy.

This text also differs from others in that it attempts to frame the energy story in a broader context of other limitations facing humanity in the form of growth (physical, economic, population) and also limitations of people themselves (psychology, political barriers). In the end, students are given quantitative guidance for adaptation and encouraged to find any number of ways to reduce consumption of resources as an effective hedge against uncertainty this century. Such advice is "bad for business" and may be seen as risky in a textbook subject to financial interests and coached by market analysis—which might explain why many textbooks come off as anodyne.

The tenor of this textbook might be characterized as being pessimistic, intoning that the coming century will present many difficulties that may not be dispatched by tidy "solutions," but instead borne with resigned adaptation. We entered this century graced by a few-hundred year run of mounting prosperity—and resulting sugar high—unlike anything previously experienced in human history, but may not exit this century in such a state of privilege. This sort of message may be off-putting to some (see also the Epilogue). But the stakes are important enough that it may be worth challenging assumptions and making students uncomfortable in a way that other texts might purposefully avoid. By the time students reach the end of Chapter 8, they are perhaps a little alarmed, and desperate to know "what's the answer?" Even though the book does not completely satisfy on that front—because it *can't*, in good faith—this is arguably exactly where an instructor would like students to be: attentive and eager. Having them carry the tension into the world is one way to help humanity take its challenges seriously and work to find a better way. Soothing their discomfort so they can emerge thinking it's all in hand is perhaps at best a wasted opportunity to create a better possible future for humanity, and at worst only contributes to humanity's fall by failing to light a fire equal to the challenge.

Tom Murphy
December, 2020
San Diego, CA

How to Use This Book

This version of the book—available for free in digital form at https://escholarship.org/uc/energy_ambitions—is prepared for electronic viewing: not differentiating between left and right pages, and thus not ideal for printing. A two-sided version better-suited for printing is also available at the aforementioned site. Most graphics in this version are vector-based and can be safely magnified to alarming proportions.

This book makes extensive use of margin space for notes,[1] citations, figures, tables, and captions. Hopefully, the utilization of margins leads to a smooth reading experience, preventing parenthetical, contextual information from interrupting the flow of primary text.

1: Numbered notes point to a specific location in the text.

Non-numbered notes are general asides relating to the paragraph.

References to figures and tables, glossary items, citations, chapters and sections, definitions, boxes, equations, etc. are hyperlinked within the electronic PDF document[2] allowing easy navigation, appearing as blue text. To help navigate in print versions, references to material outside the chapter also include page numbers. Note that the page numbering in this electronic version differs from that in the two-sided print versions.[3]

2: A link to the electronic document is in the first paragraph, above.

3: . . . due to the two-side rule that new parts or chapters start on right-side pages, often leaving empty (even-numbered) pages

The electronic version of this book is *far* easier to use once figuring out how to navigate "back" within the PDF viewer,[4] so that a link—or even several in a row—may be clicked/tapped, followed by a painless return to the starting point that may be hundreds of pages away. It is therefore **strongly recommended**[5] that you figure out how to navigate using your viewing platform. For instance, viewing the document in Adobe Acrobat on a Mac, the back and forward functions are accomplished with cursor/arrow keys as ⌘-← and ⌘-→ In Mac Preview, ⌘-[and ⌘-] go back and forward. Some mouses—especially those for gaming—have additional buttons that map to forward/backward navigation.

4: . . . much like the "back" button in a web browser

5: ⚠ Learn how to navigate "back" in your PDF viewer to make hyperlinks within the document more attractive.

Call-out boxes indicate places where a student might enhance their understanding by engaging in personal exploration. An information symbol (ⓘ) in the margin is occasionally used—mostly in Problems sections—to denote supplemental content that build useful contextual links to the real world but may otherwise obscure what the problem seeks. A caution symbol (⚠) appears when an argument is being put forth that is likely faulty or nuanced. Students should be careful to avoid literal acceptance of these points and work to understand the subtleties of the argument.

invitation to explore

ⓘ contextual information

⚠ Beware of dog.

This electronic version contains a Changes and Corrections section that lists any modifications since the original release—the relevant dates being provided on page ii. Within the text, a red square ■ marks the location of a change, the square being hyperlinked to the corresponding entry describing the change, which itself has a hyperlinked page reference

Try it! Click on the red square!

facilitating a quick return to the text. Corrections and other feedback can be left at https://tmurphy.physics.ucsd.edu/energy-text/.

A number of the examples (set apart in yellow boxes) are not posed in the form of a traditional question, but rather appear as a statement that captures a quantitative instance of the concept at hand. From this, students can construct multiple different questions that omit one piece of information and then solve for it using the remaining numbers. In this way, the example becomes a versatile guide to understanding. Note also the convention that $\approx$ is used to indicate "quantitatively very close to," while $\sim$ represents a less precise numerical value that might be read "in the neighborhood of," or "roughly."

After the primary presentation of material, an Epilogue attempts to put the tone of the textbook in context as it relates to the challenges ahead.

Appendices include supplemental information on math, chemistry, worthy tangents, and a partial answer key for problems. The answer key is constructed to discourage shortcuts. Rather than give direct numerical answers, for instance, it often provides a range within which the answer is expected to lie. In this sense, it acts more like an intuition transplant, guiding more likely correct analysis in a manner closer to the way an expert operates.[6] This approach allows students to at least catch any glaring errors in approach, like dividing rather than multiplying, for instance.

6: Experts often have a pretty good sense of the answer even before a detailed analysis, which serves as a way to identify possible mistakes and force a closer look.

The Bibliography at the end applies to the whole book. Clicking on a green reference number navigates to the reference entry. Each reference indicates the (hyperlinked) page or pages on which the reference is cited. Many of the entries contain hyperlinks to online resources. The ability to navigate "back" from these explorations is very useful.

References often provide convenient links to data collections in the form of Wikipedia pages. Primary references are easy enough to track from there.

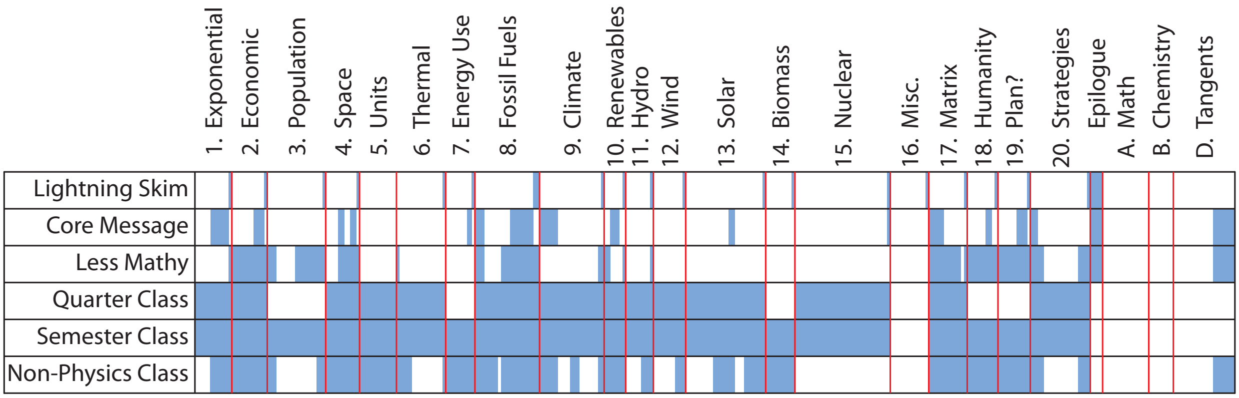
A Notation section after the bibliography provides an overview of physical constants and symbols used in the text, unit pre-factors, and a table of Greek letters.

A Glossary provides a collection of important terms encountered in the text. Hyperlinked words in the text (appearing blue) lead to the glossary, whose entries are also often internally linked to other glossary items. Additionally, page numbers where entries appear in the text are linked to easily navigate to them. The "back" navigation feature is extremely helpful in this context for returning from a rabbit hole of exploration.

Finally, a full alphabetical Index appears at the end to facilitate finding information in non-electronic versions of this text.

The graphic below reflects all chapters by page count,[7] suggesting a few different ways to approach this book. Just reading the "Upshot" for each chapter, and possibly the first few introductory paragraphs for each will convey a decent overview of the book's contents. The core message can also be picked out based on the guide below. The book has a definite quantitative slant, so it will not be possible to avoid math-laden

7: . . . and in some cases singles out individual sections by size and position within the chapter



sections completely, but the guide can help find the less numerically intense sections. Finally, different class formats might take various routes through the material, as suggested above.

Acknowledgments

I thank Brian Siana, David Crossley, and Barath Raghavan for extensive review and substantive suggestions on elements of the text. Allegra Swift offered tremendous help on permissions and open access publishing.

The first draft of this book was written in late 2019 and early 2020 for use in the Physics 12 general-education course on Energy and the Environment at UC San Diego in Winter quarter 2020. Students were presented with one chapter at a time in PDF form, lacking margin notes, having fewer figures and tables, and containing untested content. As part of weekly assignments, students were to read each chapter and offer substantive feedback on: confusing elements; missing content that would be interesting to include; ideas for additional graphics or tables; anything that seemed wrong or of questionable veracity; a critique of problems (for clarity, value, level of difficulty); and any typographical errors or wording suggestions. The voluminous feedback that resulted has been *extremely* valuable in shaping the current iteration of the book, so that each student in the class deserves acknowledgment here for hard work and keen insights.

In alphabetical order: J. Abaya, B. Arce, M. Beltran, D. Bothe, P. Buljat, T. Chung, J. Cordero, B. Espinoza Diaz, G. Fu, J. Funes, C. Goodale, T. Huang, K. Ito, A. Jhaveri, D. Kim, S. Mahlman, J. Mckinley, J. Murgo, H. Musa, J. Nam, J. Ngo, K. Nguyen, H. Ordonez, G. Osuna, C. Peach, H. Potts, P. Puno, E. Ramos, J. Song, A. Strube, D. Sun, N. Teuton, E. Tian, Q. C. Tu, Z. Tuazon, J. Wen, Z. Xu, A. Yoshinaga, S. Zweifler.

Contents

Preface: Before Taking the Plunge v

How to Use This Book viii

Contents xi

I Setting the Stage: Growth and Limitations 1

1 Exponential Growth 2
- 1.1 Bacteria in a Jar . . . 2
  - 1.1.1 Exponential Math . . . 4
- 1.2 Energy Extrapolation . . . 7
- 1.3 Thermodynamic Absurdity . . . 10
- 1.4 Upshot: Limits to Growth . . . 13
- 1.5 Problems . . . 13

2 Economic Growth Limits 17
- 2.1 Historical Coupling . . . 17
- 2.2 Decoupling and Substitution . . . 20
- 2.3 Forced Economic Limits . . . 24
- 2.4 No-Growth World . . . 26
- 2.5 Upshot: Growth Will End . . . 27
- 2.6 Problems . . . 28

3 Population 30
- 3.1 Population History . . . 31
- 3.2 Logistic Model . . . 33
  - 3.2.1 Overshoot . . . 35
  - 3.2.2 Logistic Projection . . . 38
- 3.3 Demographic Transition . . . 38
  - 3.3.1 Geographic Considerations . . . 41
  - 3.3.2 Transition Cost . . . 44
- 3.4 Touchy Aspects . . . 46
  - 3.4.1 It's Personal . . . 46
  - 3.4.2 Population Policy . . . 47
- 3.5 Upshot: It Depends on Us . . . 48
- 3.6 Problems . . . 48

4 Space Colonization 54
- 4.1 Scale of Space . . . 54
- 4.2 The Wrong Narrative . . . 58
- 4.3 A Host of Difficulties . . . 59
- 4.4 Exploration's Role . . . 61
- 4.5 Upshot: Earth First . . . 63

4.6 Problems . . . 64

II Energy and Fossil Fuels **67**

5 Energy and Power Units **68**
5.1 Energy (J) . . . 68
5.2 Energy Forms & Conservation . . . 69
5.3 Power (W) . . . 71
5.4 Kilowatt-hour (kWh) . . . 72
5.5 Calories (kcal) . . . 73
5.6 British Thermal Unit (Btu) . . . 75
5.7 Therms . . . 76
5.8 Electrical Power . . . 76
5.9 Electron Volt (eV) . . . 78
5.10 Light Energy . . . 79
5.11 Upshot on Units . . . 79
5.12 Problems . . . 80

6 Putting Thermal Energy to Work **84**
6.1 Generating Heat . . . 84
6.2 Heat Capacity . . . 85
6.3 Home Heating/Cooling . . . 86
6.4 Heat Engines . . . 88
6.4.1 Entropy and Efficiency Limits . . . 90
6.5 Heat Pumps . . . 95
6.5.1 Consumer Metrics . . . 97
6.6 Upshot on Thermal Energy . . . 99
6.7 Problems . . . 99

7 The Energy Landscape **102**
7.1 The Annual Energy Review . . . 102
7.1.1 Energy Flow . . . 103
7.1.2 Source and Sector . . . 104
7.1.3 Detailed Mix . . . 106
7.1.4 Energy Trends . . . 107
7.2 Global Energy . . . 108
7.2.1 U.S. Global Share . . . 110
7.3 Upshot: Go to the Source . . . 111
7.4 Problems . . . 111

8 Fossil Fuels **114**
8.1 The Most Important Plot Ever . . . 114
8.2 Overview: Coal, Oil, and Gas . . . 117
8.2.1 Coal . . . 117
8.2.2 Petroleum (Oil) . . . 118
8.2.3 Natural Gas . . . 120
8.3 Chemical Energy . . . 121

8.4 Fossil Fuel Pros and Cons . . . 122
8.4.1 Fossil Fuels are Amazing . . . 122
8.4.2 Fossil Fuels are Terrible . . . 123
8.4.3 On Balance? . . . 125
8.5 The Future of Fossil Fuels . . . 125
8.5.1 Scenarios . . . 125
8.5.2 Timescales . . . 126
8.5.3 Clues in the Data . . . 128
8.5.4 Geopolitics . . . 131
8.6 Upshot: Good, Bad, Limited . . . 132
8.7 Problems . . . 134

9 Climate Change **138**
9.1 The Source of CO_2 . . . 138
9.1.1 CO_2 Measurements . . . 139
9.1.2 CO_2 Expectations . . . 140
9.1.3 Chief Contributors . . . 143
9.2 Warming Mechanism . . . 143
9.3 Possible Trajectories . . . 148
9.4 Consequences . . . 150
9.4.1 Heating Up . . . 151
9.4.2 Heating Earth's Skin . . . 152
9.4.3 Sea Level Rise . . . 155
9.5 What Can Be Done? . . . 157
9.6 Upshot: A Serious Problem . . . 158
9.7 Problems . . . 158

III Alternative Energy **163**

10 Renewable Overview **164**
10.1 The Players . . . 164
10.2 Alternatives vs. Renewables . . . 165
10.3 Renewable Energy Budget . . . 167
10.4 Renewable Snapshot . . . 169
10.5 Upshot: Our Path Forward . . . 170
10.6 Problems . . . 171

11 Hydroelectric Energy **173**
11.1 Grav. Potential Energy . . . 173
11.1.1 Comparison to Other Forms . . . 174
11.2 Hydroelectric Energy . . . 175
11.2.1 Theoretical Potential . . . 177
11.3 Hydropower in the U.S. . . . 178
11.4 Global Hydropower . . . 180
11.5 Upshot: Pros and Cons . . . 180
11.6 Problems . . . 181

12 Wind Energy **184**
12.1 Kinetic Energy . . . 184
12.2 Wind Energy . . . 185
12.2.1 Wind Turbines . . . 187
12.3 Wind Installations . . . 191
12.4 Upshot: Not Overblown . . . 193
12.5 Problems . . . 194

13 Solar Energy **197**
13.1 The Energy of Light . . . 197
13.2 The Planck Spectrum . . . 199
13.3 Photovoltaics . . . 201
13.3.1 Theoretical Efficiency . . . 202
13.4 Insolation . . . 206
13.5 Incredible Solar Potential . . . 211
13.6 Residential Considerations . . . 213
13.6.1 Configurations . . . 213
13.6.2 Sizing and Cost . . . 213
13.7 Photovoltaic Installations . . . 215
13.7.1 Pros and Cons of PV . . . 217
13.8 Solar Thermal . . . 218
13.8.1 Passive Solar Heat . . . 218
13.8.2 Solar Thermal Electricity . . . 219
13.8.3 Pros and Cons of ST . . . 221
13.9 Upshot for Solar . . . 221
13.10 Problems . . . 222

14 Biological Energy **227**
14.1 Photosynthesis . . . 227
14.2 Biomass . . . 229
14.3 Biofuels . . . 230
14.3.1 EROEI . . . 231
14.3.2 EROEI of Biofuels . . . 234
14.4 Upshot: Nature's Burn . . . 235
14.5 Problems . . . 236

15 Nuclear Energy **239**
15.1 The Nucleus . . . 239
15.2 Radioactive Decay . . . 242
15.3 Mass Energy . . . 246
15.4 Fission . . . 249
15.4.1 The Basic Idea . . . 250
15.4.2 Chain Reaction . . . 251
15.4.3 Fission Accounting . . . 251
15.4.4 Practical Implementations . . . 255
15.4.5 Nuclear Waste . . . 260
15.4.6 Weapons and Proliferation . . . 262
15.4.7 Nuclear Safety . . . 263

15.4.8 Pros and Cons of Fission . . . 264
15.5 Fusion . . . 264
15.5.1 Fuel Abundance . . . 266
15.5.2 Fusion Realities . . . 267
15.5.3 Pros and Cons of Fusion . . . 268
15.6 Upshot on Nuclear . . . 269
15.7 Problems . . . 269

16 Small Players **275**
16.1 Geothermal Energy . . . 275
16.1.1 Sustainable Harvest . . . 276
16.1.2 Geothermal Depletion . . . 277
16.1.3 Geothermal Reality . . . 278
16.2 Tidal Capture . . . 280
16.3 Ocean Currents . . . 281
16.4 Wave Energy . . . 282
16.5 Hydrogen . . . 286
16.6 Upshot on Small Players . . . 286
16.7 Problems . . . 287

17 Comparison of Alternatives **289**
17.1 Alternative Energy Matrix . . . 289
17.2 Individual Tally . . . 294
17.3 Student Rankings . . . 299
17.4 Upshot: Fossil Fuel Gap . . . 300
17.5 Problems . . . 301

IV Going Forward **303**

18 Human Factors **304**
18.1 Personality . . . 304
18.1.1 Consequences and Coping . . . 307
18.2 Policy vs. Individual Action . . . 309
18.3 The Energy Trap . . . 310
18.4 Fermi Paradox Explained? . . . 312
18.5 Upshot on Humanity . . . 313
18.6 Problems . . . 313

19 A Plan Might Be Welcome **316**
19.1 No Master Plan . . . 316
19.1.1 The Growth Imperative . . . 318
19.2 No Prospect for a Plan . . . 319
19.2.1 Who Makes the Plan? . . . 321
19.3 Economic Regimes . . . 322
19.3.1 Steady State Economy . . . 324
19.4 Upshot on the Plan . . . 326
19.5 Problems . . . 326

20 Adaptation Strategies **328**
20.1 Awareness . . . 328
20.2 Communication . . . 330
20.2.1 Predicament, not Problems . . . 332
20.3 Guidelines for Adaptation . . . 332
20.3.1 Overall Framing . . . 333
20.3.2 Energy Assessment Rules . . . 334
20.3.3 Quantitative Footprint . . . 337
20.3.4 Dietary Energy . . . 339
20.3.5 Flexitarianism . . . 341
20.3.6 Discretionary Summary . . . 342
20.4 Values Shifts . . . 343
20.5 Flexibility in Uncertainty . . . 344
20.6 Upshot on Strategies . . . 346
20.7 Problems (Predicaments?) . . . 347

Epilogue **350**

Image Attributions **354**

Changes and Corrections **356**

V Appendices **358**

A Math and Equations **359**
A.1 Relax on the Decimals . . . 359
A.2 Forget the Rules . . . 361
A.3 Areas and Volumes . . . 362
A.4 Fractions . . . 363
A.5 Integer Powers . . . 365
A.6 Fractional Powers . . . 366
A.7 Scientific Notation . . . 367
A.8 Equation Hunting . . . 368
A.9 Equation Manipulation . . . 369
A.10 Units Manipulation . . . 370
A.11 Just the Start . . . 373

B Chemistry Primer/Refresher **375**
B.1 Moles . . . 375
B.2 Stoichiometry . . . 376
B.3 Chemical Energy . . . 379
B.4 Ideal Gas Law . . . 381

C Selected Answers **383**

D Alluring Tangents **392**
D.1 Edge of the Universe . . . 392

D.2 Cosmic Energy Conservation . . . 393
D.2.1 Cosmological Exception . . . 394
D.2.2 Convoluted Conservation . . . 395
D.2.3 Lost to Heat . . . 396
D.3 Electrified Transport . . . 397
D.3.1 Airplanes . . . 398
D.3.2 Shipping . . . 399
D.3.3 Long-haul trucking . . . 399
D.3.4 Buses . . . 400
D.3.5 Passenger Cars . . . 400
D.3.6 Wired Systems . . . 402
D.3.7 Collected Efficiencies . . . 402
D.4 Pushing Out the Moon . . . 402
D.5 Humanity's Long View . . . 404
D.5.1 Success vs. Failure . . . 404
D.5.2 Sustainable Living . . . 406
D.5.3 Time to Grow Up . . . 407
D.5.4 Frameworks . . . 407
D.6 Too Smart to Succeed? . . . 408
D.6.1 Evolution's Biggest Blunder? . . . 411

Bibliography **412**

Notation **420**

Glossary **422**

Alphabetical Index **439**

Part I

Setting the Stage: Growth and Limitations

We arrived late.
All the good stuff was gone.
Oh, what we would have paid.
But no amount of money could bring it back.

1 Exponential Growth

Humans have amazing strengths, but also significant weaknesses. Chief among them, perhaps, is our collective difficulty in grasping the mathematical consequences of exponential growth.[1] This is an ironic state, given that our economic and political goals are often geared explicitly to support continued growth. The degree to which an expectation and desire for continued growth is woven into our society makes it important to examine the phenomenon carefully, so that we might avoid building upon a shaky foundation. In this chapter, we explore the general nature of exponential growth, in order to understand the *impossibility* of its long-term continuance by way of exposing various absurd consequences that uninterrupted growth prescribes. The upshot[2] is that our societal framework eventually must face a mandatory departure from the current model—a piece of knowledge we should all lodge into the backs of our minds. Subsequent chapters will address applications to economic and population growth—including more realistic logistic growth curves, then pivot toward nailing down limits imposed by our finite planet.

1.1 Bacteria in a Jar 2
Exponential Math 4
1.2 Energy Extrapolation 7
1.3 Thermodynamic Absurdity . 10
1.4 Upshot: Limits to Growth . . 13
1.5 Problems 13

1: . . . a nod to Al Bartlett, who worked to raise awareness about exponential growth

2: The word "upshot" means final result or bottom-line. Each chapter has an Upshot at the end.

1.1 Bacteria in a Jar

One hallmark of exponential growth is that the time it takes to double in size, or the doubling time, is constant. An important and convenient concept we will repeatedly use in this chapter is the rule of 70:

Note that *any* growth, however slow, can be characterized by a doubling time, even if the process does not involve discrete steps of doubling.

> **Definition 1.1.1** ***Rule of 70****: The doubling time associated with a percentage growth rate is just 70 divided by the percentage rate. A 1% growth rate doubles in 70 years, while a 2% rate doubles in 35 years, and a 10% rate doubles in 7 years. It also works for other timescales: if pandemic cases are increasing at a rate of 3.5% per day, the doubling time is 20 days.*

NGC 253 photo credit: Dylan O'Donnell.

We will see how the rule of 70 arises mathematically later in this chapter. But first, it is more important to understand the *consequences*. To make the math simple, let's say that a town's size doubles every 10 years (which by the rule of 70 corresponds to a 7% growth rate, incidentally). Starting in the year 1900 at 100 residents, we expect town population to be 200 in 1910, 400 in 1920, 800 in 1930, eventually climbing to over 100,000 by the year 2000 (see Table 1.1). Unabated 7% growth would result in the town reaching the current world population just 260 years after the experiment began.

Table 1.1: Example 7% growth progression.

Year	Population
1900	100
1910	200
1920	400
1930	800
1940	1,600
⋮	⋮
2000	102,400

But let's explore an example that often reveals our faulty intuition around exponential growth. Here, we imagine a jar rich in resources, seeded with just the right number of bacteria so that if each bacterium splits every 10 minutes, the jar will become full of bacteria in exactly 24 hours. The experiment starts right at midnight. The question is: at what time will the jar be half full?

10 minutes is perhaps a little fast for biology, but we're looking for easy understanding and picking convenient numbers. In practice, 20–30 minutes may be more realistic. We will also ignore deaths for this "toy" example, although the net effect only changes the rate and not the overall behavior.

Think about this on your own for a minute. Normal intuition might suggest a half-full jar at noon—halfway along the experiment. But what happens if we work backwards? The jar is full at midnight, and doubles every ten minutes. So what time is it half full?

The answer is one doubling-time before midnight, or 11:50 PM. Figure 1.1 illustrates the story. At 11 PM, the jar is at one-64th capacity, or 1.7% full. So, for the first 23 of 24 hours, the jar looks basically empty. All the action happens at the end, in dramatic fashion.

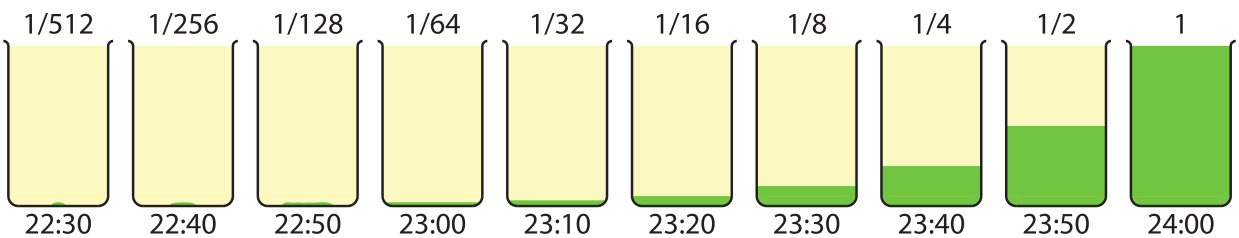


Figure 1.1: The last 90 minutes in the sequence of bacteria (green) growing in a jar, doubling every 10 minutes. For the first 22.5 hours, hardly anything would be visible. Note that the upward rise of green "bars" makes an exponential curve.

Now let's imagine another illustrative scenario in connection with our jar of bacteria. The time is 11:30 PM: one-half hour before the end. The jar is one-eighth full. A thoughtful member of the culture projects the future and decides that more uninhabited resource-laden jars must be discovered in short order if the culture is to continue its trajectory. Imagine for a second the disbelief expressed by probably the vast majority of other inhabitants: the jar is far from full, and has served for 141 generations—a seeming eternity. Nonetheless, this explorer returns reporting three other equal-sized food-filled jars within easy reach. A hero's welcome! How much longer will the culture be able to continue growing? What's your answer?

The population doubles every ten minutes. If the original jar is filled at 12:00, the population doubles to fill the second jar by 12:10. Another doubling fills all four by 12:20. The celebration is short-lived.

Now we draw the inevitable parallels. A planet that has served us for countless generations, and has seemed effectively infinite—imponderably large—makes it difficult for us to conceive of hitting limits. Are we half-full now? One-fourth? One-eighth? All three options are scary, to different degrees. At a 2% rate of growth (in resource use), the doubling time is 35 years, and we only have about a century, even if at 1/8 full right now.[3]

3: If we're at 1/8 right now and double every 35 years, we will be at 1/4 in 35 years, 1/2 in 70 years, and full in 105 years.

In relation to the bacteria parable, we've already done a fair bit of exploring. We have no more jars. One planet rhymes with jars, but it is hostile to human life, has no food, and is not within easy reach. We have no meaningful outlet.[4] And even if we ignore the practical hardships, how much time would a second planet buy us anyway for uninterrupted growth? Another 35 years?

4: Chapter 4 addresses space realities.

1.1.1 Exponential Math

Box 1.1: Advice on Reading Math

This section is among the most mathematically sophisticated in the book. Don't let it intimidate you: just calmly take it in. Realize that exponential growth obeys an unchanging set of rules, and can be covered in just a few pages. Your brain can absorb it all if you give it a chance. Read paragraphs multiple times and find that each pass can add to your comprehension. Equations are just shorthand sentences[5] capturing the essence of the concepts being covered, so rather than reading them as algorithms to file and use later when solving problems, work to comprehend the *meaning* behind each one and its reason for being a part of the development. In this way, what follows is not a disorganized jumble, recklessly bouncing between math and words, but one continuous development of thought expressed in two languages at once. The Preface offers additional thoughts related to this theme, and Appendix A provides a math refresher.

Experts habitually read complicated passages multiple times before the material sinks in. Maybe it's this calm habit that turns them into experts!

5: Unlike words/language, the symbols chosen for equations are just labels and carry no intrinsic meaning—so electing to use x, n, t, b, M, etc. reflect arbitrary choices and can be substituted at will, if done consistently. The content is in the *structure* of the equation/sentence.

The essential feature of exponential growth is that the scale goes as the *power* of some base (just some number) raised to the time interval. In the doubling sequence, we start at 1× the original scale, then go to 2×, then 4×, then 8×, etc. At each time interval, we multiply by 2 (the base). After 5 such intervals, for instance, we have $2 \times 2 \times 2 \times 2 \times 2$, or $2^5 = 32$. More generally, after n doubling times, we have increased by a factor of 2^n, where 2 is the base, and n is the number of doubling times. We might formalize this as

$$M = 2^n = 2^{t/t_2}, \quad (1.1)$$

where M represents the multiplicative scale, t is the elapsed time, and t_2 is the symbol we choose to represent the doubling time—so that $n = t/t_2$ is just "counting" the number of doubling times.

For instance, doubling has $M = 2$, tripling has $M = 3$, and increasing by 29% would mean $M = 1.29$.

Box 1.2: Interest Example

The same process happens in a bank account accumulating interest. Let's consider that you deposit $100 into a bank account bearing 2% annual interest. At the end of one year, you'll have $102, which is 1.02 times the original amount. For the next year, it's 1.02 times $102, or $104.04, which is the original $100 times 1.02×1.02. Then in three years it will be $106.18, or $100 times 1.02^3. Having sussed out the pattern, after 35 years it would be $100 times 1.02^{35}, which happens to come to $199.99. Notice that doubling in 35 years at 2% exactly obeys the rule of 70. Table 1.2 summarizes this example.

Table 1.2: Interest example (2% rate).

year	b^n	dollars
0	1.00	$100.00
1	1.02	$102.00
2	1.0404	$104.04
3	1.0612	$106.12
⋮	⋮	⋮
10	1.2190	$121.90
⋮	⋮	⋮
35	1.9999	$199.99

The pattern—whether doubling, or applying interest as in Box 1.2—is that we multiply a chain of the same number, the base, over and over. This is the same as raising the base to some power—the power equaling how many times the base appears in the chain to get our overall factor. Therefore, if we designate the base as b and the number of times it appears as n, we have

$$M = b^n. \tag{1.2}$$

Now we're going to play a math trick that will help us compute various useful attributes of growth. The exponential and natural logarithm are inverse functions, each undoing the other. So $\ln(e^x) = x$ and $e^{\ln x} = x$. We can use this trick to express the number 2 as $e^{\ln 2}$, or any base number $b = e^{\ln b}$. For the special case of $b = 2$ (doubling), we then have:

By "trick," we do not mean to imply anything devious or untoward: just a cute manipulation that can bring additional insight or make something easier.

Try it on a calculator for several examples of b that you concoct (make it real for yourself!).

$$M = 2^{t/t_2} = \left(e^{\ln 2}\right)^{t/t_2} = e^{t \ln 2/t_2}, \tag{1.3}$$

where we started with Eq. 1.1, re-expressed the number 2, and then applied the rule that raising a power to another power is the same as multiplying the powers to form a single one.[6] By employing such tricks, we could cast any base to a power, like b^x as some exponential function $e^{x \ln b}$, and thus can transform any "power" relationship into an exponential using base $e \approx 2.7183$. Casting Eq. 1.2 in this form:

6: As an example, think of $(5^3)^4$ as $(5\times5\times5)^4 = (5\times5\times5)\times(5\times5\times5)\times(5\times5\times5)\times(5\times5\times5)$, which is just 12 fives multiplied, or 5^{12}. So we effectively just multiplied the two exponents—3 and 4—to get the 12. It always works. Often, one need not memorize math rules: quick experimentation reveals how and why it works.

$$M = b^n = e^{n \ln b}. \tag{1.4}$$

If we want to go backwards, and compute the time to reach a certain M factor, we can take the natural logarithm of both sides to learn that

$$\ln M = n \ln b, \tag{1.5}$$

so that the number of applications of base, b, needed to achieve multiplicative factor M is found by solving the equation above for n, in which case we get: $n = \ln M / \ln b$.

Example 1.1.1 The time it would take to increase by a factor of 1,000 ($M = 1000$) at a rate of 1.07 (annual growth rate of 7%; $b = 1.07$) is $n = \ln M / \ln 1.07 = 102$ years.

The same result happens if using log instead of ln: try it!

The rule of 70 can be recovered[7] by setting the multiplicative factor, M, to 2. Comparing to interest accumulation described by $(1+p)^t$, where p is the annual interest (0.02 for 2%, e.g.) and t is the number of years, Eq. 1.4 can be re-expressed by substituting $b = 1+p$ and $n = t$ as the number of years, then equating the result to the doubling time representation in Eq. 1.3 to form

$$M = e^{t \ln(1+p)} = e^{t \ln 2 / t_2}. \tag{1.6}$$

7: What follows is a high-brow symbolic approach, but the same effective result can be achieved by setting $M = 2$ in Eq. 1.5 and solving for n.

From this expression, we can gather that $\ln(1+p) = \ln 2 / t_2$ by equating the exponents, and then see that the doubling time, t_2, can be solved as

$$t_2 = \ln 2 / \ln(1+p). \tag{1.7}$$

For small values of p (much smaller than 1), the natural log of $1+p$ is approximately p. In other words, when $p = 0.02$, $\ln 1.02 \approx 0.02 \approx p$. This is part of the reason why we chose e as our base, as it is mathematically "natural." Since $\ln 2 \approx 0.693 \approx .70$, the doubling time, t_2, is approximately 70 divided by the annual growth rate, p, in percent. So the reason it's a rule of 70 for doubling (and not a rule of 60 or 80) is basically because the natural log of 2 (representing doubling) is roughly 0.70.

Try it yourself to verify on a calculator, by sticking in various small amounts for p.

Example 1.1.2 To tie some things together, let's look at a quantitative case that can be used to validate how various pieces relate to each other. We will describe a 5% annual growth rate.

The rule of 70 (Definition 1.1.1) indicates a 14 year doubling time, so that we could define t_2 appearing in Eqs. 1.1, 1.3, 1.6, and 1.7 to be 14 years. Calculating exactly using Eq. 1.7 yields 14.2 years.

To evaluate growth in 10 years, we could use Eq. 1.1 with $t = 10$ and $t_2 = 14.2$ to suggest $M = 1.63$, meaning a 63% increase in size (1.63 times as large as at the start). Or we could apply Eq. 1.2 using $b = 1.05$ and $n = 10$ to get the exact same result. Note that we have freedom to define the base as 1.05 or 2, and the corresponding number of steps (n) as 10 or $t/t_2 = 0.704$, respectively, and get the same answer. In terms of the exponential form in Eq. 1.4, either pair of b and n produces $e^{0.488}$.

If we wanted to "work backwards" and ask when the amount is 3 times the original ($M = 3$), we could use Eq. 1.5 to find that n is 22.5 steps at $b = 1.05$ (thus 22.5 years, since this base is the yearly increase). Had we used $b = 2$, we would compute $n = 1.58$, meaning that the scale would reach 3× after 1.58 doubling times, or $1.58 t_2 = 22.5$ years.

Don't view this as a recipe for solving problems, but as a way to romp through the section and help piece it together.

More generally, we are not confined to any particular base, b, having just seized upon two convenient and relevant possibilities. If we wanted $b = 10$, we would have $n = 0.211$, for example. In this case, the interpretation is that our ten-year point is 21.1% of the way to a factor-of-ten multiplication, so that 47.4 years at 5% growth results in a factor of 10 growth.

We can check the result using Eq. 1.6 by putting in $t = 22.5$ and $p = 0.05$ or $t_2 = 14.2$ in the latter form.

1.2 Exponential Energy Extrapolation

Having established some basic principles of exponential growth, it's time for a first look at how we can use the math to argue about limits to our expectations. We'll concentrate on energy use. The United States Energy Information Administration (EIA) provides information on energy use from 1949 to the present. An appendix (E1: [3]) presents an approximate account of energy use from 1635–1945. Figure 1.2 displays the more recent portion of this history.

Lacking comparable data for the world, we use U.S. data simply to illustrate the more broadly applicable global growth trend. Even countries far behind are *growing* energy use—often faster than the 3% characteristic of U.S. history.

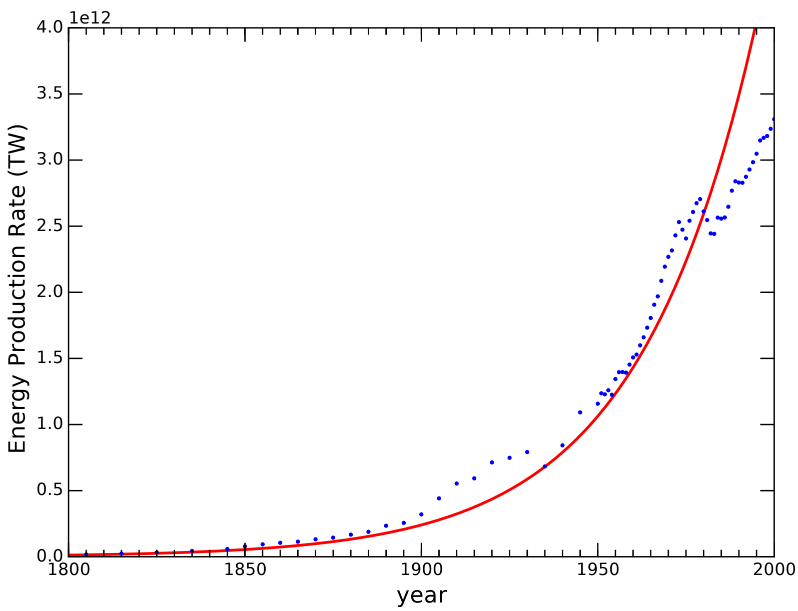


Figure 1.2: U.S. energy over 200 years, showing a dramatic rise due almost entirely to fossil fuels. The red curve is an exponential fit tuned to cover the broader period shown in Figure 1.3.

Note that the energy rate at the left edge of Figure 1.2 becomes almost invisibly small. Presenting the data on a logarithmic plot, as in Figure 1.3, we can better see the entire trajectory. On such a plot, exponentials become straight lines. The trend is remarkably consistent with an exponential (red line) for most of the history, at a rate just shy of 3% per year. Note that this total effect includes population growth, but population has not grown as fast as energy, so that per-capita energy has also risen. This makes sense: our lives today are vastly more energetically rich than lives of yesteryear, on a per-person basis.

The astute reader might note a departure from the exponential fit in recent years. This only reinforces the primary point of this chapter that sustaining exponential growth indefinitely is absurd and will not happen. If growth is destined to stop, perhaps we are beginning to experience its limits well before the theoretical timescales developed in this chapter.

Having established that energy growth over the past several centuries is well-described by an exponential, we can explore the implications of continuing this trend forward. Starting at a present-day global energy production rate of 18×10^{12} Watts (18 TW), we adopt a convenient growth rate of 2.3% per year for this exercise. We pick this for two reasons: 1) it is more modest than the historical trend, so will not over-exaggerate the

Watts is a unit of *power*, which is a rate of energy. Chapter 5 will cover the concept and units more thoroughly.

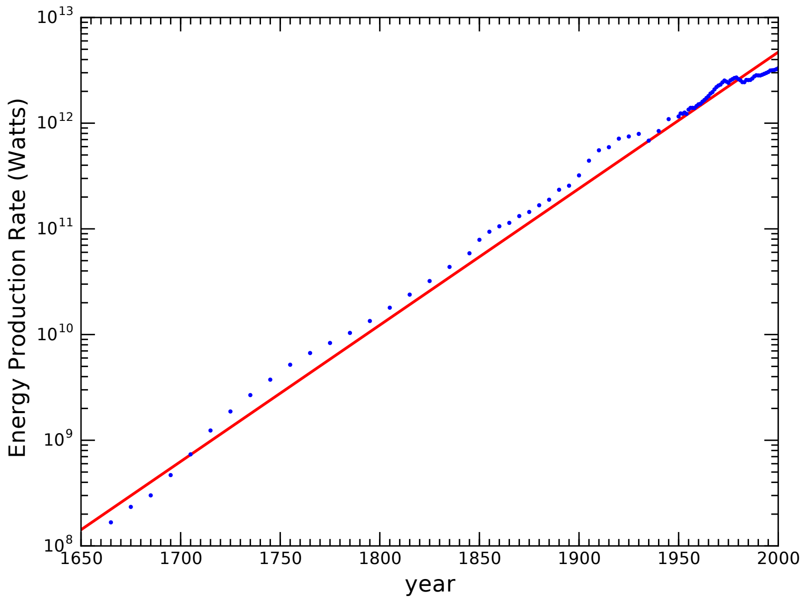


Figure 1.3: Energy trajectory in the U.S. over a long period. The red line is an exponential at a 2.9% growth rate, which appears linear on a logarithmic plot.

result; 2) this rate produces the mathematical convenience of a factor of 10 increase every century.[8]

What follows is a flight of fancy that quickly becomes absurd, but we will chase it to staggering levels of absurdity just because it is fun, instructive, and mind-blowing. Bear in mind that what follows should not be taken as predictions[9] of our future: rather, we can use the absurdity to predict how our future will *not* look!

The sun deposits energy at Earth's surface at a rate of about 1,000 W/m^2 (1,000 Watts per square meter; we'll reach a better understanding for these units in Chapter 5). Ignoring clouds, the projected area intercepting the sun's rays is just $A = \pi R_{\oplus}^2$, where $R_{\oplus}$ is the radius of the earth, around 6,400 km. Roughly a quarter of the earth's surface is land, and adding it all up we get about 30×10^{15} W hitting land. If we put solar panels on every square meter of land converting sunlight to electrical energy at 20% efficiency,[10] we keep 6×10^{15} W. This is a little over 300 times the current global energy usage rate of 18 TW. What an encouraging number! Lots of margin. How long before our growth would get us here? After one century, we're 10 times higher, and 100 times higher after two centuries. It would take about 2.5 centuries (250 years) to hit this limit. Then no more energy growth.

But wait, why not also float panels on all of the ocean, and also magically improve performance to 100%? Doing this, we can capture a whopping 130×10^{15} W, over 7,000 times our current rate. Now we're talking about maxing out in just under 400 years. Each factor of ten is a century, so a factor of 10,000 would be four factors of ten (10^4), taking four centuries.

So within 400 years, we would be at the point of using every scrap of

8: Fundamentally, this relates to the fact that the natural log of 10 is 2.30. The analog of Eq. 1.7 using 10 in place of 2 and $p = 0.023$ for 2.3% growth rate will produce a factor-of-ten timescale $t_{10} \approx 100$ years.

9: ⚠ Do not interpret this section as predictions of how our future *will* go.

Approximate numbers are perfectly fine for this exercise.

10: 20% is on the higher end for typical panels.

The merits of various alternative energy sources will be treated in later chapters, so do not use this chapter to form opinions on the usefulness of solar power, for instance.

⚠ In defiance of physical limits.

10,000 is not too different from 7,000, and the "rounding up" helps us conveniently make sense of the result, since a factor of 10,000 is easier to interpret as four applications of 10×, and thus 400 years.

solar energy hitting the planet at 100% efficiency. But our planet is a tiny speck in space. Why not capture *all* the light put out by the sun, in a sphere surrounding the sun (called a Dyson sphere; see Box 1.3)? Now we're talking some real power! The sun puts out 4×10^{26} W. If it were a light bulb, this would be its label (putting the 100 W standard incandescent bulb to shame). So the number is enormous. But the math is actually pretty easy to grasp.[11] Every century gets another factor of ten. To go from where we are now (18×10^{12} W) to the solar regime is about 14 orders-of-magnitude. So in 1,400 years,[12] we would be at 18×10^{26} W, which is about 4.5 times the solar output. Therefore we would use the entire sun's output in a time shorter than the 2,000-year run of our current calendar.

11: Math becomes easier if you blur your vision a bit and do not demand lots of precision. In this case, we essentially ignore everything but the exponent, recognizing that each century will increment it by 1, at our chosen 2.3% rate.

12: In this case, the "real" answer would be 1,335 years, but why fret over the details for little gain or qualitative difference in the outcome?

Box 1.3: Dyson Sphere Construction

If we took the material comprising the entire Earth (or Venus) and created a sphere around the sun at the current Earth-Sun distance, it would be a shell less than 4 mm thick! And it's not necessarily ideal material stock for building a high-tech sphere and solar panels. The earth's atmosphere distributed over this area would be 0.015 m thick. Don't hold your breath waiting for this to happen.

Bypassing boring realism, we recognize that our sun is not the only star in the Milky Way galaxy. In fact, we estimate our galaxy to contain roughly 100 billion stars! This seems infinite. A billion seconds is just over 30 years, so no one could count to 100 billion in a lifetime. But let's see: 100 billion is 10^{11}. Immediately, we see that we buy another 11 centuries at our 2.3% rate. So it takes 1,100 years to go from consuming our entire sun to all the stars in our galaxy! That's 2,500 years from now, adding the two timescales, and still a civilization-relevant time period. Leave aside the pesky fact that the scale of our galaxy is 100,000 light years, so that we can't possibly *get* to all the stars within a 2,500 year timeframe. So even as a mathematical exercise, physics places yet another limit on how long we could conceivably expect to maintain our current energy growth trajectory.

Table 1.3: Energy limit timescales.

Utilizing	years until
Solar, land, 20%	250
Solar, earth, 100%	390
Entire Sun	1,400
Entire Galaxy	2,500
Light in Universe	3,600
Mass in Universe	5,000

The unhinged game can continue, pretending we could capture all the light put out by all the stars in all the galaxies in the visible universe. Because the visible universe contains about 100 billion galaxies, we buy another 1,100 years. We can go even further, imagining converting all matter (stars, gas, dust) into pure energy ($E = mc^2$), not limiting ourselves to only the light output from stars as we have so far. Even playing these unhinged games, we would exhaust all the matter in the visible universe within 5,000 years at a 2.3% rate. The exponential is a cruel beast. Table 1.3 summarizes the results.

By coincidence, the visible universe has about as many galaxies as our galaxy has stars. By "visible" universe, we mean everything within 13.8 billion light years, which is as far as light has been able to travel since the Big Bang (see Sec. D.1; p. 392).

The point is not to take seriously the timescales for conquering the sun or the galaxy. But the very absurdity of the exercise serves to emphasize the impossibility of our continuing exponential growth in energy. All kinds of reasons will preclude continued energy growth, including the fact that human population cannot continue indefinite growth on this planet. We will address space colonization fantasies in Chapter 4.

1.3 Thermodynamic Consequences

Physics places another relevant constraint on growth rate, and that concerns waste heat. Essentially all of our energy expenditures end up as heat. Obviously many of our activities directly involve the production of heat: ovens, stoves, toasters, heaters, clothes dryers, etc. But even cooling devices are net heat generators. Anything that uses power from an electrical outlet ends up creating net heat in the environment, with very few exceptions. A car moving down the road gets you from place A to place B, but has stirred the air,[13] heated the engine and surrounding air, and deposited heat into the brake pads and rotors, tires and road. Our metabolic energy mostly goes to maintaining body temperature. But even our own physical activity tends to end up as heat in the environment. The only exceptions would be beaming energy out of the earth environment (e.g., light or radio) or putting energy into storage (eventually to be converted to heat). But such exceptions do not amount to much, quantitatively.

Some time, go feel the exhaust air from an air-conditioning unit, or the heat produced at the back and bottom of a refrigerator. Even though these devices perform a cooling function, they make more heat than cool.

13: Stirred-up air eventually grinds to a halt due to viscosity/friction, becoming heat.

What happens to all of this waste heat? If it all stayed on Earth, the temperature would climb and climb. But the heat does have an escape path: infrared radiation[14] to space. The earth is in an approximate thermodynamic equilibrium: solar energy is deposited, and infrared radiation balances the input to result in steady net energy. As we will see in Chapter 5, the *rate* at which energy flows is called power, so that we can describe energy flows into and out of the earth system in terms of power. Physics has a well-defined and simple rule for how much power a body radiates, called the Stefan–Boltzmann law:[15]

14: . . . a form of electromagnetic radiation

15: . . . leaving out something called *emissivity*, not relevant for our purposes

$$P = A_{\mathrm{surf}}\sigma(T_{\mathrm{hot}}^4 - T_{\mathrm{cold}}^4). \tag{1.8}$$

P is the power radiated, A_{surf} is the surface area, T_{hot} is the temperature of the radiating object in Kelvin[16] (very important!), T_{cold} is the temperature of the environment (also Kelvin), and σ is the Stefan–Boltzmann constant: $\sigma = 5.67 \times 10^{-8}\ \mathrm{W/m^2/K^4}$.[17] Note that the law operates on the *difference* of the fourth powers of two temperatures.

16: Conversions to Kelvin from Celsius (or Fahrenheit) go like:
$T(\mathrm{K}) = T(\mathrm{C}) + 273.15$;
$T(\mathrm{C}) = (T(\mathrm{F}) - 32)/1.8$

17: It's actually an easy constant to remember: 5-6-7-8 (but must remember the minus sign on the exponent).

Example 1.3.1 A table in a room in which the table and walls are all at the same temperature does not experience net radiation flow since the two temperatures to the fourth power subtract out. In this case, as much radiation leaves the table for the walls as arrives from the walls to the table. But a room-temperature object at 300 K radiates approximately 450 W per square meter to the coldness of space.

Because space is so cold (tens of Kelvin, effectively, unless exposed to the sun), the fourth power of such a small number pales *so much* in comparison to the fourth power of a number like 300 that we can safely ignore it for radiation to space:

Try it yourself on a calculator!

$$P_{\text{space}} \approx A_{\text{surf}} \sigma T^4, \quad (1.9)$$

Temperature must be in Kelvin.

where we now just have a single temperature: that of the warm body in space.

Earth reaches an equilibrium so that power–in equals power–out.[18] If more power is dumped onto the planet, then the temperature rises until σT^4 climbs to match. The relation in Eq. 1.9 is fundamentally important to Earth's temperature balance, and applies pretty universally, as highlighted in Box 1.4.

18: Climate change is due to greenhouse gases blocking the escape of some radiation to space, presently causing a ~0.1% imbalance that Chapter 9 will address.

Box 1.4: Everything Radiates

The same relation (Eq. 1.8) governs the surface of the sun, light bulb filaments, glowing coals, and even the human body. While the human body expends metabolic energy at a similar rate to an incandescent light bulb (about 100 W), one is *much* hotter than the other because the surface areas are vastly different.

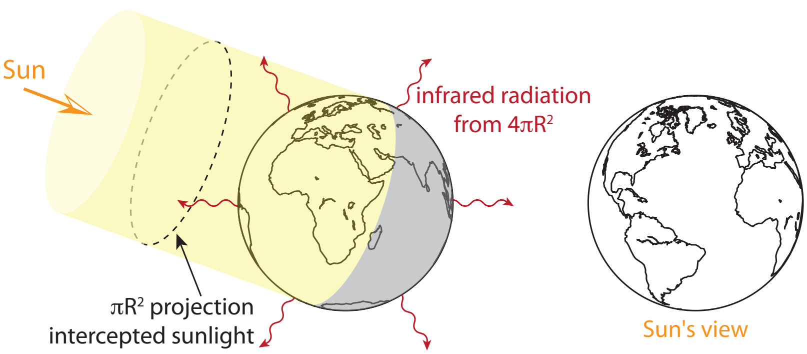


Figure 1.4: Earth—shown here in northern hemisphere summer—intercepts sunlight across the projected area of the Earth's disk (πR^2), while radiating from the entire surface area, which is four times larger ($4\pi R^2$).

To evaluate the expected temperature of the earth, we know that the sun delivers 1,360 W/m^2 to the *top* of the earth's atmosphere [4] (a bit less reaches the ground). We also know that about 29.3% of this is reflected by clouds, snow, and to a lesser extent water and terrain. So the earth system *absorbs* about 960 W/m^2. It absorbs this energy onto the area facing the sun: a projected disk of area $A_{\text{proj}} = \pi R_{\oplus}^2$. But the total surface area of the earth is four times this, all of it participating in the radiation to space (Figure 1.4). Equating the input and output for equilibrium conditions:

This 1,360 W/m^2, known as the solar constant, is the *incident* energy rate (power), or the *flux*, of sunlight incident on Earth.

$$P_{\text{in}} = 0.707 \times 1360\,\text{W/m}^2 \times \pi R_{\oplus}^2 = P_{\text{out}} = 4\pi R_{\oplus}^2 \sigma T^4, \quad (1.10)$$

The 0.707 factor represents absorbed fraction after 29.3% is reflected.

which we can rearrange to isolate temperature, satisfying

$$T^4 = \frac{0.707 \times 1360\,\mathrm{W/m^2}}{4\sigma}. \tag{1.11}$$

Solving for T yields $T \approx 255$ K, or −18°C (about 0°F). This is cold—too cold. We observe the average temperature of Earth to be about 288 K, or 15°C (59°F). The difference of 33°C is due to greenhouse gases—mostly H_2O—impacting the thermal balance by preventing most radiation from escaping directly to space. We'll cover this more extensively in Chapter 9.

Armed with Eq. 1.11, we can now estimate the impact of waste heat on Earth's equilibrium temperature. Using the solar input as a baseline, we can add increasing input using the exponential scheme from the previous section: starting today at 18 TW and increasing at 2.3% per year (a factor of 10 each century). It is useful to express the human input in the same terms as the solar input so that we can just add to the numerator in Eq. 1.11. In this context, our current 18 TW into the projected area $\pi R_\oplus^2$ adds 0.14 W/m² to the solar input (a trivial amount, today), but then increases by a factor of ten each century. Taking this in one-century chunks, the resulting temperatures—adding in the 33 K from greenhouse gases—follow the evolution shown in Table 1.4. At first, the effect is unimportant, but in 300 years far outstrips global warming, and reaches boiling temperature in a little over 400 years! If we kept going (not possible), Earth's temperature would exceed the surface temperature of the sun inside of 1,000 years!

A potential inconsistency in our treatment is that we based our exploration of energy scale on solar energy as a prelude to stellar energy capture. But in the thermodynamic treatment, we implicitly *added* our power source to the existing solar input. If the sun *is* the source, we should not double-count its contribution. Nonetheless, continued, relentless growth would eventually demand a departure from solar capture on Earth and drive the same thermodynamic challenges regardless. Synthesizing the messages: we can't continue 2.3% growth for more than a few centuries using sunlight on Earth. And if we invent something new and different to replace the fully-tapped solar potential, it too will reach thermodynamic limits within a few centuries.

Years	Power Density (W/m²)	T (K)	ΔT (C)
100	1.4	288.1	0.1
200	14	288.9	~1
300	140	296.9	~9
400	1,400	344	56
417	2,070	373	100
1,000	1.4×10^9	8,600	8,300

Table 1.4: At a constant energy growth rate of 2.3% per year, the temperature climb from waste heat (not CO_2 emissions) is slow at first, but becomes preposterous within a few-hundred years. Water boils in just over 400 years, and by 900 years Earth is hotter than the sun! The scenario of continued growth is obviously absurd.

Connecting some ideas, we found in the previous section that we would be consuming the sun's entire output in 1,400 years at the 2.3% growth rate. It stands to reason that if we used a sun's worth of energy confined to the surface of the earth, the (smaller) surface would necessarily be hotter than the sun (in 1,400 years), just like a light bulb filament is hotter than human skin despite putting out the same power—owing to the difference in area.[19]

19: This can be gleaned from Eq. 1.8 or Eq. 1.9.

One key aspect of this thermal radiation scenario is that it *does not depend on the form of power source*. It could in principle be fossil fuels, nuclear fission, nuclear fusion, or some form of energy we have not yet realized and may not even have named! Whatever it is, it will have to obey thermodynamics. Thus, thermodynamics puts a time limit on energy growth on this planet.

1.4 Upshot: Physics Limits Physical Growth

We saw in this chapter that unabated growth leads to absurd results. First, we calibrated our intuition in the context of bacteria in jars. The key point is that the jar is half full one doubling time before it is full. While this seems obvious, it delays the drama to the very end, acting fast to impose hard limits and catch the inhabitants by surprise. The conditions that persisted for many generations—thus taken for granted—suddenly change completely.

Next, we found that continuing a modest growth rate in energy becomes hopelessly absurd in a matter of centuries. Then we saw another side to this coin, in the context of thermal consequences on the surface of the earth if energy growth continues.

Was the exercise pointless, since the math leads to absurdity? Is the math wrong? No—it's immensely valuable to learn that our assumption of continued growth (and application of the corresponding correct math) fails to make sense, ultimately. The logical conclusion is that growth cannot continue indefinitely.

In the end, physics puts a timeline on expectations with respect to growth in energy on Earth. Maybe the ~300 year scale is not alarming enough. But it imposes a hard barrier against preserving our historical growth rate. In reality, other practicalities are likely to assert themselves before these hard limits are reached. We can therefore expect our growth phase to end well within a few hundred years. Given that the growth phase has lasted for far longer than that, we can say that we are closer to the end of the saga than to the beginning, yet the world is not collectively preparing for such a new reality. This seems unwise, and we will evaluate related concerns in subsequent chapters.

Note that a deviation from the assumed steady 2.3% growth rate changes all the numbers, and therein may lie the solution: ramp down growth!

Many factors will intercede to limit growth in both population and resource use: resource scarcity, pollution, aquifer depletion and water availability, climate change, warfare, fisheries collapse, a limited amount of arable land (declining due to desertification), deforestation, disease, to name a few. The point is only reinforced. By some means or another, we should view the present period of physical growth as a temporary phase: a brief episode in the longer human saga.

A number of these issues will be addressed in subsequent chapters.

1.5 Problems

Hint: for problems that require solving temperature when it appears as T^4, you'll need to take the fourth root, which is the same as raising to the $\frac{1}{4}$ power. So use the y^x button (or equivalent) and raise to the 0.25 power. You can check this technique by comparing the square root of a number to the result of raising that number to the 0.5 power. Another technique for the fourth root is to take the square root twice in a row.

1. Verify the claim in the text that the town of 100 residents in 1900 reaches approximately 100,000 in the year 2000 if the doubling time is 10 years.

2. Fill out Table 1.1 for the missing decades between 1940 and 2000.

3. Our example town from the text (page 3) starting at 100 people in the year 1900 and doubling every 10 years was said to take about 260 years (26 doubling times) to reach world population. Verify that the population indeed would approach 7 billion in 260 years (when the year would be 2160), by any means you wish.[20]

20: E.g., brute force doubling 26 times or using math to get straight at the answer.

4. Use Eq. 1.5 with $b = 2$ to figure out exactly how many years—via a computation of doubling times, which may not be an integer—our example town from the text (page 3) would take to reach 7 billion people.

Hint: ln is the natural log function found on scientific calculators (sometimes as LN).

Hint: M is the ratio of the final population to the initial population.

5. If our example town from page 3, doubling every 10 years, reaches a population of 7 billion in 260 years, how many years before it reaches 14 billion?

6. In a classic story, a king is asked to offer a payment as follows: place one grain of rice on one square of a chess board (64 squares), then two on the next square, four on the next, 8 on the next, and double the previous on each subsequent square. The king agrees, not comprehending exponential growth. But the final number (adding all the grains) is one less than 2^{64}. How many grains is this?

ⓘ To get 2,000 kcal of metabolic content per day, a person would need to eat 30,000 grains of rice each day. The amount of rice computed for this problem would feed the current world population for 240 years, which you are encouraged to check for yourself!

7. In the bacteria example of Section 1.1, how many "doubling times" are present in the 24 hour experiment (how many times did the population double)?

8. A one-liter jar would hold about 10^{16} bacteria. Based on the number of doubling times in our 24-hour experiment, show by calculation that our setup was woefully unrealistic: that even if we started with a single bacterium, we would have far more than 10^{16} bacteria after 24 hours if doubling every 10 minutes.

9. If a one-liter jar holds 10^{16} bacteria, how many bacteria would we start in the jar so that the jar reaches full capacity after 24 hours if we increase the doubling time to a more modest/realistic 30 minutes?

ⓘ Roughly 10 bacteria fit within a cubic micron (tiny), so you would not be able to see this tiny starting amount.

10. A more dramatic, if entirely unrealistic, version of the bacteria–jar story is having the population double every minute. Again, we start the jar with the right amount of bacteria so that the jar will be full 24 hours later, at midnight. At what time is the jar half full now?

11. In the more dramatic bacteria–jar scenario in which doubling happens every minute and reaches single-jar capacity at midnight, at what time will the colony have to cease expansion if an explorer finds three more equivalent jars in which they are allowed to expand without interruption/delay?

12. What is the doubling time associated with 3.5% annual growth?

13. Using Eq. 1.5 and showing work, what annual growth rate, in percent, leads to the mathematically convenient factor-of-ten growth every century?

Hint: the exponential, e^x, "undoes" the natural logarithm.

14. Use Eq. 1.5 with $b = 1 + p$ to figure out how long it takes to increase our energy by a factor of 10 if the growth rate is closer to the historical value of 2.9% ($p = 0.029$). Using 2.3% as we did in the examples (starting on page 7) puts this at 100 years.[21]

21: Hint: a good way to check your math. Note that if we were to use 2.9% instead of 2.3%, *all* of the time estimates in Section 1.2 are reduced by the ratio of this question's answer to 100 years.

15. Extrapolating a constant growth rate in energy is motivated by historical performance. During this period, population was also growing, albeit not as fast. If population were to double every 50 years,[22] how many people would Earth host when we hit the energy/thermodynamic limits in roughly 300 years?

22: This corresponds to a 1.4% growth rate, but you don't need to use this number in your calculation.

ⓘ We are unlikely to reach such a number for a host of other reasons.

16. In extrapolating a 2.3% growth rate in energy, we came to the absurd conclusion that we consume all the light from all the stars in the Milky Way galaxy within 2,500 years. How much longer would it take to energetically conquer 100 more "nearby" galaxies, assuming they are identical to our own?

ⓘ Ignoring the fact that it impossible to get to them fast enough, even at light speed.

17. In the spirit of outlandish extrapolations, if we carry forward a 2.3% growth rate (10× per century), how long would it take to go from our current 18 TW (18×10^{12} W) consumption to annihilating an entire earth-mass planet every year, converting its mass into pure energy using $E = mc^2$? Things to know: Earth's mass is 6×10^{24} kg; $c = 3 \times 10^8$ m/s; the result is in Joules, and one Watt is one Joule per second.

Hint: Dividing the number of Joules associated with Earth's mass by the number of seconds in a year gives the number of Watts being consumed. You may wish to compare the result to the timescale before we would use the power output of all stars in the Milky Way galaxy?

18. Taking cues from the discussion of waste heat channels on page 10, describe some of the ways that all your energy output turns to heat when you go on a bicycle ride.

19. Your skin temperature is about 308 K, and the walls in a typical room are about 295 K. If you have about 1 m^2 of outward-facing surface area, how much power do you radiate as infrared radiation, in Watts? Compare this to the typical metabolic rate of 100 W.

ⓘ Air convection also takes some heat away, but then clothing reduces both to bring us to equilibrium/comfort.

20. The moon absorbs 90% of the solar energy incident on it.[23] How hot would you expect the surface to get under full sun? You don't need the factor of four here[24] because the moon rotates very slowly under the sun and we're considering a patch experiencing overhead sunlight (rather than averaging over the sphere). Compare the result to boiling water temperature.

23: . . . incident at the same rate/flux as at Earth

24: Referring to the 4 that shows up in Eqs. 1.10 and 1.11.

21. Venus is, ironically, colder than Earth as an infrared radiator. This is because Venus is covered in bright clouds, absorbing only 25% of the incident solar flux. Sunlight is more intense there due to it's being closer to the sun: it's almost double, at 2,620 W/m^2. Adapting Eq. 1.11, calculate the equilibrium temperature of Venus in the infrared and compare it to the Earth value of 255 K.

ⓘ The surface of Venus is *much* hotter than that of Earth owing to a runaway greenhouse condition. On Earth, the greenhouse boost is only 33 K, but on Venus it's hundreds of degrees.

22. Adapt Eq. 1.11 to Mars to find its equilibrium temperature. The solar flux averages 590 $\mathrm{W/m^2}$ there, and it absorbs 75% of incident sunlight. Express the answer in both Kelvin and Celsius, and put in context.

23. If a human body having an outward surface area of 1 $\mathrm{m^2}$ continued to put out 100 W of metabolic power in the form of infrared radiation in the cold of space (naked; no sun), what would the equilibrium temperature be? Would this be comfortable (put in context)?

24. Verify the total solar power output of 4×10^{26} W based on its surface temperature of 5,800 K and radius of 7×10^{8} m, using Eq. 1.9.

25. Verify that Earth would reach a temperature far in excess of boiling point of water[25] after 500 years if today's power output (18 TW) increased by a factor of 10 each century.

25: Water boils at 100°C, or 373 K.

2 Economic Growth Limits

2.1 Historical Coupling 17
2.2 Decoupling and Substitution 20
2.3 Forced Economic Limits . . . 24
2.4 No-Growth World 26
2.5 Upshot: Growth Will End . . 27
2.6 Problems 28

Chapter 1 demonstrated that the laws of physics and mathematical logic render a constant few-percent growth trajectory to be absurd and untenable even a few hundred years into the future. In this chapter, we develop implications for the less physics-bound concept of *economic* growth, which is a cornerstone of modern society. Investment, loans, retirement and social security schemes all assume the march of growth. The conclusion of this chapter is that economic growth will not be able to continue at any significant rate in the absence of physical growth—which we have seen cannot continue indefinitely. The suggestion usually evokes quick criticism from economists, but they can be talked down, with patience.[1] This is how it goes.

1: See, for instance [5].

2.1 Historical Coupling

In subsistence times, esthetics held little value compared to physical goods: you couldn't eat a sculpture, for instance—nor would it help keep you warm.[2] Food, tools, resources like wood, pack or draft animals carried primary value. When basic subsistence requirements were met, gold or jewelry may have warranted some expenditure—but even these were physical resources.

2: Life, it turns out, is a struggle against thermodynamics.

Agriculture freed some individuals in society to think and create. The economy found more room to value arts and performance: things that fueled the mind, if not the body. During the Renaissance, patrons would support artists and scientists whose output had few other channels of economic support. In today's world, a complex economy distributes financial assets to a wide variety of practitioners in general accordance with society's values.

Banner photo pokes fun at what physics (gravity) finds more valuable: a silver dollar (real silver) vs. a $20 bill; Credit: Tom Murphy

But resources are still paramount. The United States prospered largely because it possessed a frontier rich in natural resources. Mining and animal pelts dominated early on, as well as agriculture (tobacco, cotton, corn, wheat). In the middle of the 20th century, the United States was the dominant oil exporter worldwide (first developed in Pennsylvania, then California and Texas). Escaping the World Wars largely unscathed in terms of domestic infrastructure, the country had tooled-up a massive manufacturing capability. Together with a can-do attitude, Americans set out to rack up superlatives in virtually every category. As a manufacturing powerhouse during the middle of the 20th century, American raw materials joined a well-educated workforce to drive innovation and production. It is no accident that many in the U.S. yearn to return to these "glory days;" however we cannot possibly do so, having permanently depleted some of the original stocks.

One might say that the U.S. was the Saudi Arabia of the day.

It is important to recognize that the past was not "glorious" for all people.

What was true in the past is largely still true today: resources like oil, steel, metals, agricultural products, and heavy machinery continue to fetch a significant price on the market. Resources fuel prosperity. It is not the *only* source, but remains a reliable and bedrock component. Figure 2.1 shows the tight correlation between economic scale and energy use, which is often expressed by saying that the two tend to be coupled.

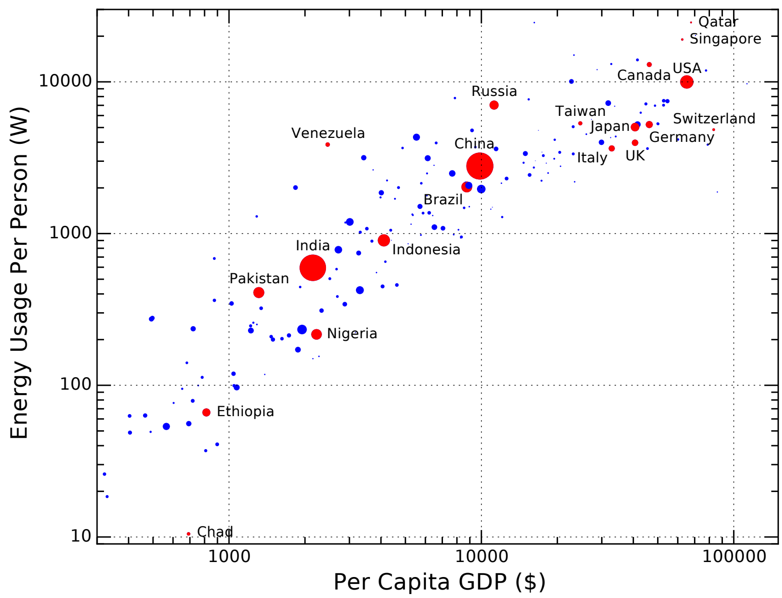


Figure 2.1: Per capita energy use as a function of GDP on a logarithmic scale. Per capita GDP is the sum total of a country's economy divided by population, effectively indicating average annual income. The rate at which an individual uses energy is expressed as a power, in Watts. A strong correlation exists here across many orders-of-magnitude: rich countries use more energy, per person [6–8]. A few instructive cases (red dots) are labeled. ■ The dot areas are scaled to population.

One way to capture the physical connection to economic activity is to represent the energy expenditure associated with each dollar[3] spent. This economic *energy intensity* (see Definition 2.1.1) of a country is just the energy expenditure of society divided by the gross domestic product (GDP).[4]

3: Or converted monetary equivalent.

4: GDP is a measure of total monetary value of goods and services produced in a country within a year.

Definition 2.1.1 ***Energy intensity*** *is a measure of how much energy a society uses relative to its economic scale—sort of like an efficiency. It can be a proxy for resource use in general, and is calculated as:*

$$\text{Energy Intensity} = \frac{\text{Energy Expended}}{\text{Money Spent}}. \tag{2.1}$$

In a resource-constrained world (limited material and energy supplies), a lower energy intensity translates to less energy consumption for a certain economic output, or conversely allows higher economic output for a fixed energy consumption rate. On a smaller scale, we can say, for instance, that spending $100 on an airplane trip is far more energy intense than spending the same amount of money on legal advice.

Energy intensity therefore provides a measure of how resource-heavy a country is in relation to the size of its economy. For instance, the U.S. uses about 10^{20} Joules of energy per year and has a GDP of approximately 20 trillion dollars. Dividing these gives an intensity of 5×10^6 J/$, or 5 MJ/$ (many variants are possible in terms of units). The world as a whole uses about 4.5×10^{20} J in a year at an estimated $90 trillion gross world product, also resulting in 5 MJ/$. The variation among developed countries is not especially large—generally in the single-digit MJ/$ range.

We will cover units of energy in Chapter 5. For now, it is sufficient to know that a Joule (J) is a unit of energy, and that MJ means megajoules, or 10^6 J.

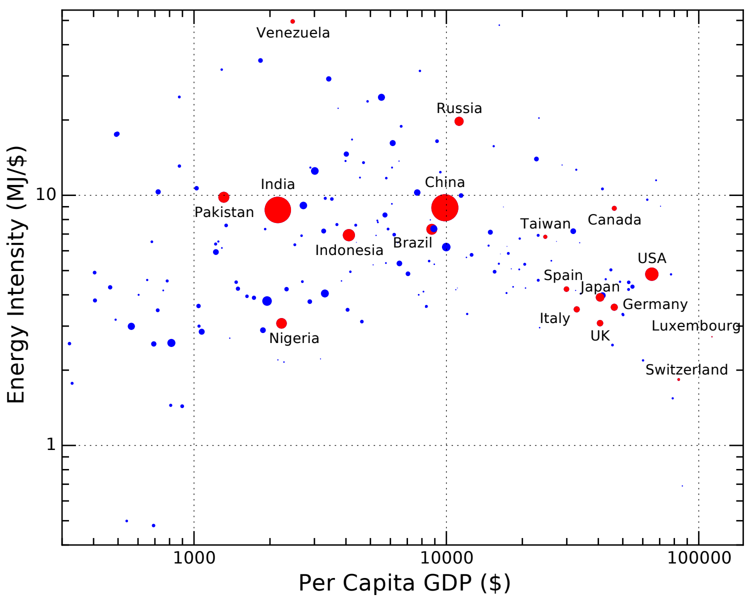


Figure 2.2: Energy intensity of countries, on a log–log plot. The vertical axis shows how energetically "hungry" each country is in relation to its economic output, while the horizontal axis sorts countries by economic output per person. A few instructive cases (red dots) are labeled. ■ The dot areas are scaled to population. Prosperous countries tend to have lower intensity than developing countries, but part of this may relate to moving manufacturing from the former to the latter [6–8].

Figure 2.2 illustrates the range of intensities for all the countries in the world. Among the factors driving energy use are geographical extent (large countries require more long-haul transportation), climate (cold countries require more heating), efficiency, and lifestyle. Russia, Canada, and the U.S. have large territories, and the former two require more heating than most. By contrast, Switzerland is geographically small and outsources much of its heavy industry. Somebody should probably check on what's happening in Venezuela.[5]

5: Maybe they left the oven on by mistake?

2.2 Decoupling and Substitution

As economies expand beyond subsistence level, a larger fraction of the total activity can go to "frivolous" elements, such as art and entertainment. The intensity of such activities can be quite low. An art collector may pay $1 million for a coveted painting. Very little energy is required. The painting was produced long ago. It may even remain on display in the same location—only the name of the owner changing. Financial transactions that require no manufacture, transport, and negligible energy are said to be "decoupled" from physical resources. Plenty of examples exist in society, and are held up by economists as illustrating how we can continue to expand the economy without a commensurate expansion of resource needs.[6]

6: This is the hope, anyway.

Definition 2.2.1 ***Decoupling*** *is the notion that economic activities need not be strongly tied to physical (e.g., energy) requirements, so that energy intensity might become arbitrarily small. The degree to which some economic activity is decoupled forms a continuous scale, where intense utilization of energy and physical resources (e.g., steel production) are on one end and fine art dealing on the other.*[7] *The only way for an economy to maintain growth in the event that physical sector growth becomes limited is to increase the degree of decoupling in the society.*

7: Services, like plumbing, journalism, or marketing fall in between, using some physical resources, but not as much as heavy industry.

The dream is that as development progresses, economic energy intensity may decline (greater decoupling) so that more money is made per unit of energy expended. If the economy can decouple from energy needs, we are not constrained in our quest to continue economic growth, bringing smiles to the faces of investors and politicians. Such a transition would mean less emphasis on energy and resource-intensive industrial development/manufacturing, and more on abstract *services*,[8] broadly speaking.

8: Such services might include things like singing lessons, life coaching, psychotherapy, financial planning, and other activities that demand little physical input.

Because the world is a sort of "experiment," representing many countries having adopted many policies and in various states of development, Figure 2.2 can be viewed as a potential roadmap to decoupling.

Part of the reason prosperous countries demonstrate a lower intensity is that manufacturing moves overseas. Driving the whole world toward lower intensity is a more difficult prospect, as the physical processes must still happen *somewhere*.

The question is: as countries develop and become more prosperous, does intensity *decrease*, as we would want it to do as a signal of decoupling? On the large scale, any effect is modest. Going from India to the U.S. affords only a factor-of-two improvement in intensity, while spanning most of the horizontal extent in personal prosperity (a factor of 30 in per capita GDP).[9] That's pretty weak tea.

9: $65,000 vs. $2,100 for the U.S. and India, respectively.

At the upper end of personal income (right side of Figure 2.2), we might detect a downward slope. But we must be careful about cherry-picking in the face of non-replicable circumstances. Not every country can assume the geography and financially-focused nature of Switzerland. And at the same time, if the U.S. imagines itself providing a model that other countries might emulate, the intensity of many European countries could actually increase if adopting U.S. habits. But more broadly, we don't have evidence that a country on the prosperous end of the distribution can operate at even a factor-of-four lower intensity than the 4 MJ/$ level typical of developed countries. In the present context of assessing the future of growth, in which we are concerned with order-of-magnitude scales and limits (as in Chapter 1), it does not appear that decoupling has very much to offer.[10]

10: That is, no orders-of-magnitude that will allow us to continue growth for centuries more after physical resources limit growth.

Definition 2.2.2 ***Substitution*** *refers to the ability to switch resources when one becomes scarce or a better/superior alternative is found. Substitution is often invoked to counter concerns about scarcity. A common and cute way to frame it is that the stone age did not end because we ran out of stones—we found bronze.*

The past is full of examples of substitution (Definition 2.2.2). Consider the progression in lighting technology from open fires to beeswax candles to whale oil lanterns to piped gas lanterns to incandescent electric bulbs to fluorescent lights to LED (light emitting diode) technology. Every step seems to be an improvement, and it is very natural to assume the story will continue developing along these lines.

Through this example, we can see how substitution and decoupling might be connected: efficiency improves through substitution, requiring less energy for the same lighting value.

Box 2.1: A Story of Lighting Efficiency

One way to quantify lighting progress is in the *luminous efficacy* of light, in units of lumens per Watt. In the 20th century, incandescent bulbs were so ubiquitous for so long that we fell into the bad habit of characterizing brightness in terms of the electrical power consumed by the bulb, in Watts. Thus we have generations of people accustomed to how bright a "100 W" or "60 W" bulb is. As technology changes, we should migrate to "lumens," which accurately captures how bright a source is perceived by the human eye.

Bulb packaging still refers to the "equivalent wattage" of a bulb, even though a "60 W equivalent" bulb may only consume 12 W of power.

Table 2.1 and Figure 2.3 present the evolution of luminous efficacy as sources improved. Can this trend continue indefinitely? No. Every photon of light carries a minimum energy[11] requirement based on its wavelength. For photons spread across the visible spectrum (creating light we perceive as white), the theoretical limit is about 300 lm/W [9]. At this extreme, no energy is wasted in the production of light, putting 100% of the energy into the light itself. Engineering rarely reaches theoretical limits, due to a host of practical challenges. It would therefore not be surprising if lighting efficiency does not improve over where it is today by another factor of two, ending yet another centuries-long trend.

11: We will see this in Sec. 5.10 (p. 79).

Table 2.1: Luminous efficacies [10, 11].

Light Source	lm/W
Candles	~0.3
Gas Lamp	1–2
Incandescent	8–15
Halogen	15–25
CFL	45–75
LED	75–120

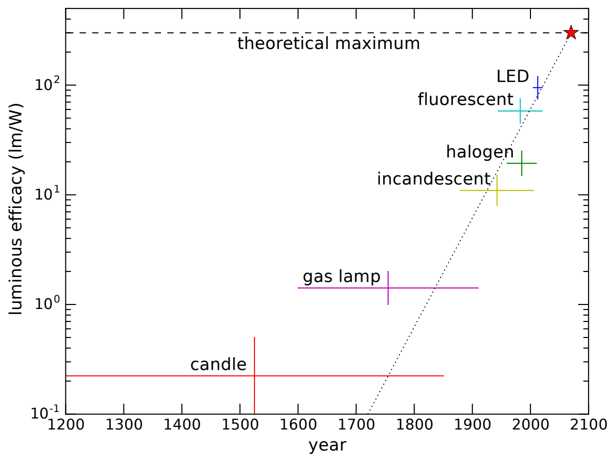


Figure 2.3: Historical progress of lighting efficiency on a logarithmic plot, using bars to indicate the approximate range of time and performance. The dashed line at top represents the maximum theoretical luminous efficacy for white light (no waste heat). The dotted line rises by our customary factor of ten per century (2.3% annual rate). Note that the guiding line reaches the theoretical maximum mid-century (red star), indicating that this centuries-long ride cannot continue much longer [10, 11].

The historical progress can fool us into thinking that we can expect a continued march to better substitutes. Having witnessed a half-dozen rabbits come out of the hat[12] in the example of lighting technology (Box 2.1), we are conditioned to believe more are forthcoming. It will be true until it isn't any more (e.g., see Figure 2.3) One way to put it is that 6 rabbits does not imply an infinite number. We should welcome each new rabbit, but not hinge our future on a continual stream of new rabbits.

12: . . . magician reference

We will return to this theme in the context of fossil fuels, which might be termed the *mother of all rabbits*, in this context. Having pulled such a stupendous rabbit out of the hat once, many assume we're set from now on. In this case, equating one to infinity is even more dubious.

For financially secure individuals at the top end of the wealth distribution, it is easier to buy into the allure of substitution as a way forward. Many have achieved wealth from humble beginnings, and have therefore lived a life of continual upgrades in terms of housing, transportation, clothing, food, travel, etc. Even those who have been surrounded by wealth their whole lives have been in a position to afford new upgrades as they become available. Yet, it is not always possible to export the capabilities of those at the top to a significant sector of the population. Not everything can scale.

An electric car having hundreds of kilometers of range seems like an obvious path forward beyond fossil fuels. But at a price tag above \$40,000, it does not look like much of a solution to most people, and we can't be sure prices will fall steeply. Section D.3 covers electrified transportation in more detail.

Box 2.2: The Fate of the Concorde

The fate of the Concorde—which offered supersonic transatlantic passenger service between 1976–2003—may offer a useful lesson here: just because it is *possible* to construct a supersonic passenger airplane does not mean that enough people can afford it to result in an economically *viable* reduction in trans-oceanic travel time for all. Consumers no longer have the option for supersonic flight, even though 50 years ago it was assumed that this was the future. Sometimes we go backwards, when our dreams don't line up to practical reality.

More generally, sometimes the best possible solution and "peak" technology arrives at some early point in history. As much as we mess around with elements on the Periodic Table, we are never going to beat H_2O

as a vital substance.[13] Marketers might sell H_2O_2 as superior, having one more beneficial oxygen atom, but *please don't drink hydrogen peroxide*! Some technologies in use today would be recognized by pre-industrial people: wheels, string, bowls, glass, clothing. We won't always find better things, though we may make a series of incremental improvements over time. Not everything will experience game-changing developments.

13: Relatedly, consider that the Periodic Table is finite and fits easily on a single sheet of paper (Fig. B.1; p. 375). We don't have an unlimited set of substitute elements/compounds available. Astrophysical measurements validate that the whole universe is limited to the same set of elements.

In summary, decoupling and substitution are touted as mechanisms by which economic growth need not slow down as energy and other resources become constrained. We can make money using less of the resource (decoupling) or just find alternatives that are not constrained (substitution), the thinking goes. And yes, this is backed up by loads of examples where such things *have* happened. It would be foolish to claim that we have reached the end of the line and can expect *no* more gains from decoupling or substitution. But it would be equally foolish to imagine that they can produce dividends eternally so that economic growth is a permanent condition.

Box 2.3: Efficiency Limits

Efficiency improvements would seem to offer a way to tolerate a stagnation or decline in available energy resources. Getting more from less is very appealing. Yes, efficiency improvements are good and should be pursued. But they are no answer to growth limits, for the following reasons.

1. For the most part, realized efficiencies are already within a factor-of-two of theoretical limits.[14] A motor or generator operating at 90% efficiency has little room to improve. If efficiencies were typically far smaller than 1%, it would be reasonable to seek improvements as a "resource" for some time to come, but that is not the lay of the land.
2. Efficiency improvements in energy use tend to creep along at ~1% per year,[15] or sometimes 2%. Doubling times are therefore measured in decades, which combined with the previous point suggests an end to this train ride within the century.[16]
3. Efficiency improvements can backfire, in a process called the Jevons paradox or the rebound effect. Increased demand for the more efficient technology results in *greater* demand for the underlying resource. For example, improvements in refrigerator efficiency resulted in larger refrigerators and more of them,[17] for a net increase in energy devoted to refrigeration. Consider that per-capita global energy and material resource use has climbed inexorably amidst a backdrop of substantial efficiency improvements over the last century [12].

Efficiency improvements are not capable of resolving resource demand.

14: Chapter 6 covers theoretical efficiency limits for thermal sources like fossil fuels.

15: . . . meaning 30% one year might be 30.3% next year (not 31%, which would be a ~3% improvement)

16: . . . similar to lighting technology, as per Box 2.1 and Figure 2.3

17: . . . e.g., in basements or garages or offices

[12]: Garret (2014), *Rebound, Backfire, and the Jevons Paradox*

2.3 Physically Forced Economic Limits

Let us now consider a thought experiment. We will use Figure 2.4 as a guide as we go along. Colored numbers in the following text point to similarly-colored labels in the figure. We start by positing a constant growth rate for the entire economy (point 1; red curve in Figure 2.4) following the familiar 2.3% annual growth rate, picked for its convenient factor of 10 each century. Meanwhile, the scale of physical resources (energy, materials) in the economy also climbs at the same rate, starting at point 2. The vertical gap between the curves at the left-hand edge conveys that the economy is not 100% physical in the beginning: the total economy is larger than the physical piece.[18]

Verify this for yourself in Figure 2.4!

18: . . . thus some room for services.

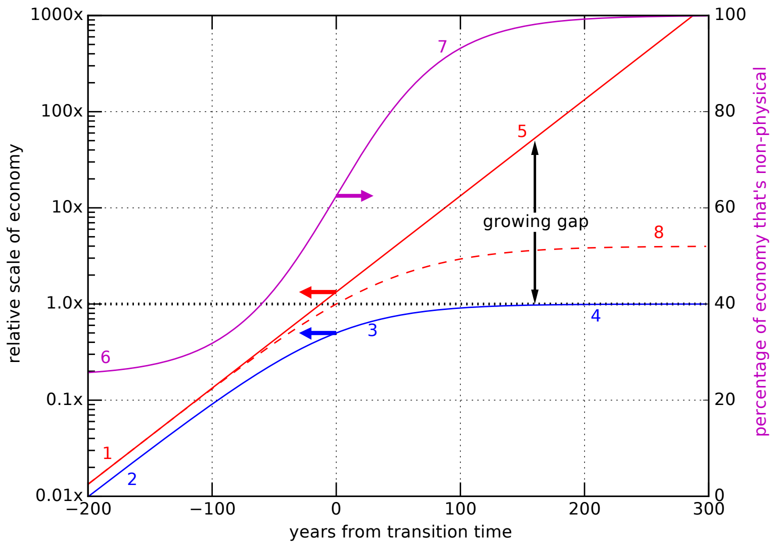


Figure 2.4: Model evolution of the economy after physical resources saturate. The blue curve is the scale of the physical economy (leveling out, or saturating). The solid red curve is the total economic scale, which we force to adhere to a constant growth rate (10× per century, or 2.3% annual rate). The magenta curve is the percentage of the economy in non-physical sectors, and the red dashed curve is a more realistic reaction of the economy to a saturating physical sector. Colored arrows point to the scale that each curve should use—logarithmic on the left for economic scales and linear on the right for the percentage curve. This model is constructed simply to illustrate the overall behavior: time scales and other quantitative details should not be taken literally.

Fast-forward to a time when physical resources have stopped growing, starting at point 3. Chapter 1—using energy and thermodynamics as the basis—made the case that we cannot expect physical growth to continue indefinitely, ending on a few-century timescale at the longest.[19] In this scenario, the scale of energy in our society flat-lines at a steady scale (point 4).

19: It is assumed here (optimistically) that we have managed to find a renewable alternative that can satisfy a constant demand effectively indefinitely. If not, the story is even worse and we are forced to *ramp down* the scale of the physical sector, which would force the blue curve in Figure 2.4 to descend in later years.

If we demand continued economic growth in the context of fixed energy, decoupling becomes increasingly necessary, shown as a growing gap in Figure 2.4. In other words, if the gross domestic product (GDP; as an indicator of economic activity) is to continue rising[20] (point 5), then overall intensity (energy per dollar) must continually decrease. For this to happen, less-energetic activities must assume increasing importance in the economy. So far, economists are on board: this is precisely what inspires an affinity for decoupling—a way forward in the face of physical limits. One might expect more abstract services, virtual experiences, art dealing, enhanced presentation: all requiring little or no additional

20: . . . and not artificially via inflation, but in terms of real value energy expenditure, or perhaps even less than before. In this way, the economic scale could keep rising while physical resources are held flat.

If the economy is to continue to expand on the basis of decoupled activities, a greater fraction of it must go toward these non-physical sectors. This means more monetary flow is associated with low-impact activities. In practical terms, then, a greater fraction of one's income is directed toward experiences not tied to energy or other physical demands. In Figure 2.4, we see, at point 6, the percentage of the economy in the non-physical sector starting at 25%: not dominant, but not negligible. The magenta curve must rise as the red and blue lines separate, until at point 7 it approaches 100% non-physical and continues to drive arbitrarily close to 100%.

During this process, the obvious converse consequence is that the energetically or physically costly activities—like transportation, food, heating, cooking, manufactured items—become an ever-smaller fraction of the economy, or an ever smaller fraction of monthly expenses, to put it more personally. In other words, they become cheap.

⚠ This, we will argue, is unrealistic.

Now, in our imagined scenario of continued economic growth, the ruthlessness of the exponential grabs the reins and drives the gulf ever wider, so that physical goods become arbitrarily cheap and demand an ever-smaller fraction of income. By the time we reach the right side of Figure 2.4, the economic scale is over 1,000 times as large as the physical scale, meaning that the physical component is less than 0.1% of the total economy. Table 2.2 illustrates the progression under the foregoing growth rate of 2.3%. If in the year 2000, 50% of one's income (and thus about half of one's work hours) goes toward physically intense products, this becomes ever smaller until by the end of the table it only takes 6 minutes of your annual work to earn enough for the physically intense goods: all your food, clothing, transportation, heating, cooking, manufactured goods.

⚠ Again, seems unrealistic.

⚠ Clearly absurd result.

Table 2.2: Cost of physical goods.

Year	% income	hours
2000	50%	1,000
2100	5%	100
2200	0.5%	10
2300	0.05%	1
2400	0.005%	0.1

If this is starting to feel like unrealistic fantasy, then good: your intuition is serving you well. How can essential, non-negotiable, life-sustaining commodities *that are in finite supply* become essentially free? The idea goes against another, more fundamental economic principle of **supply and demand**. A limited life-essential resource will always carry a moderately high value. Limited supply and inflexible demand dictate a floor to the price.

Box 2.4: Monopoly Made Easy

One way to highlight the absurdity of the scenario is that if the physically-limited but essential (life sustaining) resources became arbitrarily cheap in the fullness of time, a single person could buy them *all* for a pittance, and then charge a hefty price for anyone who wants to keep living. We simply will not find ourselves in the situation where precious and limited resources become arbitrarily cheap. Alternatively, if people only needed to work an hour per year to accommodate basic needs, expect a lot less work to be done, acting as a drag on economic productivity and thus preventing inexorable growth—one way or another.

Once the price floor is reached, the cost of physical resources will not be able to fall any further. This happens pretty soon after physical resources cease to grow in scale. Indeed, it seems unlikely (to the author) that limited resources essential for survival would fall much below 10% of the total economic scale, which happens within a century of physical saturation in our 2.3% growth scenario. Point 8 in Figure 2.4 depicts a more realistic trajectory for the economy (red dashed line) in reaction to a saturated physical scale. In this case, the economy keeps growing a bit more than the physical sector, but eventually settles down itself into a non-growth phase.

What seems like a reasonable lower limit to you? How economically insignificant can essentials be and still make sense?

We therefore have a logical sequence providing a few-century timescale for an end to economic growth. Thermodynamics limits us to at most a few centuries of energy growth on Earth, and economic growth will cease within a century or so thereafter, assuming a target rate of a few percent per year. In practice, growth may come to an end well before theoretical extremes are reached.

2.4 No-Growth World

The foregoing arguments spell out why economic growth cannot be expected to continue indefinitely—contrary to prevalent assumptions. When a mathematically-framed model delivers nonsense results—like the one we used to extrapolate energy use to absurd extremes—it does not mean the math itself is wrong, just that it has been misapplied or layered onto faulty assumptions. In this case, the breakdown indicates that the assumption of indefinite growth is untenable.

The growth regime is woven deeply into our current global society. And why wouldn't it be? We've enjoyed its benefits for many generations. We celebrate the myriad advantages it has brought, and therefore align our political and economic institutions toward its robust preservation. Community planning, interest rates, investment, loans, the very role of banks, social safety net systems,[21] and retirement plans all hinge on the assumption of growth.[22] Shock waves of panic reverberate at signs of (even temporary) recession, given the importance of growth to our institutions. Yet the message here is that we cannot expect its unfaltering continuance—implying that many things will have to change.

21: In the U.S., Social Security and Medicare are examples.

22: Growth in both workforce and investments are essential ingredients of these schemes that pay out more than an individual's past contributions to the program.

Returning to the roots of economic theory, the earliest thinkers—Adam Smith, David Ricardo, Thomas Malthus, John Stuart Mill—had foundations in natural philosophy[23] and saw growth as a temporary phase, ultimately limited by a prime physical resource: **land**. In that time, land held the key to outputs from farming, timber, mining, and game—thereby dictating economic development. What these pioneering economic thinkers did not foresee was the arrival of fossil fuels, and the technological developments that accompanied this energy explosion.

23: . . . closer to modern-day physics than to modern-day economics, rooted in the natural world

Now, we have fallen into something of a lulled complacency: having rescued ourselves so far from the end-of-growth predictions of the early economists, the temptation is to conclude that they were just wrong,[24] and we have outsmarted natural limits. This is dangerous thinking. In the end, nature is indifferent to how smart we imagined ourselves to be. If we were truly clever, we would *start thinking about a world that does not depend on growth, and how to live compatibly within planetary limits*. Chapter 19 touches on this theme, after intervening chapters paint a more complete picture of energy constraints.

24: The classic example is Thomas Malthus, who warned of limits over 200 years ago based on finite resource limits before fossil fuels ripped the narrative apart. The lasting association is that "Malthus equals wrong," leading to the dangerous takeaway that all warnings in this vein are discredited and can be ignored. Note that the most consequential and overlooked lesson from the story about "the boy who cried wolf" is that a *real* wolf *did* appear.

2.5 Upshot: Economic Growth Will End

It is worth re-iterating the recipe for an end to economic growth in summary form, as spelled out in Box 2.5. Make sure you can trace the logic and connections from one point to the next—not to be memorized as disconnected facts.

As was stated before, experts frequently read complex sections more than once to fully absorb the arguments; feel free to do so here.

Box 2.5: Economic Growth Limit

1. Physical resources (energy in our example) ultimately stabilize to a fixed annual amount.
2. Non-physical sectors of the economy must assume responsibility for continued economic growth, *if* growth is to continue.
3. The economy comes to be dominated by non-physical sectors.
4. Physical sectors are relegated to an ever-smaller fraction of the economy, ultimately vanishing if exponential growth is to hold.
5. In this scenario,[25] physical goods (energy among them) become arbitrarily cheap, requiring only one week's worth of earnings, then a day's worth, then an hour, a minute, a second.
6. This situation is impossible and does not respect common-sense supply/demand notions: a finite, limited but absolutely vital resource will never become arbitrarily cheap in a market system.
7. At some point, physical resources will "saturate" to a minimum fraction of the economy, at which point overall growth in non-physical sectors must also cease.

25: . . . which, let's be clear, we're arguing is ultimately not at all viable. . .

Just because we can point to some completely legitimate *examples* of decoupled activities and many impressive substitution stories does not mean that an entire economy can be based on indefinite continuance of such things. We are physical beings in a physical world and have non-negotiable minimum requirements for life. The activities and commodities that support critical functions cannot continue to expand indefinitely, and will not become arbitrarily cheap once their expansion hits physical limits. The finite nature of our world guarantees that such limits will be asserted, committing economic growth to stall in turn. Nothing, in the end, escapes physics.

So, while acknowledging that growth in the past has brought uncountable benefits to the human endeavor, we have to ask ourselves: If the end of growth is inevitable, why does it remain our prevailing plan?

2.6 Problems

1. At a 3.5% growth (interest) rate, how much would $1,000 invested at the time Columbus sailed to America be worth today (hint: use the rule of 70)? Put this in context (compare to richest individuals or find a similar GDP for some country).

   ⓘ Retention of Chapter 1 material is assumed. The real world is not partitioned into chapters, and neither should your brain be.

2. As an indication of how sensitive the accumulation is to interest rate, compare the result from Problem 1 to what would happen for interest rates of 4% and 5%—again putting into context.

3. Find the energy intensity for at least four countries spanning a range of development levels. For each country, look up the GDP, and find energy consumption at: Wikipedia page on Primary Energy Consumption. [13] In order to compare to Figure 2.2, multiply the number in quadrillion Btu (qBtu) by 1.055×10^{18} J/qBtu. Also note that a trillion is 10^{12}.

   Look at the column for total energy consumption in units of quadrillion Btu.

4. Estimate the energy intensity of the UCSD campus, based on an annual electricity expenditure around 10^{15} J.[26] For the financial side, assume that student payments (tuition, fees, room and board) accounts for 40% of the total budget.[27] Use your knowledge of typical tuition/fees and enrollment to come up with a number. Compare your result to global figures for energy intensity.

   26: Based on a 30 MW electrical load times the number of seconds in a year; this won't account for all energy expenditures, missing transportation, for instance.

   27: Federal grants comprise most of the rest, and a small amount from state taxes

   Don't get hung up on in-state fraction; just make a crude guess (maybe guess an average) and clearly state assumptions.

5. Typical energy costs are in the neighborhood of $0.10 per kilowatt-hour (kWh), and 1 kWh is 3.6 MJ (megajoules). Take the ratio of these two figures to form an economic energy intensity *of energy itself*, in units of MJ/$.

   The result should be larger than the typical energy intensity for *all* economic activities within a country, since not all monetary expenditure goes toward energy.

6. If a country clocks in at 5 MJ/$ for its *overall* energy intensity, and its energy costs work out to an energy intensity of 30 MJ/$,[28] what percentage of the economy do we infer constitutes the energy sector?

   28: In the same sense as was calculated for Problem 5

   The answer should be well less than 100%.

7. Come up with some of your own examples (at least three; not listed in the text) of economic activities[29] that have little resource footprint and are therefore fairly decoupled. These are transactions for which the intensity (energy or resource expenditure) is very low compared to the dollar amount.

29: . . . things that cost money

8. If a candle has a luminous efficacy of 0.3 lm/W and a modern LED light bulb achieves 100 lm/W, by what factor[30] have we improved lighting efficiency? If the theoretical limit is around 300 lm/W, what factor do we still have to go?

30: A factor is just a multiplicative scale: e.g., 24 is a factor of 6 larger than 4.

9. In going from 0.3 lm/W candle technology to the theoretical maximum luminous efficacy, we see a factor of 1,000 increase. Taking about 300 years to do this, we might recognize that we are following along our familiar factor-of-ten each century trajectory. Approximately how long[31] might we expect it to take to achieve the final factor of three to go from our current technology to the theoretical limit, at this same rate? Is it within your lifetime that we hit the limit?

31: You can just estimate, use appropriate math, or refer to the guiding line in Figure 2.3 to arrive at a rough number.

10. Provide your own example of a sequence of substitutions comprised of at least two qualitatively superior[32] substitutions over time (thus three steps: original, first replacement, second replacement).

32: . . . perhaps defined by widespread or universal adoption or replacement.

11. List three substances or critical concepts we rely on that have no superior substitutes in the universe.

12. Based on your present state of knowledge, detail what you think an optimist might say about the superiority of post-fossil energy substitutes?

You might pick any subset of solar, wind, hydroelectric, geothermal, nuclear, etc. to guide your thinking.

13. Based on your present state of knowledge, detail what you think a pessimist might say about the lack of superiority of post-fossil energy substitutes?

14. Justify what, In your mind, is a reasonable lower limit to the percentage of the economy that could be based on decoupled (not energy or resource heavy) activities? Make an argument for what leads you to this "floor."

15. One form of decoupled activity that some will bring up is virtual reality: you can travel the world (or solar system?) without resource-hogging transportation and other material costs. Do you see this as a viable alternative that is likely to largely supplant physical travel? Why or why not?

16. Are you sold on the argument that the physics-imposed limit to resource/energy growth demands an ultimate cessation of economic growth as well? If so, highlight the persuasive elements. If not, why not?

3 Population

3.1 Population History 31
3.2 Logistic Model 33
Overshoot 35
Logistic Projection 38
3.3 Demographic Transition . . . 38
Geographic Considerations . 41
Transition Cost 44
3.4 Touchy Aspects 46
It's Personal 46
Population Policy 47
3.5 Upshot: It Depends on Us . . 48
3.6 Problems 48

Underlying virtually every concern relating to our experience on this planet is the story of human population. The discussion of continued energy growth in Chapter 1 was based on the historical growth rate of energy, which is partly due to growing population and partly due to increased use per capita. But the notion that population will continue an exponential climb, as is implicit in the Chapter 1 scenario, is impractical—one of many factors that will render the "predictions" of Chapter 1 invalid and prohibit "growth forever."

So let's add a dose of reality and examine a more practical scenario. Americans' per-capita use of energy is roughly *five times* the global average rate. If global population eventually doubles, and the average global citizen advances to use energy at the rate Americans currently do,[1] then the total scale of energy use would go up by a factor of 10, which would take 100 years at our mathematically convenient 2.3% annual rate (see Eq. 1.5; p. 5). This puts a more realistic—and proximate—timescale on the end of energy growth than the fantastical extrapolations of Chapter 1.

1: . . . so that global average energy use per capita increases by a factor of five from where it is today

Although the focus of this chapter will be on the alarming rate of population growth, we should keep the energy and resource context in mind in light of the overall theme of this book. To this end, Figure 3.1 shows the degree to which energy demand has outpaced population growth, when scaled vertically to overlap in the nineteenth century. From 1900 to 1950, per-capita energy consumption increased modestly, but then ballooned dramatically after 1950, so that today we have the equivalent of 25 billion people on the planet operating at nineteenth century energy levels.

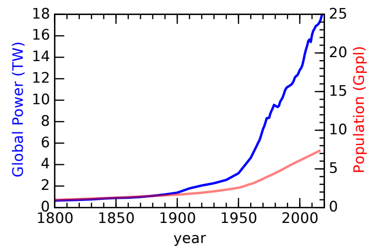


Figure 3.1: Population (red) and energy demand (blue) on the same plot, showing how much faster energy demand (power) has risen compared to population, which translates to increasing per-capita usage. The vertical axes are scaled so that the curves overlap in the nineteenth century. [14–16].

Since population plays a giant role in our future trajectory, we need to better understand its past. We can also gain some sense for theoretical

Photo Credit: Tom Murphy

expectations, then discuss the heralded "demographic transition" and its implications.

3.1 Population History

Figure 3.2 shows a history of global population for the last 12,000 years. Notice that for most of this time, the level is so far down as to be essentially invisible. It is natural to be alarmed by the sharp rise in recent times, which makes the current era seem wholly unusual: an aberration. But wait—maybe it's just a plain exponential function. All exponential functions—ruthless as they are—would show this alarming rise at some point, sometimes called a "hockey stick" plot. In order to peer deeper, we plot population on a logarithmic vertical axis (Figure 3.3). Now we bring the past into view, and can see whether a single exponential function (which would have a constant slope in a logarithmic plot) captures the story.

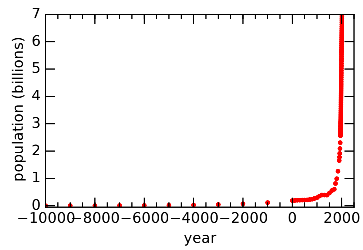


Figure 3.2: Global population estimate, over the modern human era, on a linear scale. Figure 3.1 offers a recent close-up. [14, 15].

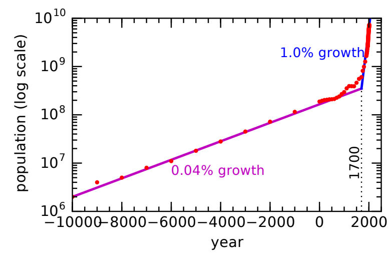


Figure 3.3: Global population estimate, over the modern human era, on a logarithmic scale. [14, 15].

Wait, what? It still looks somewhat like a hockey stick (even more literally so)! How can that be?! This can't be good news. Peering more closely, we can crudely break the history into two eras, each following exponential growth (straight lines on the plot), but at different rates. The early phase had a modest 0.044% growth rate. By the "rule of 70," the corresponding doubling time is about 1,600 years. In more recent times, a 1% rate is more characteristic (70 year doubling). Indeed, we would be justified in saying that recent centuries are anomalous compared to the first 10,000 years of the plot. If we extend the the 0.04% line and the 1% line, we find that they intersect around the year 1700, which helps identify the era of marked transition.

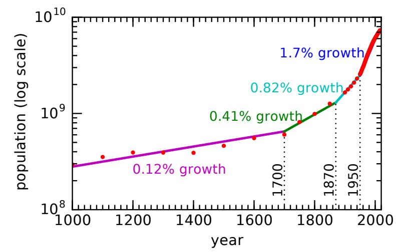


Figure 3.4: Global population estimate, over recent centuries. On the logarithmic plot, lines of constant slope are exponential in behavior. Four such exponential segments can be broken out in the plot, having increasing growth rates. [14, 15].

The recent rapid rise is a fascinating development, and begs for a closer look. Figure 3.4 shows the last ~1,000 years, for which we see several exponential-looking segments at ever-increasing rates. The doubling times associated with the four rates shown on the plot are presented in Table 3.1.

An interpretation of the population history might go as follows. Not much changed during the period following the Dark Ages.[2] The Renaissance (~1700) introduced scientific thinking so that we began to conquer diseases, allowing an uptick in population growth. In the mid-19th century (~1870), the explosive expansion of fossil fuel usage permitted industrialization at a large scale, and mechanized farming practices. More people could be fed and supported, while our mastery over human health continued to improve. In the mid-20th century (~1950), the Green Revolution [17] introduced a fossil-fuel-heavy diet of fertilizer and large-scale mechanization of agriculture, turning food production into an industry. The combination of a qualitative change in the availability of cheap nutrition and the march of progress on disease control cranked the population rate even higher.

2: . . . except that famine and plague took a toll in the 14th century

Table 3.1: Doubling times for Fig. 3.4.

Years	% growth	t_2
1000–1700	0.12%	600 yr
1700–1870	0.41%	170 yr
1870–1950	0.82%	85 yr
1950–2020	1.70%	40 yr

In more recent years, the rate has fallen somewhat from the 1.7% fit of the last segment in Figure 3.4, to around 1.1%. Rounding down for convenience, continuation at a 1% rate would increase population from 7 billion to 8 billion people in less than 14 years. The math is the same as in Chapter 1, re-expressed here as

$$P = P_0 e^{\ln(1+p)(t-t_0)}, \tag{3.1}$$

where P_0 is the population at time t_0, and P is the population at time t if the growth rate is steady at p. Inverting this equation,[3] we have

$$t - t_0 = \frac{\ln\left(\frac{P}{P_0}\right)}{\ln(1+p)}. \tag{3.2}$$

3: . . . recalling that that the natural log and exponential functions "undo" each other (as inverse functions)

Example 3.1.1 We can use Eq. 3.1 to determine how many people we will have in the year 2100 if we continue growing at a 1% rate, starting from 7 billion in the year 2010. We set $P_0 = 7$ Gppl,[4] $t_0 = 2010$, $p = 0.01$, then compute the population in 2100 to be $P = 7e^{\ln 1.01 \cdot 90} = 17$ Gppl.

Eq. 3.2 is the form that was used to conclude that increasing from 7 to 8 Gppl takes less than 14 years at a 1% rate. The computation looks like: $\ln(8/7)/\ln 1.01 = 13.4$. Note that we need not include the factors of a billion in the numerator and denominator, since they cancel in the ratio.

4: Gppl is giga-people, or billion people

The actual time for adding one billion people has lately been 12 years, as we have been growing at a rate slightly higher than 1%.

Year	Population	Time	Rate	Doubling
1804	1 Gppl	—	0.4%	170
1927	2 Gppl	123	0.8%	85
1960	3 Gppl	33	1.9%	37
1974	4 Gppl	14	1.9%	37
1987	5 Gppl	13	1.8%	39
1999	6 Gppl	12	1.3%	54
2011	7 Gppl	12	1.2%	59
2023	8 Gppl	12	1.1%	66

Table 3.2: Population milestones: dates at which we added another one billion living people to the planet. The Time and Doubling columns are expressed in years. Around 1965, the growth rate got up to 2%, for a 35 year doubling time.

Table 3.2 and Figure 3.5 illustrate how long it has taken to add each billion people, extrapolating to the 8 billion mark (as of writing in 2020). The first billion people obviously took tens of thousands of years, each new billion people taking less time ever since. Growth rate peaked in the 1960s at 2% and a doubling time of 35 years. The exponential rate is moderating now, but even 1% growth continues to add a billion people every 13 years, at this stage. A famous book by Paul Ehrlich called *The Population Bomb* [18], first published in 1968, expressed understandable alarm at the 2% rate that had only *increased* to that point. The moderation to 1% since that period is reassuring, but we are not at all out of the woods yet. The next section addresses natural mechanisms for curbing growth.

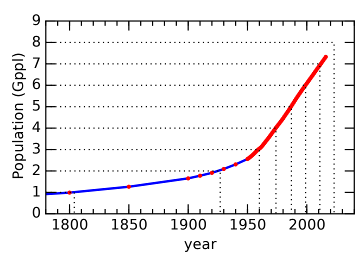


Figure 3.5: Graphical representation of Table 3.2, showing the time between each billion people added [14, 15].

3.2 Logistic Model

Absent human influence, the population of a particular animal species on the planet might fluctuate on short timescales (year by year) and experience large changes on very long timescales (centuries or longer). But by-and-large nature finds a rough equilibrium. Overpopulation proves to be temporary, as exhaustion of food resources, increased predation, and in some cases disease (another form of predation, really) knock back the population.[5] On the other hand, a small population finds it easy to expand into abundant food opportunities, and predators reliant on the species have also scaled back due to lack of prey.

5: For reference, the SARS-CoV2 pandemic of 2020 barely impacted global population growth rates. When population grows by more than 80 million each year, a disease killing even a few million people barely registers as a hit to the broader trend.

We have just described a form of negative feedback: corrective action to remedy a maladjusted system back toward equilibrium.

Definition 3.2.1 ***Negative feedback*** *simply means that a correction is applied in a direction* opposite *the recent motion. If a pendulum moves to the right, a restoring force pushes it back to the left, while moving too far to the left results in a rightward push. A mass oscillating on a spring demonstrates similar characteristics, as must all equilibrium phenomena.*

The word *negative* may sound like something we would not want, but its cousin—positive feedback—leads to disastrous runaway conditions. An example of positive feedback is the bacteria example from Chapter 1: having more bacteria only increases the rate of growth. Exponentials are the hallmark of positive feedback, while equilibrium signals negative feedback.

We can make a simple model for how a population might evolve in an environment hosting negative feedback. When a population is small and resources are abundant, the birth rate is proportional to the population.

Example 3.2.1 If a forest has 100 breeding-aged deer, or 50 couples, we can expect 50 fawns in a year (under the simplifying and unimportant assumption of one fawn per female per year). If the forest has 200 deer, we can expect 100 fawns. The birth rate is simply *proportional* to the population capable of giving birth.[6]

6: . . . no negative feedback yet

If the setup in Example 3.2.1 were the only element to the story, we would find exponential growth: more offspring means a larger population, which ultimately reaches breeding age to produce an even larger population.[7] But as the population grows, negative feedback will begin to play a role. We will denote the population as P, and its rate of change as $\dot{P}$.[8] We might say that the growth rate, or $\dot{P}$, is

$$\dot{P} = rP, \tag{3.3}$$

7: We have just described a state of positive feedback: more begets more.

8: $\dot{P}$ is a time derivative (note the dot on top), defined as $\dot{P} = dP/dt$. But don't panic if calculus is not your thing: what we describe here is still totally understandable.

where r represents the birth rate in proportion to the population (e.g., 0.04 if 4% of the population will give birth in a year).[9] This equation just re-iterates the simple idea that the rate of population growth is dependent on (proportional to) the *present* population. The solution to this differential equation is an exponential:

$$P = P_0 e^{r(t-t_0)}, \tag{3.4}$$

9: In terms of the growth rate we used before, p, as in Eq. 3.1, $r = \ln(1 + p)$. So for instance, if growing at 2%, $p = 0.02$ and r also is 0.02 ($r \approx p$ for small values of p).

which is really just a repeat of Eq. 3.1, where r takes the place of $\ln(1+p)$.

Example 3.2.2 Paralleling the deer population scenario from Example 3.2.1, if we set $r = 0.5$, and have a population of $P = 100$ adult deer (half female), Eq. 3.3 says that $\dot{P} = 50$, meaning the population will change by 50 units.[10]

We could then use Eq. 3.4 to determine the population after 4 years: $P = 100e^{0.5 \cdot 4} \approx 739$.

10: A more adorable term for "units" is fawns, in this case.

We ignore death rate here, but it effectively reduces r in ways that we will encounter later.

Let's say that a given forest can support an ultimate number of deer, labeled Q, in steady state, while the current population is labeled P. The difference, $Q - P$ is the "room" available for growth, which we might think of as being tied to available resources. Once $P = Q$, no more resources are available to support growth.

Definition 3.2.2 *The term "**carrying capacity**" is often used to describe Q: the population supportable by the environment. The carrying capacity (Q) for human population on Earth is not an agreed-upon number, and in any case it is a strong function of lifestyle choices and resource dependence.*

$Q - P$ quantifies a growth-limiting mechanism by representing available room. One way to incorporate this feature into our growth rate equation is to make the rate of growth look like

$$\dot{P} = \frac{Q-P}{Q} rP. \tag{3.5}$$

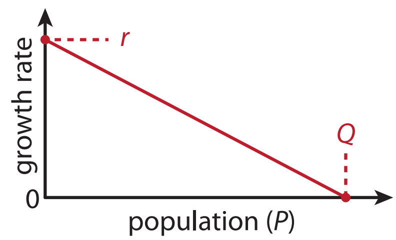


Figure 3.6: The rate of growth in the logistic model decreases as population increases, starting out at r when $P = 0$ and reaching zero as $P \to Q$ (see Eq. 3.5).

We have multiplied the original rate of rP by a term that changes the effective growth rate $r \to r(Q-P)/Q$. When P is small relative to Q, the effective rate is essentially the original r. But the effective growth rate approaches zero as P approaches Q. In other words, growth slows down and hits zero when the population reaches its final saturation point, as $P \to Q$ (see Figure 3.6).

Try it yourself: pick a value for Q (1,000, maybe) and then various values of P to see how the effective growth rate will be modified.

The mathematical solution to this modified differential equation (whose solution technique is beyond the scope of this course) is called a logistic curve, plotted in Figure 3.7 and having the form

$$P(t) = \frac{Q}{1 + e^{-r(t-t_0)}}. \tag{3.6}$$

The first part of the curve in Figure 3.7, for very negative values[11] of $t - t_0$, is exponential but still small. At $t = t_0$ (time of inflection), the population is $Q/2$. As time marches forward into positive territory, P approaches Q. As it does so, negative feedback mechanisms (limits to resource/food availability, predation, disease) become more assertive

11: The parameter t_0 is the time when the logistic curve hits its halfway point. Times before this have negative values of $t - t_0$.

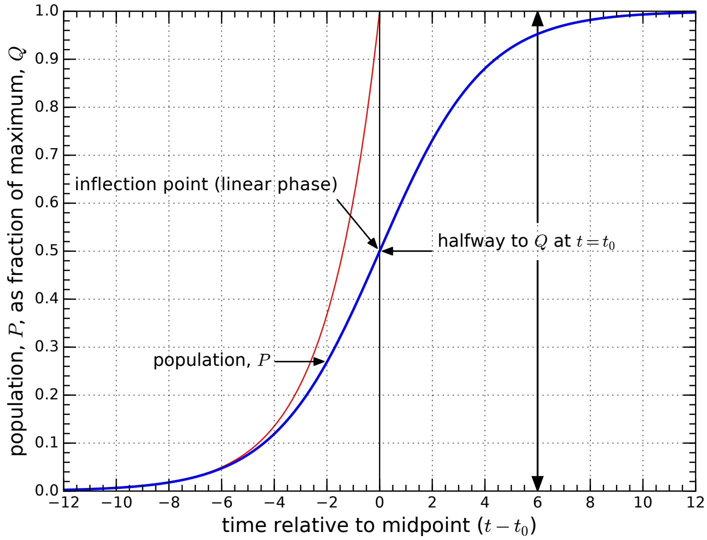


Figure 3.7: Logistic population curve (blue), sometimes called an S-curve, as given in Eq. 3.6, in this case plotting for $r = 0.5$ to match examples in the text. The red curve is the exponential that would result without any negative feedback.

and suppress the rate of growth until it stops growing altogether when P reaches Q.

Example 3.2.3 Continuing the deer scenario, let's say the forest can ultimately support 840 adults,[12] and keep $r = 0.5$ as the uninhibited growth rate. Using these numbers, Eq. 3.6 yields 100 adults at $t = t_0 - 4$ years (effectively the initial state in Example 3.2.1). One year later, at $t = t_0 - 3$, Eq. 3.6 yields 153—very close to the nominal addition of 50 members. But now four years in ($t = t_0$), we have 420 instead of the 739 we got under unrestricted exponential growth in Example 3.2.2.[13]

12: . . . tuned for a convenient match to the numbers we have used in the foregoing examples

13: Not coincidentally, $P = Q/2$ at the halfway point, $t = t_0$.

The logistic curve is the *dream scenario*: no drama. The population simply approaches its ultimate value smoothly, in a tidy manner. We might imagine or hope that human population follows a similar path. Maybe the fact that we've hit a linear phase—consistently adding one billion people every 12 years, lately—is a sign that we are at the inflection, and will start rolling over toward a stable endpoint. If so, we know from the logistic curve that the linear part is halfway to the final population.

Three consecutive 12-year intervals appear in Table 3.2. If the middle one is the midpoint of a logistic linear phase—in 2011 at 7 billion people—it would suggest an ultimate population of 14 billion.

3.2.1 Overshoot

But not so fast. We left out a crucial piece: feedback delay. The math that leads to the logistic curve assumes that the negative feedback[14] acts instantaneously in determining population rates.

14: . . . based on remaining resources, $Q - P$, at the moment in Eq. 3.5

Consider that human decisions to procreate are based on present conditions: food, opportunities, stability, etc. But humans live for many decades, and do not impose their full toll on the system until many years after birth, effectively delaying the negative feedback. The logistic curve and equation that guided it had no delay built in.

Definition 3.2.3 ***Overshoot*** *is a generic consequence of delaying negative feedback. Since negative feedback is a "corrective," stabilizing influence, delaying its application allows the system to "get away" from the control, thereby exceeding the target equilibrium state.*

By "generic consequence," we just mean an outcome that is characteristic of the situation, independent of details.

This is a pretty easy concept to understand. The logistic curve of Figure 3.7 first accelerates, then briefly coasts before decelerating to arrive smoothly at a target. Following an example from [1], it is much like a car starting from rest by accelerating before applying the brakes to gently come to a stop when the bumper barely kisses a brick wall. The driver is operating a negative feedback loop: seeing/sensing the proximity to the wall and slowing down accordingly. The closer to the wall, the slower the driver goes until lightly touching the wall. Now imagine delaying the feedback to the driver by applying a blindfold and giving voice descriptions of the proximity to the wall, so that decisions about how much to brake are based on conditions from a delayed communication process. Obviously, the driver will crash into the wall if the feedback is delayed, unless slowing down the whole process dramatically. Likewise, if the negative consequences—signals that we need to slow down population growth—arrive decades after the act of producing more humans, we can expect to exceed the "natural" limit, Q—a condition called overshoot.

[1]: Meadows et al. (1972), *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*

Another example of feedback delay leading to overshoot: let's say you are holding down the space bar and trying to position the cursor in the middle of the screen. But your connection is lagging and even though you release the space bar when you *see* the cursor reach the middle, it keeps sailing past due to the delay: overshooting.

Example 3.2.4 We did not detail the mechanisms of negative feedback operating on the deer population in Example 3.2.3 that act to stabilize the population at Q, but to illustrate how delayed negative feedback produces overshoot, consider predation as one of the operating forces. To put some simple numbers on it, let's say that steady state can support one adult (hunting) mountain lion for every 50 deer. Initially, when the population was 100 deer, this means two predators. When the deer population reaches $Q = 840$, we might have ~17 predators. But it takes time for the predators to react to the growing number of prey, perhaps taking a few years to produce the requisite number of hunting adults. Lacking the full complement of predators, the deer population will sail past the 840 mark until the predator population rises to establish the ultimate balance. In fact, the predators will likely also exceed their steady population in a game of catch-up that leads to oscillations like those seen in Figure 3.8.

We can explore what happens to our logistic curve if the negative feedback is delayed by various amounts. Figure 3.8 gives a few examples of overshoot as the delay increases. To avoid significant overshoot, the delay (τ) needs to be smaller than the natural timescale governing the problem: $1/r$, where r is the rate in Eqs. 3.5 and 3.6. In our deer example using $r = 0.5$, any delay longer than about 2 years causes overshoot. For more modest growth rates (human populations), relevant delays are in decades (see Box 3.1).

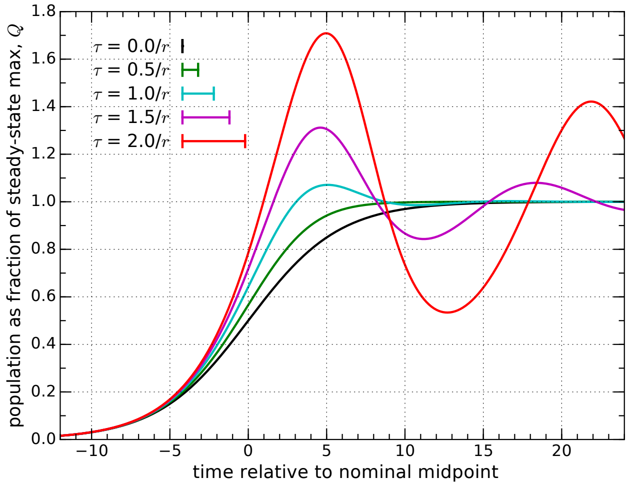


Figure 3.8: Feedback delay generally results in overshoot and oscillation, shown for various delay values, τ. The black curve ($\tau = 0$) is the nominal no-delay logistic curve. As the delay increases, the severity of overshoot increases. Delays are explored in increments of 0.5 times the characteristic timescale of $1/r$ (using $r = 0.5$ here to match previous examples, so that a delay of $\tau = 1.5/r$ equates to 3 time units on the graph, for instance). The delay durations are also indicated by bar lengths in the legend.

Eventually all the curves in Figure 3.8 converge to the steady state value of 1.0,[15] but human population involves complexities not captured in this bare-bones mathematical model.[16] All the same, the generic phenomenon of overshooting when negative feedback is delayed is a robust attribute, even if the oscillation and eventual settling does not capture the future of human population well.

15: . . . meaning that population P arrives at Q

16: For instance, a dramatic overshoot and collapse could be disruptive enough to take out our current infrastructure for fossil-fuel-aided agriculture so that the Q value essentially resets to some lower value.

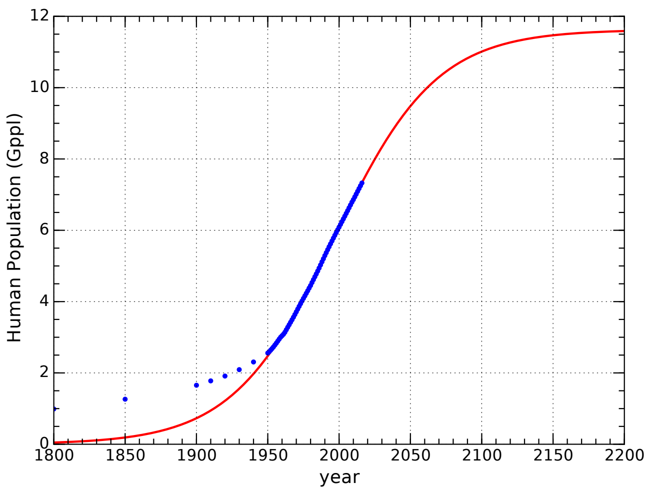


Figure 3.9: Human population data points (blue) and a logistic curve (red) that represents the best fit to data points from 1950 onward. The resulting logistic function has $Q \approx 12$ Gppl, $r = 0.028$, and a midpoint at the year 1997. The actual data sequence has a sudden bend at 1950 (Green Revolution?) that prevents a suitable fit to a larger span of data. In other words, the actual data do not follow a single logistic function very well, which is to be expected when conditions change suddenly (energy and technology, in this case) [14, 15].

Box 3.1: Will Human Population Overshoot?

Are humans in danger of population overshoot? What is our r value? It is tempting to take $r = 0.01$ corresponding to the present 1% growth rate. This would imply that any delay shorter than 100 years will not produce significant overshoot, which seems reassuring. But if human population is following a logistic curve rather than an exponential, resource availability is already exerting a moderating influence, now appearing to be in the linear "cruise" phase roughly halfway to the limiting value. A fit to the data (Figure 3.9) suggests that $r \approx 0.028$, corresponding to a timescale of 36 years ($1/r$). This puts the overshoot-prone delay squarely into relevant timescales for human lifetimes, generations, and societal change—thus leaving the door open for an overshoot scenario.

3.2.2 Logistic Projection

As suggested by Figure 3.9, Human population is *not* following a strict logistic curve. If it were, the early period would look exponential at the ~2.8% rate corresponding to the best-fit logistic matching our recent trajectory, but growth was substantially slower than 2.8% in the past. Technology and fossil fuels have boosted our recent growth well beyond the sub-percent rates typical before ~1950. The point is that while reference to mathematical models can be extremely helpful in framing our thinking and exposing *robust, generic modes* of interest, *we should seldom take any mathematical model* ***literally***, as it likely does not capture the full complexity of the system it is trying to model. In the present case, it is enough to note that:

1. exponentials relentlessly drive toward infinity (ultimately unrealistic);
2. logistic curves add a sensible layer of reality, capping growth in some steady-state outcome;
3. other dynamical factors such as delays can prevent a smooth logistic function, possibly leading to overshoot; and
4. many other factors (medical and technological breakthroughs, war, famine, climate change) can muddy the waters in ways that could make the situation better or worse than simple projections.

3.3 Demographic Transition

Perhaps not surprisingly, the rate of a country's population growth is correlated to its wealth, as seen in Figure 3.10. An attractive path to reducing population growth would be to have poor countries slide down this curve to the right: becoming more affluent and transforming societal values and pressures accordingly to produce a lower net population growth rate.

Population growth happens when the birth rate exceeds the death rate.

Definition 3.3.1 ***Birth rate****, typically expressed in births per 1,000 people per year, minus* ***death rate*** *(also in deaths per 1,000 people per year) is the net population rate.*[17] *If the difference is positive, the population grows, and it shrinks if the difference is negative.*

17: This ignores immigration, which just shifts living persons around.

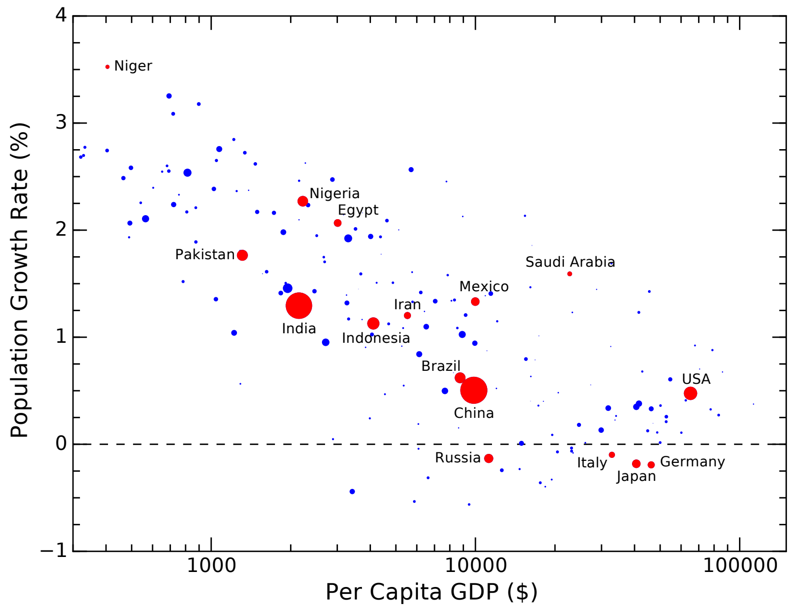


Figure 3.10: Net population rate, in percent, as a function of per-capita GDP. A clear trend shows wealthier countries having lower growth rates. A win–win solution would seem to present itself, in which everyone arrives at the lower right-hand side of this graph: more money for all and a stable population! Dot size (area) is proportional to population [6, 8, 19, 20].

Example 3.3.1 The U.S. has a birth rate of about 12 people per 1,000 per year, and a death rate of 8.1 people per 1,000 per year. The net rate is then roughly +4 per 1,000 per year, translating to 0.4% net growth.[18]

Niger has a birth rate of 46 per 1,000 per year and a death rate of 11, resulting in a net of positive 35, or 3.5%.

18: 4 per 1,000 is 0.4 per 100, which is another way to say 0.4 percent.

As conditions change, birth and death rates need not change in lock-step. Developed countries tend to have low birth rates *and* low death rates, balancing to a relatively low net population growth rate. Developing countries tend to have high death rates and even higher birth rates, leading to large net growth rates. Figure 3.11 depicts both birth rates and death rates for the countries of the world. A few countries (mostly in Europe) have slipped below the replacement line, indicating declining population.[19]

19: Note that immigration is not considered here: just birth rate and death rate within the country.

The general sense is that developed countries have "made it" to a responsible low-growth condition, and that population growth is driven by poorer countries. An attractive solution to many[20] is to bring developing countries up to developed-country standards so that they, too, can settle into a low growth rate. This evolution from a fast-growing poor country to a slow (or zero) growth well-off country is called the demographic transition.

20: . . . but unsolicited "preaching" to others

Definition 3.3.2 *The **demographic transition** refers to the process by which developing countries having high death rates and high birth rates adopt technologies, education, and higher standards of living that result in low death rates* and *low birth rates, more like advanced countries.*

In order to accomplish this goal, reduced death rates are facilitated by

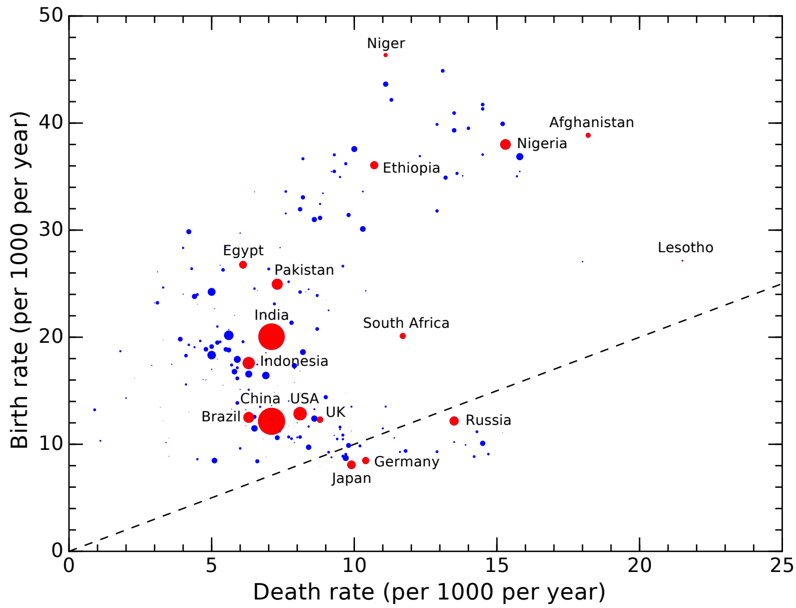


Figure 3.11: Birth rates and death rates for countries, where dot size is proportional to population. The diagonal line indicates parity between birth and death rates, resulting in no population growth. Countries above the line are growing population, while countries below are shrinking. A few countries fall a bit below this line [8, 19, 20].

introducing modern medicine and health services to the population. Reduced birth rates are partly in response to reduced infant mortality—eventually leading to fewer children as survival is more guaranteed. But also important is better education—especially among women in the society who are more likely to have jobs and be empowered to exercise control of their reproduction (e.g., more say in relationships and/or use of contraception). All of these developments take time and substantial financial investment.[21] Also, the economy in general must be able to support a larger and better-educated workforce. The demographic transition is envisioned as a transformation or complete overhaul, resulting in a country more in the mold of a "first-world" country.[22]

21: Better hospitals and schools are not free.

22: One may justifiably question whether this is the "correct" goal.

Figure 3.11 hints at the narrative. Countries are spread into an arc, one segment occupying a band between 5–10 deaths per 1,000 people per year and birth rates lower than 20 per 1,000 people per year. Another set of countries (many of which are in Africa) have birth rates above 20 per 1,000 per year, but also show higher death rates. The narrative arc is that a country may start near Lesotho, at high death and birth rates, then migrate over toward Nigeria as death rates fall (and birth rates experience a temporary surge). Next both death and birth rates fall and run through a progression toward Pakistan, India, the U.S., and finally the European steady state. Figure 3.12 schematically illustrates the typical journey.

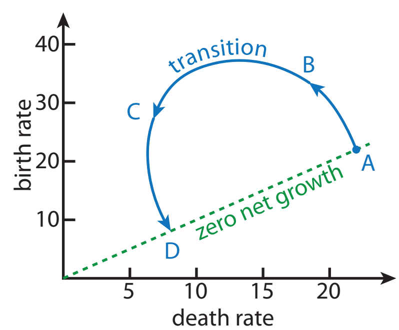


Figure 3.12: Schematic of how the demographic transition may play out in the space plotted in Figure 3.11. At points A and D, birth rates and death rates are equal, resulting in no population growth. Typically, death rates decline while birth rates increase (point B), and eventually death rates reach a floor while birth rates begin to fall (at C).

The demographic transition receives widespread advocacy among Western intellectuals for its adoption, often coupled with the sentiment that it can't come soon enough. Indeed, the humanitarian consequences appear to be positive and substantial: fewer people living in poverty and hunger;

empowered women; better education; more advanced jobs; and greater tolerance in the society. It might even seem condemnable *not* to wish for these things for all people on Earth.

However, we need to understand the consequences. Just because we *want* something does not mean nature will comply. Do we have the resources to accomplish this goal? If we fail in pursuit of a global demographic transition, have we unwittingly unleashed even *greater* suffering on humanity by increasing the total number of people who can no longer be supported? It is possible that well-intentioned actions produce catastrophic results, so let us at least understand what is at stake. It may be condemnable *not to wish* for a global demographic transition, but failing to explore potential downsides may be equally ignoble.

3.3.1 Geographic Considerations

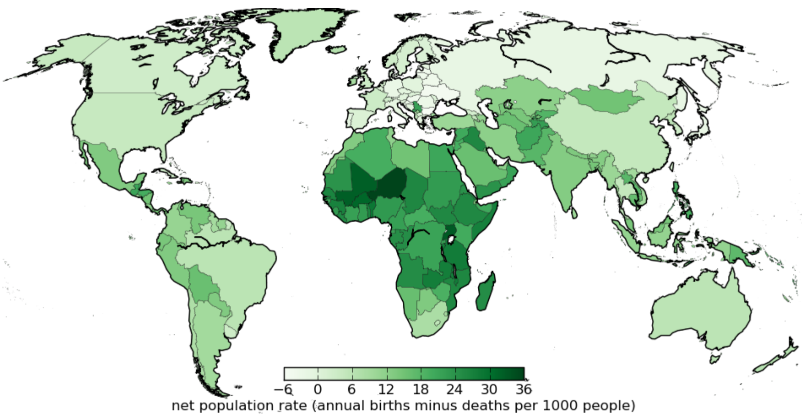


Figure 3.13: Net population growth rate by country: birth rate minus death rate per 1,000 people per year. The highest net growth (darkest shading) is Niger, in Saharan Africa [19, 20].

Figure 3.13 shows the net population rate (birth minus death rate) on a world map. Africa stands out as the continent having the largest net population growth rate, and has been the focus of much attention when discussing population dynamics.

But let us cast population rates in different countries in a new light. Referring to Figure 3.13, it is too easy to look at Niger's net population rate—which is about ten times higher than that of the U.S. (see Example 3.3.1)—and conclude that countries similar to Niger present a greater risk to the planet in terms of population growth. However, our perspective changes when we consider absolute population levels. Who cares if a country's growth rate is an explosive 10% if the population is only 73 people?[23]

23: But check back in 100 years!

Figure 3.14 multiplies the net rate by population to see which countries contribute the most *net new people* to the planet each year, and Table 3.3 lists the top ten. Africa no longer appears to be the most worrisome region in this light.[24] India is the largest people-producing country at present, adding almost 18 million per year. Far behind is China, in second

24: Although, the continent as a whole accounts for 35% of the total added population each year.

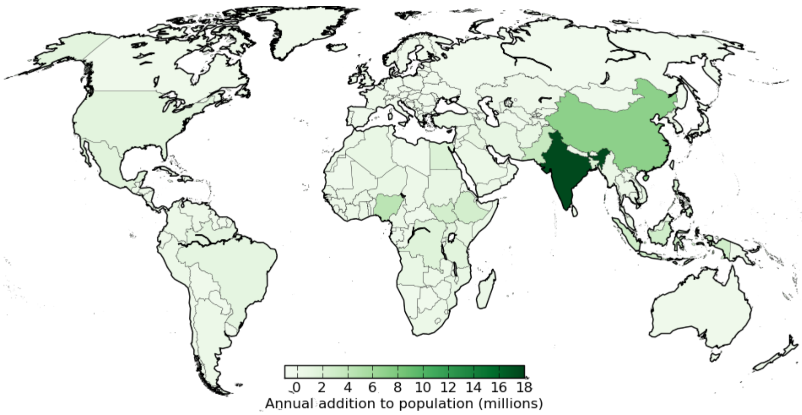


Figure 3.14: Absolute population growth rate by country: how many millions of people are added per year (birth rate minus death rate times population) [8, 19, 20].

place. The U.S. adds about 1.6 million per year, a little beyond the top ten. This exercise goes to show that context is important in evaluating data.

Country	Population (millions)	Birth Rate	Death Rate	Annual Millions Added
India	1,366	20.0	7.1	17.7
China	1,434	12.1	7.1	7.2
Nigeria	201	38.0	15.3	4.6
Pakistan	216	24.9	7.3	3.8
Indonesia	271	17.6	6.3	3.1
Ethiopia	112	36.1	10.7	2.8
Bangladesh	163	20.2	5.6	2.3
Philippines	108	24.2	5.0	2.1
Egypt	100	26.8	6.1	2.1
DR Congo	87	36.9	15.8	1.8
Whole World	7,711	19.1	8.1	86

Table 3.3: Top ten populators [8, 19, 20], in terms of absolute number of people added to each country. Birth rates and death rates are presented as number per 1,000 people per year. These ten countries account for 55% of population growth worldwide.

Adding another relevant perspective, when one considers that the per-capita energy consumption in the United States is more than 200 times that of Niger,[25] together with the larger U.S. population, we find that the *resource impact* from births is almost 400 times higher for the U.S. than for Niger.[26] On a per capita basis, a citizen of the U.S. places claims on future resources at a rate 28 times higher than a citizen of Niger via population growth.[27] On a finite planet, the main reason we *care* about population growth is in relation to limited resources. Thus from the resource point of view, the problem is not at all confined to the developing world. Table 3.4 indicates how rapidly the top ten countries are creating energy demand (as a proxy to resource demands in general) based on population growth alone. Figure 3.15 provides a graphical perspective of the same (for all countries). For reference, one gigawatt (GW) is the equivalent of a large-scale nuclear or coal-fired power plant. So China, the U.S., and India each add the equivalent of 10–20 such plants per year just to satisfy the demand created by population growth.[28]

25: The average American rate of energy use is 10,000 W vs. 50 W for Niger.

26: In other words, for every additional kilogram of coal, steel, or whatever required by Niger's added population, the U.S. will require 400 kg of the same to satisfy its population growth.

27: 28 is smaller than 400 by the ratio of populations in the two countries.

28: This does not even consider rising standards placing additional burdens.

Country	Population ($\times 10^6$)	Annual Growth ($\times 10^6$)	Per Capita Power (W)	Power Added Annually (GW)	Power Added Per Citizen (W)
China	1,434	7.2	2,800	20.2	14
United States	329	1.6	10,000	15.6	48
India	1,366	17.7	600	10.5	8
Saudi Arabia	34	0.54	10,100	5.5	160
Iran	83	1.0	4,300	4.3	52
Mexico	128	1.7	2,000	3.3	26
Indonesia	271	3.1	900	2.8	10
Brazil	211	1.3	2,000	2.7	13
Egypt	100	2.1	1,200	2.5	25
Turkey	83	0.85	2,100	1.8	21
Whole World	7,711	86	2,300	143	18.4

Table 3.4: Top ten countries for growth in energy demand. Populations are in millions. Power is in Watts or 10^9 W (GW). The power added annually is the absolute increase in demand *due to population growth*, and is a proxy for resource demands in general. The last column provides some measure of an individual citizen's share of the responsibility in terms of increasing pressure on resources. The top three contributors to new power demand via population growth alone (China, the U.S., and India) account for about a third of the global total. [7, 8, 19, 20]

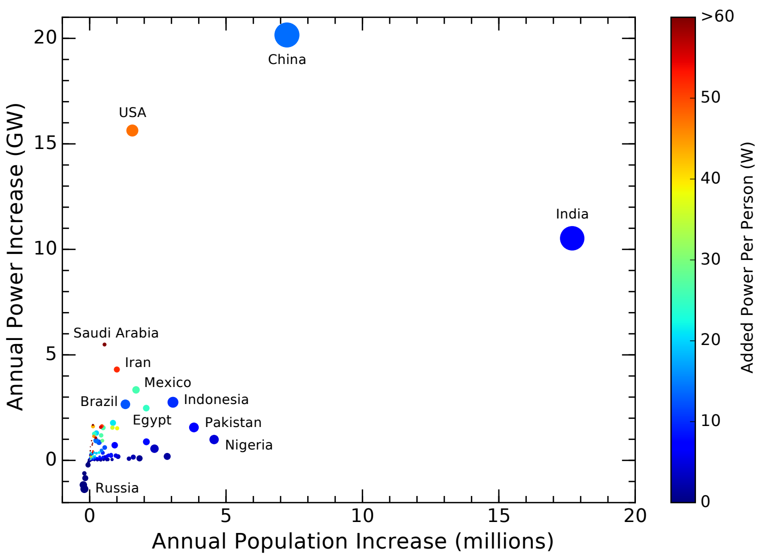


Figure 3.15: Graphical representation of Table 3.4, for all countries. Dots, whose size is proportional to population, indicate how many people are added per year, and how much additional energy demand is created as a consequence. Color indicates the added population-growth-driven power demand an individual citizen is responsible for generating each year as a member of the society. Negative cases (contracting) include Russia, Japan, Germany, and Ukraine [7, 8, 19, 20].

The last column in Table 3.4 is the per-citizen cost, meaning, for instance that each person in the U.S. adds about 50 Watts per year of energy demand via the country's net population growth rate.[29] In this sense, the last column is a sort of "personal contribution" an individual makes to the world's resource demands via net population rates and consumption rates in their society. Those having high scores should think twice about assigning blame externally, and should perhaps tend to their own house, as the saying goes.

29: A citizen of Niger, by comparison, only adds 1.7 W of demand per year on energy resources via population growth.

Before departing this section, let us look at continent-scale regions rather than individual countries in terms of adding people and resource demands. Table 3.5 echoes similar information to that in Table 3.4, in modified form. What we learn from this table is that Asia's demands are commensurate with their already-dominant population; North America creates the next largest pressure despite a much smaller population;

Country	Population (%)	Annual Growth (%)	Per Capita Power (W)	Power Added Annually (%)	Power Added Per Citizen (W)
Asia	59.7	55.1	1,800	60.5	18.9
N. America	7.6	5.5	7,100	23.0	56.1
Africa	16.9	34.7	500	9.9	10.8
S. America	5.5	4.4	2,000	5.4	18.1
Oceania	0.5	0.5	5,400	1.5	49.5
Europe	9.7	−0.1	4,900	−0.3	−0.6
Whole World	7,711 M	86 M	2,300	143 GW	18.4

Table 3.5: Population pressures from regions of the world, ranked by added power demand. Some of the columns are expressed as percentages of the total. The bottom row has totals in millions of people or total GW in place of percentages. [7, 8, 19, 20]

Africa is significant in terms of population growth, but constitutes only 10% of resource pressure at present. Finally, Europe holds 10% of the globe's people but lays no claim on added resources via population growth, resembling the target end-state of the demographic transition.[30]

30: Note that European countries are nervous about their decline in a growing, competitive world.

3.3.2 Cost of the Demographic Transition

A final point relates to the trajectory depicted in Figure 3.12 for demographic transitions: death rate decreases first while birth rates remain high—or rise even higher—before starting to come down. An example sequence is illustrated in Figure 3.16: initially the rates are high (at r_1), and the same (resulting in steady population); then the death rate transitions to a new low rate (r_2) over a time T; and the birth rate begins to fall some time τ later before matching the death rate and stabilizing population again. The yellow-shaded area between the curves indicates the region where birth rate exceeds death rate, leading to a net population growth (a surge in population).

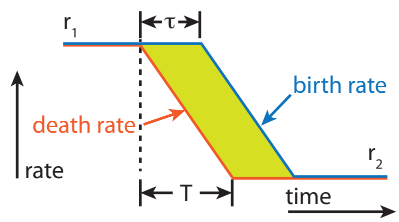


Figure 3.16: Schematic demographic transition time sequence.

The amount of growth in the surge turns out to be proportional to the exponential of the area between the curves. For this trapezoid cartoon, the area is just the base (τ) times the height (rate difference), so that the population increase looks like $e^{(r_1-r_2)\tau}$, where r_1 is the initial rate per year and r_2 is the final rate. The actual curves may take any number of forms, but the key point is that delayed onset of birth rate decrease introduces a population surge, and that magnitude of the surge grows as the area between the curves increases.

Note in the cartoon example of Figure 3.16, the area between the curves is only dependent on the rate difference (height) and the delay, τ. The time it takes to complete the transition, T, is irrelevant, as the area of the parallelogram is just the base times the height. Thus the population surge associated with a demographic transition is primarily sensitive to the rate difference and the delay until birth rate begins to decline.

Example 3.3.2 If we start at a birth/death rate of 25 per 1,000 per year ($r_1 = 0.025$), end up at 8 ($r_2 = 0.008$; verify that these numbers are reasonable according to Figure 3.11), and have a delay of $\tau = 50$ years for the birth rate to start decreasing, we see the population increasing by a factor of

$$e^{(r_1-r_2)\tau} = e^{(0.025-0.008)\cdot 50} = e^{0.85} = 2.34.$$

This means that the population more than doubles, or increases by 134%.

So to effect a demographic transition means to increase the population burden substantially. Meanwhile, the transitioned population consumes resources at a greater rate—a natural byproduct of running a more advanced society having better medical care, education, and employment opportunities. Transportation, manufacturing, and consumer activity all increase. The net effect is a double-whammy: the combined impact of a *greater* population using *more* resources per capita. The resource impact on the planet soars.

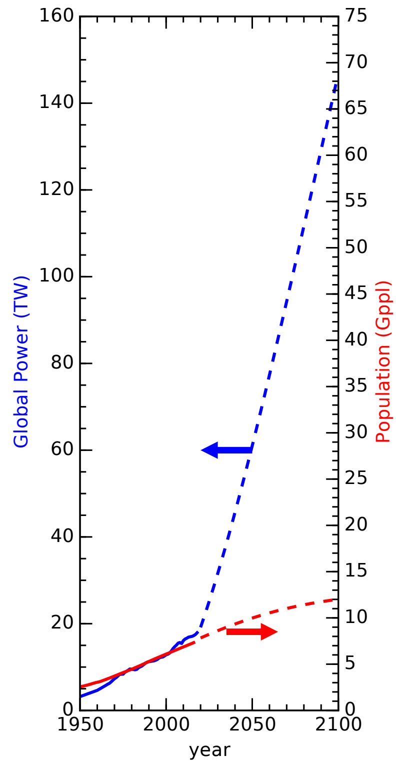


Figure 3.17: What our energy demand would have to do (blue-dashed line) if the growing global population (here projected as a red-dashed logistic curve) grew its per-capita energy consumption to current U.S. standards by the year 2100 (a factor-of-five increase). Historical energy and population are represented as solid curves. The departure from past reality would have to be staggering [15, 16].

The pertinent question is whether the Earth is prepared to host a dramatic increase in resource usage. Just because we might find appealing the idea that all countries on Earth could make it through the demographic transition and live at a first-world standard *does not mean nature has the capacity to comply*. The U.S. per-capita energy usage is roughly five times the current global average. To bring 7 billion people to the same standard would require five times the current scale. Completion of a global demographic transition would roughly double the current world population so that the total increase in energy would be a factor of ten. The blue-dashed projection in Figure 3.17 looks rather absurd as an extension of the more modest—but still rather remarkable—energy climb to date. As we are straining to satisfy current energy demand, the "amazing dream" scenario seems unlikely to materialize.

Energy in this context is a proxy for other material resources. Consider the global-scale challenges we have introduced today: deforestation, fisheries collapse, water pressures, soil degradation, pollution, climate change, and species loss, for instance. What makes us think we can survive a global demographic transition leading to a consumption rate many times higher than that of today? Does it not seem that we are already approaching a breaking point?

If nature won't let us realize a particular dream, then is it morally responsible to pursue it? This question becomes particularly acute if the very act of *pursuing* the dream increases the pressure on the system and *makes failure even more likely*. Total suffering might be maximized if the population builds to a point of collapse. In this sense, we cleverly stack the most possible people into the stadium to witness a most spectacular event: the stadium's collapse—which only happened because we packed the stadium. You see the irony, right?

The drive to realize a global demographic transition is strong, for the obvious set of reasons discussed above (improved quality of life, educational opportunity, greater tolerance, dignity, and fulfillment). Challenging the vision may be an uphill battle, since awareness about resource limits is not prevalent. This may be an example of the natural human tendency to extrapolate: we have seen the benefits of the demographic transition in many countries over the last century, and may expect this trend to continue until all countries have completed the journey. But bear in mind that earlier successes transpired during times in which global resource availability was not a major limitation. If conditions change, and we reach a "full" earth, past examples may offer little relevant guidance.

3.4 Touchy Aspects

3.4.1 Population Discussions Quickly Get Personal

Some of the decisions we make that translate into impact on our physical world are deeply personal and very difficult to address. No one wants to be told what they should eat, how often they should shower, or what temperature they should keep their dwelling. But the touchiest of all can be reproduction. It can be tricky to discuss population concerns with someone who has kids. Even if not intentional, it is too easy for the topic to be perceived as a personal attack on their own choices. And we're not talking about choices like what color socks to wear. Children are beloved by (most) parents, so the insinuation that having children is bad or damaging quickly gets tangled into a sense that their "angel" is being attacked—as is their "selfish" decision to have kids (see Box 3.2). The disconnect can be worse the larger the number of kids someone has. Couples having two kids take some solace in that they are exercising net-zero "replacement."

Having two kids is not a *strict* replacement, in that parents and children overlap (double-occupancy) on Earth. But the practice is at least consistent with a steady state.

One common side-step is to focus attention on the high birthrates in other countries, so that the perceived fault lies elsewhere. As pointed out above, if stress on the planet—and living within our means—is what concerns us, undeveloped countries are not putting as much pressure on global resources as many of the more affluent countries are. So while pointing elsewhere offers a bit of a relief, and is a very natural tendency, it does not get the whole picture.

The overall point is to be aware of the sensitive nature of this topic when discussing with others. Making someone feel bad about their choices—even if unintentionally—might *in rare cases* result in an appreciation and greater awareness. But it is *more likely* to alienate a person from an otherwise valuable perspective on the challenges we face.

Box 3.2: Which is More Selfish?

Parents, many of whom sacrifice dearly in raising kids—financially, emotionally, and in terms of time investment—understandably view their tireless commitment as being selfless: they often give up their own time, comfort, and freedom in the process. It is understandable, then, that they may view those not having kids as being selfish: the opposite of selfless. But this can be turned on its head. Why, exactly, did they decide to have kids and contribute to the toll on our planet? It was their choice (or inattention) that placed them in parental roles, and the entire planet—not just humans—pays a price for their decision, making it seem a bit selfish.[31] In the end, almost any decision we make can be called selfish, since we usually have our own interests at least partly in mind. So it is pointless to try assigning more or less selfishness to the decision to have kids or not to have them. But consider this: if the rest of the Earth—all its plants and creatures—had a say, do you think they would vote for adding another human to the planet? Humans have the *capacity*, at least, to consider a greater picture than their own self interest, and provide representation to those sectors that otherwise have no rights or voice in our highly human-centric system.

31: Reasons for having children are numerous: genetic drive; family name/tradition; labor source; care in old age; companionship and love (projected onto not-yet-existing person). Note that adoption can also satisfy many of these aims without contributing additional population.

3.4.2 Population Policy

What could governments and other organizations do to manage population? Again, this is touchy territory, inviting collision between deeply personal or religious views and the state. China initiated a one-child policy in 1979 that persisted until 2015 (exceptions were granted depending on location and gender). The population in China never stopped climbing during this period, as children born during prior periods of higher birth rates matured and began having children of their own—even if restricted in number. The population curve in China is not expected to flatten out until sometime in the 2030–2040 period.[32] Such top-down policies can only be enacted in strong authoritarian regimes, and would be seen as a severe infringement on personal liberties in many countries. Religious belief systems can also run counter to deliberate efforts to limit population growth. In addition, shrinking countries are at a competitive disadvantage in global markets, often leading to policies that incentivize having children.

The net effect of the various exceptions meant that for most of this period half of Chinese parents could have a second child.

32: This is another case of delay in negative feedback resulting in overshoot.

One striking example of rarely-achieved sustainable population control comes from the South Pacific island of Tikopia [21]. Maintaining a stable population for a few thousand years on this small island involved not only adopting food practices as close to the island's natural plants as possible, but also invoking strict population controls. The chiefs in this egalitarian society routinely preached zero population growth, and also prevented overfishing. Strict limits were placed on family size, and cultural taboos kept this small island at a population around 1,200.[33] Population control methods included circumventing insemination, abortion, infanticide, suicide, or "virtual suicide," via embarking on dangerous sea voyages unlikely to succeed. In this way, the harshness of nature was replaced by harsh societal norms that may seem egregious to us. When Christian missionaries converted inhabitants in the twentieth century, the practices of abortion, infanticide and suicide were quenched and the population

[21]: Diamond (2005), *Collapse: How Societies Choose to Fail or Succeed*

33: A group size of 1,200 is small enough to prevent hiding irresponsible actions behind anonymity.

began to climb, leading to famine and driving the population excess off the island.

Nature, it turns out, is indifferent to our belief systems.

In the end, personal choice will be important, if we are to tame the population predicament. Either conditions will be too uncertain to justify raising children, or we adopt values that place short term personal and human needs into a larger context concerning ecosystems and long-term human happiness.

3.5 Upshot: Everything Depends on Us

We would likely not be discussing a finite planet or limits to growth or climate change if only one million humans inhabited the planet, even living at United States standards. We would perceive no meaningful limit to natural resources and ecosystem services. Conversely, it is not difficult to imagine that 100 billion people on Earth would place a *severe* strain on the planet's ability to support us—especially if trying to live like Americans—to the point of likely being impossible. If we had to pick a single parameter to dial in order to ease our global challenges, it would be hard to find a more effective one than population.

A common knee-jerk reaction to a statement that we would be better off with a smaller population is to demand an answer to who, exactly, we propose eliminating. Ideally, we should be able to discuss an important topic like population without resorting to accusations of advocating genocide. Of course we need to take care of those already alive, and address the problem via future reproductive choices.

Maybe we need not take any action. Negative feedback *will* assert itself strongly once we have gone too far—either leading to a steady approach to equilibrium or producing an overshoot/collapse outcome. Nature *will* regulate human population one way or another. It just may not be in a manner to our liking, and we have the opportunity to do better via awareness and choice.

Very few scholars are unconcerned about population pressure. Yet the issue is consistently thorny due to both its bearing on personal choice and a justified reluctance to "boss" developing nations to stop growing prior to having an opportunity to naturally undergo a demographic transition for themselves. Conventional thinking suggests that undergoing the demographic transition ultimately is the best solution to the population problem. The question too few ask is whether the planet can support this path for all, given the associated population surge and concomitant demand on resources. If not, pursuit of the transition for the world may end up causing more damage and suffering than would otherwise happen due to increased populations competing for dwindling resources.

3.6 Problems

1. The text accompanying Figure 3.1 says that Earth currently hosts the equivalent of 25 billion nineteenth-century-level energy consumers. If we had maintained our nineteenth-century energy appetite but

followed the same population curve, what would our global power demand be today, in TW? How does this compare to the actual 18 TW we use today?

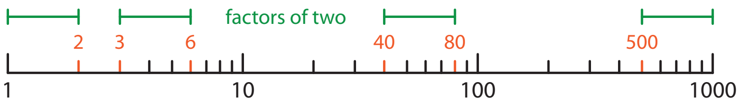
Hint: It is perfectly acceptable to hold a (preferably transparent) straight-edge up to a graph!

2. Notice that on logarithmic plots,[34] factors of ten on the logarithmic axis span the *same distance*. This applies for *any* numerical factor—not just ten.[35] Shorter (minor) tick marks between labeled (major) ticks multiply the preceding tick label by 2, 3, 4, 5, 6, 7, 8, 9. The graphic below illustrates the constant distance property for a factor of two.[36] Now try a different multiplier (not 2 or 10), measuring the distance between tick marks, and report/draw how you graphically verified that your numerical factor spans the same distance no matter where you "slide" it on the axis.

34: See, for example, Figures 3.3 and 3.4.

35: This is due to the property of logarithms that $\log\left(\frac{a}{b}\right) = \log a - \log b$. The property applies for any base, so $\log_{10}$ and ln behave the same way.

36: The green bars indicate that the same distance from 1 to 2 applies to 3–6, 40–80, and 500–1,000.



3. Looking at Figure 3.3, if humans had continued the slow growth phase characteristic of the period until about 1700, what does the *graph* suggest world population would be today, approximately, if the magenta line were extended to "now?"[37] Put the answer in familiar terms, measured in millions or billions, depending on what is most natural.[38]

37: Determine graphically (may need to zoom in). See Problem 2 and the associated graphic to better understand how the tick marks work.

38: I.e., don't say 0.01 billion if 10 million is more natural, or 8,000 million when 8 billion would do.

4. Looking at Figure 3.4, if humans had continued the moderate growth phase characteristic of the period from the year 1000 to 1700, what does the *graph* suggest world population would be today, approximately, if the magenta line were extended to "now?"

See margin notes for Problem 3.

5. If we were to continue a 1% population growth trajectory into the future, work out how many years it would take to go from 7 billion people to 8 billion, and then from 8 billion to 9 billion.

6. At present, a billion people are added to the planet every 12 years. If we maintain a 1% growth rate in population, what will global population be in 2100 (use numbers in Table 3.2 as a starting point), and how quickly will we add each new billion at that point?

7. A decent approximation to recent global population numbers using a logistic function is[39]

$$P = \frac{14}{1 + \exp[-0.025 \times (year - 2011)]}$$

in billions of people. First verify that inserting the year 2011 results in 7 (billion), and then add a column to Table 3.2 for the "prediction" resulting from this function. Working back into the past, when does it really start to deviate from the truth, and why do you

39: See Eq. 3.6.

think that is (hint: what changed so that we invalidated a single, continuous mathematical function)?

8. Using the logistic model presented in Problem 7, what would the population be in the year 2100? How does this compare to the exponential result at 1% growth as in Problem 6?

Comparison of this problem and Problem 6 highlights the difference the choice of mathematical model can make.

9. Which of the following are examples of positive feedback, and which are examples of negative feedback?
   a) a warming arctic melts ice, making it darker, absorbing more solar energy
   b) if the earth's temperature rises, its infrared radiation to space increases, providing additional cooling
   c) a car sits in a dip; pushing it forward results in a backward force, while pushing it backward results in a forward force
   d) a car sits on a hill; pushing it either way results in an acceleration (more force, thanks to gravity) in that direction
   e) a child wails loudly and throws a tantrum; to calm the child, parents give it some candy: will this encourage or discourage similar behaviors going forward?

10. Think up an example from daily life (different from examples in the text) for how a delay in negative feedback can produce overshoot, and describe the scenario.

11. Pick five countries of interest to you not represented in any of the tables in this chapter and look up their birth rate and death rate [19, 20], then find the corresponding dot on Figure 3.11, if possible.[40] At the very least, identify the corresponding region on the plot.

[19]: (2016), *List of Sovereign States and Dependent Territories by Birth Rate*
[20]: (2011), *List of Sovereign States and Dependent Territories by Mortality Rate*

40: Numbers may change from when the plot was made; population can help settle based on dot size.

12. A country in the early stages of a demographic transition may have trimmed its death rate to 15 per 1,000 people per year, but still have a birth rate of 35 per 1,000 per year. What does this amount do in terms of net people added to the population each year, per 1,000 people? What rate of growth is this, in percent?

13. If the population of the country in Problem 12 is 20 million this year, how many people would we expect it to have next year? How many were born, and how many died during the year?

14. Figure 3.11 shows Egypt standing well above China in terms of excess birth rate compared to death rate.[41] Yet Table 3.3 indicates that China contributes a much larger annual addition to global population than does Egypt. Explain why. Then, using the first four columns in Table 3.3, replicate the math that produced the final column's entries for these two countries to reinforce your understanding of the interaction between birth and death rates and population in terms of absolute effect.

41: . . . much farther from dashed line

Show work and add one more decimal place to the answer as a way to validate that you did more than copy the table result.

15. In a few clear sentences, explain why the maps in Figure 3.13

and Figure 3.14 look so different, in terms of which countries are shaded most darkly?

16. Table 3.4 indicates which countries place the highest population-driven new demand on global resources using energy as a proxy. Which countries can American citizens regard as contributing more total resource demand? At the individual citizen-contribution level, what other citizens can Americans identify as being responsible for a greater demand on resources via population growth?

The point is that the U.S. is a major contributor to increased resource demand via population growth.

17. The last two columns in Table 3.4 were computed for this book from available information on population, birth and death rates, and annual energy usage for each country (as represented in the first four columns; references in the caption). Use logical reasoning to replicate the calculation that produces the last two columns from the others and report how the computation goes, using an example from the table.

Careful about 10^6 factors and GW = 10^9 W.

18. The bottom row of Table 3.4 is important enough to warrant having students pull out and interpret its content. What is world population, in billions? How many people are added to the world each year? What is the typical power demand for a global citizen (and how does it compare to the U.S.)? If a typical coal or nuclear plant puts out 1 GW of power, how many power-plant-equivalents must we add each year to keep up with population increase? And finally, how much power (in W) is added per global citizen each year due to population growth (and it is worth reflecting on which countries contribute more than this average)?

Some students may see this as free/easy points, but consider the value in internalizing the associated information.

19. If you were part of a global task force given the authority to make binding recommendations to address pressures on resources due to population growth, which three countries stand out as having the largest impact at present? Would the recommendations be the same for all three? If not, how might they differ?

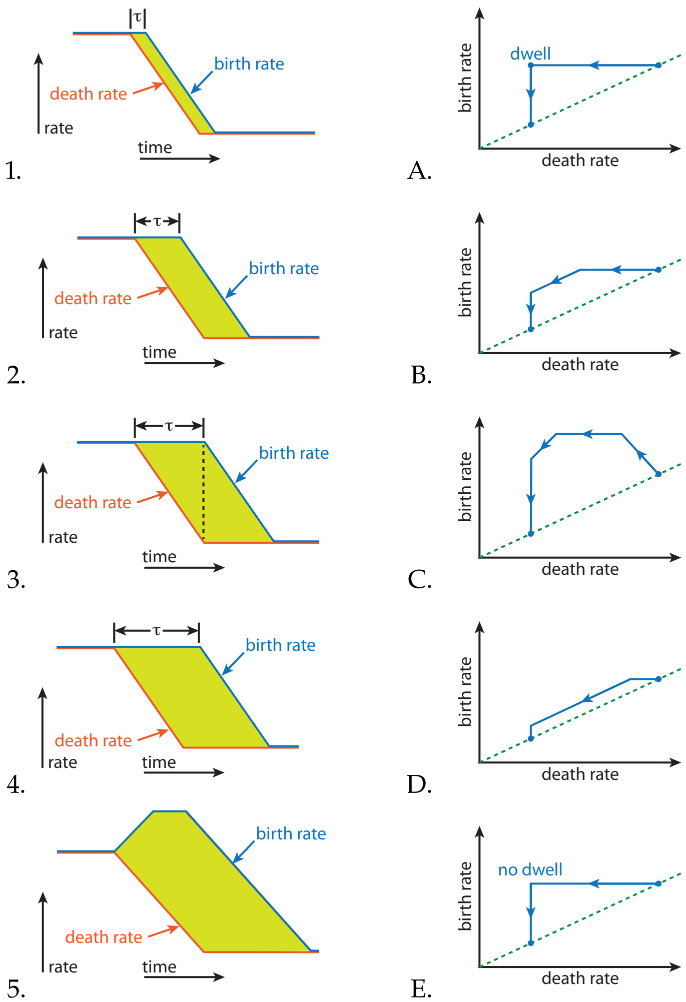
20. Table 3.5 helps differentiate concerns over which region contributes pressures in raw population versus population-driven resource demand. By taking the ratio of population growth (in %) to population (as % of world population), we get a measure of whether a region is “underperforming” or “overperforming” relative to its population. Likewise, by taking the ratio of the added power (in %) to population, we get a similar measure of performance in resource demand. In this context, which region has the highest ratio for population pressure, and which region has the highest ratio for population-induced pressure on energy resources?

For instance, Oceania has a ratio of 1.0 for population growth (0.5% of population growth and 0.5% of global population), meaning it is not over- or under-producing relative to global norms. But in terms of power, it is 3 times the global expectation (1.5 divided by 0.5).

21. If a country starting out at 30 million people undergoes the demographic transition, starting at birth/death rates of 35 per 1,000 per year and ending up at 10 per 1,000 per year, what will the final population be if the delay, τ, is 40 years?

22. The set of diagrams below show five different time sequences on the left akin to Figure 3.16, labeled 1–5. The first four on the left have increasing τ (delay until birth rate begins falling), and the last increases birth rate before falling again. On the right are five trajectories in the birth/death rate space (like Figure 3.12), scrambled into a different order and labeled A–E.[42] Deduce how the corresponding trajectory for each time sequence would appear in the birth/death rate plot on the right, matching letters to numbers for all five.

42: Note that figures A. and E. differ only by whether the transition pauses (dwells) at the corner for some time.



23. Referring to the figures for Problem 22 (and described within the same problem), which pair[43] corresponds to the largest population surge, and which pair produces the smallest? Explain your reasoning, consistent with the presentation in the text.

43: . . . number and associated letter; not necessarily arranged next to each other

24. Referring to the figures for Problem 22 (and described within the same problem), which pair[44] is most similar to the actual trajectory we witness (i.e., Figure 3.11), and what does this say about the

44: . . . number and associated letter; not necessarily arranged next to each other population cost of the demographic transition in the context of Problem 23?

25. Considering Figure 3.11 in the context of a trajectory (as in Figure 3.12), would it appear that most countries in the world have begun the demographic transition? Have very few of them started? Is it about half-and-half? Justify your answer.

Hint: think about what the graph would look like in these scenarios.

26. Express your view about what you learn from Figure 3.17. Do you sense that the prescribed trajectory is realistic? If so, justify. If not, what about it bothers you? What does this mean about the goal of bringing the (growing) world to "advanced" status by the end of this century? Are we likely to see this happen?

27. Make as compelling an argument as you can for why we should promote the demographic transition worldwide for those countries who have not yet "arrived" at the lower-right corner of Figure 3.10. What are the positive rewards?

28. Make as compelling an argument as you can for why pursuit of the demographic transition may be ill-advised and potentially create rather than alleviate hardship. What are the downsides?

29. List the pros and cons a young person without children might face around the decision to have a biological child of their own[45] Consider not only personal contexts, but external, global ones as well, and thoughts about the future as you perceive it. It does not matter which list is longer or more compelling, but it is an exercise many will go through at some point in life—although maybe not explicitly on paper.

45: Assume for the purpose of the question that it is biologically possible.

30. Do you think governments and/or tribal laws have any business setting policy around child birth policies? If so, what would you consider to be an acceptable form of control? If not, what other mechanisms might you propose for limiting population growth (or do you even consider that to be a priority or at all appropriate)?

4 Space Exploration vs. Colonization

4.1 Scale of Space 54
4.2 The Wrong Narrative 58
4.3 A Host of Difficulties 59
4.4 Exploration's Role 61
4.5 Upshot: Earth First 63
4.6 Problems 64

This textbook assesses the challenges and limitations imposed upon us by living on a finite planet having finite resources. If harboring expectations that we will break out into a space-faring existence as a way to mitigate our earthly challenges, then it becomes harder for us to respond earnestly to information about where things are headed on Earth. This chapter is placed where it is to "close the exit" so that the content in the rest of the book might become more relevant and worth the investment to learn. Some of the sections in this chapter offer more of an author's perspective than might be typical for a textbook. Some may disagree with the case that is made, but consider that the burden of proof for a way of life unfathomably beyond our current means should perhaps fall to the enthusiasts.[1]

1: To quote Carl Sagan, extraordinary claims require extraordinary evidence.

4.1 Scale of Space

In the span of two hours, we can sit through a movie and "participate" in interstellar travel without getting tired. Let's step out of the entertainment (fiction) industry and come to terms with the physical scale of the *real* space environment.

Describing an analogous scale model of the solar system, galaxy, and universe—as we will do momentarily—is a fraught exercise, because in order to arrive at physical scales for which we have solid intuition (driving distance in a day?) we end up with inconceivably small (invisible) specks representing familiar objects like the earth. By the time we make Earth the size of something we can hold and admire, the scales become too big for easy comprehension. Figures 4.1 and 4.2 demonstrate how awkward or impossible correctly-scaled graphics are in a textbook.

The convention is to capitalize Earth when it is used as a proper *name*, and refer to the earth when it is an *object*. Similar rules apply to Moon and Sun.

Photo credit: NASA/Bill Anders from Apollo 8 [22].

Figure 4.1: Earth and Moon (far right) to scale. On this scale, the sun would be larger than the page and about 400 pages away. Mars would be 160 to 1,100 pages away. Since 1972, humans have not traveled beyond the black outline of the earth in this figure (600 km).

Figure 4.2: Proving the point that textbooks are not conducive to correctly-scaled graphics of objects in space, by the time the Earth-Sun distance spans the page, Earth (on far right) is too small to be visible in print, at less than 1% the diameter of the orange sun at far left. The Earth–Moon distance is about the width of the arrow shaft pointing to Earth. Humans have *never* traveled more than the arrow shaft's width from Earth, and have not even gone 0.2% *that far* in about 50 years! Mars, on average, is farther from Earth than is the sun.

Let us first lay out some basic ratios that can help build suitable mental models at whatever scale we choose.

Table 4.1: Progression of scale factors.

Step	Factor
Earth diameter	(start)
Moon distance	30×
Sun distance	400×
Neptune distance	30×
Nearest Star	9,000×
Milky Way Center	6,000×
Andromeda Galaxy	100×
Universe Edge	6,000×

Definition 4.1.1 *Scale models of the universe can be built based on these approximate relations, some of which appear in Table 4.1 and Table 4.2:*

1. *The moon's diameter is one-quarter that of Earth, and located 30 earth-diameters (60 Earth-radii) away from Earth, on average (see Figure 4.1).*
2. *The sun's diameter is about 100 times that of Earth, and 400 times as far as the moon from Earth (see Figure 4.2).*
3. *Mars' diameter is about half that of Earth, and the distance from Earth ranges from 0.4 to 2.7 times the Earth–Sun distance.*
4. *Jupiter's diameter is about 10 times larger than Earth's and 10 times smaller than the sun's; it is about 5 times farther from the sun than is the earth.*
5. *Neptune orbits the sun 30 times farther than does Earth.*
6. *The Oort cloud[2] of comets ranges from about 2,000 to 100,000 times the Earth–Sun distance from the sun.*
7. *The nearest star[3] is 4.2 light years from us, compared to 500 light-seconds from Earth to the sun—a ratio of 270,000.*
8. *The Milky Way galaxy has its center about 25,000 light years away away,[4] and is a disk about four times that size in diameter.*
9. *The next large galaxy[5] is 2.5 million light years away, or about 25 Milky Way diameters away.*
10. *The edge of the visible universe[6] is 13.8 billion light years away, or about 6,000 times the distance to the Andromeda galaxy.*

2: The Oort cloud marks the outer influence of the sun, gravitationally.

3: . . . Proxima Centauri

A light year is the *distance* light travels in a year.

4: That's 6,000 times the distance to the closest star.

5: . . . the Andromeda galaxy

6: The "edge" is limited by light travel time since the Big Bang (13.8 billion years ago), and is called our cosmic horizon. See Sec. D.1 (p. 392) for more.

We will construct a model using the set of scale relations in Definition 4.1.1, starting local on a comfortable scale.

As we build up our model, pause on each step to lock in a sense of the model: visualize it or even recreate it using objects around you!

We'll make Earth the size of a grain of sand (about 1 mm diameter). The moon is a smaller speck (dust?) and the diameter of its orbit would span the separation of your eyes. On this scale, the sun is 100 mm in diameter (a grapefruit) and about 12 meters away (40 feet). Mars could be anywhere from 4.5 meters (15 feet) to 30 meters (100 feet) away. Reflect

Body	Symbol	Approx. Radius	Distance (AU)	Alt. Distance
Earth	⊕	$R_\oplus \approx 6,400$ km	—	
Moon	☾	$\frac{1}{4}R_\oplus$	$\frac{1}{400}$	$60R_\oplus \approx 240R_☾$
Sun	⊙	$100R_\oplus$	1	$240R_\odot$
Mars	♂	$\frac{1}{2}R_\oplus$	0.4–2.7	
Jupiter	♃	$10R_\oplus \approx \frac{1}{10}R_\odot$	4–6	
Neptune	♅	$4R_\oplus$	~30	
Proxima Centauri	—	$0.15R_\odot$	270,000	4.2 light years

Table 4.2: Symbols, relative sizes, and distances in the solar system and to the nearest star. An AU is an Astronomical Unit, which is the average Earth–Sun distance of about 150 million kilometers. The fact that both the sun and moon are 240 of their radii away from Earth is why they appear to be a similar size on the sky, leading to "just so" eclipses.

for a second that humans have never ventured farther from Earth than the moon, at 3 cm (just over an inch) in this scale.[7] Mars is outlandishly farther. Neptune is about four-tenths of a kilometer away (on campus at this scale), and the next star is over 3,000 km (roughly San Diego to Atlanta). So we've already busted our easy intuitive reckoning and we haven't even gotten past the first star. Furthermore, this was starting with the earth as a tiny grain of sand. We've only ever traveled two-finger-widths away from Earth on this scale,[8] and the next star is like going on a giant trip across the country. For apples-to-apples, compare how long it takes to walk a distance of two-finger-widths (3 cm) to the time it would take to walk across the U.S. The former feat of traveling to the moon was super-hard; the latter is comparatively impossible.

7: For this, picture a grain of sand sitting on the bridge of your nose representing the earth, and a speck of dust in front of one eye as the moon.

Glance over to where Mars would be if the earth is a grain of sand on your nose.

8: The last time we went this far was 1972.

Box 4.1: When Will We Get There?

It took 12 years for Voyager 2 to get to Neptune, which is "in our back yard." The only spacecraft to date traveling fast enough to leave the solar system are the two Voyagers, the two Pioneers, and the New Horizons probe [23]. The farthest and fastest of this set is Voyager 1 at about 150 times the Earth–Sun distance after 43 years. The closest star is about 2,000 times farther. At its present speed of 17 km/s, it would reach the *distance* to the nearest star[9] in another 75,000 years.

The fastest spacecraft on record as yet is the Parker Solar Probe, which got up to a screaming 68.6 km/s, but only because it was plunging (falling) around the sun. Because it was so close to the sun, even this amount of speed was not enough to allow it climb out of the sun's gravitational grip and escape, as the five aforementioned probes managed to do. Even if Voyager 1 ended up with 70 km/s left over after breaking free of the solar system,[10] it would still take 20,000 years to reach the distance to the nearest star. Note that human lifetimes are about 200 times shorter.

Pushing a human-habitable spacecraft up to high speed is *immensely* harder than accelerating these scrappy little probes, so the challenges are varied and extreme. For reference, the Apollo missions to the very nearby Moon carried almost 3,000 tons of fuel [24], or about 80,000 times the typical car's gasoline tank capacity. It would take

[23]: (2020), *List of artificial objects leaving the Solar System*

9: It does not happen to be aimed toward the nearest star, however.

10: It only had 17 km/s left.

[24]: (2020), *Saturn V*

a typical car 2,000 years to spend this much fuel. Do you think the astronauts argued about who should pay for the gas?

In fairness, fuel requirements don't simply scale with distance for space travel, unlike travel in a car. Still, just getting away from Earth requires a hefty fuel load.

Let's relax the scale slightly, making the sun a chickpea (garbanzo bean). Earth is now the diameter of a human hair (easy to lose), and one meter from the sun. The moon is essentially invisible and a freckle's-width away from the earth. The next star is now 300 km away (a 3-hour drive at freeway speed), while the Milky Way center is 1.5 million kilometers away. Oops. This is more than four times the *actual* Earth–Moon distance. We busted our scale *again* without even getting out of the galaxy.

So we reset and make the sun a grain of sand. Now the earth is 10 cm away and the next star is 30 km.[11] Think about space this way: the swarm of stars within a galaxy are like grains of sand tens of kilometers apart. On this scale, solar systems are bedroom-sized, composed of a brightly growing grain of sand in the middle and a few specks of dust (planets) sprinkled about the room.[12] It gets even emptier in the vast tracts between the stars. The Milky Way extent on this scale is still much larger than the actual Earth, comparable to the size of the lunar orbit.

11: . . . a long day's walk

12: Even a solar system, which is a sort of local oasis within the galaxy, is mostly empty space.

Box 4.2: Cosmic Scales

It is not necessary to harp further on the vastness of space, but having come this far some students may be interested in completing the visualization journey.

As mind-bogglingly large as the solar system is, not to mention that it itself is dwarfed by interstellar distances, which in turn are minuscule compared to the scale of the galaxy, how can we possibly appreciate the largest scales in the universe? Let's start by making galaxies manageable. If galaxies are like coins (say a U.S. dime at approximately 1 cm diameter), they are typically separated by meter-like scales. The edge of the visible universe (see Sec. D.1; p. 392) would be only 1.5 km away. Finally, the picture is easy to visualize: coins as galaxies separated by something like arm's length and extending over an area like the center of a moderately-sized town. We can even imagine the frothy, filamentary arrangement of these galaxies, containing house-sized (5–50 m) voids empty of coins (galaxies). See Figure 4.3 for a visual explanation.

But penetrating the nature of the individual galaxies (coins, in the previous example scale) is extremely daunting: they are mostly empty space, and by the time we reduce the galaxy to a manageable scale (say 10 km, so that we can picture the whole thing as city-sized), individual stars are a few tenths of a meter apart and only about 50 atoms across (roughly 10 nm). Cells and bacteria are about 100–1,000 times larger than this. So it's nearly impossible to conceive of the scale of the galaxy while simultaneously appreciating the sizes of the stars and just how much space lies between.

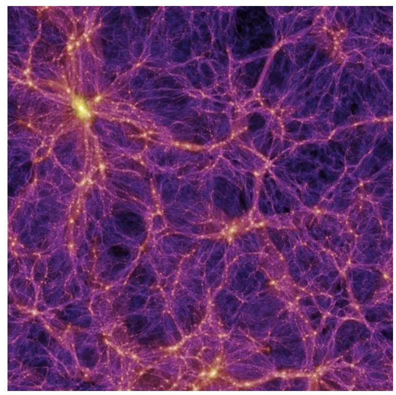


Figure 4.3: Galaxies are actually distributed in a frothy foam-like pattern crudely lining the edges of vast bubbles (voids; appearing as dark regions in the image). This structure forms as a natural consequence of gravity as galaxies pull on each other and coalesce into groups, leaving emptiness between. This graphic shows the bubble edges and filaments where galaxies collect. The larger galaxies are bright dots in this view—almost like cities along a 3-dimensional web of highways through the vast emptiness. From the Millennium Simulation [25].

Given the vastness of space, it is negligent to think of space travel as a "solution" to our present set of challenges on Earth—challenges that operate on a much shorter timescale than it would take to muster any meaningful space presence. Moreover, space travel is enormously expensive energetically and economically (see Table 4.3). As we find ourselves competing for dwindling one-time resources later this century, space travel will have a hard time getting priority, except in the context of escapist entertainment.[13]

Table 4.3: Approximate/estimated costs, adjusted for inflation (M = million; B = billion). [26–29]

Effort	Cost
Apollo Program	$288B
Space Shuttle Launch	$450M
Single Seat to ISS	$90M
Human Mars Mission	$500B

13: . . . which is great stuff as long as it does not dangerously distort our perceptions of reality

4.2 The Wrong Narrative

Humans are not shy about congratulating themselves on accomplishments, and yes, we have done rather remarkable things. An attractive and common sentiment casts our narrative in evolutionary terms: fish crawled out of the ocean, birds took to the air, and humans are making the next logical step to space—continuing the legacy of escaping the bondage of water, land, and finally Earth. It is a compelling tale, and we have indeed learned to escape Earth's gravitational pull and set foot on another body.

But let's not get ahead of ourselves. Just because we can point to a few special *example* accomplishments does not mean that such examples presage a new normal. A person can climb Mt. Everest, but it is not ever likely to become a commonplace activity. We can build a supersonic passenger airplane for trans-atlantic flight, but it does not mean it will be viable to sustain.[14] One can set up a backyard obstacle course for squirrels and generate viral videos, but the amusing demonstration does not signal a "new normal" in backyard design. We need to separate the *possible* from the *practical*. The moon landings might then be viewed as a nifty stunt—a demonstration of capability—rather than a path to our future. We encountered similar arguments in Chapter 2 in relation to decoupling: just because it *can* happen in certain domains of the economy does not mean that the entire economy can decouple and "defy gravity."

14: . . . or even still available today (see the story of the Concorde; Box 2.2; p. 22)

The attractive evolutionary argument misses two critical facets of reality. When fish crawled out of the sea, they escaped predation (as the first animals on land) and found new food sources free of competition. That's a win-win: less dangerous, more sustenance.[15] Likewise, when birds took flight (or we could discuss insects, which beat the birds to it), it was a similar story: evade ground-based predators who could not fly, and access a whole new menu of food—another win-win.

15: Evolution works on exploiting advantages, favoring wins and letting the "lose" situations be out-competed.

Going to space could easily be cast as a lose-lose. It's an extremely hostile environment offering no protection or safe haven,[16] and there's nothing to eat.[17] Think about it: where would you go to grab a bite in our solar system at present, outside of Earth? And a solar system is an absolute *oasis* compared to the vast interstellar void. The two factors that jointly

16: Earth **is** the safe haven.

17: Amusingly, consider that no cheeseburgers have ever smacked into a space capsule.

promoted evolution onto land and into the air *will not operate* to "evolve" us into space. It's a much tougher prospect. Yes, it could be possible to grow food on a spacecraft or in a pressurized habitat, but then we are no longer following the evolutionary meme of stumbling onto a good deal.

One "win" some imagine from space is access to materials. Yet Earth is already well-stocked with elements from the Periodic Table, and the economics of retrieval from space are prohibitive in any case.

Box 4.3: Accomplishments in Space

Before turning attention to what we have not yet done in space, students may appreciate a recap of progress to date. The list is by no means exhaustive, but geared to set straight common misconceptions.

- 1957: Sputnik (Soviet) is the first satellite to orbit Earth.
- 1959: Luna 3 (Soviet; unmanned) reaches the moon in a fly-by.
- 1961: Yuri Gagarin (Soviet), first in space, orbits Earth once.
- 1965: Alexei Leonov (Soviet) performs the first "space walk."
- 1965: Mariner 4 (U.S.; unmanned) reaches Mars.
- 1968: Apollo 8 (U.S.) puts humans in lunar orbit for the first time.
- 1969: Apollo 11 (U.S.) puts the first humans onto the lunar surface.
- Pause here to appreciate how fast all this happened. It is easy to see why people would assume that Mars would be colonized within 50 years. Attractive narratives are hard to retire, even when wrong.
- 1972: Apollo 17 (U.S.) is the last human mission to the moon; only 12 people have walked on another solar system body, the last about 50 years ago.
- 1973–now: as of this writing (2020), humans have not ventured farther than about 600 km from Earth's surface (called low earth orbit, or LEO; see Figure 4.4) since the end of the Apollo missions.
- 1981–2011: U.S. operates the Space Shuttle, envisioned to make space travel routine. After 135 launches (two ending in catastrophe), the shuttle was retired, leaving the U.S. with no human space launch capacity.
- 1998–now: The International Space Station (ISS) [30] provides an experimental platform and maintains a presence in space. It is only 400 km from Earth's surface (4-hour driving distance), and—despite its misleading name—is not used as a space-port hub for space travel. It *is* the destination.

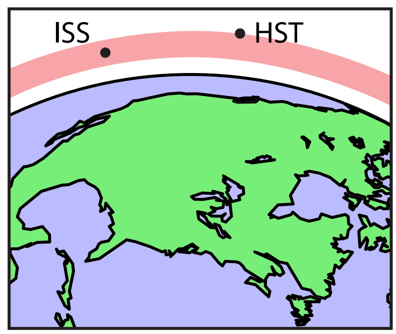


Figure 4.4: The pink band indicates the farthest humans have been from the surface of the earth for the last ∼ 50 years. The Hubble Space Telescope (HST) orbits at the top of this band at 600 km altitude, and the International Space Station in the middle at 400 km. Beyond the thin black line outlining the globe, Earth's atmosphere is too tenuous to support life.

[30]: (2020), *International Space Station*

4.3 A Host of Difficulties

If undeterred by the vast emptiness, hostile conditions, or lack of human-supporting resources in space, then maybe it's because you believe human ingenuity can overcome these challenges. And this is correct to a degree. We *have* walked on one other solar system body.[18] We *have* had individuals spend a year or so in earth orbit. Either these represent first baby steps to a space future, or just rare feats that we can pull off at great effort/expense. How can we tell the difference?

18: The last Apollo landing was in 1972.

Box 4.4: Comparison to Backpacking

The way most people experience backpacking is similar to how we go about space exploration: carry on your back all the food, clothing, shelter, and utility devices that will be needed for a finite trip duration. Only air and water are acquired in the wild. For space travel, even the air and water must be launched from Earth. So space travel is like a glorified and hyper-expensive form of backpacking—albeit offering breathtaking views!.

One way to probe the demonstration vs. way-of-the-future question is to list capabilities we have not yet demonstrated in space that would be important for a space livelihood, including:

1. Growing food used for sustenance;
2. Surviving long periods outside of Earth's magnetic protection from cosmic rays;[19]
3. Generating or collecting propulsive fuel away from Earth's surface;
4. Long-term health of muscles and bones for periods longer than a year in low gravity environments;
5. Resource extraction for in-situ construction materials;
6. Closed-system sustainable ecosystem maintenance;
7. Anything close to terraforming (see below).

19: The ISS (space station) remains within Earth's protection.

It would be easier to believe in the possibility of space colonization if we first saw examples of colonization of the ocean floor.[20] Such an environment carries many similar challenges: native environment unbreathable; large pressure differential; sealed-off self-sustaining environment. But an ocean dwelling has several *major* advantages over space, in that food is scuttling/swimming just outside the habitat; safety/air is a short distance away (meters); ease of access (swim/scuba vs. rocket); and all the resources on Earth to facilitate the construction/operation (e.g., Home Depot not far away).

20: Even just 10 meters under the surface!

⚠ This is not to advocate ocean floor habitation as a *good* idea; it is merely used to illustrate that space habitation is an even *less* practical idea, by far.

Building a habitat on the ocean floor would be vastly easier than trying to do so in space. It would be *even easier* on land, of course. But we have not yet successfully built and operated a closed ecosystem on land! A few artificial "biosphere" efforts have been attempted, but met with failure [31]. If it is not easy to succeed on the surface of the earth, how can we fantasize about getting it right in the remote hostility of space, lacking easy access to manufactured resources?

[31]: (2020), *Biosphere 2*

On the subject of terraforming, consider this perspective. Earth right now has a problem of excess CO_2 as a result of fossil fuel combustion (the subject of Chapter 9). The problem has flummoxed our economic and political systems, so that not only do we seem to be powerless to revert to pre-industrial CO_2 levels, but even arresting the annual increase in emissions appears to be beyond our means. Pre-industrial levels of CO_2 measured 280 parts per million (ppm) of the atmosphere, which we will treat as the *normal* level. Today's levels exceed 400 ppm, so that the modification is a little more than 100 ppm, or 0.01% of our atmosphere.[21] Meanwhile, Mars' atmosphere is 95% CO_2. So we might say that Earth has a 100 ppm problem, but Mars has essentially a *million* part-per-million problem. On Earth, we are completely stymied by a 100 ppm CO_2 increase while enjoying access to all the resources available to us on the planet. Look at all the infrastructure available on this developed world and still we have not been able to reverse or even stop the CO_2 increase. How could we possibly see transformation of Mars' atmosphere into habitable form as realistic, when Mars has zero infrastructure to support such an undertaking? We must be careful about proclaiming notions to be *impossible*, but we can be justified in labeling them as *outrageously impractical*, to the point of becoming a distraction to discuss. Figure 4.5 further illustrates the giant gap between tolerable conditions and actual atmospheres on offer in the solar system.

Terraforming is the speculative idea of transforming a planet so its atmosphere resembles that of Earth (chemical makeup, temperature, pressure) and can support human life.

21: While the increase from 280 to 400 is about 50%, as a fraction of Earth's total atmosphere, the ~ 100 ppm change is 100 divided by one million (from definition of ppm), or 0.01%.

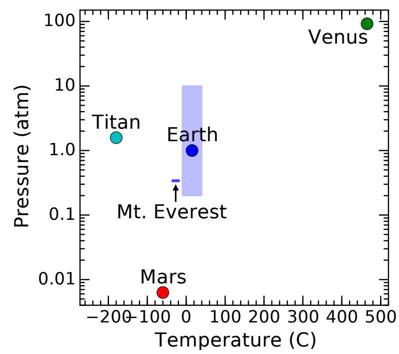


Figure 4.5: Rocky-body atmospheres in the solar system, showing *average* temperature (Celsius) and pressure (atmospheres). The range of "comfort" for Earth is shown as a blue rectangle going from −10°C to 40°C and 0.2 atm (where the atmosphere would need to be 100% oxygen) to (arbitrarily) 10 atm. Not only are the other bodies *far* outside our comfort range, the compositions are noxious, and lack oxygen. Bear in mind that a change of even a few degrees—as in climate change—is a big deal. Even Mt. Everest, where humans can survive for only a few hours with supplemental oxygen is *substantially* more hospitable than Mars. ■

We also should recall the lesson from Chapter 1 about exponential growth, and how the addition of another habitat had essentially no effect on the overall outcome, aside from delaying by one short doubling time. Therefore, even if it is somehow misguided to discount colonization of another solar system body, who cares? We still do not avoid the primary challenge facing humanity as growth slams into limitations in a finite world (or even finite solar system, if it comes to that).

4.4 Exploration's Role

It is easy to understand why people might latch onto the idea that we will likely leverage our exploration of space into ultimate colonization. Much as early explorers of our planet opened pathways for colonization of "new worlds," the parallels in exploring literal new worlds like planets are obvious.[22] In short, it is a familiar story, and therefore an easy "sell" to primed, undoubting minds. Plus, we're captivated by the novelty and challenge space colonization represents—as attested by a vibrant entertainment industry devoted to stories of eventual life in space. But not all exploration leads to settlement, and entertainment is not truth.

22: Reaching the Americas involved a leap across a span of (life-supporting) ocean about twice the size of Europe. Reaching Mars involves a leap across inhospitable space 5,000 times the diameter of Earth—not very similar at all.

Humans have explored (a small portion of) the crushing deep ocean, scaled Earth's highest and wholly inhospitable peaks, and visited the harsh ice cap at the north pole. In such instances, we had zero intention of establishing permanent residence in those locations. They represented places to test our toughness and also learn about new environments. We do not view these sorts of explorations as *mistakes* just because they did not pave the way for inhabitation. Rather, we speak fondly of such excursions as feathers in our collective cap: feats that make us proud as a species. Space might be viewed in a similar way: superlative in terms of challenge and wonderment, reflecting positively on our curiosity, drive, ingenuity, and teamwork. We also derive benefits[23] in the way of technological advancement propelled by our quest to explore, and in furthering our scientific understanding of nature.

23: . . . among them a deeper appreciation for the rare and precious Earth

So even if space does not fulfill the fantasy of continued human expansion across the cosmos, it is in our nature to at least *explore* it. We would do well to put space exploration in the category of conquering Mt. Everest rather than that of Europeans stumbling upon the West Indies (one is as imminently uninhabitable as the other is inhabitable). Let us not make the mistake of applying the wrong narrative to space.

Many positive things might be said about space exploration, and hopefully we continue poking into our outer environment indefinitely. Yet hoping that such exploration is a pathway to human colonization of space is probably wrong and almost certainly counterproductive at present, given the short timescale on which human expansion is likely to collide with Earth's limits.

Despite the pessimistic tone of this chapter, the author is himself captivated by space, and has built a life around it: *Star Wars* was a transformative influence as a kid, and later *Star Trek*. The movie *The Right Stuff* is still a favorite. He has peered to the edge of the universe—first through a 10-inch telescope he built in high school, and later using the largest telescopes in the world. He has worked on a Space Shuttle experiment, met astronauts, knew Sally Ride, and spent much of his career building and operating a laser system to bounce and detect individual photons off the reflectors placed on the lunar surface by the Apollo astronauts (as a test of the fundamental nature of gravity), which directly inspired part of a *Big Bang Theory* episode via personal interactions with the show's writers. So a deep fondness for space? Yes. Would volunteer to go to the moon or Mars? Yes. Believes it holds the key to humanity's future? No.

If, in the fullness of time, we do see a path toward practical space colonization, then fine. But given the extreme challenge and cost—both energetically and economically, and for what could only be a tiny footprint in the near term—it seems vastly more prudent to take care of our relationship with Planet Earth first, and *then* think about space colonization in due time, if it ever makes sense. Otherwise, not only do we spend precious resources unwisely, but (even worse) our mindset is tainted by unrealistic dreams that diminish the importance of confronting the real challenge right here on the ground. We need to have our heads in the real game. Perhaps twenty øne piløts said it best in the song *Stressed Out*:

> We used to play pretend, give each other different names
> We would build a rocket ship and then we'd fly it far away
> Used to dream of outer space but now they're laughing at our face
> Saying, "Wake up, you need to make money."
> Yeah.

Space colonization might be treated as a pretend fantasy for the moment. We would be better off waking up to face real here-and-now challenges. In some sense, perhaps the only way to achieve the dream of migration to space—should that be in the cards at all—is to first pretend that it is impossible and turn attention to the pressing matters on Earth. Otherwise we risk failing at *both* efforts.

Box 4.5: Q&A on State of Exploration

After reading the first draft of this chapter, students had a number of remaining questions. Here are some of them, along with the author's responses.

1. How long before we live on other planets?

   Maybe never.[24] The staggering distances involved mean that our own solar system is effectively the only option. Within the solar system, Mars is the most hospitable body—meaning we might live as long as two minutes without life support. By comparison, Antarctica and the ocean floor are millions of times more practical, yet we do not see permanent settlements there.[25]

2. What is the status of searches for other planets to colonize?

   We understand our solar system pretty well. No second homes stand out. We have detected evidence for thousands of planets around other stars [32], but do not yet have the sensitivity to detect the presence of earth-like rocks around most stars. It is conceivable that we will have identified Earth analogs in the coming decades, but they will give new meaning to the words "utterly" and "inaccessible."

3. Haven't we benefited from space exploration in the technology spin-offs, like wireless headsets and artificial limbs?

   No doubt! The benefits have been numerous, and I would never characterize our space efforts to date as wasted effort. It's just that what we have done so far in space does not mean that colonization is in any way an obvious or practical next step. Actually, the banner image for this chapter from the Apollo 8 mission captivated the world and made our fragile shell of life seem all the more precious. So perhaps the biggest benefit to our space exploration will turn out to be a profound appreciation for and attachment to Earth!

24: True, never is a long time. The notion that we may *never* colonize space may seem preposterous to you now. Check back at the end of the book. The odds favor a more boring slog, grappling with our place in nature.

25: Staffed research stations are not the same as human settlements, in the case of Antarctica.

[32]: IPAC/NASA (2020), *NASA Exoplanet Archive*

4.5 Upshot: Putting Earth First

The author might even go so far as to label a focus on space colonization in the face of more pressing challenges as disgracefully irresponsible. Diverting attention in this probably-futile[26] effort could lead to *greater total suffering* if it means not only mis-allocation of resources but perhaps

26: . . . at least on relevant time scales more importantly lulling people into a sense that space represents a viable escape hatch. Let's not get distracted!

The fact that we do not have a collective global agreement on priorities or the role that space will (or will not) play in our future only highlights the fact that humanity is not operating from a master plan[27] that has been well thought out. We're simply "winging it," and as a result potentially wasting our efforts on dead-end ambitions. Just because some people are enthusiastic about a space future does not mean that it can or will happen.

27: Prospects for a plan are discussed in Chapter 19.

It is true that we cannot know for sure what the future holds, but perhaps that is *all the more reason* to play it safe and not foolishly pursue a high-risk fantasy.[28] From this point on, the book will turn to issues more tangibly relevant to life and success on Planet Earth.

28: Tempted as we may be by the e-mail offer from the displaced Nigerian prince to help move his millions of dollars to a safe account, most of us know better than to bite. The promise of wealth can lead the gullible to ruin.

Box 4.6: Survey Says?

It would be fascinating to do a survey to find out how many people think that we will have substantial populations living off Earth 500 years from now. It is the author's sense that a majority of Americans believe this to be likely. Yet, if such a future is not to be—for a host of practical reasons, including the possibility that we falter badly and are no longer in a position to pursue space flight—we would find ourselves in a situation where most people may be completely wrong about the imagined future. That would be a remarkable state of affairs in which to find ourselves—though not entirely surprising.

4.6 Problems

1. If the sun were the size of a basketball, how large would Earth be, and how far away? How large would the moon be, and how far from Earth, at this scale?
2. Find objects whose sizes *approximately* match the scales found in Problem 1 and place them in your environment at the scaled/appropriate distances. Submit a personalized/unique picture of your arrangement.

   ⓘ This kind of exercise might seem like a hassle, but it can really help internalize the scales in a way words never will.

3. How far would the nearest star be at the scale from Problem 1, and how big is this in relation to familiar objects?
4. Find an Earth globe and an object about one-fourth its size to represent the moon, then place at the appropriate distance apart. Report on how far this is. Take a personalized/unique picture to document, and take some time appreciating how big Earth would look from the moon.

   ⓘ By doing this, you can get maybe 2% of the enjoyment of a trip to the moon for less than one-billionth the cost: a real bargain! Can you make out Florida? Japan?

5. Highway 6563 in New Mexico has signs along a roughly 30 km stretch of road corresponding to the solar system scale from the Sun to Neptune. On this scale, how large would Earth, Sun, and Jupiter be, in diameter? Express in convenient units appropriate to the scale.

6. Using the setup in Problem 5, how fast would you have to travel on the road to match the speed of light, for which it takes 500 seconds to go from Earth to the sun? Express in familiar/convenient units.

7. Note that the size of the moon in Figure 4.1 is about the same size as the sun in Figure 4.2. Explain how this is related to the fact that they appear to be about the same size in our sky. Hint: imagine putting your eye at the earth location in each figure and looking at the other body.

ⓘ This is why eclipses are special on Earth.

8. Use Table 4.1 to accumulate (multiplicatively combine) scale factors and ask: which is a bigger ratio: the distance to the nearest star compared to the diameter of Earth, or the distance to the edge of the universe compared to that to the nearest star? Compared to the large numbers we are dealing with, is one much smaller or much bigger than the other, or are they roughly the same?

9. It may be tempting to compare Earth to a life-sustaining oasis in the desert—maybe spanning 100 m. But this is a pretty misleading view. One way to demonstrate this is to consider that in a real desert, the next oasis might be a perilous 100 km journey away. Using the ratio of distance to size (diameter of planet or oasis), how close would another Earth have to be (and compare your answer to other solar system scales) to hold the analogy?[29] How long would it take to drive this distance? Do we have another oasis or potential oasis within this distance?

29: In other words, oasis size is to distance between oases as Earth's diameter is to *how far*?

10. Another way to cast Problem 9 is to imagine that the actual distance between Earth and a comparable oasis is more like the distance between stars.[30] In this case, how far would the next oasis be in the desert if we again compare the 100 m scale of the oasis to the diameter of the Earth?[31] How long would it take to drive between oases at freeway speeds (cast in the most informative/intuitive units)?

30: . . . since Earth is the only livable "oasis" in our own solar system

31: In other words, Earth diameter is to interstellar distances as a 100 m oasis is to *how far*?

ⓘ The insight you develop will not depend on exact choices for distance and speed, as long as they are reasonable.

11. On the eighth bullet of Box 4.3 (the one that asks you to pause), imagine someone from the year 2020 traveling back 50 years and explaining that we have not been to the moon since 1972, and that Americans get to space on Russian rockets. How believable do you think they would be, and what assumptions might be made to reconcile the shock?

12. Prior to exposure to this material, what would you honestly have said in response to: How far have humans been from the planet in

the last 45 years;

a) 600 km (about $\frac{1}{10}$ Earth radius; low earth orbit)
b) 6,000 km (roughly Earth radius)
c) 36,000 km (about 6 $R_\oplus$; around geosynchronous orbit)
d) 385,000 km (approximate distance to the moon)
e) beyond the distance to the moon

Explain what led you to think so (whether correct or not).

13. List at least three space achievements that impress you personally, even if they do not bear directly on colonization aims.

14. In the enumerated list beginning on page 60, which item is most surprising to you as not-yet accomplished, and why?

15. List three substantive challenges that prevented successful long-term operation of the artificial biosphere project [31].

[31]: (2020), *Biosphere 2*

16. Since Figure 4.5 spans the range of atmospheres found in our solar system, we can imagine how likely it would be that a random planet might happen to be livable for humans.[32] Imagine throwing a dart at the diagram to get a random instance. How likely are we to hit the comfort zone of Earth, by your estimation?

32: . . . just in terms of temperature and pressure; ignoring composition and a host of other considerations!

17. Come up with three examples (not repeating items in the text) of feats that are technically possible, but not common or practical.

18. For some perspective, imagine you were able to drive your car up a ramp to an altitude characteristic of low-earth orbit (about 320 km, or 200 miles). It takes about 5×10^{10} J of energy[33] to win the fight against gravity. Meanwhile, each gallon of gasoline can do about 25×10^6 J of useful work. How many gallons would it take to climb to orbital height in a car? Roughly how many miles per gallon is this (just counting vertical miles)?

33: We'll encounter gravitational potential energy later, but this quantity is computed as mgh with $m \approx$1,500 kg.

19. In Problem 18, we ignored the energy required to provide the substantial orbital speed (~ 8 km/s, but will not need), which essentially *doubles* the total energy.[34] How much gasoline will it now take, and how massive is the fuel if gasoline is 3 kg per gallon, compared to the 1,500 kg mass of the car?

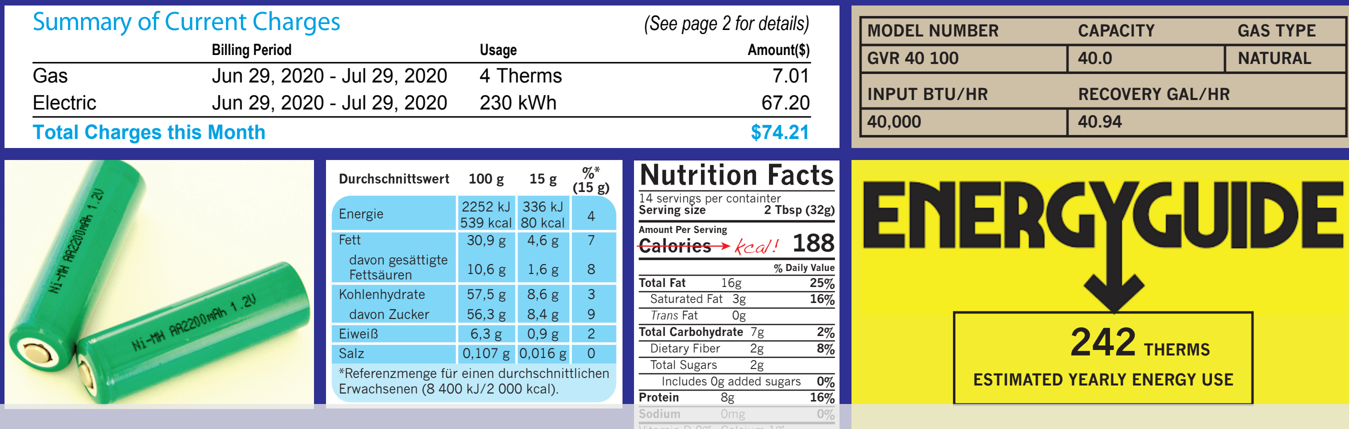
34: The same amount of energy to climb against gravity must go into accelerating to speed (kinetic energy).

ⓘ Given this mass of fuel, and its significance compared to the car's mass, we can see that we've underestimated the fuel required since we'll have to put substantial additional energy into lifting the fuel away from the earth.

Part II

Energy and Fossil Fuels

What worked for us in the past
cannot work for us in the future.
We must learn the language of our old friends
in order to say a proper goodbye.



5 Energy and Power Units

5.1 Energy (J) 68
5.2 Energy Forms & Conservation 69
5.3 Power (W) 71
5.4 Kilowatt-hour (kWh) 72
5.5 Calories (kcal) 73
5.6 British Thermal Unit (Btu) . 75
5.7 Therms 76
5.8 Electrical Power 76
5.9 Electron Volt (eV) 78
5.10 Light Energy 79
5.11 Upshot on Units 79
5.12 Problems 80

This chapter provides a baseline for understanding the rest of the content in this book, so that students may learn to interpret and convert units, while building a useful intuition in the process. Sec. A.10 (p. 370) in the Appendices offers some tips on manipulating units and performing unit conversions.

Unlike most chapters, this one does not tell a single story or advance our perspective on the world. But it builds a foundation, putting us in a position to start looking at consequential matters of energy use in our society in chapters to come. Hopefully, patience will be rewarded.

5.1 Energy (J)

First, what is energy?

Definition 5.1.1 ***Energy*** *is defined as the* capacity to do work. *Work is well-defined in physics as the* ***application of force through a distance***.[1] *The colloquial use of the word "work" matches relatively well, in that pushing a large couch across the floor (applying force through a distance) or lifting a heavy box up to a shelf feels like work and can tire you out.*

1: This definition applies to the common circumstance when the motion is aligned with the direction of force, like pushing a box across a level floor, propelling a car along the road, or lifting a weight.

The SI unit of force is the Newton (N), breaking down more fundamentally to $\text{kg} \cdot \text{m/s}^2$. The best way to remember this is via Newton's Second Law: $F = ma$ (force equals mass times acceleration). Mass has units of kg, and acceleration[2] is measured in meters per second squared.

2: Acceleration is the rate of change of velocity. Since velocity is measured in meters per second, the rate at which it changes will be meters per second *per second*, or m/s/s, or m/s^2. Some students may know that gravitational acceleration on Earth's surface is $9.8\ \text{m/s}^2$, which is another way to remember.

Since work is force times distance, the unit for work (and thus energy) is Newtons times meters, or $\text{N} \cdot \text{m}$. We give this unit its own name: the

Energy units from everyday life. Clockwise from upper left: a utility bill (kWh and Therms); a hot water heater label (Btu/hr); EnergyGuide for same hot water heater (Therms); U.S. nutrition label for peanut butter (Calories; should be kcal); a German nutrition label for Nutella (kJ, kcal); and rechargeable AA batteries (2200 mAh, 1.2 V).

Joule (J). Thus, the application of 1 N of force across a distance of 1 m constitutes 1 J of work, requiring 1 J of energy to perform. Table 5.1 offers contextual examples (unit prefixes are on page 420).

Table 5.1: Approximate energy for familiar activities. The first freeway example is just kinetic energy; the second is the energy cost of a whole trip.

Action	Energy
Nerf football toss	15 J
Lift loaded bookbag	100 J
Fast-pitch baseball	120 J
Speeding bullet	5 kJ
Charge cell phone	30 kJ
Car on freeway (K.E.)	675 kJ
human daily diet	8 MJ
1 hour freeway drive	250 MJ

Example 5.1.1 Several examples[3] illustrate force times distance, the first two amounting to one Joule of energy:

- Pushing a book across a table, applying 2 N of force and sliding it 0.5 m amounts to 1 J of work.
- Pushing a matchbox toy car across the floor might require only 0.1 N of force. One would have to push it through a distance of 10 m to make up one Joule of energy.
- A car on level ground may require 150 N of force to roll against friction. Pushing a car 5 m would then require 750 J of work.

3: For examples like these, framed as statements and not questions, you can practice solving several types of problems by covering up one number and then solving for it using still-available information. So each statement can be seen as several examples in one!

Writing out Newtons as $\text{kg}\cdot\text{m/s}^2$, we find that the unit of energy amounts to $\text{J} = \text{N}\cdot\text{m} = \text{kg}\cdot\text{m}^2/\text{s}^2$. Notice that this looks like mass times velocity-squared. Box 5.1 explores how this makes a lot of sense.

Box 5.1: The Units Make Sense!

Think about the famous equation $E = mc^2$. Energy is mass times the speed of light squared. The units work!

More on mass-energy in Chapter 15.

Also, kinetic energy is K.E. $= \frac{1}{2}mv^2$, telling a similar story in terms of units: mass times velocity-squared.

More on kinetic energy in Chapter 12.

Gravitational potential energy is just the weight of an object times the height it is lifted through.[4] The weight (force) is mass (m) times the acceleration due to gravity[5] (g), so that lifting (applying a force equal to the weight) through a height (h) results in a potential energy gain of P.E. $= mgh$. The units again check out as

$$mgh \rightarrow \text{kg}\cdot\frac{\text{m}}{\text{s}^2}\cdot\text{m} = \frac{\text{kg}\cdot\text{m}\cdot\text{m}}{\text{s}^2} = \frac{\text{kg}\cdot\text{m}^2}{\text{s}^2} = \text{J}.$$

More on gravitational potential energy in Chapter 11.

4: Another example of work (energy) being force times distance.

5: The force needed to hold against gravity is just $F = ma = mg$

Sec. A.10 (p. 370) in the Appendices provides additional guidance on manipulating units.

We'll encounter other ways to describe energy in this book, but *any energy unit can always be cast into units of Joules*, if desired. Later sections in this chapter detail alternative units whose acquaintance we must make in order to interpret energy information in our lives.

5.2 Energy Forms and Conservation

Energy manifests in a variety of forms, which we will treat in greater detail in application-specific chapters in Part III of this text. For now we just want to name them and point to related chapters and applications, as is done in Table 5.2.

Energy Form	Formula	Chapter(s)	Applications
gravitational potential	mgh	11, 16	hydroelectric, tidal
kinetic	$\frac{1}{2}mv^2$	12, 16	wind, ocean current
photon/light	$h\nu$	13	solar
chemical	$H - TS$	8, 14	fossil fuels, biomass
thermal	$c_p m\Delta T$	6, 16	geothermal, heat engines
electric potential	qV	15	batteries, nuclear role
mass (nuclear)	mc^2	15	fission and fusion

Table 5.2: Energy forms. Exchange is possible between all forms. Chemical energy is represented here by Gibbs free energy.

A bedrock principle of physics is conservation of energy, which we take to *never* be violated in any system, ever.[6] What this means is that energy can *flow* from one form to another, but it is *never created or destroyed*.

6: The only exception is on cosmological scales and times. But across scales even as large as the Milky Way galaxy and over millions of years, we are on solid footing to consider conservation of energy to be inviolate. It is fascinating to note that conservation of energy stems from a symmetry in time itself: if the laws and constants of the Universe are the same across some span of time, then energy is conserved during such time—a concept we trace to Emmy Noether. See Sec. D.2 (p. 393) for more.

Box 5.2: Energy: The Money of Physics

A decent way to conceptualize energy conservation is to think of it as the *money* of physics. It may change hands, but is not created or destroyed in the exchange. A large balance in a bank account is like a potential energy: available to spend. Converting to another form—like heat or kinetic energy—is like the act of spending money. The *rate* of spending energy is called *power*.

Example 5.2.1 traces a few familiar energy conversions, and Figure 5.1 provides an example illustration. A more encompassing narrative connecting cosmic sources to daily use is provided in Sec. D.2.2 (p. 395).

Example 5.2.1 Various illustrative examples:

- A rock perched on the edge of a cliff has gravitational potential energy. When it is pushed off, it trades its potential energy for kinetic energy (speed) as it races toward the ground.
- A pendulum continually exchanges kinetic and potential energy, which can last some time in the absence of frictional influences.
- A stick of dynamite has energy stored in chemical bonds (a form of potential energy). When ignited, the explosive material becomes very hot in a small fraction of a second, converting chemical energy into thermal energy.
- The fireball of hot material from the exploding dynamite expands rapidly, pushing air and nearby objects out of the way at high speed, thus converting thermal energy into kinetic energy.
- Light from the sun (photons) hits a black parking lot surface, heating it up as light energy is converted to thermal energy.
- A uranium nucleus splits apart, releasing nuclear (potential) energy, sending the particles flying off at high speed (kinetic energy). These particles bump into surrounding particles transferring kinetic energy into thermal energy.
- Thermal energy from burning a fossil fuel or from nuclear fission

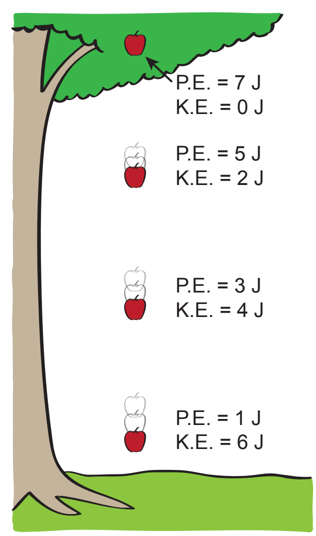


Figure 5.1: Example exchange of potential energy (P.E.) into kinetic energy (K.E.) as an apple drops from a tree. The total energy always adds to the same amount (here 7 J). The apple speeds up as it gains kinetic energy (losing potential energy). When it comes to rest on the ground, the energy will have gone into 7 J of heat (the associated temperature rise is too small to notice).

can be used to make steam that drives a turbine (kinetic energy) that in turn generates electrical energy (voltage, current).

Any of the forms of energy (e.g., in Table 5.2) can convert into the other, directly or indirectly. In each conversion, 100% of the energy is accounted for. In the general case, the energy branches into multiple paths, so we do not get 100% efficiency into the channel we want. For instance, the pendulum example above will eventually bleed its energy into stirring the air (kinetic energy) and friction (heat) at the pivot point. The stirring air eventually turns to heat via internal (viscous) friction of the air.

One useful clarification is that thermal energy is really just random motions—kinetic energy—of individual atoms and molecules. So in the case of nuclear fission in Example 5.2.1, the initial kinetic energy of the nuclear fragments is already thermal in nature, but at a higher temperature (faster speeds) than the surrounding material. By bumping into surrounding atoms, the excess speed is diffused into the medium, raising its temperature while "cooling" the fragments themselves as they are slowed down.

The differences between kinetic and thermal energy is about *coherence*, in that we characterize the kinetic energy of a raindrop by its bulk motion or velocity. Meanwhile, water molecules *within* the drop are zipping about in *random* directions and at very high speeds exceeding 1,000 meters per second.

If accounting for all the possible paths[7] of energy, we are confident that they always add up. Nothing is lost.[8] Energy is never created or destroyed in any process we study. It just sloshes from one form to another, often branching into multiple parallel avenues. The sum total will always add up to the starting amount. Sec. D.2.3 (p. 396) provides a supplement for those interested in better understanding where energy ultimately goes, and why "losing energy to heat" is not actually a loss but just another reservoir for energy.

7: . . . sometimes called channels

8: Actually, the principle is so well established that new particles (like the neutrino) have been discovered by otherwise unaccounted energy in nuclear processes.

5.3 Power (W)

Before getting to the various common units for energy, we should absorb the very important concept and units of power.

Definition 5.3.1 ***Power*** *is simply defined as energy per time: how much energy is expended in how much time. The SI unit is therefore J/s, which we rename* ***Watts*** *(W).*

One Watt is simply one Joule per second.

While energy is the capacity to do work, it says nothing about how quickly that work might be accomplished. Power addresses the *rate* at which energy is expended. Figure 5.2 provides a sense of typical power levels of familiar animals and appliances.

Example 5.3.1 Lifting a 10 kg box, whose weight is therefore about 100 N, through a vertical distance of 2 m requires about 200 J of energy. If performed in one second, the task requires 200 W (200 Joules in one second). Stretching the same task out over four seconds requires only 50 W.

Weight is mg. In this case, m is 10 kg. If we're being sticklers, $g = 9.8\ \mathrm{m/s^2}$, but for convenience we can typically use $g \approx 10\ \mathrm{m/s^2}$ without significant loss of precision.

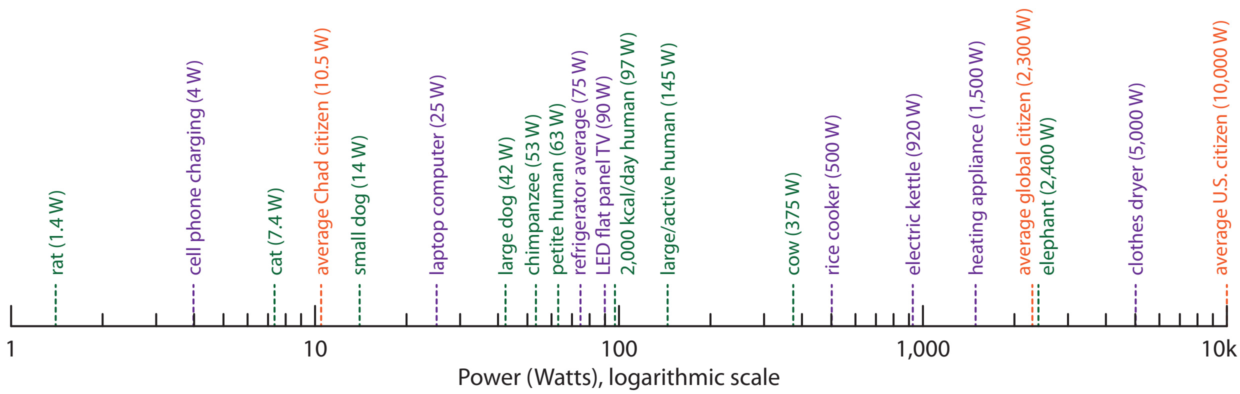


Figure 5.2: Various power levels for comparison and intuition-building. Green entries correspond to metabolic power [33]. Purple entries are devices and appliances. Orange entries are per-capita totals for societal (non-metabolic) energy use. Note that appliances whose job it is to create heat demand the greatest power. The "heating appliance" entry stands for things like microwave ovens, toaster ovens, space heaters, or hair dryers plugged into electrical outlets. Do not take the numbers provided as definitive or exact, as almost everything in the figure will vary somewhat from one instance to another.

Of course, we commonly apply the usual multipliers of factors of 10^3 to the unit to make it more useful. Thus we have the progression W, kW, MW, GW, TW, etc. For reference, a large college campus will require several tens of MW (megawatts) for electricity. A large power plant is typically in the 1–4 GW range. See Table 5.3 for scales at which we are likely to use the various multiplying factors, and a more complete set of multipliers on page 420.

Table 5.3: Power multipliers and contexts

Factor	Unit	Context
1	W	phones; computers
10^3	kW	microwave oven
10^6	MW	campus; community
10^9	GW	power plant; city
10^{12}	TW	societal scale

Although it won't come up too often in this course, it is worth mentioning that the common unit of horsepower equates to 745.7 W. Thus a 100 hp car is capable of delivering about 75 kW of power.

It is usually sufficient to remember that 1 hp is about 750 W.

5.4 Kilowatt-hour (kWh)

Definition 5.4.1 *The* ***kilowatt-hour*** *is an amount of energy (not a power) resulting from an expenditure of energy at a* rate *of 1 kW for a* duration *of one hour, and is the unit of choice for residential electricity usage.*

This unit causes no end of confusion, but it's really pretty straightforward. The kilowatt-hour is a kilowatt *times* an hour. Thus it is power multiplied by time, which is energy (since power is energy over time).

It may help to think of the sequence: kilowatt × hour; kW×h; kW-h; kWh.

Example 5.4.1 Let's say you plug in a space heater rated at 1,000 W (1 kW) and run it for one hour. Congratulations—you've just spent 1 kWh.

Or maybe you turn on a 100 W incandescent light bulb (0.1 kW) and leave it on for 10 hours: also 1 kWh!

What if you run a 500 W rice cooker (0.5 kW) for half an hour? That's 0.25 kWh.

It is straightforward to convert back to Joules, because 1 kW is 1,000 J/s and one hour is 3,600 s. So 1 kWh is 1 kW times 1 hr, which is 1,000 J/s times 3,600 s, and is therefore equal to 3,600,000 J, or 3.6 MJ. A related measure sometimes comes up: the watt-hour (Wh). In much the same vein, this is equivalent to 1 J/s for 3,600 seconds, or 3,600 J.[9]

9: A Wh is one-thousandth of a kWh, not surprisingly.

Box 5.3: Don't be one of those people. . .

If you ever hear someone say "kilowatts per hour," it's likely a mistake,[10] and has the side effect of leading people to erroneously think that kilowatts is a unit of energy, not a power. Kilowatts is already a *rate* (speed) of energy use: 1,000 Joules per second.

10: Literally, kW/hr would be a sort of *acceleration* through energy. It's a real thing that can happen, but it's usually not what people mean.

One tendency some people have is to mix up kW and kWh.[11] Kilowatts is a unit of *power*, or how fast energy is being used. Think of it like a speedometer: how fast are you moving (through space or energy)? Kilowatt-hours is a multiplication of power times time, becoming an *energy*. It's more like the odometer: how much have you accumulated (distance or energy)? Just like distance is rate (speed) times time, energy is rate (power) times time.

11: ⚠ Perhaps related to Box 5.3.

Example 5.4.2 We will explore kWh using a light bulb for an example. Let's say the light bulb is labeled as 100 W.[12] How much energy does it use?

Well, it depends on how long it's on. If it is never turned on, it uses *no* energy. If it is on for 10 seconds, it uses far less than if it's on for a day.

The characteristic quality of the light bulb is the power it expends when it's on—in this case 100 W. It only has one speed. In analogy to a car and speedometer, it's similar to saying that a car travels at a constant speed,[13] and asking how far it travels. Well, it depends on how much *time* it spends traveling at speed.

So view kWh (energy) as an *accumulated* amount that increases with time. On the other hand, kW is a *rate* of energy expenditure.

12: . . . an incandescent, for instance

13: . . . maybe 30 m/s; 67 m.p.h.; 108 k.p.h.

5.5 Calories (kcal)

A common unit for describing chemical and thermal processes is the calorie and its siblings.

Definition 5.5.1 *A* ***calorie*** *is defined as the amount of energy it takes to heat one gram of water (thus also 1 mL, or 1 cm^3, or 1 cc) by one degree Celsius (Figure 5.3). One calorie (note the small "c") is 4.184 J of energy.*

One ***Calorie*** *(note the capital "C")*[14] *is 1,000 calories, or 1* ***kilocalorie***

14: This might win the prize for the dumbest convention in science: never define a unit as case-sensitive, as it cannot be differentiated in spoken language!

(1 kcal), equating to 4,184 J. Most memorably, it is the amount of energy it takes to heat one kilogram *(or one liter; 1 L) of water by 1°C. Due to the tragic convention of Calorie, we will opt for* ***kcal*** *whenever possible.*

Food labels in the U.S. are in Calories, describing the energy *content of the food we eat.*[15] *We would all do ourselves a favor by calling these kcal instead of Calories (same thing). Many other countries sensibly use either kJ or kcal for quantifying food energy.*

15: Human metabolism is not the same as heating water, but the energy involved can still be counted in an energy unit that is defined in terms of heating water. It's still just energy.

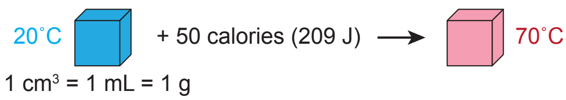


Figure 5.3: Following the definition of a calorie, adding 50 cal to one gram of water raises its temperature by 50°C.

Example 5.5.1 To change 30 mL (30 g) of water by 5°C requires 150 cal, or a little over 600 J.

Injecting 40 kcal of energy into a 2 L (2 kg) bottle of water will heat it by 20 degrees.

Drinking 250 mL of ice-cold water and heating it up to body temperature (thus raising its temperature by approximately 35 degrees) will take about 8,750 cal, or 8.75 kcal, or a bit over 36 kJ of energy.

It is usually sufficient to remember that the conversion factor between calories and Joules is about 4.2—or just 4 if performing a crude calculation.

No deep significance attaches to the fact that 1 cal happens to equate to 4.184 J, other than to say this describes a property of water (called specific heat capacity).

$$\begin{aligned} 1\text{ cal} &= 4.184\text{ J} \approx 4.2\text{ J} \sim 4\text{ J} \\ 1\text{ kcal} &= 4{,}184\text{ J} \approx 4.2\text{ kJ} \sim 4\text{ kJ} \end{aligned}$$

Two examples will help cement use of the kcal (a more useful scale in this class than the much smaller calorie).

Example 5.5.2 A typical diet amounts to a daily intake of about 2,000 kcal of food energy. If you think about it, 2,000 kcal/day is a *power* (energy per time). We can convert to Watts by changing kcal to J and one day to seconds. 2,000 kcal is 8.368 MJ. One day has 86,400 seconds. The division of the two is very close to 100 W.[16]

16: It would serve little purpose to perform exact math here—producing 96.85 W in this case—since the idea that someone's daily diet is exactly 2,000.00 kcal is pretty preposterous. It will likely vary by at least 10% from day to day, and by even larger amounts from individual to individual, so that 100 W is a convenient and approximate representation.

A second example hews closely to the definition of the kcal: heating water.

Example 5.5.3 Let's say you want to heat a half-liter (0.5 kg) of water from room temperature (20°C) to boiling (100°C). Since each kcal can heat 1 kg by 1°C, that same energy will raise our half-kg by 2°C.[17] So raising the temperature by 80°C will require 40 kcal, or 167 kJ.

If the water is heated at a rate of 1,000 W (1,000 J/s), it would take 167 seconds for the water to reach boiling temperature.

17: Make sure this is clear to you; by understanding, we are installing concepts instead of formulas, which are more powerful and lasting.

Notice that we did not apply an explicit formula in Example 5.5.3. By proceeding stepwise, we attempt to keep it intuitive. We *could* write a

Appendix Sec. A.8 (p. 368) addresses this philosophy in a bit more detail.

formula, but we implicitly create the formula on the fly by recognizing that the amount of energy required should scale with the mass of water and with the amount of temperature increase. Hopefully, this approach leads to a deeper understanding of the concept, while printing a formula on the page might short-circuit comprehension.

This would be an excellent opportunity to *create your own* formula to capture the idea, like an expert!

5.6 British Thermal Unit (Btu)

Why would we waste our time talking about the arcane British thermal unit (Btu)? It's because data provided by the U.S. Energy Information Administration on global energy use is based on the Btu. More specifically, country-scale annual energy expenditures are measured in units of quadrillion (10^{15}) Btu (see Box 5.4). Also, heating appliances in the U.S.[18] are rated in Btu/hour—a unit of power that can be converted to Watts.

We need to cover the unit in this chapter in order to be energy-literate in the U.S., and because it will come up later in this book.

18: . . . hot water heaters, furnaces, air conditioners, ovens and stoves

Definition 5.6.1 *The* ***Btu*** *is the Imperial analog to the kcal.*[19] *One Btu is the energy required to heat one pound of water one degree Fahrenheit.*

In terms of Joules, 1 Btu is about 1,055 J, or not far from 1 kJ.

19: Recall that 1 kcal is the energy it takes to heat one kilogram of water by 1°C.

We can make sense of the conversion to Joules in the following way: a pound is *roughly* half a kilogram and one degree Fahrenheit is *approximately* half a degree Celsius. So a Btu should be roughly a quarter of a kcal. Indeed, 1,055 J is close to one quarter of 4,184 J.

Box 5.4: Quads: qBtu

The U.S. uses quadrillion Btu to represent country-scale annual energy expenditures. It is denoted as qBtu, or informally "quads." One qBtu is approximately 10^{18} J.[20]

The U.S. uses about 100 quads per year. Since a year is about 3.16×10^7 seconds,[21] dividing energy in Joules by time in seconds tells us that the U.S. *power* is about 3×10^{12} W (3 TW), working out to about 10,000 W per person as a per-capita *rate* of energy use.

20: 1.055×10^{18} J, more precisely.

21: A cute and convenient way to remember this, approximately, is $\pi \times 10^7$ seconds per year.

Example 5.6.1 For appliances characterized by Btu/hr, we can relate to power in Watts via 1 Btu/hr as 1,055 J per 3,600 s, working out to 0.293 W.

Thus, a hot water heater rated at 30,000 Btu/hr is effectively 8,800 W.

Let's also pause to understand how long it will take to heat a shower's worth of hot water at this rate. We'll do it two ways:

1. Heating 15 gallons[22] (125 pounds) from a cool 68°F to a hot 131°F at 30,000 Btu/hr will take how long? We must put in

22: Typical shower flow is about 2 gallons, or ~8 L, per minute.

$125 \times 63 = 7,900$ Btu of energy at a rate of 30,000 Btu/hr, so it will take 7,900/30,000 of an hour, or just over 15 minutes.

2. In metric terms, the equivalent to 15 gallons is 57 L (57 kg), and we heat from 20°C to 55°C at 8,800 W.[23] Since one kcal heats one kilogram of water 1°C, heating 57 kg by 35°C will require 57×35 kcal, or $57 \times 35 \times 4,184\,\text{J} = 8.35\,\text{MJ}$, which at 8,800 W will take 950 seconds, also just over 15 minutes (reassuringly, the same answer).

23: 30,000 Btu/hr is equivalent to 8,800 W, as worked out above.

5.7 Therms

We will rarely encounter this unit, but include it here because natural gas utility bills[24] in the U.S. often employ Therms. Since part of the goal of this book is to empower a personal understanding of energy and how to compare different measures of energy (e.g., on a utility bill), conventions in the U.S. demand that we cover the unit here.[25]

24: See, for instance, the banner image for this chapter on page 68.

25: Chapter 20 will explore what might be learned from utility bills.

Definition 5.7.1 *One **Therm** is 100,000 Btu, or* 1.055×10^8 *J, or 29.3 kWh.*

Box 5.5: Why Therms?

The Therm is partly adopted for the near-convenience that 100 cubic feet of natural gas (CCF or 100 CF), which meters measure directly, equates to 1.036 Therms. Relatedly, one gallon (3.785 L) of liquid propane gas[26] contains 91,500 Btu, which is 0.915 Therms. Thus the Therm very closely matches convenient measures of natural gas (100 cubic feet) or liquid propane (a gallon).

26: Propane is often used in more remote locations as a substitute for natural gas when the pipeline infrastructure for natural gas is absent.

Example 5.7.1 It might take approximately 10,000 kcal of energy[27] to heat a fresh infusion of cold water into a hot water heater tank. How many Therms is this?

We do a two-step conversion: first, 10,000 kcal is 41.84 MJ, which at 1,055 J per Btu computes to about 40,000 Btu, which is the same as 0.4 Therm, requiring approximately 40 cubic feet of natural gas, or a little less than half-a-gallon (about 2 L) of liquid propane.

If we learn that the hot water heater is rated at 30,000 Btu per hour, it will take an hour and 20 minutes to complete the job.

27: Based on a capacity of 200 L, pulling in chilly water at 5°C and heating it to 55°C, thereby requiring $200\,\text{kg} \times 50\,\text{C}$ kcal.

It is interesting to reflect on the notion that 200 L of water can be heated by 50°C for only 2 L of liquid fuel: 1% of the water volume in fuel. If heating to boiling, it would take twice as much fuel, so 2% of the water volume. Seems like a good bargain—especially for backpackers who want to boil water and have to lug the fuel around to do so. Inefficiencies in getting heat into the water might require more like 10% fuel volume.

5.8 Electrical Power

Electronic interactions are governed by charges pushing on each other. For the purposes of this course, we need only understand a few concepts. The first is voltage.

Voltage is a measure of *electric potential*, in Volts, and can be thought of as analogous to how high something is lifted.[28] A higher voltage is like sitting higher on the shelf, and can do more work if allowed to be released.

28: . . . making electric potential a lot like gravitational potential energy in flavor

Charge is moved around by electrical forces, and the amount of charge moved plays a role similar to that of mass in gravitational settings. The unit of charge is the Coulomb (C), and the smallest unit of charge we encounter in normal situations is from the proton ($+1.6 \times 10^{-19}$ C) or the electron (-1.6×10^{-19} C).

Definition 5.8.1 *The amount of energy in a charge, q, at a voltage, V, is*

$$E = qV. \tag{5.1}$$

One Coulomb of charge at a potential of 1 V has an energy of 1 J.

Current is the rate at which charge flows, and is usually symbolized by the letter I. Imagine setting up a toll booth in a conducting wire and counting how many charges (or how much cumulative charge) pass the gate per unit time. This gives rise to the Definition 5.8.2.

Definition 5.8.2 *Current is measured in Amps,[29] which is defined as one Coulomb per second.*

29: Amperes, formally

Moving one Coulomb through one Volt every second would constitute one Joule of energy every second, which is the definition of one Watt. Putting the concepts of Definition 5.8.1 and Definition 5.8.2 together, we find ourselves able to define electrical power.

Definition 5.8.3 *Electrical power is simply current multiplied by voltage:*

$$P = IV. \tag{5.2}$$

Current, I, is in Amps, and voltage, V is in Volts.

Example 5.8.1 Households in the U.S. often have circuit breakers allowing maximum currents of 15 or 20 Amps for regular power outlets. At a voltage of 120 V,[30] this corresponds to a maximum power of 1,800 W or 2,400 W, respectively.[31]

30: The alternating current nature is already accommodated in this measure of voltage.

31: Safety regulations limit continuous use to 80% of the breaker current capacity, so that realistically the limits are 1,400 W and 1,920 W, respectively. This is why "heating appliances" in Figure 5.2 top out around 1,500 W: circuit/safety limits.

Finally, we are in a position to understand how much energy a battery will hold. Batteries are rated by two numbers: voltage, and charge capacity. Since current is charge per time, multiplying current and time results in just charge.[32] Therefore, charge capacity in batteries is characterized as Amp-hours (Ah) or milli-amp-hours (mAh). Since Amps times Volts is Watts (Eq. 5.2), Amp-hours times Volts is Watt-hours, a familiar unit of energy from Section 5.4.

32: For example, 0.1 Amps (0.1 Coulombs per second) of current sustained for a duration of 100 seconds results in 10 Coulombs of charge flow.

Example 5.8.2 A typical 9-volt battery has a capacity of 500 mAh. How much energy is this?

500 mAh is 0.5 Ah. Multiplying by 9 V produces 4.5 Wh. Recall that 1 Wh is 1 J/s times 3,600 s (one hour), or 3,600 J. So 4.5 Wh is 16.2 kJ.

How long can we power a 1 W LED array from this battery? We can go the long way (16.2 kJ divided by 1 J/s) and say 16,200 seconds, or recognize that a 4.5 Wh battery can dispense 1 W for 4.5 hours. It's the same either way.[33]

33: Approaching a problem from multiple directions provides validation and also promotes greater flexibility.

5.9 Electron Volt (eV)

The electron-volt (eV) is the unit of choice for energy at the atomic scale. This makes it ideal for discussing individual chemical bond strengths, the energy of individual photons of light emitted from atoms, and thermal energy per atom or molecule.[34] We also use the eV for nuclear physics, but must increase the scale one million-fold and therefore speak of the mega-electron-volt, or MeV.

34: Really, this is just the kinetic energy of the particle.

We have already hit all the relevant concepts for understanding the eV in Section 5.8. The main reason to have its own section is so that it appears separately in the table of contents, making it easier to find and reference. The definition follows Definition 5.8.1 closely.

Definition 5.9.1 *One* ***electron-volt*** *is the energy associated with pushing one fundamental unit of electron charge,* $|e| = 1.6\times10^{-19}$ *Coulombs, through an electric potential of 1 V:*

$$1\,\text{eV} = 1.6\times10^{-19}\,\text{C}\cdot1\,\text{V} = 1.6\times10^{-19}\,\text{J} \tag{5.3}$$

The electron-volt, at 1.6×10^{-19} J, is a tiny amount of energy. But it's just the right level for describing energetic processes for individual atoms.

Example 5.9.1 When 12 grams of carbon (one mole, or 6×10^{23} atoms[35]) reacts with oxygen to form CO_2, about 394 kJ of energy is released.[36] How much energy is this *per carbon atom* in electron-volts?

Since we have one mole, or 6×10^{23} carbon atoms, we divide our total energy (3.94×10^5 J) by the number of atoms to get 6.5×10^{-19} J per atom. This is just a bit larger than 1 eV (1.6×10^{-19} J), and the division leads to something very close to 4 eV per atom.

Because CO_2 has a total of four bonds between the carbon atom and the two oxygen atoms,[37] we see that each bond accounts for about 1 eV. Chemical bonds are often in this range, highlighting the usefulness of the eV unit at the atomic level.

35: See Appendix B for a primer/refresher on chemistry.

36: Tables in chemistry books contain this type of information.

37: Each carbon-to-oxygen link is a *double* bond, meaning that two electrons participate in the link, for a total of four.

5.10 Light Energy

Light energy and its spectrum will be explored more extensively in Chapter 13, but the main concepts are covered here for completeness.

Light can be used to describe any part of the electromagnetic spectrum, from radio waves and microwaves, through infrared, visible, ultraviolet, and on to X-rays and gamma rays. Like atoms, light is "quantized" into smallest indivisible units—in this case particles called photons. An individual photon's energy is characteristic of its wavelength, λ (Greek lambda), or frequency, ν (Greek nu).[38]

38: The two are related by the speed of light, c, via $\lambda\nu = c$.

Definition 5.10.1 *The energy of a photon is given by*

$$E = h\nu = \frac{hc}{\lambda}, \tag{5.4}$$

where $h = 6.626 \times 10^{-34}\,\mathrm{J \cdot s}$ *is Planck's constant and* $c \approx 3.0 \times 10^{8}$ *m/s is the speed of light.*

Example 5.10.1 Visible light has a wavelength of 0.4–0.7 μm,[39] corresponding to 2.8–5.0×10^{-19} J for each photon.

39: A micron (μm, or micrometer) is another way to say 10^{-6} m.

We also routinely express photon energy in electron-volts (eV) according to Definition 5.10.2.

Definition 5.10.2 *Given the wavelength in microns (μm), the energy of a photon in eV units is*

$$E_{\mathrm{eV}} = \frac{1.24}{\lambda(\mu\mathrm{m})}\,\mathrm{eV}. \tag{5.5}$$

Example 5.10.2 The red-end of the visible spectrum, around 0.7 μm, corresponds to photon energies around 1.8 eV, while the blue-end, around 0.4 μm, corresponds to 3.1 eV.

5.11 Upshot on Units

Every chapter has an upshot, usually distilling key lessons from the chapter or offering final thoughts. Such a treatment is not necessary here, although we could reinforce the idea that energy can always be expressed in Joules, or converted into any of the units described in the chapter. Also critical is the notion that energy is conserved—only exchanging from one form to another but never truly disappearing or coming from nowhere.

Students may wish to see a master table of conversions between all the units discussed—and what a glorious table this would be! But it is intentionally left out for three reasons:

1. It could short-circuit your effort to learn the material;
2. Problems will ask you to do some of this;
3. This would be a fantastic opportunity for you to design and populate your *own* master conversion table. Then you'll really own it.

Great idea! Go for it!

5.12 Problems

1. A typical textbook may have a mass of 1 kg, and thus a weight of about 10 N. How high could the textbook be lifted (against the force of gravity) by supplying one Joule of energy?

Weight is mg, where $g \approx 10\,\mathrm{m/s^2}$.

2. If you look in your "energy wallet" and only have 24 J of energy available to spend, how far can you expect to slide an empty box across the floor if it takes 6 N of force to move it along?
3. A 50 kg crate might require 200 N to slide across a concrete floor. If we must slide it 10 m along the floor and then lift it 2 m into a truck, how much energy goes into each action, and what fraction of the total energy expenditure is each?
4. Come up with your own scenario (a force and a distance) that would result in 100 J of energy expenditure.
5. The numbers in Table 5.1 are *reasonable* but should not be thought of as *right*.[40] You can make your own table by using mgh for lifting and $\frac{1}{2}mv^2$ for kinetic energy. For this exercise, pick three familiar activities or situations that allow you to estimate an energy scale in Joules and compute/estimate the results.

40: Every nerf toss is not 15 J; the bookbag lift depends on how heavy and how high the lift; every example would have a range of reasonable numbers.

6. Just for fun, compute the energy associated with the mass of a tiny bit of shaving stubble having a mass of 0.01 mg[41] using $E = mc^2$. Make sure you use the correct units to put the result in Joules. The speed of light, c, is approximately 3×10^8 m/s.

41: . . . based on 0.1 mm diameter and 1 mm long

7. What exchanges of energy (between what forms) happens when a hand grenade explodes and sends pieces of its casing flying away from the explosion at high velocity? You may wish to describe more than one step/exchange.
8. Follow the evolution of energy exchanges for a wad of clay that you throw high into the air. Describe what is happening as the clay moves upward, as it reaches its apex, as it falls back down, and finally hits the ground with a thud. Where does the initial energy you put into the clay end up?
9. A couch might take 100 N to slide across a floor. If someone slides the couch 4 meters and does it in 8 seconds, how much power did they expend?

ⓘ The equivalent force to lifting 10 kg, or 22 lb of pushing force.

10. If a 70 kg person (weight: 700 N) is capable of putting out energy at a rate of 500 W in short bursts, how long will it take the person to race up a flight of stairs 4 m high, considering only the vertical energy[42] required?

42: Ignoring inefficiencies of moving legs, rounding flights, etc.

11. If asked to compute the power associated with performing a pull-up,[43] what specific information would you need to solve the problem (and what are the units of each)? Write out the math that would give the final answer.

43: . . . or chin-up, lifting your entire body up to a bar using your arms

12. How many kcal will it take to heat 1 liter of water (e.g., in a pot) from room temperature (20°C) to boiling (100°C)? How many Joules is this?

13. If a microwave operates at a power of 1,600 W (1,600 J/s), how long will it take to heat 0.25 L of water from room temperature to boiling (changing temperature by 80°C) if 50% of the microwave energy is absorbed by the water?

ⓘ 50% is typical for microwave efficiency.

14. A smaller or less active person may require only 1,300 kcal per day of food intake, while a larger or more active person might demand 3,000 kcal per day. Approximately what range of power does this spread translate to, in Watts?

ⓘ The result can help inform your sense for the typical range of human metabolic power.

15. If a typical metabolic intake is 2,000 kcal each day, approximately how much energy does this translate to for one day, in units of kWh? Compare this to a typical American household's electricity usage of 30 kWh in a day.

Hint: it may be convenient to first get power in Watts and round to a nice number before proceeding.

16. The chapter banner image (page 68) shows food labels for peanut butter and Nutella. The former indicates 188 Calories in a 32 g serving, while Nutella is 539 kcal in 100 g. To compare, we must adjust to the same serving size. Using 100 g as a sensible reference, which of the two is more energetic for the same serving size, and by how much (as a percentage)?

17. Based on the peanut butter label in the chapter banner image (page 68), showing 188 Cal per 32 g serving, how much mass of peanut butter would need to be consumed daily to constitute a 2,000 kcal/day diet? If a baseball has a mass of 145 g, how many baseballs of peanut butter would need to be consumed each day?

18. A generic $10 pizza might contain about 2,500 kcal. What is this in kWh? Electricity typically costs $0.15 per kWh,[44] so how much would a pizza's amount of energy cost in electrical terms? Which of the two is a cheaper form of energy?

ⓘ Comparable to a full day's intake.

44: . . . regionally variable

19. A refrigerator cycles on and off. Let's say it consumes electrical power at a rate of 150 W *when it's on*, and (essentially) 0 W *when it's off*. If it spends half of its time in the on-state, what is its *average* power? How much energy does it consume in a 24-hour day, in

kWh? At a typical electricity cost of $0.15 per kWh, about how much does it cost per *year* to run the refrigerator?

Hint: either treating it as if it is only on for 12 hours, or operating at half-power for 24 hours will yield equivalent answers.

20. The chapter banner image (page 68) shows data from the author's utility bill, indicating 230 kWh of electrical usage for a 30-day period in 2020. What does this rate of energy usage translate to, in Watts?

21. Heating a typical house might require something like 200 W of power for every degree Celsius difference between inside and outside temperatures. If the inside temperature is kept at 20°C and the outside temperature holds steady all day and night at 0°C, how much power is required to maintain the temperature?

ⓘ For instance, if the temperature difference is 10°C, the house will require 2,000 W of steady input to maintain temperature.

22. If Problem 21 had resulted in 5,000 W,[45] how much energy is used in a 24-hour day, in Joules? Express in the most natural/convenient multiplier (i.e., J, kJ, MJ, GJ, etc.) depending on the scale.

45: ⚠ it does not, exactly

23. If Problem 21 had resulted in 5,000 W,[46] how many kilowatt-hours (kWh) are expended in a 24-hour period? At an electricity cost of around $0.15 per kWh,[47] about how much will it cost, per day, to maintain heat?

46: ⚠ it does not, exactly

47: . . . regionally variable

24. If Problem 21 had resulted in 5,000 W,[48] how many Btu are required in a day to maintain temperature? How many Therms is this? At a typical cost of around $1.25 per Therm, about how much does it cost per day to heat the home?

48: ⚠ it does not, exactly

25. If Problem 21 had resulted in 5,000 W,[49] how many gallons of liquid propane[50] would be consumed in heating the home for a day? At a cost of around $2.50 per gallon, about how much does it cost per day to heat the home?

49: ⚠ it does not, exactly

50: Hint: related to Problem 24 and the fact that a gallon of propane contains 0.915 Therms of energy.

26. The chapter banner image (page 68) shows data from the author's utility bill, reflecting 230 kWh of electricity and 4 Therms of gas usage. Annoyingly, the units are different. How do the actual energies compare, if expressed in the same units?[51] How would you capture in a simple sentence the approximate comparison of energy use for each?

51: . . . recommend kWh as common basis

27. The chapter banner image (page 68) shows part of the hot water heater label in the author's home, showing a rating of 40,000 Btu/hr. How much power is it capable of putting out, in Watts?

28. The chapter banner image (page 68) shows the energy label associated with the author's hot water heater, estimating that it will use 242 Therms per year. If the estimated energy cost is distributed evenly across 12 months, what would the utility bill be expected to report? Based on an actual utility bill in the same image, the usage for one billing period was 4 Therms. How does actual usage compare to estimated usage, as an approximate percentage?

29. The chapter banner image (page 68) has two panels relating to the same hot water heater. One indicates the rate of gas usage when the heater is on (ignited, heating water) as 40,000 Btu/hr, and the other anticipates 242 Therms per year will be used. How many hours per day is the heater expected to be on (heating water) based on these numbers?

Hint: useful to convert to Therms/hr

30. Gather up or compute conversion factors from the chapter to start your own conversion table (empty version below). Express kWh, cal, kcal, Btu, and Therms in terms of Joules.

From →	kWh	cal	kcal	Btu	Therms
To: J					

31. Perhaps assisted by Problem 30, create a table for conversions between kWh, kcal, Btu, and Therms in terms of one another. The table is started out below, populating the diagonal (no conversion necessary) and also providing a start that 1 Therm is 29.3 kWh.

Not all these conversions are likely to be useful, but a few will come up in practice.

From →	kWh	kcal	Btu	Therms
To: kWh	1			29.3
To: kcal		1		
To: Btu			1	
To: Therms				1

32. A car headlight using light emitting diodes (LEDs) operates at about 15 W. If drawing from the car's 12 V battery, how much current, in Amps, flows to the headlight?

33. Houses in the U.S. are equipped with circuit protection rated to 100 or 200 Amps, typically. If a 100 A house is operating at 80% of its rated capacity,[52] how much power is it consuming (at 120 V)? If sustained for a month, how many kWh will show up on the bill? At $0.15/kWh, what is the cost?

52: All circuits blazing at the safety limit!

34. The chapter banner image (page 68) shows a rechargeable AA battery, operating at 1.2 V and holding 2,200 mAh of charge. How many Joules is this, and how long could it power a 1 W LED array?

35. If we have 6×10^{23} molecules,[53] and each molecule releases 1 eV in a chemical reaction, how many kJ (per mole, as it turns out) is this reaction?

53: This is one mole, as covered in Sec. B.1 (p. 375).

36. Considering the typical wavelength of light to be 0.55 μm, what is a typical photon energy, in Joules, and how many photons per second emerge from a 1 W light source?[54]

54: Assuming 100% efficiency

37. At what wavelength, in microns (μm), is the corresponding photon energy in eV the same number? A deliberately wrong example to illustrate would be if a 2.6μm wavelength corresponded to 2.6 eV (it doesn't').

A locomotive engine as an example heat engine. Photo credit: South Australian Government Photographer.

6 Putting Thermal Energy to Work

We have already encountered thermal energy in two contexts. The first was infrared radiation (Eq. 1.8; p. 10), and the second was in the definition of the kilocalorie (Sec. 5.5; p. 73). Otherwise, heat has often been treated as a form of "waste" in a chain of energy conversion: friction, air resistance, etc. The insinuation was that heat is an unwanted byproduct of no value.

6.1 Generating Heat 84
6.2 Heat Capacity 85
6.3 Home Heating/Cooling 86
6.4 Heat Engines 88
Entropy and Efficiency Limits 90
6.5 Heat Pumps 95
Consumer Metrics 97
6.6 Upshot on Thermal Energy . 99
6.7 Problems 99

Yet 94% of the energy we use today is thermal in nature [34]: we *burn* a lot of stuff for energy![1] Sometimes heat is what we're after, but how can we use it to fly airplanes, propel cars, and light up our screens? This chapter aims to clarify how heat is used, and explore limits to the efficiency at which heat can perform non-thermal work.

[34]: U.S. Energy Inform. Administration (2011), *Annual Energy Review*

1: The exceptions are wind, hydroelectricity, and solar.

Like the previous chapter, this topic represents a slight detour from the book's overall trajectory, which otherwise aims to build a steady narrative of what we *can't* expect to continue doing, what options we *might* use to change course, and finally how to bring about such change. Nonetheless, the way we utilize thermal energy is a key piece of the story, and relates to both current and future pathways to satisfying our energy demands.

6.1 Generating Heat

Before diving in to thermal issues, let's do a quick run-down of the various ways we can generate heat.

Example 6.1.1 Ways to Generate Heat: Roughly arranged according to degree of sophistication:

1. Rub your hands together (or other forms of friction).
2. Harvest sunlight, possibly concentrating it, for heat; drying clothes outside and letting sunlight warm a room through a window are examples.
3. Access geothermal heat in select locations.
4. Burn wood in a fireplace or stove.
5. Burn a fossil fuel for direct heat; gas is often used in homes for space heating, as well as for heating water and cooking.
6. Run electrical current through a coil of wire that glows orange; seen in toaster ovens, hair dryers, space heaters.
7. Use electricity to run a heat pump (Section 6.5).
8. Allow nuclear material to undergo fission in a controlled chain reaction.
9. Contrive a plasma hot enough to sustain nuclear fusion—as the sun has done for billions of years.

6.2 Heat Capacity

First, we'll connect a basic thermal concept to something we already covered in Sec. 5.5 (p. 73) in the context of the calorie. The statement that it takes 1 kcal to heat 1 kilogram of H_2O by 1°C is in effect defining the *specific heat capacity* of water. In SI units, we would say that H_2O has a specific heat capacity of 4,184 J/kg/°C.[2] Very few substances top water's specific heat capacity. Most liquids, like alcohols, tend to be in the range of 2,000 J/kg/°C. Most non-metallic solids (and even air) come in around 1,000 J/kg/°C. Metals are in the 130–900 J/kg/°C range—lighter metals at the top, and heavier ones at the lower end.[3] Table 6.1 provides a sample of specific heat capacities for common substances.

Table 6.1: Specific heat capacities of common materials.

Substance	J/kg/°C
steel	490
rock, concrete	750–950
glass	840
aluminum	870
air	1,005
plastic	1,100–1,700
wood	1,300–2,000
alcohol	2,400
flesh	3,500
water	4,184

2: For temperature changes, it is always possible to interchange per-degrees-Celsius and per-Kelvin because the two are only different by a constant offset, so that any *change* in temperature is the same measure in both.

3: The pattern here is that substances like water or alcohols containing light atoms like hydrogen have higher heat capacities than substances like metals containing heavier atoms.

Knowing the specific heat capacity of a substance allows us to compute how much energy it will take to raise its temperature. A useful and approximate guideline is to treat water as 4,000 J/kg/°C and *all other stuff* (air, furniture, walls) as 1,000 J/kg/°C. Mixtures, like food, might be somewhere between, at 2,000–3,500, due to high water content. If in doubt, 1,000 J/kg/°C is never going to be *too* far off. For estimation purposes, deviate from this only for high water-content[4] or for metals.[5]

4: . . . go as high as 4,000 J/kg/°C in this case

5: . . . 500 for heavier metals like steel; although light metals like aluminum are not far from 1,000 J/kg/°C

Example 6.2.1 A 2,000 kg pick-up truck is transporting a one-cubic-meter container of water. How much energy will it take to raise the temperature of the whole ensemble by 5°C?

A cubic meter of water (1,000 L) is 1,000 kg and has a heat capacity around 4,000 J/kg/°C; the truck is mostly steel, so we might guess 500 J/kg/°C. Multiply each specific heat capacity by the respective mass and the 5 degree change to get 20 MJ to heat the water and 5 MJ to heat the truck for a total of 25 MJ.[6]

6: Notice that the water demands far more energy to heat, even though it is half the mass.

To perform computations using specific heat capacity, try an intuitive approach rather than some algorithmic formula.[7] The following should just make a lot of sense to you, and can guide how to put the pieces together: it takes *more energy* to heat a *larger* mass **or** to raise the temperature by a *larger* amount. It's all proportional. The units also offer a hint. To go from specific heat capacity in J/kg/°C to energy in J, we need to multiply by a mass and by a temperature change.

7: Although, this would be a good opportunity for a student to make up their *own* formula, driving home the concept and the fact that equations simply capture a concept. Also, the choice of symbols is arbitrary, which the experience would reinforce.

Example 6.2.2 To compute the amount of energy it will take to heat a 30 kg piece of furniture[8] by 8°C, we will multiply the specific heat capacity by the mass—to capture the "more mass" quality—and then multiply by the temperature change—to reflect the "more temperature change" element. In this case, we get 240 kJ.

8: . . . assuming 1,000 J/kg/°C

6.3 Home Heating/Cooling

Our personal experience with thermal energy is usually most connected to heating a living space and heating water or food. Indeed, about two-thirds of the energy used in residential and commercial spaces[9] relate to thermal tasks, like heating or cooling the spaces, heating water, refrigeration, drying clothes, and cooking.

9: . . . in the form of natural gas, electricity, and fuel oil

When it comes to heating (or cooling) a home, we might care about two things:

- how long will it take to change its temperature by some certain amount; and
- how much energy it will take to keep it at the desired temperature.

The first depends on how much stuff is in the house,[10] how much ΔT you want to impart, and how much power is available to create[11] the heat. The energy required is mass times ΔT times the catch-all 1,000 J/kg/°C specific heat capacity. The time it takes is then the energy divided by the available power.

10: . . . including walls, furniture, air

11: . . . or to remove, if cooling

Example 6.3.1 How long will it take to heat up the interior of a mobile home from 0°C to 20°C using two 1,500 W space heaters? We'll assume that we must heat up about 6,000 kg of mass.[12]

The first job is to find the energy required and then divide by power to get a time. We'll use the good-for-most-things specific heat capacity of 1,000 J/kg/°C.

Multiplying the specific heat capacity by mass and temperature change results in 120 MJ of energy. At a rate of 3,000 W, it will take 40,000 s to inject this much energy, which is about 11 hours.

12: Only 300 kg is in the form of air: most of the mass to be heated is in the walls, floor, and ceiling.

How much it takes to *maintain* temperature depends on how heat flows out of (or into) the house through the windows, walls, ceiling, floor, and air gaps. But it also depends *linearly* on ΔT—the difference between inside and outside temperatures—that is being maintained. A house can be characterized by its heat loss rate in units of Watts per degree Celsius.[13] This single number then indicates how much power is needed to maintain a certain ΔT between inside and outside. Box 6.1 explores an example of how to compute the heat loss rate for a house, and Example 6.3.2 applies the result to practical situations.

13: ...or equivalently, Watts per degree Kelvin

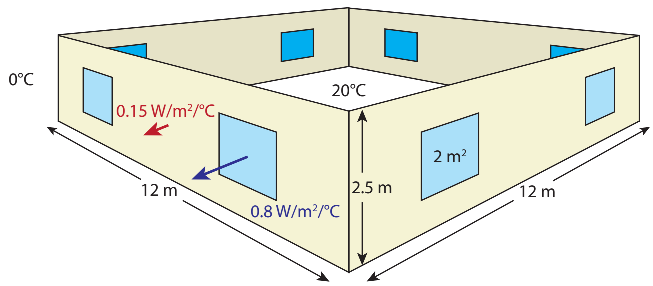


Figure 6.1: External walls and windows for the house modeled in Box 6.1. The floor and ceiling are not shown. The numbers in $W/m^2/°C$ are U-values, and in this case represent the very best engineering practices. Most houses will have larger values by factors as high as 2–6. Don't forget the door in a real house!

Box 6.1: House Construction

The very best practices result in a snugly-built house qualified as a "Passive House," achieving 0.15 W/°C *for each square meter* of external-interfacing surface[14] and 0.8 W/°C per square meter of windows.

Let's imagine a house having a square footprint 12 m by 12 m, walls 2.5 m high, each of the four walls hosting two windows, and each window having an area of 2 m^2 (Figure 6.1). The floor and the ceiling are both 144 m^2, and the wall measures (perimeter times height) $48 \times 2.5 = 120\,m^2$. But we deduct 16 m^2 for the eight windows, leaving 104 m^2 for the walls. The resulting heat loss measure for the house is 13 W/°C for the windows ($0.8\,W/m^2/°C \times 16\,m^2$), plus 59 W/°C for the walls/floor/ceiling for a total of 72 W/°C.

The loss rate for a *decently*-constructed house might be about twice this, while a *typically*-constructed house (little attention to efficiency) might be 3–6 times this—several hundred W/°C. Of course, smaller houses have smaller areas for heat flow and will have smaller loss rates.

14: ...outer walls, ceiling-to-unconditioned attic, floor-to-crawl-space

The numbers used to characterize heat loss properties of walls and windows are called U-values, in units of $W/m^2/°C$, where low numbers represent better insulators. In the U.S., building materials are described by an inverse measure, called the R-value, in ugly units of °F $\cdot$ft^2 $\cdot$ hr/Btu. The two are numerically related as $R = 5.7/U$, so that our Passive House wall has $R \approx 38$ and the windows have $R \approx 7$—both rather impressive and hard to achieve.

Example 6.3.2 Let's compare the requirements to keep three different houses at 20°C while the temperature outside is 0°C (freezing point). The first is a snugly-built house as described in Box 6.1, where we round the heat loss rate to a more convenient 75 W/°C. We'll then imagine a decently built house at 150 W/°C, and a more typical[15] house at 300 W/°C.

The temperature difference, ΔT, is 20°C, so that our super-snug house

15: ...not "thermally woke"

will require 75 W/°C times 20°C, or 1,500 W[16] to keep it warm, while the decent house needs 3,000 W and the shoddy house needs 6,000 W.

16: . . . a single space heater

Once we understand how much power it takes to maintain a certain temperature (ΔT) in a house, we can anticipate the behavior of the house's heater. Heaters are typically either on full-blast or off. Regulation is achieved by turning the heat on and off—usually controlled by a thermostat. Given the *rating* of a heater,[17] it is then straightforward to anticipate the duty cycle: the percentage of time it has to be on to produce an *average* output meeting the power requirement for some particular ΔT.

17: The rating is effectively the power delivered when operating at full capacity.

In a sensible world, heaters are characterized by W (or kW). In the U.S., the measure for many appliances is Btu/hr. Since 1 Btu is 1,055 J and one hour is 3,600 s, one Btu/hr equates to 0.293 W.[18] A whole house heater—sometimes in the form of a furnace—might be rated at 30,000 Btu/hr (about 10 kW), in which case the three outcomes in Example 6.3.2 would require the heater to be on about 15%, 30%, or 60% of the time[19] to maintain $\Delta T = 20$°C in the three houses.

18: 1,055 J in 3,600 s is 0.293 J/s.

19: These are duty cycles.

It is also possible to assess how much ΔT the foregoing heater could maintain in the three houses. It should stand to reason that a house requiring 100 W/°C and having a 10,000 W heater can support a ΔT as high as 100°C.[20] Thus, the three houses from Example 6.3.2 could support ΔT values of 133°C, 67°C, and 33°C if equipped with a 10 kW (~30,000 Btu/hr) heater. The snug house does not need such a powerful heater installed. The poorly built house can maintain a $\Delta T = 33$°C differential at full-blast, which means that if the temperature drops below −13°C (8.6°F) outside, it will not be able to keep the inside as high as 20°C. So a house in a cold climate should either be built to better thermal standards, or will require a bigger heater—costing more to heat the home.[21]

20: First, this is a ridiculously high number! Second, rather than rely on an equation, or memory about whether the 100 W/°C and 10,000 W should be divided or multiplied, try to internalize the meaning of each, or at least use the units as a guide. Then, the appropriate math manipulation becomes more obvious.

21: Other possible options are to tolerate a lower internal temperature or move someplace warmer.

Cooling a home (or refrigerator interior, or whatever) is also a thermal process, but in this case involves *removing* thermal energy from the cooler environment. Removing heat is harder to do, as witnessed by the length of human history that has utilized heating sources—starting with fire—compared to the relatively short amount of time when we have been able to produce cooling on demand.[22] Section 6.5 will get into how this is even possible, in principle. For now, just be aware that the rating on air conditioners uses the same units as heaters: how much thermal energy can be moved (out of the cooler environment) per unit time. In SI units, we know this as the Watt. In the U.S., it's Btu/hr.

22: In fact, we've had the word "warmth" for a long time, but have not even gotten around to inventing the word "coolth" yet.

6.4 Heat Engines

Now we get to the part where thermal energy can be used to do something other than just provide direct heat to a home. It may seem odd to always characterize burning fuel as a purely *thermal* action, since what transpires within the cylinder of a gasoline-burning internal combustion engine seems like more of a little *explosion* than just the generation of heat. This is not wrong, but neither is it the whole story. The process still begins as a fundamentally thermal event. When the fuel-air mixture ignites, the temperature in the cylinder increases dramatically. To appreciate what happens as an immediate consequence, we turn to the ideal gas law:

$$PV = Nk_B T. \tag{6.1}$$

This is the physicist's version, which looks a little different than the chemist's $PV = nRT$. For a comparison, see Sec. B.4 (p. 381).

P, V and T are pressure, volume, and temperature (in N/m^2, m^3, and Kelvin). N is the number of atoms or molecules, and $k_B = 1.38\times10^{-23}$ J/K is the Boltzmann constant, which we will see again in Sec. 13.2 (p. 199). The temperature rise upon ignition is fast enough that the cylinder volume does not have time to change.[23] Eq. 6.1 then tells us that the pressure must also spike when temperature does, all else being held constant. The increase in pressure then pushes the piston away, increasing the cylinder volume and performing work.[24] But it all starts thermally, via a sharp increase in temperature.

23: The moving piston allows the volume to change, but on slower timescales.

24: Work is measured as pressure times the change in volume. Pressure is force per unit area, so the units work out to force times distance, as they should given the definition of work.

In the most general terms, thermal energy tries to flow from hot to cold—out of a pot of hot soup; or into a cold drink from the surrounding air; or into your feet from hot sand. Some part of this flow can manifest as physical work, at which point the system can be said to be acting as a heat engine.

Definition 6.4.1 *A **heat engine** is loosely defined as any system that turns heat, or thermal energy into mechanical energy: moving stuff.*

Example 6.4.1 Example heat engines: when heat drives motion.

1. Hot air over a car's roof rises, gaining both kinetic energy and gravitational potential energy;
2. Wind is very similar, in that air in contact with the sun-heated ground rises and gains kinetic energy on an atmospheric scale;
3. The abrupt temperature increase in an internal combustion cylinder drives a rapid expansion of gas within the cylinder;
4. Steam in a power plant races though the turbine because it is flowing to the cold condenser.

The last example deserves its own graphic, as important as this process is in our lives: almost all of our electricity generation—from all the fossil fuels and even from nuclear fission—follows this arrangement. Figure 6.2 illustrates the basic scheme. Table 6.2 indicates that 98% of our electricity involves turning a turbine on a shaft connected to a generator, and 84% involves a thermal process as the motive agent for the turbine—most often in the form of steam.

Table 6.2: Schemes for electricity generation. Most are thermal in nature, and nearly all employ a turbine and generator. Data for 2018 from Table 8.2a of [34].

Source	% elec. in U.S.	therm.	turb./ gen.
Nat. Gas	35.3	✓	✓
Coal	27.3	✓	✓
Nuclear	19.2	✓	✓
Hydroelec.	7.0		✓
Wind	6.6		✓
Solar PV	2.2		
Biomass	1.5	✓	✓
Oil	0.6	✓	✓
Geotherm.	0.4	✓	✓
Sol. Therm.	0.09	✓	✓

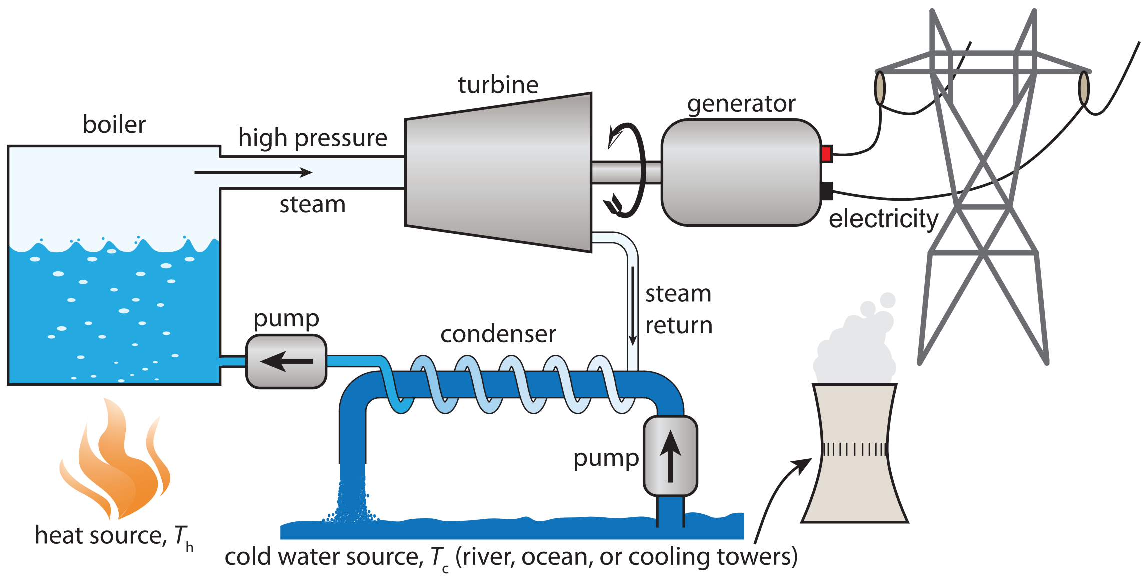


Figure 6.2: Generic power plant scheme, in which some source of heat at T_h generates steam that flows toward the condenser—where the steam cools and reverts to liquid water, by virtue of thermal contact to a cool source at T_c provided by a body of water or evaporative cooling towers. Along the way, the rushing steam turns a turbine connected to a generator, exporting electricity. This basic arrangement is employed for most power plants using fossil fuels, nuclear, solar thermal, or geothermal sources of heat.

6.4.1 Entropy and Efficiency Limits

A deep and powerful piece of physics intervenes to limit how much useful work may be extracted out of a flow of heat from a hot source at temperature T_h to a cold source at temperature T_c. That piece is entropy. You don't need to fully grasp the deep and subtle concept of entropy in order to follow the development in this chapter and understand the role entropy plays in limiting heat engine efficiency. All the same, it is a stimulating topic that we'll dip a toe into for some appreciation.

Definition 6.4.2 ***Entropy*** *is a measure of how many ways a system might be organized at the microscopic level while preserving the same internal energy.*[25]

25: E.g., at constant temperature, pressure, volume.

This definition may be an obscure disappointment to those expecting entropy to be defined as a measure of *disorder*.[26] Consider a gas maintained at constant pressure, volume, and temperature—thus fixing the total energy in the gas. The atoms/molecules comprising the gas can arrange into a staggeringly large number of configurations: any number of positions, velocities, rotational speeds and axis orientations, or vibrational states of each molecule, for instance—all while keeping the same overall energy.

26: Entropy is indeed *related* to disorder, in that there are many more ways to configure matches in a mess than there are ways to neatly stack them.

Example 6.4.2 To illustrate, consider a tiny system containing 3 molecules labeled A, B, and C, having a total energy of 6 units split between them in some way. They can all have exactly 2.0 units of energy apiece, or can have individual energies of 1.2, 1.8, and 3.0 units; or 3.2, 0.4, and 2.4; or any other of myriad combinations adding to the same thing. Entropy provides a measure of how many combinations[27] are possible.

27: It is far beyond the scope of this book to detail the counting scheme, but it is perhaps important to appreciate that energy levels are discrete—or *quantized*—which prevents an infinite number of possible energy combinations.

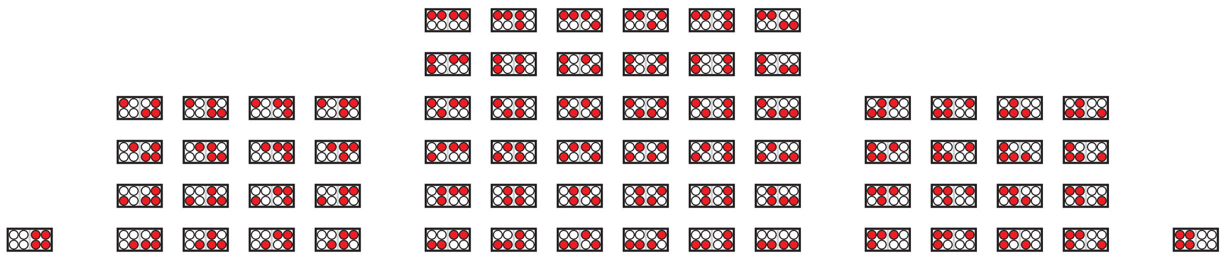


Figure 6.3: A box containing 4 atoms or molecules of one type (white) and 4 of another type (red) has many more configurations available (number in parentheses) when species are equally distributed so that left and right sides both have two of each. Entropy is related to the number of ways a system can distribute itself (at the same energy level), acting to favor disordered mixing over (improbable) orderly separation.

Example 6.4.3 To better elucidate the connection between entropy and disorder, imagine a box of air, containing both N_2 and O_2 molecules. As Figure 6.3 illustrates, a thoroughly-mixed arrangement has a larger number of possible configurations, thus the highest entropy. Nature does not give rise to spontaneous organization in a closed system.[28]

28: It is, however, possible to see lowered entropy in one place if balanced by an increase elsewhere: life organizes matter, but at the expense of increased entropy in the wider universe.

The First Law of Thermodynamics is one we already encountered as conservation of energy:

Definition 6.4.3 ***First Law of Thermodynamics****: the energy of a closed system is* conserved*, and cannot change if nothing—including energy—enters or leaves the system boundaries.*

Now we are ready for the Second Law.

Definition 6.4.4 ***Second Law of Thermodynamics****: the total entropy of a closed system* may never decrease.

It is entropy that governs *which way* heat flows (hot to cold, if left alone) and in a deep sense defines the "arrow of time."

Box 6.2: The Arrow of Time

Consider that if you were shown videos of a rock splashing into water, a coffee mug shattering on the floor, or an icicle melting, you would have no difficulty differentiating between the forward and reverse playbacks of the video.

The reverse action, you would conclude, is preposterous and can never happen. Pieces of ceramic strewn about the floor will never spontaneously assemble into a mug and leap from the floor! Energy is not the barrier, because the total energy in all forms is the same[29] before and after. It's entropy: the more ordered states are less likely to spontaneously emerge. To appreciate how pervasive entropy is, imagine how easy it is to spot a "fake" video run backwards.

29: . . . provided that the system boundary is drawn large enough that no energy escapes

These two laws of thermodynamics, plus a way to quantify entropy changes that we will see shortly, are all we need to figure out the maximum efficiency a heat engine can achieve in delivering work. If we draw an amount of heat, ΔQ_h from a hot bath[30] at temperature T_h, and allow part of this energy to be "exported" as useful work, ΔW, then we must have the remainder flow as heat (ΔQ_c) into the cold bath at temperature T_c. Figure 6.4 offers a schematic of the process. The First Law of Thermodynamics[31] requires that $\Delta Q_h = \Delta Q_c + \Delta W$, or that all of the extracted heat from the hot bath is represented in the external work and flow to the cold bath: nothing is lost.

30: By "bath," we mean a large reservoir at a constant temperature that is large enough not to appreciably change its temperature upon extraction of some amount of thermal energy, ΔQ.

31: . . . conservation of energy

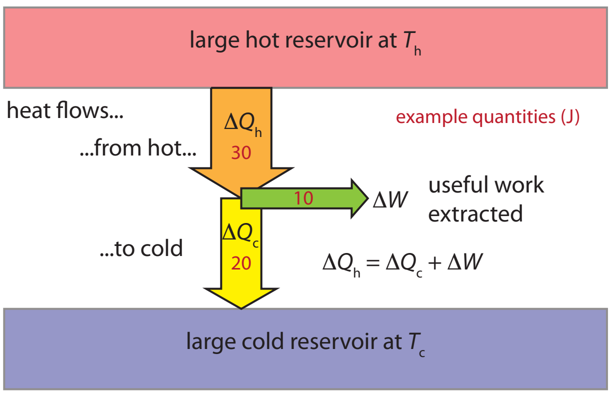


Figure 6.4: Heat engine energy balance. Heat flowing from the hot bath to the cold bath can perform useful work, ΔW, in the process—subject to conservation of energy ($\Delta Q_h = \Delta Q_c + \Delta W$), where ΔQ is a heat flow. Entropy constraints limit how large ΔW can be. Arrow widths are proportional to energy, and red numbers are example energy amounts, for use in the text.

So where does entropy come in? Extracting heat from the hot bath in the amount ΔQ_h results in an entropy change in the hot bath according to Definition 6.4.5.

Definition 6.4.5 ***Entropy Change***: *when energy (heat, ΔQ, in J) is moved into or out of a thermal bath at temperature T, the accompanying change in the bath's entropy, ΔS, obeys the relation:*

$$\Delta Q = T\Delta S. \tag{6.2}$$

When heat is removed, entropy is reduced. When heat is added, entropy increases. The temperature, T, must be in Kelvin, and entropy is measured in units of J/K.

Table 6.3: Thermodynamic symbols.

Symbol	Describes (units)
T	temperature (K)
ΔT	temp. change (K,°C)
ΔQ	thermal energy (J)
ΔW	mechanical work (J)
ΔS	entropy change (J/K)
ε	efficiency
η	entropy ratio

So the extraction of energy from the hot bath results in a *decrease* of entropy in the hot bath of ΔS_h according to $\Delta Q_h = T_h \Delta S_h$. Meanwhile, ΔS_c of entropy is *added* to the cold bath according to $\Delta Q_c = T_c \Delta S_c$. The Second Law of Thermodynamics enforces that the *total* change in entropy may not be negative (it can't decrease). In equation form (symbol definitions in Table 6.3):[32]

$$\Delta S_{\text{tot}} = \Delta S_{\text{c}} - \Delta S_{\text{h}} \geq 0, \tag{6.3}$$

where we have subtracted ΔS_{h} since it was a *deduction* of entropy, while ΔS_{c} is an addition. We therefore require that

$$\Delta S_{\text{c}} \geq \Delta S_{\text{h}}. \tag{6.4}$$

Now we are in a position to ask what fraction of ΔQ_{h} can be diverted to useful work (ΔW) within the constraints of the Second Law. We express this as an efficiency,[33] denoted by the Greek epsilon:

$$\varepsilon = \frac{\Delta W}{\Delta Q_{\text{h}}} = \frac{\Delta Q_{\text{h}} - \Delta Q_{\text{c}}}{\Delta Q_{\text{h}}}. \tag{6.5}$$

The second step applies conservation of energy: $\Delta Q_{\text{h}} = \Delta Q_{\text{c}} + \Delta W$.

32: Remember: treat equations as *sentences* expressing important concepts in precise ways—not merely as algorithmic machines to memorize for plugging in while solving problems. What does it *say*?

33: This definition of efficiency captures what we care about: what fraction of the extracted heat can be turned into useful work.

Example 6.4.4 Actual Efficiency: If a heat engine is observed to remove 30 J from the hot bath and deposit 20 J into the cold bath, as in Figure 6.4, what is the efficiency of this heat engine in producing useful work?

Whether we deduce that $\Delta W = 10$ J and use the first form in Eq. 6.5 or apply the second form using the given heat flows, the answer is 1/3, or 33%.

We can add a step to Eq. 6.5 to express it in terms of entropy changes:

$$\varepsilon = \frac{\Delta W}{\Delta Q_{\text{h}}} = \frac{\Delta Q_{\text{h}} - \Delta Q_{\text{c}}}{\Delta Q_{\text{h}}} = \frac{T_{\text{h}}\Delta S_{\text{h}} - T_{\text{c}}\Delta S_{\text{c}}}{T_{\text{h}}\Delta S_{\text{h}}}, \tag{6.6}$$

where we have re-expressed each ΔQ as an equivalent $T\Delta S$ withdrawal/deposit of entropy. Now we can divide both numerator and denominator by ΔS_{h} to be left with

$$\varepsilon = \frac{T_{\text{h}} - T_{\text{c}}\eta}{T_{\text{h}}}, \tag{6.7}$$

where we create η (eta) to represent the ratio of entropies: $\eta = \Delta S_{\text{c}}/\Delta S_{\text{h}}$, which we know from Eq. 6.4 cannot be smaller than one:[34]

$$\eta \geq 1. \tag{6.8}$$

34: If $A \geq B$, then we know that $A/B \geq 1$.

Looking at Eq. 6.7, if we want the highest possible efficiency in extracting work from a flow of heat, we want the numerator to be as large as possible. To achieve this, we want to subtract as little as possible from T_{h}. If η were allowed to be very large, then the numerator would be reduced. So we want the *smallest possible* value for η, which we know from Eq.

6.8 happens when $\eta = 1$. We therefore derive the maximum physically allowable efficiency of a heat engine as

$$\varepsilon_{\max} = \frac{T_h - T_c}{T_h} = \frac{\Delta T}{T_h}, \tag{6.9}$$

⚠ Temperature must be in Kelvin. Recall that $T(\text{K}) \approx T(°\text{C}) + 273$.

where we have designated $\Delta T = T_h - T_c$ as the temperature difference between hot and cold baths. A major takeaway is that efficiency improves as ΔT gets bigger, and becomes vanishingly small for small values of ΔT.

Example 6.4.5 If operating between a hot bath at 800 K and ambient temperature around 300 K,[35] a heat engine could produce a maximum efficiency of 62.5%.

35: 300 K is a convenient and reasonable temperature for "normal" environments, corresponding to 27°C or 80.6°F.

Example 6.4.6 A heat engine operating between boiling and freezing water has $T_h \approx 373$ K and $\Delta T = 100$ K, for a maximum possible efficiency of $\varepsilon_{\max} = 0.268$, or 26.8%.

Example 6.4.7 A heat engine operating between human skin temperature at 35°C and ambient temperature at 20°C has a maximum efficiency of $\varepsilon_{\max} = 15/308 \approx 0.05$, or 5%.

If the cold bath is fixed,[36] the maximum possible efficiency improves as the temperature of the hot source goes up. Conversely, for a given T_h, the efficiency improves as the cold temperature decreases and thus ΔT increases.

36: This is a common situation, as T_c is usually set by the ambient temperature of the air or of a body of water.

Box 6.3: At the Extreme Limit...

If T_c approaches 0 K[37], the maximum efficiency approaches 100%. We can trace this back to the relation $\Delta Q = T\Delta S$, which implies that when T is very small, it does not take much heat (ΔQ) to meet the requirement for the amount of entropy added to the cold bath (ΔS_c) to be large enough to satisfy the prohibition on net entropy decrease, so the arrow width in Figure 6.4 for ΔQ_c can be rather thin (small) allowing ΔW to be about as thick (large) as ΔQ_h, meaning that essentially all the energy is available to do work and the efficiency can be very high. In practice, Earth does not provide baths cold enough for this effect to kick in, but discussing it is a means to better understand how Eq. 6.9 works.

37: . . . absolute zero temperature, −273°C

Real heat engines like power plants (Figure 6.2) or automobile engines tend to only get about halfway to the theoretical efficiency due to myriad practical challenges. A typical efficiency for an electrical power plant is 30–40%, while cars are typically in the 15–25% range. In contrast, combustion temperatures around 700–800°C suggest a maximum theoretical efficiency around 60%.

6.5 Heat Pumps

We can flip a heat engine around and call it a heat pump. In this case, we *apply* some external work to *drive* a heat flow opposite its natural direction—like pushing heat uphill. This is how a refrigerator[38] works, for instance. Figure 6.5 sets the stage.

38: . . . or a freezer, or air conditioner

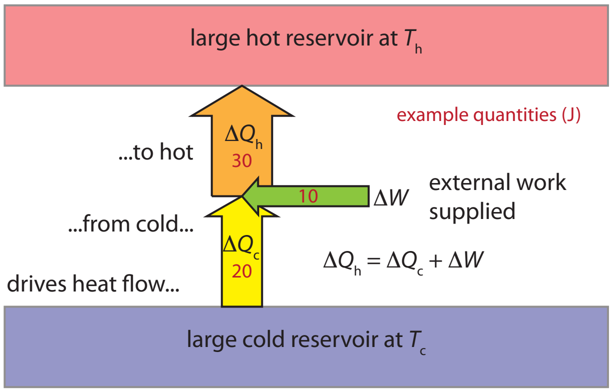


Figure 6.5: Heat pump energy balance. The application of work (ΔW; from an electrical source, for instance) can drive heat to flow—counterintuitively—from a cold reservoir (like the interior of a freezer) to a hotter environment. Example $T_c \rightarrow T_h$ pairs might include freezer-interior → room-air; cooled-inside → summer-outside; winter-outside → warmed-inside. We still must satisfy conservation of energy ($\Delta Q_h = \Delta Q_c + \Delta W$), where ΔQ is a heat flow. Entropy constraints limit how large ΔQ_c can be for a given ΔW input. Arrow widths are proportional to energy, and red numbers are example energy amounts, for use in the text.

A very similar chain of logic can be applied to this configuration, invoking the Second Law to guarantee no entropy decrease. We define the efficiency according to the application and what we care about, giving rise to two different figures of merit.

Definition 6.5.1 $\varepsilon_{\text{cool}}$: *For cooling applications,*[39] *we care about how much heat can be removed from the cooler environment (ΔQ_c) for a given input of work (ΔW). The efficiency is then characterized by the ratio* $\varepsilon_{\text{cool}} = \Delta Q_c / \Delta W$.

39: . . . freezer, refrigerator, air conditioner

Definition 6.5.2 $\varepsilon_{\text{heat}}$: *For heating applications,*[40] *we care about the heat delivered to the hot bath (ΔQ_h) for a given amount of input work (ΔW). The efficiency is then characterized by the ratio* $\varepsilon_{\text{heat}} = \Delta Q_h / \Delta W$.

40: . . . home heating via heat pump

The derivation goes similarly to the one above, but now we require that the entropy *added* to the hot bath must not be smaller than the entropy *removed* from the cold bath so that the total change in entropy is not negative.[41] In this case, the maximum allowed efficiencies for cooling and heating via heat pumps are:

41: Imposing this condition has the result that $\Delta S_h \geq \Delta S_c$; opposite Eq. 6.4 since the direction of flow changed.

$$\varepsilon_{\text{cool}} \leq \frac{T_c}{T_h - T_c} = \frac{T_c}{\Delta T}, \tag{6.10}$$

and

$$\varepsilon_{\text{heat}} \leq \frac{T_h}{T_h - T_c} = \frac{T_h}{\Delta T}. \tag{6.11}$$

⚠ Temperature must be in Kelvin for these relations.

These look a lot like Eq. 6.9, but *turned upside down*. The maximum efficiencies can be larger than unity![42]

42: See Box 6.4.

Example 6.5.1 What is the limit to efficiency of maintaining a freezer at −10°C in a room of 20°C?

First, we express the temperatures in Kelvin: $T_c \approx 263$ K and $\Delta T = 30$ K.[43] The maximum efficiency, by Eq. 6.10, computes to $\varepsilon_{\text{cool}} \leq 8.8$ (880%).

43: Note that $\Delta T = 30$ in either K or °C.

Example 6.5.2 What is the limit to efficiency of keeping a home interior at 20°C when it is −10°C outside?

First, we express the temperatures in Kelvin: $T_h \approx 293$ K and $\Delta T = 30$ K.[44] The maximum efficiency, by Eq. 6.11, computes to $\varepsilon_{\text{heat}} \leq 9.8$ (980%).

44: Note that $\Delta T = 30$ in either K or °C.

Box 6.4: Is >100% Really Possible?

At first, it seems to be spooky and impossible that efficiencies can be greater than 100%. Example 6.5.1 essentially says that as many as 8.8 J of thermal energy can be *moved* for a mere 1 J *input* of work! The situation bears analogy to the martial art of Jiu Jitsu, whereby the opponent's momentum is used to their detriment, requiring little work to direct its flow. In this case, we convince a bundle of thermal energy sitting in the freezer to move outside where it is hotter (uphill; against natural flow) and in the process use less energy than the amount of thermal energy residing in the bundle.

The fact that our "efficiency" metrics come out to be greater than 100% is an illusion: an artifact of how we defined $\varepsilon_{\text{cool}}$ and $\varepsilon_{\text{heat}}$. Conservation of energy is not violated; we're just putting the small piece (ΔW) in the denominator to form the efficiency metric.[45] In this sense, it's not the usual sort of efficiency measure, which puts the *largest* quantity (total budget) in the denominator.

Maybe the situation can be compared more understandably to money transfers, where one might pay a $20 fee to wire $1,000 from account A to account B. It doesn't mean that $1,000 was created out of $20—just that $20 was spent (like ΔW) to move a much larger sum into account B. But if account A belonged to somebody else, it would seem like you just turned $20 of your own money into $1,000 at a gain of 5,000%, even though it really came from elsewhere.

45: Following the example numbers in Figure 6.5, we would say that $\varepsilon_{\text{cool}}$, defined as $\Delta Q_c/\Delta W$, is 2.0, and $\varepsilon_{\text{heat}}$ is 3.0.

In the case of heating, it is worth comparing the output of a heat pump to the application of direct heat. Let's revisit the scenarios explored in Section 6.3.

Example 6.5.3 If a house's thermal performance is 150 W/°C and we want to maintain 20°C inside while the outside temperature is a frigid −20°C, we would need to supply 6,000 W of energy[46] to the home in the form of burning fuel (natural gas, propane, firewood) or electricity for direct-heating application.[47]

But according to Eq. 6.11, a heat pump could theoretically move 6,000 W of thermal energy by only applying 820 W without violating the Second Law, since $\varepsilon_{\text{heat}} \leq 293/40 = 7.3$ and 6,000 J (ΔQ_h) divided by 7.3 (to get ΔW) is 820 J.[48]

46: 150 W/°C times 40°C.

47: . . . e.g., four space heaters each expending 1,500 W

48: We are solving for $\Delta W = \Delta Q_h/\varepsilon_{\text{heat}}$, and consider the energy moved in one second in order to go from W to J.

Engineering realities will prevent operating right up to the thermodynamic limit, but we might at least expect to be able to accomplish the 6,000 W goal of Example 6.5.3 for under 2,000 W. Thus the heat pump has shaved a factor of three (or more) off the energy required to provide heat inside. Heat pumps are very special.

As Eq. 6.10 and Eq. 6.11 imply, heat pumps are most efficient when ΔT is small. Thus a refrigerator in a hot garage must not only work harder to maintain a large ΔT, it does so less efficiently—making it a double-whammy. For home heating, heat pumps offer the most gain in milder climates where ΔT is not so brutal.

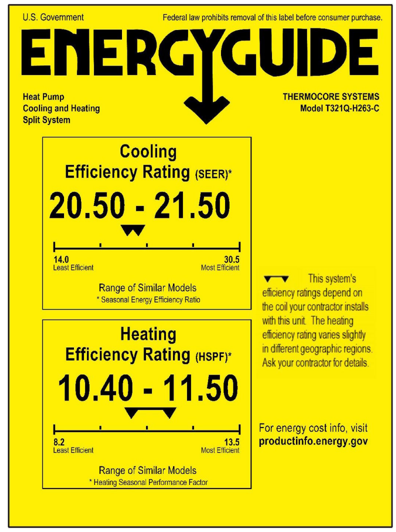


Figure 6.6: Typical heat pump energy label in the U.S., showing an EER around 21 and a HSPF around 11. From U.S. DoE.

6.5.1 Consumer Metrics: COP, EER, HSPF

When shopping for heat pumps or air conditioners (or freezers/refrigerators), products are specified by the coefficient of performance (COP) or energy efficiency ratio (EER) or heating seasonal performance factor (HSPF), as in Figure 6.6. How do these relate to our $\varepsilon_{\text{heat}}$ and $\varepsilon_{\text{cool}}$ values? The first one is easy.

Definition 6.5.3 ***COP***: *Heat pumps used for heating are specified by a coefficient of performance (COP), which turns out to be familiar already:*

$$\text{COP} = \varepsilon_{\text{heat}}. \tag{6.12}$$

Example 6.5.4 COP Example: Using the red numbers in Figure 6.5, we can compute $\varepsilon_{\text{heat}}$, the COP, and then determine the minimum T_c theoretically permissible (resulting in maximum possible efficiency) if $T_h = 300$ K.[49]

We go back to the original definition of $\varepsilon_{\text{heat}}$ as $\Delta Q_h/\Delta W$, which for our numbers works out to 30/10, or 3.0 The COP is then simply 3.0.

Setting $\varepsilon_{\text{heat,max}} = T_h/\Delta T$ equal to 3.0, we find that ΔT is 100 K, so that the minimum permissible $T_c = 200$ K in this case.

49: This corresponds to maintaining the hotter environment at 27°C, for instance in the context of heating a house.

The EER is different, and perhaps a little odd. EER is defined as the amount of heat moved (ΔQ_c), in Btu, per work input (ΔW), in watt-hours (Wh). What?! Sometimes the world is just loopy. But we can manage this. If handed an EER (Btu/Wh), we can convert it to our same/same numerator/denominator units by converting both numerator and denominator to the same units. We could convert Btu to Wh in the numerator and be done, or convert Wh to Btu in the denominator and be done, or we could convert both numerator *and* denominator to Joules[50] to get there. For illustrative purposes, we'll pick the last approach. To get from Btu to Joules, we multiply (the numerator) by 1,055. To get from Wh to Joules, we multiply the denominator (or divide the EER construct) by 3,600.[51] The net effect is highlighted in the following definition.

50: ...or any other energy unit of choice

51: 1 watt-hour (Wh) is 1 J/s times 3,600 s.

Definition 6.5.4 ***EER***: *Heat pumps used for cooling are specified by the energy efficiency ratio (EER), which modifies Eq. 6.10 as follows.*

$$\varepsilon_{\text{cool}} = \text{EER}\left(\frac{\text{Btu}}{\text{Wh}}\right)\frac{1055\,\text{J/Btu}}{3600\,\text{J/Wh}} = \text{EER}\cdot 0.293, \qquad (6.13)$$

or the converse

$$\text{EER} = \frac{\varepsilon_{\text{cool}}}{0.293} \approx 3.41 \times \varepsilon_{\text{cool}}. \qquad (6.14)$$

Example 6.5.5 EER Example: Using the red numbers in Figure 6.5, we can compute $\varepsilon_{\text{cool}}$, the EER, and then determine the maximum T_{h} theoretically permissible (resulting in maximum possible efficiency) given a target T_{c} of 260 K, as we might find in a freezer.

We go back to the original definition of $\varepsilon_{\text{cool}}$ as $\Delta Q_{\text{c}}/\Delta W$, which for our numbers works out to 20/10, or 2.0. The EER is then 3.41 times this amount, or 6.8.

Setting $\varepsilon_{\text{cool,max}} = T_{\text{c}}/\Delta T$ equal to 2.0, we find that ΔT is 130 K, so that the maximum permissible $T_{\text{h}} = 390$ K in this case.

Because the theoretical maximum efficiency depends on ΔT—according to Eq. 6.10 and Eq. 6.11—and therefore can fluctuate as outdoor temperatures change, a seasonal average is often employed, called the SEER (seasonal EER). In a similar vein, the HSPF measures the same thing as the COP, but in units of EER[52] and averaged over the heating season.

52: . . . Btu/Wh

Definition 6.5.5 ***HSPF***: *Heat pumps used for heating are sometimes specified by the heating seasonal performance factor (HSPF), which modifies Eq. 6.11 as follows.*

$$\varepsilon_{\text{heat}} = \text{HSPF}\left(\frac{\text{Btu}}{\text{Wh}}\right)\frac{1055\,\text{J/Btu}}{3600\,\text{J/Wh}} = \text{HSPF}\cdot 0.293, \qquad (6.15)$$

or the converse

$$\text{HSPF} = \frac{\varepsilon_{\text{heat}}}{0.293} \approx 3.41 \times \varepsilon_{\text{heat}} = 3.41 \times \text{COP}. \qquad (6.16)$$

Example 6.5.6 HSPF Example: Using the red numbers in Figure 6.5, we can compute $\varepsilon_{\text{heat}}$ and the HSPF.

We go back to the original definition of $\varepsilon_{\text{heat}}$ as $\Delta Q_{\text{h}}/\Delta W$, which for our numbers works out to 30/10, or 3.0. The COP is then 3.0, and the HSPF is 3.41 times this, or 10.2.

Typical COP values for heat pumps range from about 2.5–4.5.[53] This means an energy savings by a factor of 2.5 to 4.5 for heating a house via heat pump vs. direct electrical heating. Quite a bargain. An air conditioner EER rating is typically in the range 10–20, corresponding to 3–6 in terms of $\varepsilon_{\text{cool}}$—similar to the range for heat pumps in heating

53: . . . mapping to HSPF from ~8–15

mode.[54] Houses equipped with electric heat pumps can typically be run for both cooling and heating applications, making them a versatile and efficient solution for moving thermal energy in or out of a house.

54: EER and HSPF numbers are "inflated" by a factor of $1/0.293 \approx 3.41$ compared to COP due to the unfortunate choice of units for EER and HSPF.

Heat pumps leveraging the moderate-temperature ground just below the surface as the external thermal bath are called "geothermal" heat pumps, but have nothing to do with geothermal energy (as a source). Compared to heat pumps accessing more extreme outside air temperatures, geothermal heat pumps benefit from a smaller ΔT, and thus operate at higher efficiency.

6.6 Upshot on Thermal Energy

Sometimes we just want heat. Cooking, home heating, and materials processing all need direct heat. Burning fossil fuels, firewood, biofuels, extracting geothermal energy, or simply letting the sun warm our houses all directly utilize thermal energy. Specific heat capacity tells us how much thermal energy is needed to change something's temperature, using 1,000 J/kg/°C as a rough guess if lacking more specific information.[55] We also saw how to estimate home heating demand using a metric of heat loss rate, such as 200 W/°C.

55: Or we frequently use water's value at 4,184 J/kg/°C, connected to the definition of a kcal.

But it turns out that we use heat for much more than this. 84% of our electricity is produced by heat engines, using heat flow to drive a turbine to turn a generator. The maximum efficiency a heat engine can achieve is set by limits on entropy and amounts to $\varepsilon_{\mathrm{max}} < \Delta T/T_{\mathrm{h}}$, although in practice we tend to be a factor of two or more short of the thermodynamic limit.[56] In any case, thermal energy plays a giant role in how we run our society.

56: Typical efficiencies are 20% for cars and 35% for power plants—compared to 60% theoretical.

Heat pumps are like heat engines in reverse: driving a flow of thermal energy against the natural hot-to-cold direction by putting in work. Any refrigeration or cooling system is likely to use this approach.[57] Because heat pumps only need to *move* thermal energy, each Joule they move can require a small fraction of a Joule to accomplish, making them extremely clever and efficient devices.

57: A notable exception is evaporative cooling.

6.7 Problems

1. How many Joules does it take to heat your body up by 1°C if your (water-dominated) mass has a specific heat capacity of 3,500 J/kg/°C?
2. How long will it take a space heater to heat the air[58] in an empty room by 10°C if the room has a floor area of 10 m^2 and a height of 2.5 m and the space heater is rated at 1,500 W? Air has a density[59]

58: We only consider the air for this problem, and ignore other objects—including walls and furniture—that would add substantially to the time required in real life.

59: Use density to get at the mass of air.

of 1.25 kg/m^3. Express your answer as an approximate number in minutes.

3. When you put clothes on in the morning in a cool house at 15°C, you warm them up to something intermediate between your skin temperature (35°C) and the ambient environment.[60] If your clothes have a mass of 2 kg, how much energy must be deposited into the clothes? If you are emitting power at 100 W, how long will this take?

60: The inside surface of the clothing will be near skin temperature, and the outside will be near ambient temperature.

4. You score this massive 1 kg burrito but decide to put it in the refrigerator to eat later. It comes out at 5°C, and you want to heat it in the microwave up to 75°C before eating it. If the microwave puts energy into the burrito at a rate of 700 W.[61] How long should you run the microwave for a high-water-content burrito having an effective specific heat capacity of 3,000 J/kg/°C?

61: ⓘ Note that a microwave oven might be rated for 1,500 W, but not all the energy ends up in the burrito, so we pick 700 W to be realistic.

5. Let's say you come home from a winter vacation to find your house at 5°C and you want to heat it to 20°C. Let's say the house contains: 500 kg of air;[62] 1,000 kg of furniture, books, and other possessions; plus walls and ceiling and floor that amount to 6,000 kg of effective[63] mass. Using the catch-all specific heat capacity for all of this stuff, how much energy will it take, and how long to heat it up at a rate of 10 kW? Express in useful, intuitive units, and feel free to round, since it's an estimate, anyway.

62: . . . appropriate for a 150 m^2 footprint

63: ⓘ We only count half-thickness of exterior walls since they are not heated to the interior temperature all the way to the outside.

6. In a house achieving a heat loss rate of 200 W/°C equipped only with two 1,500 W space heaters, what is the coldest it can get outside if the house is to maintain an internal temperature of 20°C?

7. In a house achieving a heat loss rate of 200 W/°C equipped a 5,000 W heater, what will the internal temperature be if the outside temperature is −10°C and the heater is running 100% of the time?

8. In a super-tight house achieving 100 W/°C equipped with a 5,000 W heater, what percentage of the time will the heater need to run in order to keep the internal temperature at 20°C if the temperature outside is at the freezing point?[64]

64: Hint: compute the average power that would be needed in this case.

9. How much will it cost per day to keep a house at 20°C inside when the external temperature is steady at −5°C using direct electric heating[65] if the house is rated at 150 W/°C and electricity costs $0.15/kWh?

65: . . . no heat pump: just straight energy deposition at 100% efficiency

10. Provide at least one example not listed in the text in which heat flows into some other form of energy.[66] In the text, we mentioned hot air over a car, wind, internal combustion, and a steam turbine plant.

66: Think about motion deriving from or caused by heat or thermal release.

11. What is the only form of significant electricity production in the

U.S. that does not involve a spinning shaft?

12. If a can of soda (350 mL; treat as water) cools from 20°C to 0°C, how much energy is extracted, and how much is the entropy (in J/K) in the can reduced using the average temperature and the relation that $\Delta Q = T\Delta S$?

13. What would the maximum thermodynamic efficiency be of some heat engine operating between your skin temperature and the ambient environment 20°C cooler than your skin?

14. We can think of wind in the atmosphere as a giant heat engine[67] operating between the 288 K surface and the top of the troposphere[68] at 230 K. What is the maximum efficiency this heat engine could achieve in converting solar heating into airflow?

15. Since the sun drives energy processes on Earth, we could explore the maximum possible thermodynamic efficiency of a process operating between the surface temperature of the sun (5,800 K) and Earth's surface temperature (288 K). What is this maximum efficiency?[69]

16. A heat engine pulls 100 J out of a hot bath at 800 K, and transfers 80 J of heat into the cold bath at 300 K. What efficiency does this heat engine achieve in producing useful work, and how does it compare to the theoretical maximum?

17. Human efficiency[70] is in the neighborhood of 25%, meaning that in order to do 100 J of external work, we need to eat 400 J of energy content. To investigate whether human energy is working as a heat engine, figure out what the cold temperature, T_c, would have to be to achieve this efficiency, thermodynamically.[71] Do you conclude that our biochemistry operates as a heat engine, or no?[72]

18. A 350 mL can of soda[73] at 20°C is placed into a refrigerator having an EER rating of 10.0. How much energy will you have to spend (ΔW) to remove the thermal energy from the soda and bring it to a frosty 0°C?

19. If a refrigerator works at *half* of its theoretical ε_{cool} limit, how much more energy does it take to maintain an internal temperature of 0°C in a 40°C garage vs. a 20°C house interior? Two things are going on here: even at the same efficiency, the cooling energy scales as ΔT, but the efficiency also changes for a double-whammy.

20. Changing from direct electrical heating to a heat pump operating with a COP of 3 means spending one-third the energy for a certain thermal benefit. If a house averages 30 kWh/day in heating cost through the year using direct electrical heating at a cost of $0.15/kWh, how long will it take to recuperate a $5,000 installation cost of a new heat pump?

67: And it really is!

68: ⓘ Atmospheric wind and weather are confined to the lowest layer of the atmosphere, called the troposphere, extending to about 12 km high.

69: We would not expect any solar-derived process to exceed this limit in the Earth environment.

70: . . . in terms of converting food energy into useful work

71: ⓘ The hot temperature, T_h, would be internal body temperature of 37°C.

72: Hint: do our bodies have regular access to temperatures this cold?

73: Treat as water, and recall that the density of water is one gram per milliliter.

7 The Energy Landscape

7.1 The Annual Energy Review 102
Energy Flow 103
Source and Sector 104
Detailed Mix 106
Energy Trends 107
7.2 Global Energy 108
U.S. Global Share 110
7.3 Upshot: Go to the Source . . 111
7.4 Problems 111

Now that we have a handle on common energy units and thermal processes, we can take a look at various sources of energy data and make sense of the information, allowing meaningful cross-comparisons. In this chapter, we will do exactly that, gaining in the process a perspective on the past and present roles different energy sources play at a national and global level.

Most of the information in this chapter comes from the U.S. Energy Information Administration's (EIA's) *Annual Energy Review* [34], and from a compilation of global data owing to Vaclav Smil and the British Petroleum *Statistical Review of World Energy* [16]. Rather than laboriously citing each instance, it is sufficient to assume for this chapter that numbers for the U.S. come from the former and global numbers come from the latter, unless stated otherwise.

7.1 The Annual Energy Review

Until 2011, the Annual Energy Review (AER) was compiled for the U.S. as an annual report. Since then, a web interface provides access to many of the same products, but not as a single document. An impressive amount of detail is available in the AER products, and we will only scratch the surface in this book, looking at high-level overviews. Later chapters will sometimes rely on deeper information to provide state-by-state use of hydroelectric, solar, wind, etc. But for now, we stick mostly to section 1 of the Annual Energy Review (AER), labeled *Energy Overview*.

An oil pump dominates the foreground, while wind makes a visible presence in the background. Photo credit: Tom Murphy

7.1.1 Energy Flow

Section 1.0 of the AER is a one-page PDF graphic that conveys at a glance the flow of energy into and out of the U.S. Figure 7.1 shows the 2018 version.

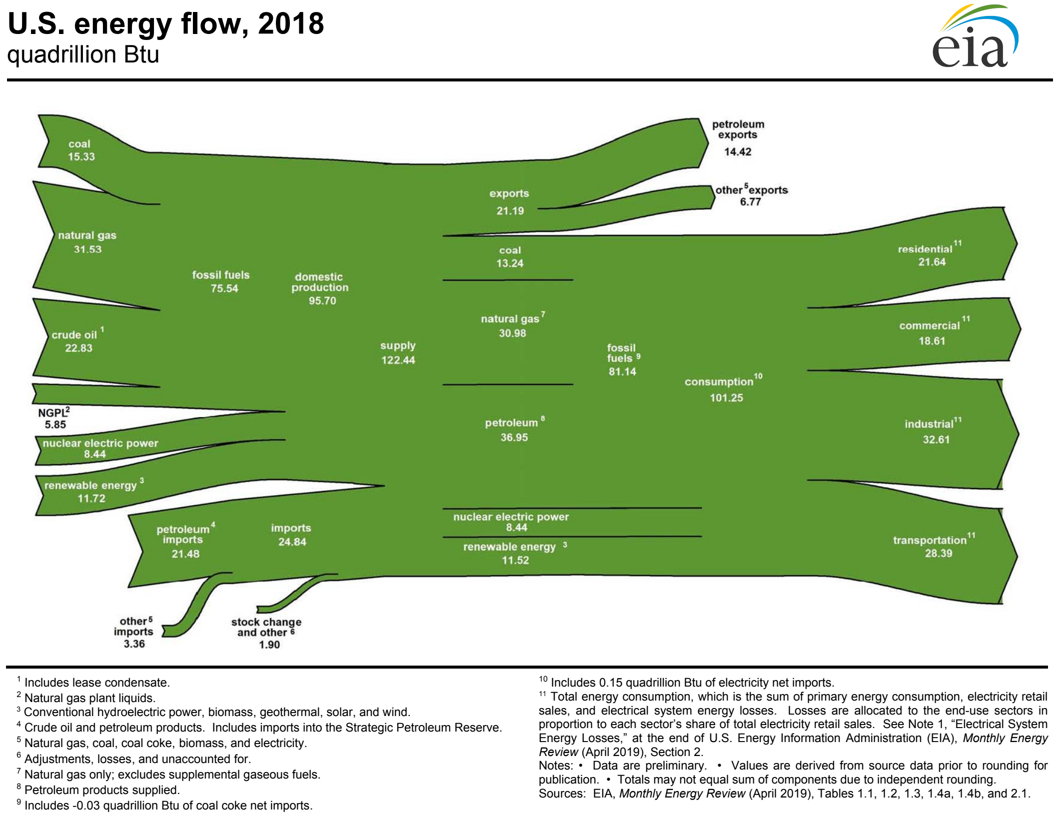


Figure 7.1: The flow of energy in the U.S. for 2018, as presented in [34]. Units are quadrillions of Btu (qBtu), unfortunately. From U.S. EIA.

From past experience, many students dislike this graphic. Firstly, it's a product of the EIA, and not a creation of this book. Secondly, it is actually not so bad, once you get the hang of it.

Resources come in from the left. Expenditures or exports go off to the right. The format guarantees that all inputs must match all outputs.[1] We also see at a glance the big players vs. small players.

To understand, let's start in the middle section. To the left of center, we see that the total supply sums to 122.44 qBtu. Of this, we consume 101.25 qBtu (right of center) and export the remaining 21.19 qBtu. Now we focus on the central column to get a powerful visual and quantitative snapshot of how our energy is partitioned.[2] From this, we see that 13% is coal, 31% is natural gas, 36.5% is petroleum (oil), 8% is nuclear energy, and 11.5% is renewable energy.[3]

1: That is, no significant amount of energy is stored or drawn from a stockpile.

2: By luck, total consumption is very nearly 100 qBtu, so the amount of each source in qBtu is already approximately a percentage!

See how they add up?

3: Think of the three forms of fossil fuels as solid (coal), liquid (petroleum/oil) and gas (natural gas; not the same as liquid gasoline, which is a petroleum product).

Now the right-hand side shows the sectors into which the energy flows, finding *roughly* equal distribution between residential (homes), commercial (businesses), industrial (manufacture), and transportation (both personal and commercial/shipping). In this graphic, we lose entirely any sense for how much of each energy source contributes to each sector,[4] but that is coming in the next section.

4: For instance, we cannot tell how much coal is used in the industrial sector.

Finally, the left-hand side indicates the inputs, grouped as domestic fossil fuel supply at top (out of our own ground), nuclear energy and renewable in the middle, and imports at bottom. From this, we can learn that we export some coal,[5] that almost all of our natural gas and 100% of our nuclear is domestic, and that 62% of our petroleum comes from domestic crude oil production.

5: For instance, the supply of 15.33 qBtu that is mined is larger than the 13.24 qBtu we consume.

Other insights are present in the graphic as well. Don't be afraid to subtract or divide numbers to aid new discoveries.

Ask yourself what else you can learn from the numbers!

Box 7.1: 100 quads? So what?

To put the scale into a bit of perspective, 100 qBtu in a year for the U.S. is about 10^{20} J in a year. A year is 3.156×10^7 seconds long,[6] meaning that the U.S. power budget is just over 3 TW (3×10^{12} W). Distributed among a little over 300 million people, the average contribution per person is about 10,000 W.[7] That's a lot. As we have seen in Sec. 5.5 (p. 73), human metabolism is about 100 W. So Americans have approximately 100 times as much energy available as their personal metabolism. The situation has been compared to each person *having 100 energy servants*! No wonder we live better than royalty of ages past. Even though the U.S. uses about 4.5 times the global energy per capita (about 20% of the world's energy and 5% of population[8]), the average citizen of Earth still has over 20 energy servants available, on average, thanks almost entirely to fossil fuels. They have been an unqualified game changer.

6: Neat trick: roughly $\pi \times 10^7$ seconds in a year.

7: This showed up in Table 3.4 (p. 43) and also in Box 5.4 (p. 75).

8: American energy usage is much higher than average because of consumerism, diet, comfort standards, prevalence of detached housing, and transportation.

7.1.2 Source and Sector

Figure 7.2 provides a more detailed breakdown of how energy flows from source to usage sectors.[9] In other words out of the 101.25 qBtu consumed in 2018, we see how much comes from each source, and within each source can track how much goes to each end-use category. For example, we learn that 91% of coal and 100% of nuclear go to electricity, and that 92% of transportation is based on petroleum.

9: Notice that the source and end-use numbers in the boxes match the numbers in Figure 7.1 within rounding error.

Notice the black and gray block at lower center, representing electricity. We derive electricity from all the sources on the left, and electricity is consumed in all sectors. Also, of the 38.3 qBtu going *into* making electricity, only 13.0 qBtu (34%) makes it out the door as electricity, due to thermodynamic losses that were covered in Chapter 6.

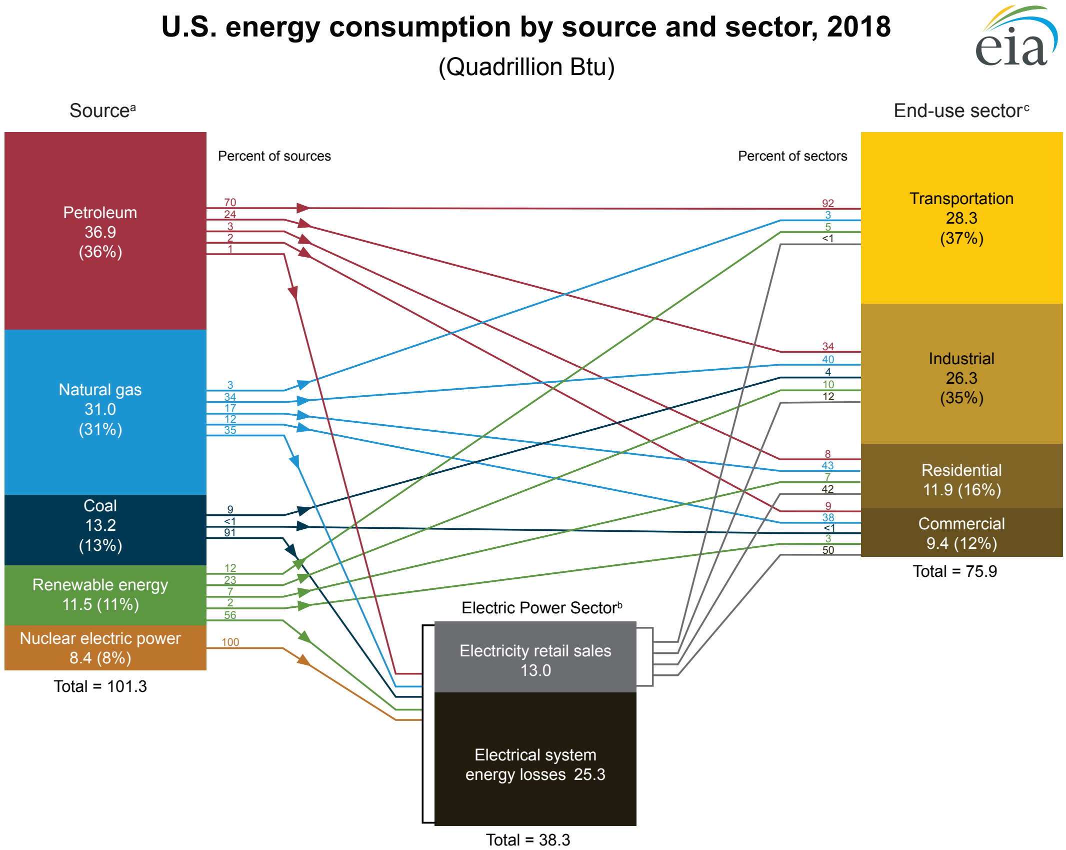


Figure 7.2: Tracking of energy sources and end-use in the U.S. for 2018, from section 2.0 of the AER. Small numbers beside the blocks represent percentages. Numbers that are not percentages are qBtu (quads). From U.S. EIA.

In principle, it is possible (and would be nice) to put percentages where the arrows enter and exit the electricity sector, but enough numbers are present to work this out, as Example 7.1.1 demonstrates. Without these numbers, the story is a little misleading. For instance, only 17% of natural gas goes *directly* to residential use, but some natural gas produces electricity, which then flows to residences. It is therefore not immediately obvious what percentage of residential energy ultimately comes from natural gas, but it's more than the 43% indicated in the figure.

A similar graphic combining some elements of both Figure 7.1 and Figure 7.2 is provided by Lawrence Livermore National Lab [35].[10]

[35]: LLNL (2019), *Energy Flow Charts*

10: See also [36] for a fascinating animated version.

Example 7.1.1 Let's work through the numbers in Figure 7.2 to elucidate what percentage of residential energy ultimately derives from natural gas. The same technique can be pursued to ask similar questions about any source-to-sector pathway, by incorporating the electricity contribution.

We start simply, by noting that 43% of the 11.9 qBtu residential energy budget comes directly from natural gas. So that's 5.1 qBtu.[11]

Now, 35% of natural gas goes toward electricity, which we can compute to be 10.9 qBtu.[12]

So of the 38.3 qBtu total energy coming into the electricity block, 10.9 qBtu (28%) is from natural gas.[13]

Assuming the 34% efficiency[14] of electricity production applies equally across all sources (close to the truth), we can say that 28% of the electricity output comes from gas: 28% of 13.0 qBtu (electricity output) is 3.7 qBtu.

But not all of this goes into homes. The home gets 42% of its 11.9 qBtu from electricity, or 5.0 qBtu. We can assume that 28% of the 5 qBtu of electricity flowing into the home derives from natural gas, as decided above. So that's 1.4 qBtu of gas-derived electricity flowing into the home.

We can add this 1.4 qBtu of gas-derived electricity to the 5.2 qBtu[15] of direct gas-to-home to learn that 6.6 qBtu of residential input is sourced from natural gas—either directly or via electricity. Compared to the 11.9 qBtu total for residences, natural gas therefore contributes 55% of the energy used in homes, not just 43% as listed. Now we know.

11: As a check, we note that the other side of the blue arrow has 17% of 31.0 qBtu, or 5.3 qBtu leaves the gas block for homes: close enough to the 5.1 qBtu we got on the other side (essentially the same, to rounding error).

12: 35% of 31.0 qBtu

13: Wouldn't it be nice if Figure 7.2 printed a blue 28 where the blue arrow comes into the electricity block?

14: 13.0 qBtu of electricity is produced from 38.3 qBtu energy input.

15: . . . averaging the two estimates from before

7.1.3 Detailed Mix

Delving a bit further into the AER, Section 1.3 provides a more detailed breakdown of consumption, now separating out the "renewable" category into its constituent parts, as seen in Table 7.1 and Figure 7.3.

In sum, 80% of the U.S. energy in 2018 came from fossil fuels. Less than 2.5% came from wind, and less than 1% was solar in origin—the other 16% mainly in the form of nuclear, biomass, and hydroelectricity. Most of the renewable energy is from biomass—like burning wood. The wider world is pretty similar, in that about 80% of energy is from fossil fuels. It's still our main squeeze. Table 7.2 breaks out electricity sources separately.

Table 7.1: U.S. energy consumption for 2018 in thermal equivalent terms.

Resource	qBtu
Petroleum	36.88
Natural Gas	31.09
Coal	13.25
Nuclear	8.44
Biomass	4.98
Hydroelectric	2.77
Wind	2.48
Solar	0.92
Geothermal	0.21
Total	101.0

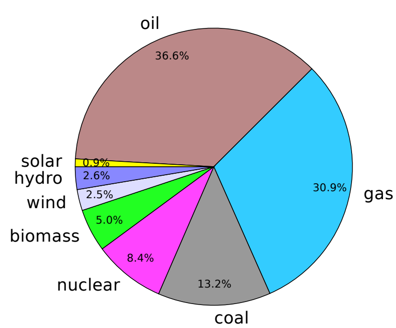


Figure 7.3: 2018 Energy sources for the U.S. Figure 7.6 shows the global distribution.

Box 7.2: Thermal Equivalent

Note that the EIA—and thus Table 7.1—habitually applies a thermal conversion factor to some energy sources in order to more meaningfully compare one source to another. Fossil fuel energy is characterized by its *thermal* content, which makes sense as they are burned for thermal energy. Often—but not always—the thermal energy is turned into electrical energy. Meanwhile, some sources, like solar, hydroelectric, wind, nuclear, and geothermal are almost exclusively used for electricity production and are most easily measured by electrical *output*, not thermal input (which is meaningless for solar, wind, and hydro).

Multiplying the electrical output by a factor of about 3 recovers the thermal equivalent.[16] The interpretation is: how much fossil fuel (thermally) would have been necessary to achieve the same result? As a consequence, when Table 7.1 says the solar contribution is 0.92 qBtu, and therefore about 1% of the total, the *actual* solar energy was smaller by a factor of three, but the practice is fair because now we can directly compare solar to the fossil fuels. Reporting electrical output alongside thermal inputs would make the renewables appear to have a smaller contribution than they effectively do, against fossil fuels.

16: The actual factor is just the inverse of the electrical conversion efficiency discussed above (34%, so 1/0.34). The conversion efficiency adopted by the EIA has slowly increased over time, and is tracked in Appendix A6 of the AER—now at 37.5%, leading to a conversion factor of 2.67.

Region	Coal	Gas	Oil	Nuclear	Hydro	Wind	Solar	Bio	Geo
U.S.	27.3	34.9	0.6	19.2	7.0	6.5	2.3	1.5	0.4
World	38.0	23.0	2.9	10.1	16.2	4.8	2.1	2.4	0.5

Table 7.2: Percentages of **electricity** derived from various sources in the U.S. and globally in 2018. Bio includes burning wood and waste, and Geo means geothermal. Data are from Table 7.2a of [34] and from [37].

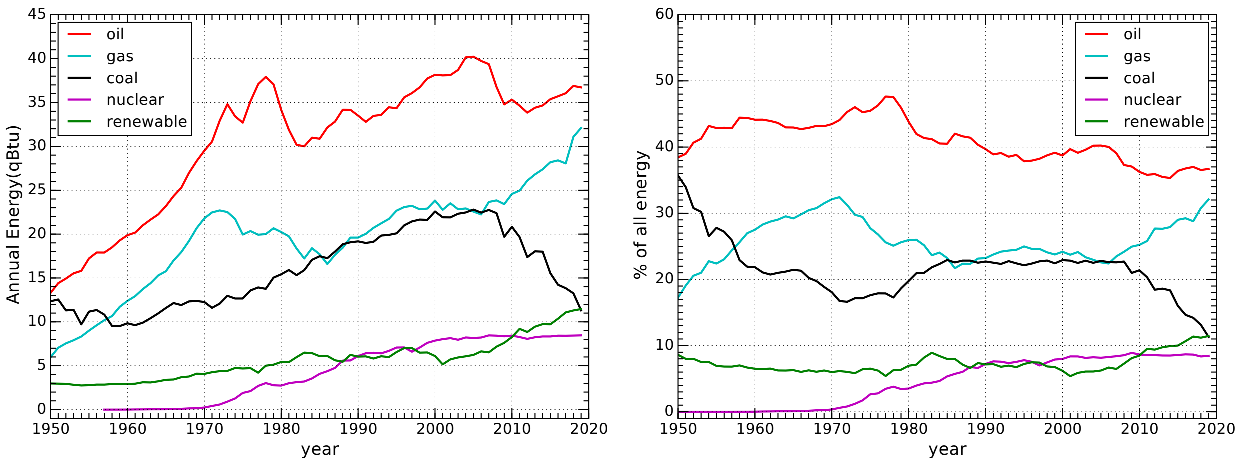


Figure 7.4: Recent history of primary energy consumption in the U.S. The three fossil fuels and nuclear are shown separately, and then all renewable sources are grouped together. Note that at the end of the plot, coal has sunk into a tie with renewable resources. The plot on the right shows percentages of total energy. Most of the lines are fairly flat, although in recent years the main story is gas replacing coal.

7.1.4 Energy Trends

It is worth looking at trends to understand not only the state of affairs today, but what happened over past decades and trends that may carry into the near future. Section 1.3 of the AER includes data going back to 1950 on the categories in Table 7.1.

Figure 7.4 shows the trends for the fossil fuels over the last 70 years, along with the slow rise of the sub-dominant non-fossil sources. Recent news touted the fact that the renewable sources[17] surpassed coal as an energy source in the U.S. Indeed, the lines basically meet on the

17: The term "renewable," will be more fully explained in Chapter 10.

right-hand side of the plot, and the trends suggest a clear reversal of rank going forward. Note, however, that this result is largely due to natural gas replacing coal at electrical power plants. The sharp rise in natural gas nearly mirrors the decline in coal, while the rise in renewable resources is more modest. So this is really more a story of trading gas for coal than renewables replacing coal. Figure 7.4 also shows each source as a percentage of all energy. For a few decades (1980–2010), coal and gas were essentially tied, while oil sat at almost double these two. Lately, gas is approaching oil while coal plummets.

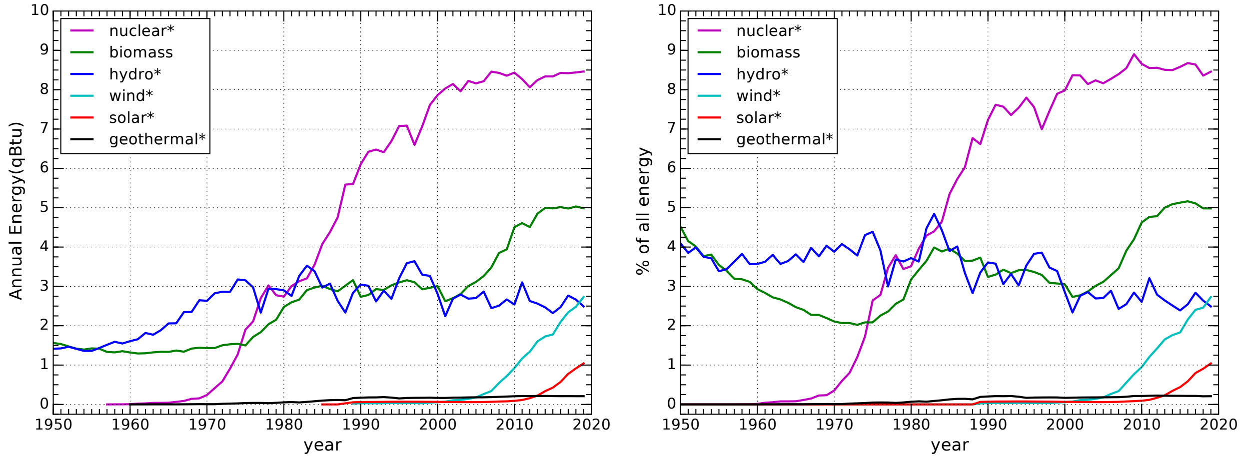


Figure 7.5: Recent history of non-fossil energy consumption in the U.S. Nuclear, hydroelectric, and biomass have dominated, while wind and solar are rising to join as players. Asterisks indicate thermal equivalents, as described in Box 7.2. The same data are plotted at right as a percentage of total energy. Aside from the rapid rise of nuclear in the middle years of the plot, the recent entry of wind and solar (though still only a few percent) are the most interesting developments.

The non-fossil consumption in Figure 7.5 clarifies the breakdown of the "renewables" curve in Figure 7.4, alongside nuclear. From this, we see that nuclear dominates non-fossil energy, rising quickly from 1970 to 2000 and holding steady since then. Hydroelectric has been pretty stable over the last 50 years as other sources surpass it and lower its rank. The surge in biofuels around 1980 appears to be largely driven by increased burning of wood, while the next surge (2000–2010) was due to biofuels—mostly ethanol. Wind is approaching a 3% contribution to our total ~100 qBtu consumption budget, edging up about 0.2% per year. Solar is also on the move, reaching the 1% level recently and rising more slowly than wind. Geothermal is and will continue to be a paltry contributor.

7.2 Global Energy

Not surprisingly, the global story is not dramatically different from the story in the U.S., as Figure 7.6 and Figure 7.7 show. Fossil fuels dominate, with oil at the top. Coal has held a lead over natural gas in the wider world, unlike the U.S. Also, while nuclear and renewables are

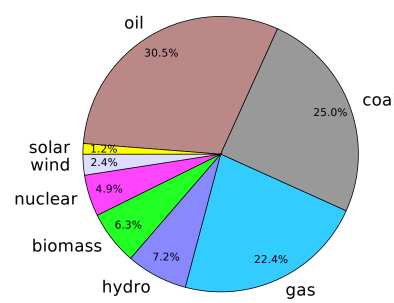


Figure 7.6: 2018 Energy sources for the world. Figure 7.3 shows the U.S.

comparable in the U.S., this is not true globally, for reasons discussed shortly. Note that different assessments of global energy may report different percentage contributions depending on whether or not thermal equivalents are used (see Box 7.2).

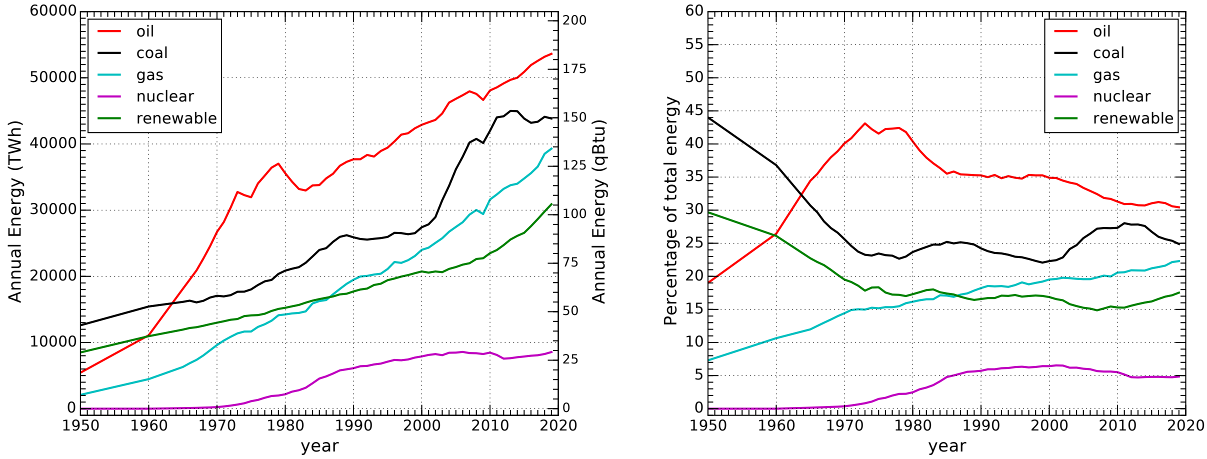


Figure 7.7: Recent history of primary energy consumption in the world. The three fossil fuels and nuclear are shown separately, while renewable sources are grouped together. The plot on the right shows the same data as a percentage of the whole.

For non-fossil contributions, Figure 7.8 shows the evolution of recent decades. Here, we see that a large part of the reason why renewables exceed nuclear energy globally is because of biomass. This makes sense, as countries having a lower standard of living are more likely to burn wood and less likely to have nuclear power.

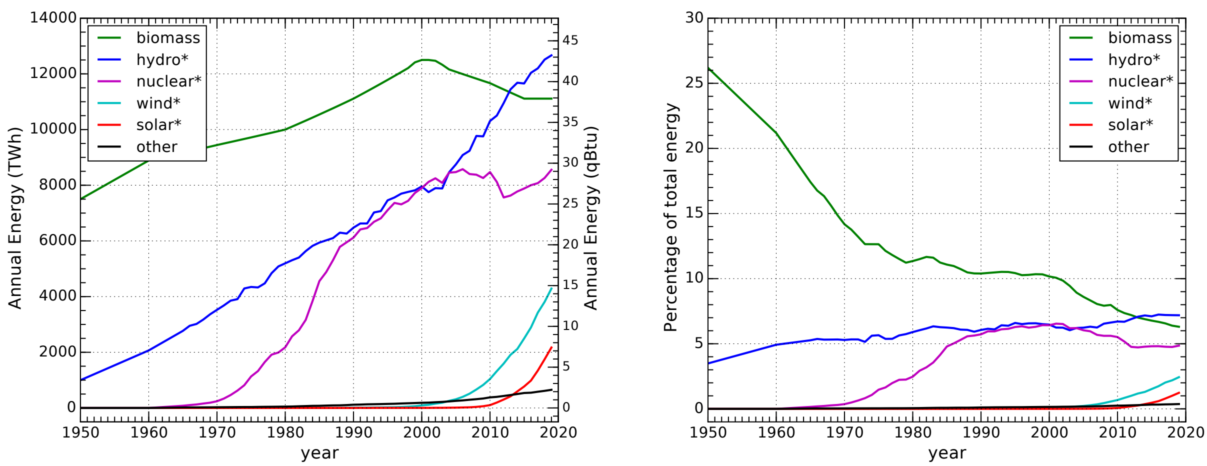


Figure 7.8: Recent history of non-fossil global energy consumption. Asterisks indicate thermal equivalents, as described in Box 7.2. The plot at right shows each source as a percentage of the total energy. Biomass accounted for a quarter of global energy in 1950.

Box 7.3: TWh vs. qBtu

You may have noticed that as soon as we departed from the AER data, which expressed energy in qBtu, the units on the plot (Fig. 7.7) changed to terawatt-hours (TWh). It means what it sounds like: tera is 10^{12}, so this is 10^{12} watt-hours (Wh). We use kWh more often than Wh, so a TWh is the same as a giga-kWh, or GkWh (can you do that?). One kWh is 3.6×10^6 J, so 1 TWh is 3.6×10^{15} J. Meanwhile, 1 qBtu is 1.055×10^{18} J, facilitating a conversion. The figures for global power also put qBtu on the right side for easier comparison between plots.

The source of numbers for this section [16] mix thermal and electrical output, so the plots have multiplied some entries (asterisks in plot legends) by 3.06 for reasons described in Box 7.2.

[16]: Smil (2017), *Energy Transitions: Global and National Perspectives*

7.2.1 U.S. Global Share

A final overview to help frame a number of discussions in this textbook looks at the U.S. share of consumption of various energy resources compared to the global total. The evolution seen on the left side of Figure 7.9 contains a crucial insight into geopolitics. In 1950, the U.S. used an astounding 84% of global natural gas and 72% of petroleum. At only 6% of the world's population at the time, Americans used more than ten times the global average oil and gas, and substantially more than the rest of the world combined. Since energy per year is the definition of power, we can understand how the U.S. was a literal *superpower* during this era. Parroting Bill Clinton: It's the resources, stupid.

This may be a factor in nostalgia for what some Americans see as the "glory days" of the 1950s. To the extent that U.S. energy share played a role, longing for a return to that era is not likely to materialize.

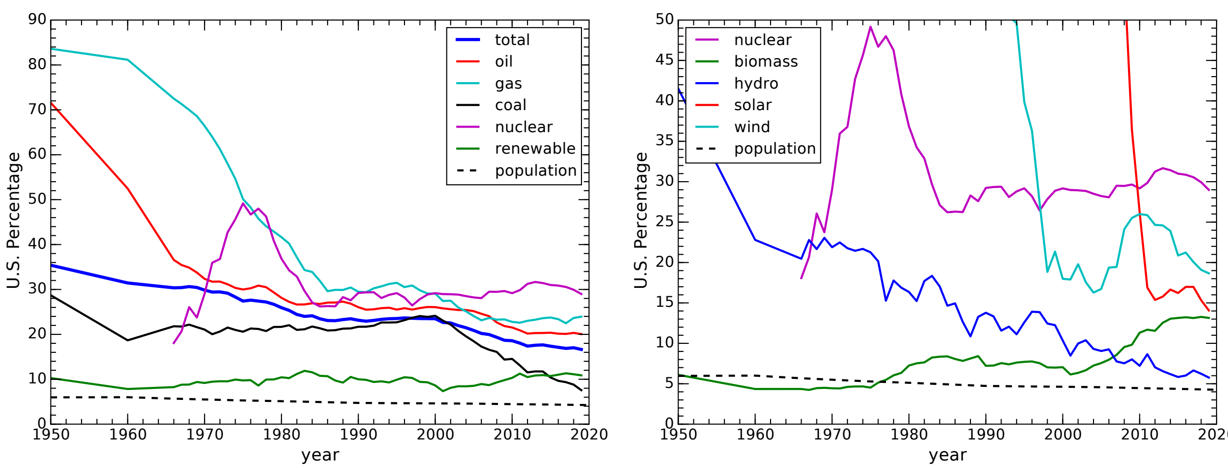


Figure 7.9: The left figure combines Figure 7.4 and Figure 7.7 to show the percentage of energy resources consumed by the U.S. over time. The overall picture is of a world catching up to an early leader. The U.S. was a literal "superpower" in the middle of the twentieth century. The dashed line at bottom represents the fraction of U.S. population in the world, so that energy use above this line means a greater-than-average share, which is true for all sources. The plot at right combines Figure 7.5 and Figure 7.8 to show the percentage of renewable and nuclear energy resources consumed by the U.S. over time. Solar and wind are characteristic of a nation known for innovation: first on the scene.

The thicker dark blue line in the left panel of Figure 7.9 represents all sources of energy, combined. Around 1950, Americans used a third of all the global energy, corresponding to almost 8 times[0mm] the global

0mm: The math is 35% over 6% of population compared to 65% over 94% of population: $(35/6)/(65/94) \approx 8.4$.

average per non-American. Today, the ratio is closer to 4.

The right side of Figure 7.9 similarly explores U.S. share of renewables. The only up-trending resource is biomass, due to mandates for ethanol usage.[18] But it is a minor player in the scheme of things. Solar and wind are interesting, in that the U.S. initially held a large global share as pioneers of the technology before the rest of the world joined in.

18: More on biofuels in Chapter 14.

7.3 Upshot: Go to the Source

The purpose of this chapter was twofold: first to introduce students to sources of reliable information on national and global energy production; and second to communicate the landscape of energy use. What emerges is a picture of a world still firmly in the grip of fossil fuels, whose annual usage continues to increase. Wind and solar are making inroads, but only at the few-percent levels thus far. The U.S. has played an outsized role in global energy relative to its population, especially in the mid-twentieth century.

7.4 Problems

1. Referring to Figure 7.1 and Figure 7.2, figure out the following measures:
    a) What percentage of energy consumption in the U.S. is from petroleum?
    b) What percentage of transportation is powered by petroleum?
    c) What percentage of petroleum goes directly to transportation?
    d) What percentage of petroleum goes directly to industrial processes (ignoring via electricity)?

⚠ Not the same as previous question.

2. The electricity block at the bottom center of Figure 7.2 is said to be 38.3 qBtu in size. Using the qBtu numbers in the sources at left, and the percentages of each going to electricity, figure out how many qBtu each line connecting to the left side of the electricity block represents.[19] What is the total, and does it match the 38.3 qBtu expectation, within reasonable rounding errors?

19: Example 7.1.1 may offer guidance.

3. Building off the result in Problem 2, calculate the percentages[20] of contributions coming into the left side of the electricity block in Figure 7.2? Which is the dominant input?

20: Verifying that they add to 100% is a good check.

4. Following a similar approach as for Problem 2,[21] concentrate on the output side of electricity production[22] and figure out how many qBtu are delivered to each sector on the right-hand side of the figure, based on input percentages to each of the four sectors

21: See also Example 7.1.1.

22: 13.0 qBtu delivered; implying 34% conversion efficiency from primary sources to delivered electricity and their total qBtu amounts. Treat "< 1%" as 0.5%. Do these add to 13 qBtu, as they should, within rounding error?[23]

5. Figure 7.2 hides contributions of sources to end sectors behind the "electric black box."[24] Following similar logic to that in the margin, and using results from Problem 4, figure out "corrected" values for what percentage of coal provides energy to each of the four end-sectors (re-distributing the 91% going to electricity into end-sectors).[25]
6. Figure 7.2 makes it look as if residential demand is satisfied without coal or nuclear, but 42% of residential demand comes from electricity, which *does* depend in part on coal and nuclear. Using numbers derived in Problem 3, and following a logic similar to that in Problem 5 and Example 7.1.1, redistribute this 42% residential contribution from electricity into its primary sources to ascertain what fraction of residential demand comes from each of the five source categories. For instance, petroleum would be the direct 8% plus 42% times the fraction (or percentage) of electricity coming from petroleum.[26]
7. While no energy source is free of environmental harm, arguably the last four entries in Table 7.1 are the cleanest, requiring no burning and no evidently problematic "waste." What percentage of the total U.S. energy is in this "clean" form, at present?
8. Let's say that in the course of one year a county in Texas produces 5 million kWh of electrical output from wind, and also pumps 100,000 barrels of oil from the ground containing a (thermal) energy content of about 6 GJ per barrel. What percentage of total energy production came from wind, if scaling wind in terms of thermal equivalent, as explained in Box 7.2?
9. Referring to Figure 7.4, what is the fastest-growing energy source in the U.S., and is it one of the fossil fuels?
10. If the approximately linear trends for recent increases in solar and wind seen in Figure 7.5 were to continue at the current (linear) pace, approximately how long would it take for the pair of them to cover our current ~ 100 qBtu per year demand?[27]
11. If the downward trend in U.S. coal use continues at its current pace, approximately what year would we hit zero?
12. Globally, do any of the resources appear to be phasing out, as coal is in the U.S. (as in Problem 11)? If so, how long before we would expect to reach zero usage, globally, based on simple extrapolation?
13. Globally, would you say that renewable energy sources are climbing faster than the combined fossil fuels, or more slowly? Can we therefore confidently project a time when renewables will overtake

23: This is a great way to check the correctness of your answers.

24: For example, the figure indicates that 17% of natural gas goes directly to residential end-users. But a substantial fraction of natural gas (35%) also goes to electricity, and 38.5% of electrical output goes toward residential use—a result of Problem 4. So the fraction of natural gas ending up satisfying residential demands is the direct 17% *plus* 38.5% of 35%, adding to 30.5%.

25: The four numbers you get should add to 100%, within rounding error.

26: Make sure your five numbers add to 100%, within rounding error.

ⓘ Keep it simple, as there is no single correct way to extrapolate this far into the future; just explain your approach.

27: We can hope to see faster-than-linear expansion in renewables, but this question asks what *would* happen without dramatic changes to the recent trends.

fossil fuels, based on trends to date?

14. As explored in Problem 11, the U.S. usage of coal is falling precipitously. According to the left plot in Figure 7.9, is the U.S. usage of coal greater or less than the global per-capita average?

15. Is the U.S. per-capita usage of *any* energy source lower than the global average, according to Figure 7.9?

8 Fossil Fuels

8.1 The Most Important Plot Ever 114
8.2 Overview: Coal, Oil, and Gas 117
Coal 117
Petroleum (Oil) 118
Natural Gas 120
8.3 Chemical Energy 121
8.4 Fossil Fuel Pros and Cons . 122
Fossil Fuels are Amazing . 122
Fossil Fuels are Terrible . . 123
On Balance? 125
8.5 The Future of Fossil Fuels . 125
Scenarios 125
Timescales 126
Clues in the Data 128
Geopolitics 131
8.6 Upshot: Good, Bad, Limited 132
8.7 Problems 134

We are now ready to dive into the core content of the book: assessing global energy demands and prospects. For most of human history, we derived energy from food-supplied muscle power of people and work animals, burning firewood, and harnessing wind and water flow (all deriving from solar energy). Then a most remarkable thing happened: the discovery and widespread utilization of fossil fuels. The abundance of energy delivered by fossil fuels profoundly changed the human condition, such that many elements of our modern world would seem like magic to someone living 200 or even 100 years ago.

Fossil fuels still completely dominate our energy usage. Every country is reliant on some amount of fossil fuels—especially for transportation. Even though fossil fuels cannot be our future—due to finite resource depletion and climate change concerns—it is critical that we look at these pillars of modern life, assessing what makes them both amazing and terrible, and what we might expect going forward. Facing the stark and underappreciated reality of fossil fuels will sharpen our desire to learn more about what might come after, as subsequent chapters address.

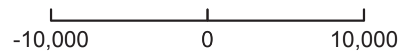
8.1 The Most Important Plot Ever

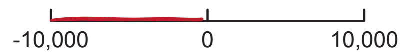
We have so far gained a few big-picture perspectives on the human endeavor. First, we illustrated the absurdity of constant growth in both physical and economic terms, concluding that growth must be confined to a temporary phase and will not be physically allowed to continue indefinitely. Next, we looked at population realities to understand how that story might develop. Then we looked at the scale of the universe, how minuscule Earth is in the vast emptiness, and explored the extreme

Oil pipelines and gas flaring on the Alaskan tundra at Prudhoe Bay. A drill rig fades into the fog at top center. Note the optical illusion that makes the photo's bottom border look crooked! Photo credit: Tom Murphy difficulties of colonization—putting the emphasis on managing our challenges right here on Earth.

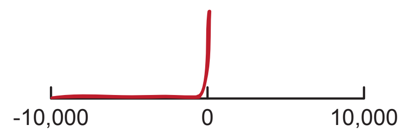
In order to frame just how important fossil fuels are and have been, we again take a broad view to put our energy trajectory in perspective before getting into the nuts and bolts of fossil fuels. The picture that emerges has the potential to reframe personal perspectives on our future.

The result may have greater impact if you are an active participant in its creation. So get some paper, the back of an envelope, or something. Draw a horizontal axis as a timeline. Label the left edge as –10,000 years (past). The right edge is +10,000 years (future). The middle is 0 (now; see the example in the margin). The vertical axis represents global energy production, on a linear scale. For ages, this was too tiny to see poking up above the floor. Only about 200 years ago did it become visible. So for the first 98% of the way from -10,000 to 0, draw a line hugging the floor. In the last 200 years, energy usage has increased exponentially.[1] So draw a smooth curve connecting the previous line into a steep rise at present (middle of the plot), using much or all of the available vertical space.





1: . . . a smooth curve peeling up off the floor and rocketing to an essentially vertical recent trajectory



What emerges is the classic "hockey stick" plot that applies to many physical attributes of our world: population, carbon dioxide, temperature, and—in the present case—energy use. In the long flat portion of the plot, our energy came from firewood and muscle (both animal and human labor). But the sudden transformative rise is really a story of fossil fuels. Even today, having added hydroelectric, nuclear, solar, wind, geothermal, and tidal power to the mix, fossil fuels still account for over 80% of the total.[2]

2: It's even worse than it sounds, since 10% is still in the form of biomass, much of which is old-technology firewood, leaving only 10% in the more modern forms of hydro, nuclear, solar, wind, geothermal, and tidal.

Let us then continue the plot *in the context of fossil fuels*. Being a finite resource, we know in broad terms what the curve *must* look like. It must drop back down to zero and ride into the future looking much as it did in the past: at zero. One may debate the exact timing of the peak of fossil fuel use, but for a variety of reasons we would be well justified in placing it sometime this century. We'll leave it to individual preference if you want to allow the curve to climb a bit more before turning down, but don't stray too far. This century ends only 1% of the way from 0 to +10,000, so don't let the peak get very far at all from the middle of the plot. Once turning down, the curve is likely to look reasonably symmetric, returning to zero in short order and staying there.

Independent of individual choices, if keeping within reason we're all looking at the same basic plot (as in Figure 8.1): fossil fuels are a blip on the time scales we associate with history. We live in a most abnormal time.[3] Because the upswing has lasted for generations, *it seems entirely normal* to most people: it's the only reality we or any person we've ever met has known. Lacking perspective, a child will view their life circumstances as *normal*, no matter how impoverished or privileged: it's the only world they've ever known or seen. Likewise, we accept and define our current world as normal—even if historical perspective

3: Social scientists are trained to not label their own time as *abnormal*, as such thinking may reflect a sloppy bias that all people through history might be tempted to adopt. Yet, neither should we declare that abnormal times can *never* happen. Any quantitative assessment of the current human scale and planetary resource impact argues that we are justified in allowing ourselves an exception for the present age.

ultimately considers the last century or two to be the most insanely unusual period of the human experience—like a fireworks show.

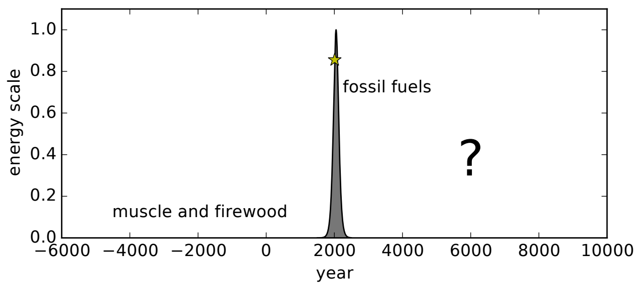


Figure 8.1: Energy over the ages, in the form of fossil fuels. Up until the present, fossil fuels capture the bulk of the human energy story. We know what it must look like in the long term as well. The huge question is how the second half of human history looks, after fossil fuels are depleted or abandoned. The yellow star is a guess as to our current position, based on evidence addressed later in the chapter suggesting that the resources are nearly halfway depleted.

Figure 8.1 should stimulate a swarm of questions. Where are we on the curve? When is the peak? Is the decline phase marked by escalating energy scarcity, or the advent of a renewable energy future? Might the far future look more like the past (muscle and firewood) than the present? Will this plot change how we interpret the world and our own plans for the future? The only fair conclusion is that we really do not know how the future will unfold.[4] We can label the left side as "muscle and firewood," and the spike as fossil fuels, but the only credible occupant of the right-hand side is a gigantic question mark.

4: We *can* rule some things out, though, like unending growth and fossil fuels lasting centuries more.

The idea of Figure 8.1 is not original to this textbook, having been portrayed in various incarnations over the last half-century or so [38]. When anyone makes a claim about what they think will happen by late-century, think about this plot. So many of our assumptions are based on the recent but abnormal past. All bets are off in defining the future. In one sense, those who rightly point out that we can't expect to be clever enough to foresee the future are correct—but perhaps in an unintentionally symmetric way. The future could be far more dismal than our dreams currently project. That would also be a surprise to many. We need to approach the future with humility, and set aside preconceived notions of where things are heading so that we can make choices now that will help define what comes next. Taking it for granted is a risky move.[5] Only by acknowledging the potential for a disastrous outcome can we take steps to mitigate that possibility. Waving it off is the most dangerous move we could make.

[38]: Hubbert (1962), "Energy resources: a report to the Committee on Natural Resources of the National Academy of Sciences; National Research Council"

5: In this sense, taking the risk seriously fits the definition of the word "conservative," even if present political alignments are mislabeled in this regard.

Box 8.1: Will Renewables Save Us?

Just because fossil fuel energy must return to pre-industrial levels in Figure 8.1 does not dictate that human society must return to pre-industrial energy levels. After all, solar, wind, nuclear, hydroelectricity are available to us now. Yet we will struggle to match today's energy levels on these resources alone. More disturbing is the notion that we may not be able to maintain high-technology approaches in a world devoid of fossil fuels. No one has demonstrated how, yet.

Also, the very disruption of losing such a critical resource without adequate advanced preparation may damage our capabilities. The short answer is: we simply do not know. The question mark in Figure 8.1 is the most fair statement we can make.

Note that Figure 8.1 is not intended to predict a particular future path. But it can serve to to counterbalance the prevailing optimism about a technologically marvelous future by providing a sanity check so that we might acknowledge that *we really do not know*. How can it be wrong to say that we do not know what the future holds? Yet, accompanying this uncertainty is a glimmer of hope: if the future is so uncertain and unscripted, then perhaps we have the power to write the script and set ourselves onto a viable and pleasant future path. If we elect to do so, it is of paramount importance that we do not ignore limitations imposed by nature in the process.

8.2 Overview: Coal, Oil, and Gas

Fossil fuels are found in three principal forms: coal, oil (petroleum), and natural gas.[6] They are essentially a form of ancient solar energy that plants once captured and stored as chemical energy to be locked away underground for many millions of years.[7] Sporadic, low-level use of fossil fuels dates back millennia, but modern use began in earnest in the eighteenth century with coal in Britain. Figure 8.2 makes clear that the use of coal did not *really* gather steam until the mid-nineteenth century, when industrialization took off. One may suspect that much of the rise in the use of fossil fuels is simply a reflection of population growth, but this turns out to be wrong. The right-hand side of Figure 8.2 divides the amount of fossil fuel use by global population to show that energy use per capita has *also* risen steeply over this time period, so that the exponential-looking phenomenon in the left panel is a *combination* of more people *and* more use per person. Today, the global average rate of use of fossil fuel use is a little over 2,000 W per person.[8] From Figure 8.2, we may say that coal really ramped up starting around 1850, oil around 1915, and natural gas around 1970.[9]

6: Think of the three forms of fossil fuels as solid (coal), liquid (petroleum) and gas (natural gas).

7: It is in this sense that the word "fossil" is appropriate: ancient remnants of life buried underground.

8: 15 TW of fossil fuel use divided by nearly 8 billion people is about 2,000 W per person. Compare to the U.S. total energy appetite of 10,000 W per person.

9: All of these sources were *first* used much earlier, but at insignificant levels. Natural gas makes a meaningful appearance starting around 1920, but heavy use began 50 years later after pipeline infrastructures were in place.

8.2.1 Coal

Coal—which looks like black rock—is the remnant of plant matter deposited, turned to peat, and heated/compressed by burial to form a mostly-carbon substance that can be combusted with oxygen to generate heat. The heat can be used to make steam, which can then power machinery or turbines for producing electricity.[10] Or the heat may be used directly for materials processing, like creating molten steel in blast furnaces.

10: . . . covered in Chapter 6

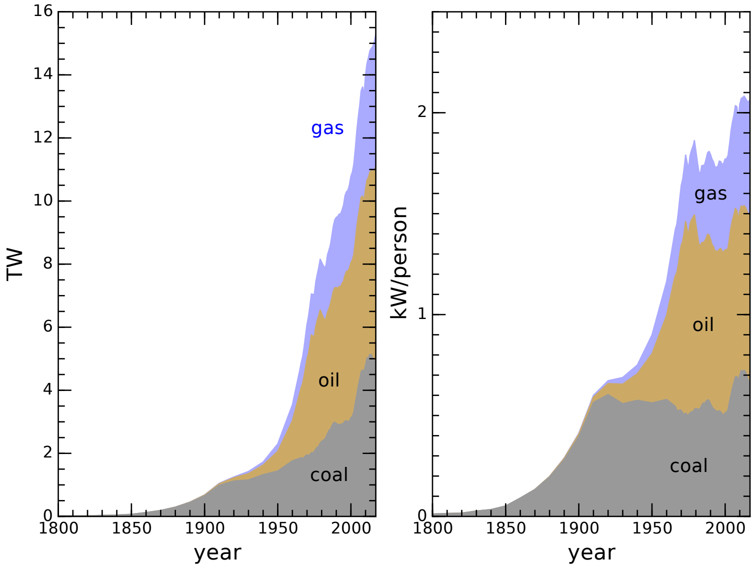


Figure 8.2: Historical use of fossil fuels worldwide, which may be viewed as a zoom-in of the left-hand side of the peak in Figure 8.1. The three types are stacked on top of one another, so that gas makes the smallest contribution, not the biggest. On the left is the raw usage rate expressed in terawatts, while the right is a per-capita measure showing that the left-hand rise is much more than just a reflection of population growth [16].

Coal opened the door on the Industrial Revolution[11] in the late eighteenth century, allowing locomotion (trains), mechanized manufacturing, large-scale materials processing, and heating applications. Somewhat circularly, the first major use of the steam engine[12] was to pump water out of coal mines to accelerate the extraction of. . . coal. This fact further highlights that from the very start, the Industrial Revolution was focused on the fossil fuel resource that enabled it.

11: . . . which history may rename the Fossil Fuel Revolution

12: The first widely adopted steam engine design is credited to James Watt, from whom we get the name of our unit for power.

Today in the U.S., coal accounts for 13% of total energy consumption[13]—down considerably from 23% in 2000.[14] For the world at large, coal still accounts for 25% of primary energy use.[15] The vast majority of coal (91%) in the U.S. goes to electricity production, the remainder fueling industrial processes requiring lots of heat. The quality of coal varies greatly. Table 8.1 presents properties of the four main coal categories. Anthracite is the king of coals, but has been largely consumed at this stage. Coal grades having lower energy content contain more non-combustible materials[16] like SiO_2, Al_2O_3, Fe_2O_3, and water.

13: Fig. 7.4 (p. 107)

14: As discussed in Chapter 7, coal's decline in the U.S. is largely due to increased reliance on natural gas for electricity.

15: Fig. 7.7 (p. 109)

16: . . . sometimes called "ash" content and volatiles

Grade	Carbon Content (%)	Energy Density (kcal/g)
Anthracite	86–97	6–8
Bituminous	45–86	5.5–8
Sub-bituminous	35–45	4.5–6.5
Lignite	25–35	2.5–5

Table 8.1: Four classes for grades of coal, in order of decreasing energy content and value. Anthracite has been largely depleted and is a rare find today. [39, 40]

8.2.2 Petroleum (Oil)

Petroleum—also called oil—is ubiquitous in our world as the source for gasoline, diesel, kerosene, lubricating oils, tar/asphalt, and even most plastics. Virtually all[17] transportation: planes, trains, automobiles, and ships run on petroleum-based energy.

17: Even electric cars may depend on fossil fuels, since > 60% of electricity in the U.S. is fossil-generated.

Petroleum first entered the modern scene around 1850, and the first drilled well[18] was in 1858 in Pennsylvania. Early uses were for kerosene lamps.[19] The first commercial internal combustion engine closely followed in 1859, arriving at an essentially modern form in 1876 at the hands of Nikolaus Otto.[20] The first production automobile using a gasoline-powered internal combustion engine was developed by Karl Benz in 1885 and Henry Ford's Model T began mass-production in 1913. In the intervening years, electric cars surprisingly were more popular, but quickly gave way to the gasoline[21] car due to superior range, quick refueling, and cost.

18: . . . using a steam engine powered by coal

19: . . . a relief from expensive and declining whale oil resources

20: Why isn't it Otto-mobile, then?

21: For clarity, *gasoline* is a liquid that derives from petroleum. Natural gas is in gaseous form, not directly related to gasoline.

Today, petroleum supplies 37% of energy consumption in the U.S.[22] 70% of petroleum goes to transportation (92% of transportation energy is in the form of petroleum), while another 24% goes to industrial processes.[23]. Globally, petroleum usage represents a slightly smaller fraction of total energy than in the U.S., at 31% of total energy consumption.[24]

22: Recall that Chapter 7 presented these breakdowns in graphical form.

23: Fig. 7.2 (p. 105)

24: Fig. 7.4 (p. 107) and Fig. 7.7 (p. 109)

The petroleum extracted from the ground is often called crude oil, and consists primarily of hydrocarbon chains of various lengths. The lighter molecules (shorter chains)—typified by octane (Figure 8.3)—are useful for gasoline, while the much heavier (longer) molecules are found in tar/asphalt, lubricants, or used as "petrochemical feedstock" for plastics. The process of refinement separates constituents by chain length, producing gasoline, kerosene, diesel, heating oil, lubricants, tar, etc. 92% of crude oil goes to energy production of some form (burned), while 8% is used to create petrochemical products, as depicted in Figure 8.4.

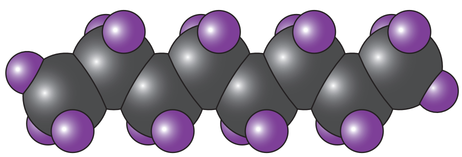


Figure 8.3: Octane (C_8H_{18}, containing 8 carbon atoms and 18 hydrogens) is among the shorter/lighter hydrocarbon chains found in oil, and is typical of gasoline. Longer chains of the same basic design are found in lubricants, tar, and as feedstock for plastics.

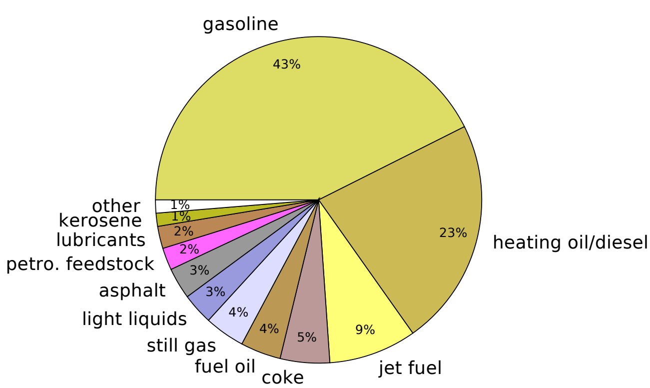


Figure 8.4: Fractional use of a barrel of petroleum, from [41]. All but asphalt, petrochemical feedstock, lubricants, and "other" are burned for energy, amounting to 92% burned. Still gases include methane, ethane, propane and butane in gaseous form, while the light liquids are also mostly propane and butane in liquid form. Coke is not the soft drink.

Petroleum is measured in barrels (bbl), equating to 159 L (42 gal). Each barrel of crude oil contains about 6.1 GJ of energy (1,700 kWh; 5.8 MBtu). For reference, the world consumes about 30 billion barrels a year (the U.S. is about 7 billion barrels per year, or 20 million barrels per day). No single country produces oil at a rate greater than about 12 million barrels per day.[25]

25: As a consequence, the U.S. is presently unable to support its petroleum needs from domestic resources alone.

To provide some perspective on how special/rare oil is, the chances of finding any by drilling a random spot on the planet is about 0.01%.[26] This is because many geological conditions must be met to make oil:

26: . . . based on a crude calculation of the total resource and assuming a typical deposit thickness of 10 m

1. Organic material must be deposited in an oxygen-poor environment to inhibit decomposition, like dead animal and plant remnants settling to the bottom of a still, shallow sea;
2. The material must be buried and spend time under at least 2 km of rock, to "crack" large organic molecules into the appropriate size, like octane, for instance (Figure 8.3);
3. The material must not go below about 4 km of rock, or the pressure will "overcrack" the molecules to form natural gas (still useful, if trapped underground);
4. An impermeable caprock structure must sit atop the permeable and porous rock (Figure 8.5) that holds the high-pressure oil to keep it from simply escaping.[27]

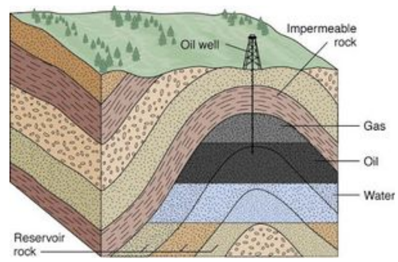


Figure 8.5: Oil and gas embedded in porous rock, under an impermeable caprock [42]. From U. Calgary.

27: Losing even a drop per second adds up to 20 million barrels over one million years, which is short on these geological timescales.

Oil deposits are rare and tend to be clustered in certain regions of the world where ancient shallow seabeds and geological activity have conspired to sequester organic material and transform it appropriately. The process takes millions of years to complete, and we are depleting the resource about 100,000 times faster than it is being replenished.[28]

28: A simple way to see this is that it took tens of millions years to create the resource that we are consuming over the course of a few centuries: a ratio of at least 100,000 (see Box 10.2; p. 169). This is like charging a phone for 3 hours and then discharging it in 0.1 seconds! Viva Las Vegas! Fireworks!

Many early oil wells were "gushers"—under enough pressure to push up to the surface under no effort. Modern extraction is not so lucky, having depleted the easy oil already. A combination of techniques is used to push or pull the oil out of its porous rock, including pumps, injecting water under high pressure, bending the drill path to travel horizontally through the deposit, or fracturing[29] the underground rock via pressurized fluids. More work is required to coax the oil out of the ground as time moves forward.

29: . . . colloquially called fracking

8.2.3 Natural Gas

Natural gas is familiar to many as a source of heat in homes (stoves, hot water, furnace), but is also a major contributor to electricity production and industrial processes (usually for direct heat in furnaces/ovens). It is also used extensively in the production of fertilizer via the Haber process.[30]

30: The Haber process uses the energetically cheap hydrogen in methane (CH_4) to produce ammonia (NH_3) as a chief ingredient in nitrogen-rich fertilizers.

Natural gas is primarily methane (CH_4). Its formation process is similar to that of oil, but deeper underground where the pressure is higher and longer-chain hydrocarbons are broken down to single-carbon methane molecules. We find natural gas trapped in underground reservoirs, often on top of oil deposits (Figure 8.5). Thus petroleum drilling operations typically also produce natural gas output.[31] The gas itself tends to flow out freely once a well is drilled, since it is under great pressure and not viscous like oil. The first commercial use of natural gas started with a well in New York in 1821, leading to a pipeline distribution for street lighting in Philadelphia in 1836. Because of its low density[32] compared to coal or petroleum, it is often impractical to collect, store and transport the gas, strongly favoring a pipeline infrastructure for its delivery. Lack of pipeline infrastructure delayed widespread use of natural gas until about 1970. It is also possible to liquefy natural gas (called LNG) by cooling to −160°C and then storing/transporting in cryogenic vessels.

31: Unless a gas pipeline is in place at the drill site, the natural gas is too voluminous to be contained/stored, so is often flared (burned; wasted) at the well-head.

32: . . . on account of its being gaseous

Natural gas constitutes 31% of energy consumption in the U.S., and 22% globally.[33]. Because of the need for pipeline infrastructure in order to deliver gas to consumers, remote areas are typically unable to take advantage of the resource. The uses for natural gas in the U.S. are more diverse than for coal or oil: 35% goes to electricity production, 34% for industrial purposes, and 29% for residential and commercial heating.[34]

33: Fig. 7.4 (p. 107) and Fig. 7.7 (p. 109)

34: Fig. 7.2 (p. 105)

8.3 Chemical Energy

Chemical energy is released as heat when combustible materials are ignited in the presence of oxygen. Sec. B.3 (p. 379) in Appendix B provides some background.

Fossil fuels all work the same way, chemically. The three key reactions for coal, methane, and octane[35] are:

35: Gasoline—the main product extracted from petroleum—is a blend of medium-sized hydrocarbon chains, and we use octane as a decent representative for oil.

$$\begin{array}{lccccccc} \text{coal}: & C & + & O_2 & \rightarrow & CO_2 & & + 32.8\,\text{kJ/g} \\ \text{gas}: & CH_4 & + & 2O_2 & \rightarrow & CO_2 & + 2H_2O & + 55.6\,\text{kJ/g} \\ \text{oil}: & C_8H_{18} & + & \frac{25}{2}O_2 & \rightarrow & 8CO_2 & + 9H_2O & + 48.0\,\text{kJ/g} \end{array} \tag{8.1}$$

The energy amounts above represent the total energy available per gram of input fuel.[36] Table 8.2 provides several key attributes of fossil fuel combustion. Energy density, in kJ per gram or often kcal/g, is a fundamentally important measure of the fuel's potency. By expressing in kcal/g, we can compare to food labels in the U.S., for which fats are around 9 kcal/g, while carbohydrates and proteins clock in around 4 kcal/g.

36: . . . not counting the oxygen; just the carbon-based fuel

Both fossil fuels and food are a type of chemical storage ultimately tracing back to photosynthesis in plants.

Fuel	Representative	molar mass	kJ/mol	kJ/g	kcal/g
coal	C	12	393.5	32.8	7.8
natural gas	CH_4	16	890.3	55.6	13.3
petroleum	C_8H_{18}	114	5,471	48.0	11.5

Table 8.2: Combustion properties of fossil fuels.

Note that fossil fuels are more like fat (near 10 kcal/g) than carbohydrates (at 4 kcal/g).[37]

37: The simplest way to understand this is that carbohydrates (sugars, such as glucose: $C_6H_{12}O_6$) already have oxygen in the molecules, and are in some sense already half-reacted (combusted) with oxygen, as elaborated in Sec. B.3 (p. 379).

Box 8.2: Superlative Energy Density

To put these energy densities into perspective and demonstrate how amazing fossil fuels are, consider that the explosive TNT[38] has an energy density of just 1.0 kcal/g. But comparing to TNT is somewhat unfair, as explosives must carry their oxygen with them.[39] Hydrogen gas tops the energy density charts, chemically, at 34 kcal/g, because hydrogen is such a light atom.[40] If having to carry the oxygen along, as rockets must, for instance, the hydrogen-plus-oxygen source is down to 3.8 kcal/g. Rocket fuels and explosives, in general, tend to be in this range of a few kcal/g for this reason. Aside from hydrogen, very few compounds outperform methane for energy density. So crudely, 15 kcal/g is about the top of the chemical scale.

38: . . . $C_6H_2(NO_2)_3CH_3$

39: An explosion is too fast—and violent—to get oxygen from the surrounding air.

40: But hydrogen is both bulky and so highly flammable as to be dangerous to store in gaseous form (look up Hindenburg), so don't get too excited.

8.4 Fossil Fuel Pros and Cons

8.4.1 What Makes Fossil Fuels Amazing

Energy Density: We have seen in Section 8.3 that the energy density of fossil fuels is quite respectable: about the best that chemistry delivers. Anything over 10 kcal/g is a "superfood" energetically. Table 8.3 compares to other substances, by which we see that fossil fuels are two orders-of-magnitude more energy-dense than battery storage.

Table 8.3: Energy densities of familiar energy substances. The Tesla Powerwall represents available lithium-ion capability. Alkaline batteries are familiar AA or AAA cells, and lead-acid batteries are the 12 V ones found in most cars. For hydroelectricity, a 50 m dam is assumed.

Substance	kcal/g
Gasoline	11
Fat (food)	9
Carbohydrates	4
Rocket Fuel	4
TNT explosive	1
Alkaline battery	0.11
Tesla Powerwall	0.10
Lead-acid battery	0.03
Hydroelectric (50 m)	0.0001

Safety: Fossil fuels have greater energy density than explosives, without being particularly explosive! The safety aspect of fossil fuels is a big selling point. Sure, gasoline burns, but really it's the vapor mixed with oxygen that goes poof. If you (foolishly; please don't do this!) throw a match onto a bowl of gasoline, you'll certainly get some lively fire, but the thing won't *explode*. Only the vapor above the pool will be on fire. Think about how many cars you've seen in your life, and how many of those have exploded.[41] How many wrecked cars have you seen, and how many of those exploded? It is not impossible to have an explosive accident from gasoline, but it's pretty rare.

41: . . . discounting dramatic events the entertainment industry prepares for us

Cheap: Fossil fuels were bestowed upon us as a byproduct of biological and geological processes on our planet. They are essentially free—at least the way we have historically viewed natural resources as ours to grab. How cheap are they? Hiring a physical laborer to exert 100 W of mechanical power (digging, for instance) for 40 hours a week at $15/hr costs $600 for a week. For that price, we receive 4 kWh of work. In electricity terms, the same 4 kWh costs $0.60 at typical rates (1,000 times cheaper than human labor). Gasoline—for which one gallon contains

37 kWh and costs $4—would be just $0.43. Efficiency differences, and the cost of the machine to perform the labor also factor in. But the point should be clear enough.

Perfect Storage: Effectively, fossil fuels represent a form of long-term storage of ancient sunlight, captured in plant matter and (sometimes via animal ingestion) ending up buried underground as chemical energy. Compared to other forms of storage, like rechargeable batteries, flywheels, or even hydroelectric reservoirs (pumped storage), fossil fuels are astoundingly superior. Fossil fuel deposits are tens or hundreds of millions of years old. Try finding a battery that will hold its charge that long! Seemingly permanent man-made dams/reservoirs are unlikely to last even one-thousandth as long. Combined with their superior energy density, fossil fuels are perhaps the best form of energy storage available to us, aside from nuclear materials.

Food Production: The Green Revolution in agriculture would not have been possible without fossil fuels. Not only did they provide the motive force for mechanized farming (plowing larger tracts of land, harvesting and processing crops quickly), but the all-important fertilizer is derived from natural gas.[42]

42: Methane (CH_4) provides an energetically favorable source of hydrogen to make ammonia (NH_3) as a way to deliver nitrogen to plants (called the Haber process). Water (H_2O) may seem like a more obvious and abundant source for hydrogen, but in this case substantial energy would have to be injected to extract hydrogen. Methane, by contrast, will give up its hydrogen more easily.

Technology Catalyst: Fossil fuels opened the door to widespread mechanization and electrification, completely transforming our way of life. As central as their role has been, it is difficult to claim that many of the benefits we enjoy today—whether health care, technology, scientific knowledge, or comfortable living standards—would have been possible without them. Much that we celebrate in this world rode on the back of fossil fuels.

8.4.2 What Makes Fossil Fuels Terrible

Climate Change: Nothing comes for free. Fossil fuels also bring many downsides. Chief on many peoples' minds today is climate change, via CO_2 emission—an unavoidable consequence of combustion (Eq. 8.1). Extracting energy from fossil fuels,[43] leaves no choice but to accept CO_2 as a byproduct, in large quantities. We will get to the details of climate change in Chapter 9, but for now will just say that increased CO_2 in the atmosphere changes the equilibrium temperature of Earth by altering how effectively the surface can radiate heat away to space through the atmosphere. The physical mechanism is *very* well understood, and the amount of CO_2 that fossil fuel combustion has produced is *more than enough* to account for the measured CO_2 increase in our atmosphere. What is less certain is how the complex, nonlinear, interconnected climate systems will react, and whether positive feedbacks that exacerbate the problem dominate[44] over negative feedbacks that act to tame the consequences. In the meantime, fossil fuels have handed us a global-scale

43: . . . that's the whole point

44: All evidence says positive dominates.

problem of uncertain magnitude and may end up costing us—and other species—dearly.

Population Enabler: Human population pressures on our planet may also be traced to fossil fuels via agricultural mechanization and fertilizer feedstock (the Green Revolution). Since so many new global challenges—deforestation, fisheries collapse, species loss, climate change—scale with the population, perhaps all of these ills can be attributed to fossil fuels—in that it is doubtful these problems would exist at the present scale had we never discovered or utilized them.

Military Conflict: Fossil fuels represent such a prize that access and control of the resources has played a key role in many armed conflicts. Put another way, how many have lost their lives to fights over these precious resources? It is hard to view the complex and fraught relationships in the middle-east as being disconnected from the fact that it is the most oil-rich region in the world.[45]

45: Countries and regions lacking important resources receive far less attention from the developed world.

Environmental Toll: Environmental effects from the extraction of fossil fuels can be pretty destructive. We have seen oil tankers crash and coat beaches and wildlife in tarry sludge. The Deepwater Horizon drill platform failure in 2010 spewed vast amounts of oil into the ocean. Coal extraction can leave mountaintops bare and contaminate local water sources from the tailings. Hydraulic fracturing (fracking) can contaminate groundwater supplies. Natural gas wells—including fracking sites—often leak methane into the atmosphere, which is 80 times more potent than CO_2 as a greenhouse gas[46] on short timescales.

46: Although, methane does not last in the atmosphere as long as CO_2. Still, this is why gas is often flared (burned) at drill sites lacking pipeline infrastructure, rather than allowing it to escape as methane.

Substance Addiction: Finally, the very fact that fossil fuels are *finite* may be viewed as a serious negative. Granted, an effectively inexhaustible supply would be devastating for the climate change story. Setting that aside, the fossil fuel inheritance might be viewed as a sort of bait-and-switch trick. We have built up to our current state wholly in the context of cheap and available fossil fuels, and simply *do not know* if we can continue to live at a similar standard in a post-fossil world. Fossil fuels have lasted long enough (several generations) to seem normal. We take them for granted, and have not formulated a master plan for a viable world devoid of these critical resources. How will air travel, ships, trains, and long-haul trucking[47] be handled without petroleum? The current situation is precarious. Failure to plan wisely for a post-fossil world would not be the fault of fossil fuels by themselves. But the fossil fuel endowment that happened to grace our planet was large enough to harm the climate and to lull us into complacency. Had it been a much smaller amount, we would be less likely to fall into the trap.[48] This is the "rabbit out of a hat" referred to in Chapter 2: just getting one conditions us to expect an eternal state of rabbits.

47: All of these modes of transportation are difficult to accomplish via electric drive (Sec. D.3; p. 397), and critical to our global supply chains for manufacture of consumer goods.

48: By the same token, it is unlikely that we would be at a comparable technological level if our inheritance had been much smaller.

8.4.3 On Balance?

Deciding whether fossil fuels have had a net-positive or net-negative influence on humanity may not be answerable (Table 8.4 provides a summary of the previous two subsections). How many lives has it saved through better technology and health care? How many lives has it destroyed through conflict, pollution, and transportation accidents? How many lives has it created, through vast increases in agricultural productivity—as well as via better medical care? How many species has it destroyed, by promoting habitat loss both directly via extraction and indirectly as a catalyst to population growth via increased agricultural productivity? Sometimes it is even hard to decide which category to put these impacts into. For instance, in the fullness of time, will we see all the lives created on the back of fossil fuels as a good thing? If the result is overshoot, collapse, and the unprecedented suffering of billions of people, then perhaps not. It's a mess.

Table 8.4: Pros and cons of fossil fuels.

Pro	Con
energy dense	climate change
safe	overpopulation
inexpensive	agent of war
long storage	environ. damage
agriculture	overdependency
technology	so yesterday

In essence, humanity is running this global-scale unauthorized experiment on the planet without a plan. Nothing like this has ever happened, so we don't know how it will turn out. We have plenty of evidence that past civilizations overextend and collapsed [43], but we can't identify a fitting analog to successful navigation of the fossil fuel phenomenon. Meanwhile, plenty of signs justify grave concern.

[43]: Diamond (2005), *Collapse: How Societies Choose to Fail or Succeed*

8.5 The Future of Fossil Fuels

8.5.1 Scenarios

Figure 8.1 provocatively asserts that fossil fuel use must fall back to essentially zero in a relatively short time (within a century or two). This fact alone does not define our future on the spectrum of dismal to glorious, but it is one we need to consider carefully given the fundamentally important role fossil fuels have played in getting us to where we are today. The return to zero fossil fuels could take a variety of forms:

These are not strictly exclusive of each other, so some combinations are also possible.

1. We discover a new form of cheap energy not yet known or appreciated that is a game changer, quickly abandoning the fossil fuels still left in the ground.
2. Known renewable energy sources (solar, wind) are developed to the point of being effectively superior to fossil fuels so that market forces naturally move us away from fossil fuels before actually running out.
3. Climate change concerns result in politically enforced financial dis-incentives to using fossil fuels, so that we migrate away—albeit likely at higher cost, politically controversial, and not globally adopted.

4. Increased difficulty in extracting fossil fuels drives their price up so that the market is ultimately forced to accept less convenient and more expensive forms of energy.
5. We fail to find suitable substitutes to this precious and unique resource, so that global geopolitics increasingly center on competition for remaining fuel, likely touching off destructive resource wars.
6. Perhaps together with the previous point, society slowly grinds to a less energy-rich state, diminishing agricultural capacity and decreasing both the number and standard of living of people on the planet.

We cannot predict which of these paths might manifest, but it is not hard to find adherents to any of these narratives. Part III of this book covers alternatives to fossil fuels, and Chapter 17 summarizes practical challenges to the various alternatives. One lesson that emerges is that fossil fuels beat out alternatives on a host of considerations, leaving a gap between the two groups. If not for the finite supply and climate ills, we would have no incentive to adopt otherwise inferior sources of energy at higher cost. But first, we should briefly look into future prospects for extraction of fossil fuels. How limiting is the physical resource?

8.5.2 Timescales

The simplest approach to evaluating a timescale for resource availability is the R/P ratio: reserves to production.[49] The idea is very intuitive: if you have $10,000 in a bank account, and tend to spend $1,000 per month on living expenses, you can predict that—absent additional income—you will be able to go for ten months. So if we have an estimate for resource remaining in the ground, and the current rate of use, we simply divide to get a timescale.

49: Here, "production" means "obtaining from the ground," not fabricating artificially.

Table 8.5 reports the proven reserves in the world and in the U.S. for the three fossil fuels, the estimated fraction used so far globally, the rate of consumption,[50] and the timescale given by the R/P ratio.

50: Consumption and production are essentially identical: no stockpiling.

Region	Resource	Remaining	% Used	Annual Use	R/P (years)
World	oil	1,700 Gbbl	~45%	30 Gbbl	~60
	gas	200 Tcm	~33%	3.5 Tcm	~60
	coal	900 Gt	~30%	8 Gt	~110
U.S.	oil	35 Gbbl		7 Gbbl	~5
	gas	8.5 Tcm		0.85 Tcm	~10
	coal	250 Gt		0.7 Gt	~360

Table 8.5: Summary of proven reserves, usage rates, and time remaining for the world and for the U.S. (if using *only* its domestic supply) [44–46]. Oil is measured in giga-barrels (Gbbl; 10^9 bbl), gas in tera-cubic-meters (Tcm; 10^{12} m^3), and coal in gigatons (Gt; 10^{12} kg; noting that 1 ton is 1,000 kg).

The world has already consumed 1.5 trillion barrels of oil, which is nearly the same amount as the 1.7 trillion barrels of proven reserves—indicating that we are roughly halfway through the resource.[51] Certainly, we can expect that additional resources will be discovered and added to the

51: This fact is one justification for believing we may be near the top of the symmetric curve in Figure 8.1.

proven reserves,[52] but the globe is pretty well explored now, and we would not expect huge surprises like another hidden middle-east-size oil deposit. Note that for natural gas, the estimated total resource in the U.S. (what we *think* we may yet find beyond proven reserves) is about 55 Tcm, which would last just over 60 years.[53]

52: Also, technological advances can make previously impractical resources available, adding to reserves.

53: Because gas is harder to transport, and typically delivered by pipelines, domestic supply is more relevant for gas than it is for the globally-traded oil resource.

It is difficult to compare the remaining resource in the three forms directly, since different units are used for each. But we can cast each in terms of energy units for comparison. Doing so, the global reserves of oil, gas, and coal correspond to 10, 8, and 20 ZJ[54] remaining, respectively. We have so far consumed 8, 4, and 8 ZJ of oil, gas, and coal (Table 8.6). These form the basis of the estimated fraction consumed in Table 8.5. Note that the amount of oil and gas remaining are roughly comparable in energy, while coal is roughly twice as much.

54: ZJ is zetta-Joules, or 10^{21} J.

Table 8.6: Proven reserves and amount used, in energy terms.

Fuel	Proven 10^{21} J	Used 10^{21} J
Coal	20	8
Oil	10	8
Gas	8	4

Coal[55] therefore seems to be our most abundant fossil fuel, which prompts two comments. The first is that it is the worst offender in terms of CO_2 emission, emitting roughly twice as much CO_2 per unit of delivered energy as the other fossil fuels (covered in Chapter 9). The second is a caution in trusting the reserves estimates for coal, having often been vastly overestimated and then reduced significantly. For instance, Britain had to downward-revise their estimated coal reserves over the period from 1970–2000 to about 1% of their original because most of the estimated resource turned out to be in seams too thin and difficult to be commercially viable [47].

55: Coal reserves estimates [46] are broken into higher-quality anthracite and bituminous (~7 kcal/g), then sub-bituminous and lignite (~4.5 kcal/g) varieties, totaling 480 Gt (gigatons) and 430 Gt, respectively. (see Table 8.1).

[47]: Rutledge (2011), "Estimating long-term world coal production with logit and probit transforms"

For some, the R/P numbers in Table 8.5 may seem alarmingly short, while for others they may signal a comfortable amount of time to devise alternative energy strategies. Either way, this century is critical. But it is also important to recognize that the story is not quite as simple as the R/P ratio. While it provides a useful *scale*,[56] we should consider these nuances:

56: If the number worked out to 5 years, we would be in a panic. If it worked out to 5,000 years, climate change would loom as the chief concern.

1. The production (thus consumption) rate is not steady, but on the whole has grown over time (continued growth would shorten timescale).
2. New exploration and discovery adds to reserves (lengthening the timescale), but with diminishing success lately.
3. Advances in oil extraction technologies increase the amount of accessible oil (lengthening timescale).
4. Geological challenges limit the rate of production (lengthening timescale but also limiting resource availability).
5. Demand (thus production) could plummet if superior substitutes are found.

Point number 4 deserves some elaboration. We should not think of fossil fuel reserves as a bank account from which we may withdraw funds at an arbitrary rate, or as a cavernous underground lake just waiting to be slurped out by whatever straw we wish to shove in. Coal, firstly, does not flow, requiring substantial physical effort to expose and remove. The

A car's gas tank is another tempting, but flawed mental model. Getting water out of wet sand is closer to the truth for oil extraction.

rate at which it can be removed depends on the thickness of the seam, how deep it is located, and how hard it is to dig out surrounding rock. Even oil is not in some sloshing reservoir, but permeates porous rock, limiting how quickly the viscous fluid can be coaxed to flow out of the rock and into the pump tube. Gas is the quickest to escape its rocky tomb, but at this stage the U.S. has moved to "tight gas" that does not so easily break free—forcing a technique of fracking the rock to open channels for gas to flow. The same technique is being used to access "tight oil" that otherwise refuses to be pumped out of the ground by conventional means.

In all cases, it is obvious that we would pursue the easiest resources first: the low-hanging fruit. As time marches on, we are forced to the more difficult resources.[57] Adding to the geological factors is the simple fact that we do not possess unlimited extraction machinery, limiting the rate at which fossil fuels can be delivered from the ground. It is also worth pointing out that drilling deeper will not continue to pay dividends, as Section 8.2.2 points out that oil buried too deep will be "cracked" into gas.

57: . . . deeper underground, under deep water, or in "tight" formations

Figure 8.6 illustrates three variants of possible trajectories for a finite resource. The left-most panel corresponds to the R/P ratio: how long can we go at *today's* rate of use, if we locked in consumption at a steady value? The second assumes we continue an upward trajectory, which shortens the time compared to the R/P ratio before the resource runs out (using it ever-faster). Both of these are unrealistic in their own ways—the second one because of the physical constraints on extraction listed above (not a free-flowing resource). The third case is more realistic: a peak and somewhat symmetric decline. This is how real fossil fuel resources behave in practice. All three scenarios could create shocks to the system, but note that the (realistic) peak scenario brings the trauma of declining supplies soonest—long before the R/P ratio would suggest.

Figure 8.6: Three scenarios for a finite resource playing out, all based on the same initial history (the red dot is "now") and the same remaining amount (blue-shaded region). The red bar over each represents the remaining time until resource decline. See text for details.

8.5.3 Clues in the Data

Despite the uncertainties listed above, we can say for sure that Earth is endowed with a finite supply of fossil fuels, and that in order to consume the resource, deposits must first be discovered via exploration and then developed into active wells. Even in areas known to have oil,[58] only about one in ten exploratory wells bears fruit. The chances of striking oil at a random location[59] on Earth is in the neighborhood of 0.01%. Section 8.2.2 indicated the chain of events that must transpire to produce oil.

58: . . . also applies to gas

59: Think about throwing a dart at the globe.

A plot of the discovery history of conventional oil is revealing, seen in Figure 8.7. In it, we see that discovery peaked over 50 years ago. Since we can't extract oil we have not yet discovered—much like we can't possess an iPhone model that hasn't even been designed yet, the area under the consumption (red) curve must ultimately be *no larger* than

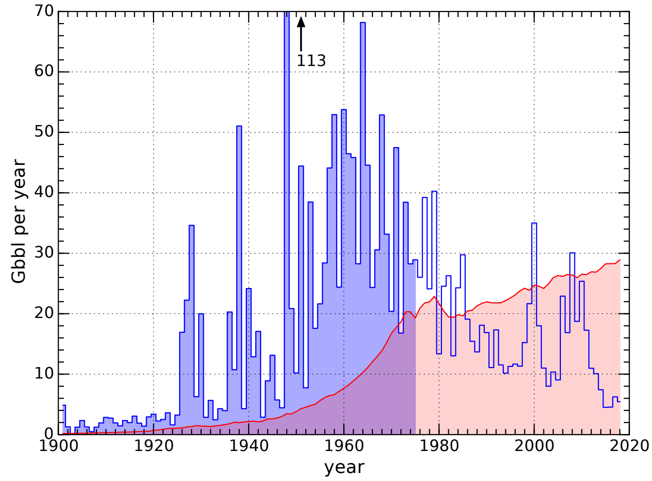


Figure 8.7: Historical discovery rate of conventional oil (blue), measured in billion barrels (Gbbl) discovered each year [48]. The red curve shows annual global consumption of conventional crude oil. Until about 1985, we tended to discover more oil than we used each year, but the rate of discovery peaked decades ago and is now in decline as we complete the job of exploring Earth's resources. The blue area is made equal to area under the red curve, which itself represents the amount of oil used to date. This effectively means that we have used all the oil discovered up to 1976, and are now left with a dwindling bank account (oil reserve)—our annual income (new discovery) being less than our spending (consumption).

the area under the discovery data (blue). It is therefore inevitable that consumption will peak and fall at some point, by whatever means. Note that a symmetric curve peaks when the resource is half-consumed.

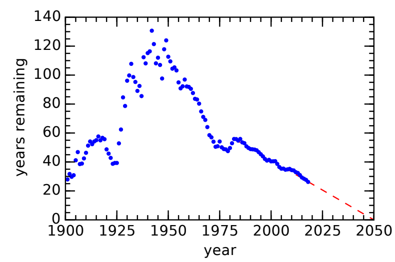


Figure 8.8: Years remaining in the global conventional oil resource as a function of time, extracted from the data in Figure 8.7. Since 1982, the world has been on a steady path toward depletion of the conventional oil resource by 2050.

The information in Figure 8.7 can also be re-cast to ask how many years remain in the resource. For any given year, the total remaining resource can be assessed as the cumulative amount discovered to date minus the cumulative amount consumed. Then dividing by that year's annual production (same as consumption) rate produces an estimate of remaining time (the R/P ratio again). Figure 8.8 shows the result.

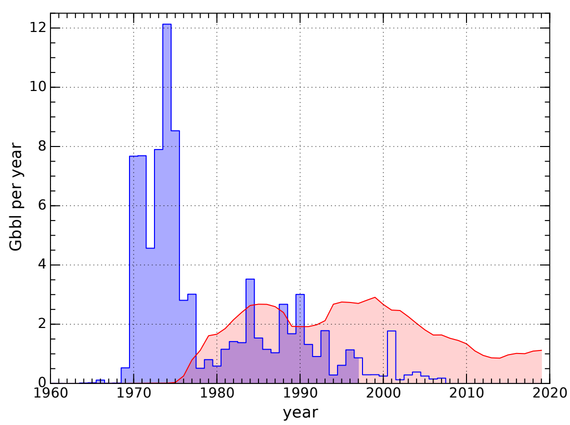


Figure 8.9: North Sea (U.K.) oil discoveries (blue, in giga-barrels per year) peaked in the 1970s and have basically ended. Production (red) lags discovery, and cannot carry on much longer as the last of the discovered oil (unshaded blue outline) is extracted. Plot conventions follow those in Figure 8.7.

We have seen this story play out numerous times within oil-producing regions. Discovery of oil in the North Sea put the U.K. into the oil business about 50 years ago (Figure 8.9). At first, the discovery rate was brisk, followed by 20 years of modest discovery. It appears that nothing is left to find, as discoveries have stopped. The production shows a double-peak structure—maybe echoing the discovery lull around 1980—but in any case is nearing the end of extraction. Only about 6% of the discovered oil (effectively that discovered after 1996; unshaded in Figure 8.9) is left: not much remains to pump out.

The U.S. experienced a similar history (Figure 8.10) in that discovery of conventional oil peaked around 1950, and production peaked two decades later, around 1970. Nobody wanted this to happen, although some oil geologists (notably M. King Hubbert) pointed out its inevitability based on the preceding discovery peak and simple logic.[60] The U.S. had been the largest oil producer since the dawn of the oil age, and was now slipping.[61] The peak and subsequent fall caused great anxiety and stimulated tremendous effort to find and develop additional oil resources, leading to the discovery of oil at Prudhoe Bay in Alaska—responsible for the second (lower) peak in the mid 80s. But then the decline resumed for another couple of decades, to the chagrin of many.[62]

60: Still, the prevailing attitude was one of denial, until it actually happened.

61: This is a large factor in the prosperity of the U.S.: it was the "Saudi Arabia" of the first half of the 20th century, leading oil exports and expansion of automotive transportation.

62: To reiterate a key point: it wasn't for lack of will or effort.

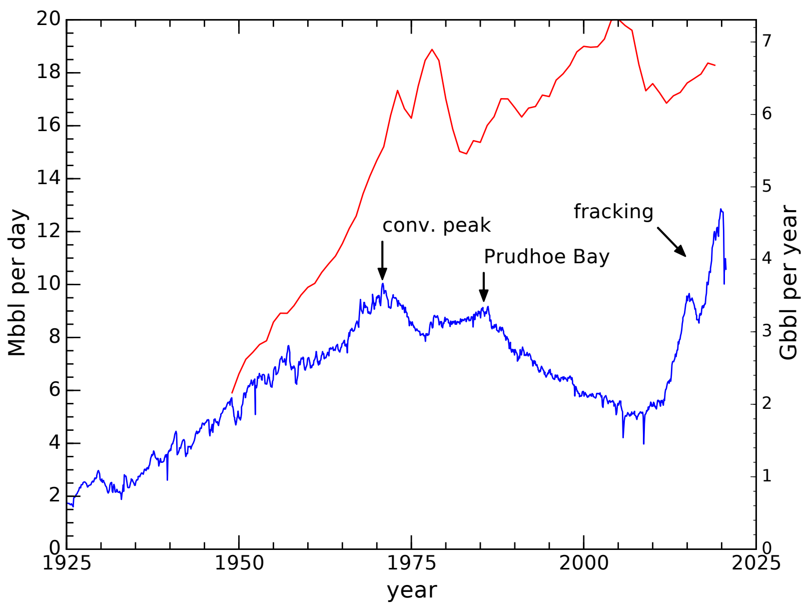


Figure 8.10: U.S. oil production history (blue; from [49]) and consumption history (red; from [34]), in both million barrels per day (left axis) and billion barrels per year (right). The conventional production peak is visible around 1970, a second peak around 1985 from Prudhoe Bay in Alaska, and finally a dramatic upturn due to hydraulic fracturing practices in the last decade. The gap between blue and red curves is made up by imports. The downturn in fracking production in 2020 coincides with the COVID pandemic, so it is not clear whether U.S. oil production will resume the climb or if we are past the peak.

Something unexpected happened next, which may serve as a cautionary tale to those who might attempt confident predictions of the future. The "fracking" boom opened access to "tight" oil deposits that were previously untenable for conventional drilling. The history is shown in Figure 8.10.

How long will the fracking boom last? One aspect to appreciate is that conventional wells take something like a decade to fully "develop,"[63] and even after individually peaking continue to deliver at diminishing rates for many years. Notice the approximate symmetry of the curve[64] in Figure 8.10 and its slow decline phase prior to 2010. Fracking "plays"[65] are fast: once the small region has been fractured and pumped, the whole process can be over in a matter of a few years. Thus it is certainly possible that the fracking boom on the right-hand-side of Figure 8.10 will end as abruptly as it started—the easy plays being exploited first, leaving less productive fields to round out the declining phase of this boom. In any case, declaring the current state of oil production in the U.S. to represent a "new normal" seems premature.

63: By develop, we mean populate the deposit with multiple drill sites and pumps.

64: Note that the curve is an aggregation of many hundreds of individual wells whose individual production rates rise and fall on shorter time scales.

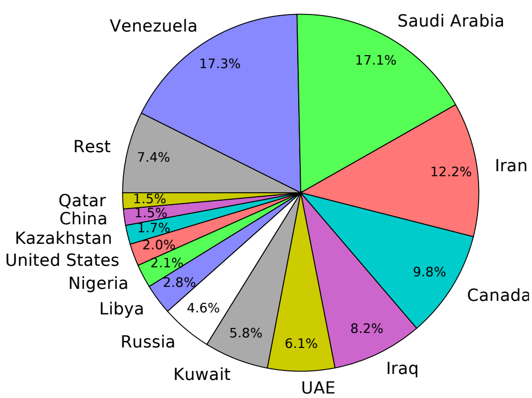
65: . . . the term for a field to be exploited

8.5.4 Geopolitics

Another wrinkle worth mentioning is the geopolitical angle. Much of the world's proven reserves are not owned by the countries having the highest oil consumption. Figure 8.11 shows which countries hold the largest stocks, with a caveat that the deposits in Venezuela and Canada are heavy oils,[66] which are harder to extract and refine into lighter forms like gasoline, making the middle-east (Saudi Arabia, Iran, Iraq, UAE, and Kuwait) the "real" leaders of light[67] crude oil dominated by the more useful shorter-chain hydrocarbon molecules like octane (Figure 8.3). One thing that should cause Americans alarm is to go around the circle looking for close allies. Aside from Canada, with its inconvenient heavy-oil, the picture is not terribly reassuring. Proven reserves of oil in the U.S. amount to 35 billion barrels. At a consumption rate of 20 million barrels per day, the math suggests only 5 years, if we only used our own supply. The proven reserve, however, is a conservative number, often short of estimated total resource: exploration can add to proven reserves. The estimated resource in the U.S. is closer to 200 Gbbl, which would last a little less than 30 years without imports at the present rate of consumption. These short timescales offer some relief for climate change concerns, but perhaps represent bad news for global economies utterly dependent on fossil fuels.

66: . . . e.g., tar sands; long-chain hydrocarbons

67: . . . sometimes called "sweet"



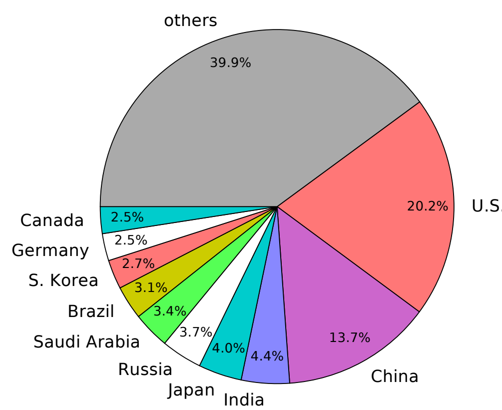


Figure 8.11: Distribution of proven oil reserves by country, on left, according to the U.S. Energy Information Administration. The oil in Venezuela and Canada is heavy oil, harder to extract and process than the light oil characteristic of the middle-east. At right is the oil consumption by country for the top ten consumers (U.S. EIA). Note that the U.S. possesses 2% of the oil, but consumes about 20% of annual production, and an overall lack of correlation between who *has* oil and who *needs* it.

Because the rate of extraction can be a limiting factor, it often happens that the rate of production begins to slow down (peaks) around the time half the resource has been exhausted,[68] producing a symmetric usage curve over time. This suggests that the peak can occur *well before* the timescales resulting from the R/P ratio, as depicted in Figure 8.6. Once the world passes the peak rate of oil production, a sequence of panic-driven damaging events could ensue, making it more difficult (less likely) for us to embark on a renewable-centered post-fossil world. Boxes Box 8.3 and Box 8.4 paint scenarios that cause concern.

68: . . . about where we appear to be on oil

Box 8.3: Resource War

Imagine the scenario in which oil prices climb from their current \$50/bbl to \$100/bbl.[69] Some major oil-producing country—seeing the writing on the wall that this precious resource is only going to become more valuable as supplies inevitably diminish—will decide that its economy was doing just fine at \$50/bbl, so can sell half as much at \$100/bbl and have the same income. Removing that oil from the market pushes oil prices up further to \$150/bbl, at which point other countries may begin playing the same game, but now selling a third as much oil for the same income. The resulting domino effect will cause international crisis, and some military power, acting as the world's police,[70] will step in to ensure continued flow of this vital global resource. Other countries possessing military strength will object to this one country's presumption and control of important segments of the global oil supply, and might potentially engage in a resource war. Sadly, this turn of events would consume massive amounts of energy and other resources to *destructive* ends, rather than channel these resources into *constructive* activities like building a post-fossil renewable energy infrastructure.

69: Oil has been as high as about \$160/bbl, in June 2008 (inflation-adjusted for 2020). A compelling argument can be made that the stress of high oil prices on many sectors of our economy provided a trigger for the financial crisis—putting an end to the growth-fueled bubble in the sub-prime housing market.

70: This hypothetical country may also have built numerous military bases in the middle-east, in anticipation of this day.

Box 8.4: The Energy Trap

If we find ourselves in a state of annual decline in energy resources—having clung too tightly to fossil fuels as a cheap and largely superior energy resource—we will have a hard time politically pulling out of the dive, because to do so means transitioning away from fossil fuels via a renewable infrastructure. But such an enormous enterprise will require substantial energy investment. And energy *is the very thing in short supply*. To embark on this transition, the society would have to voluntarily sacrifice *even more* than they already are in the energy decline crisis by diverting energy toward the decades-long initiative. The temptation to vote for a politician who would end the program and bring instant energy relief in the short term may be overwhelming. In other words, we could find ourselves in an energy trap. Witness the difficulty the world is having weaning itself off of fossil fuels despite obvious perils in the form of climate change. If it were easy, cheap, and superior to move to renewables, it would have already happened in a heartbeat. Maybe we're stuck on the flypaper.

This notion is further explored in Sec. 18.3 (p. 310).

8.6 Upshot: Amazing, Terrible, and Limited

History may very well view this time period as the Fossil Fuel Age rather than the Industrial Age. Fossil fuels are a ubiquitous and defining characteristic of this unusual time. The current level of technology, global population, or impressive state of knowledge would not have been possible without fossil fuels. We therefore owe a great debt of gratitude to these three amazing resources. Perhaps the first species on any planet to discover and use fossil fuels[71] will follow a similar madcap trajectory and even temporarily poke into space, as we have.

71: It is plausible that fossil fuels would be a common result of billions of years of evolution, resulting from buried biological matter on planets supporting rich ecosystems—Earth being the only one we know about (see Sec. 18.4; p. 312).

Yet fossil fuels bring a number of downsides, like climate change, potential population overshoot (and associated myriad pressures on the planet), pollution and environmental damage. More subtly, a near-complete dependence on fossil fuels has transformed human expectations in a way that could result in failure to adapt once they are no longer available. Superior substitutes are not guaranteed, and inferior replacements may not be gracefully adopted.

One thing we know for sure about fossil fuels is that the supply is finite. We are arguably approaching the halfway point[72] in extraction, and have naturally harvested the easiest deposits of the resources. As extraction gets harder, supply-rate (relative to demand) may become the limiting factor *well* before the R/P ratio says we will "run out" (see Box 8.5). Recall that fossil fuels are not situated in the equivalent of a single bank account permitting withdrawals of arbitrary size and speed.

72: ... especially in the hardest-to-replace oil resource

Box 8.5: Running Out One Day?

Fossil fuels will not abruptly run out one day, or even one year (see Figure 8.6). Production will taper off slowly over decades as ever-smaller deposits are harder to access and extract. In this sense, "running out" of fossil fuels will not be a sudden, jarring event in human history that sends us into a panicked chaos. Nonetheless, passing the peak and having less available with each passing year creates its own set of hardships. In the best scenario, alternatives ramp up fast enough to offset declining fossil fuel supplies. But the challenge is enormous, and success is far from guaranteed.

Given the important role the diminishing fossil resource plays in our world, today's insignificant contribution from renewable sources—as presented in Chapter 7—is all the more worrisome. This fate has been apparent to many for at least 50 years, but fossil fuel use has only continued to increase, while growth of alternatives has been lackluster. Part of the reason has to do with the low cost and amazing convenience of fossil fuels compared to alternatives.[73] Another part is lack of awareness. Sometimes old—yet no less important—stories have trouble maintaining currency in our news-oriented society.

73: One is justified in asking why prices are not raised to discourage fossil fuel use and catalyze development of alternatives. See Box 8.6.

Box 8.6: Why Not Raise the Price?

If continued reliance on fossil fuels is risky—both from resource scarcity and climate change points of view—then why do prices remain low, serving to encourage continued use and hinder adoption of alternatives? Why doesn't the government raise the price?

The rookie mistake here is assuming that adults are in charge. Markets are in charge. Governments may impose taxes and tariffs, but cannot go overboard before voters[74] object. Global competition without global government penalizes those countries self-imposing additional costs on their citizens. And finally, short-term sacrifice for long-term benefit is not a human strong suit—especially in the face of uncertainty. Convincing people of a future problem that has never surfaced for generation after generation turns out to be hard.

74: . . . in democracies, anyway

8.7 Problems

1. Make a zoom-in[75] of Figure 8.1 showing the central fossil fuel spike. You could have it "leave the floor" around 1850, reach a peak maybe at 2050 (fine for the purposes of this problem), and return to zero in symmetric fashion. Now draw—perhaps using a different color—the part of the curve you've lived through, and project out using a dotted line the part of the curve you think you'll live through (over the peak?). Now draw a segment representing your parents,[76] and do the same for your grandparents and great grandparents. You'll end up with overlapping lines.[77] Don't worry about exact dates; we're just looking for a visual impression. Has anybody you've ever met known any period but the rapid growth phase in energy you've experienced in your life?

75: Figure 8.2 is at least a good example of the left-hand side of the spike.

76: . . . or something representative of their generation

77: Just stack them a bit so you can tell them apart.

2. If you had to fill in the big question mark in Figure 8.1 with a prediction of the scenario you think is most likely to result in a few thousand years, what would you say? How do you think humans will live?[78] This is really an exercise to make us think about possibilities: no one knows the "right" answer.

78: E.g., what energy sources, primitive vs. technological, dwelling style, etc.

3. If for some reason we are grossly mistaken about the amount of fossil fuels remaining, and have 1,000 years instead of ~100 left, how qualitatively different would Figure 8.1 look?

4. Today, 20% of energy comes from non-fossil resources. Redraw Figure 8.1 under the condition that we manage to hold on to this capability indefinitely, after fossil fuels are gone.

5. Guided by Figure 8.1, how do you think humans 200 years from now will view the period from 1900–2100?

6. It is fair to say that the scientific consensus has held for a while that curtailing our use of fossil fuels would be in the best interest of the planet. From Figure 8.2, report on what has happened to

global fossil fuel use (total, not per-capita) during the last 20 years.

7. Coal usage in the U.S. has declined dramatically in the last 20 years as natural gas has replaced much of the electricity production from coal. What does Figure 8.2 say about the *global* coal trend during this period?

8. As we exploit the best coal resources first, working our way from the premium Anthracite towards Lignite (Table 8.1), will we need to mine *more* coal, or *less*, to achieve the same energy output from coal, in terms of mass removed?

9. Referring back to Fig. 7.2 (p. 105), deduce what fraction of the 38.3 qBtu electricity budget derives from coal. How much would the U.S. need to reduce its electricity dependence if we suddenly stopped using coal?

This problem has much in common with Prob. 3 (p. 111) in Chapter 7.

10. Octane is C_8H_{18}. On either side is heptane and nonane, containing 7 and 9 carbons, respectively. Referring only to Figure 8.3 and recognizing the pattern, what would the chemical formulas for heptane and nonane be, in the form of C_xH_y?

11. The U.S. uses approximately 20 million barrels of oil per day, and has a population of about 330 million people. On average, then, how many barrels *per year* is one person responsible for consuming?[79]

79: Even if not personally consuming this much, it is used on behalf of individuals to provide goods and services for them.

12. If an average American is responsible for consuming a barrel[80] of oil every 18 days, what power does this correspond to, in Watts?

80: A barrel contains about 6.1 GJ of energy.

13. Using the values in Table 8.2, compute the energy content of a gallon of gasoline assuming that octane (C_8H_{18}) is a good representative, energetically. Express your answer in both MJ and kWh. One gallon is 3.785 L and in the case of gasoline has a mass of 2.8 kg.[81]

81: The density of gasoline is about 0.75 times that of water.

14. Every day, Americans use about 9×10^8 J of energy per person. Since we know that 37%, 13%, and 31% of this comes from oil, coal, and gas, respectively, use Table 8.2 to figure out how much mass of each is used per day on American's behalfs, and take a moment to compare to equivalent-mass volumes of water to provide familiar context.

15. What if we could get our energy from drinking gasoline?[82] Referring to Table 8.2, how many grams of gasoline[83] would we have to drink daily to satisfy the typical 2,000 kcal/day diet? How much volume does this represent if gasoline is 0.75 g/mL? Relate this to a familiar container for holding liquids.[84]

82: Don't do this! It won't work!

83: . . . represented by octane, C_8H_{18}

84: For reference, 100 mL is 3.4 oz.

16. One liter of gasoline (1,000 mL) has a mass of about 750 g and contains about 9.7 kWh of energy. Meanwhile, a typical AA battery