# 4 Engines and Refrigerators

4.1 Heat Engines

A **heat engine** is any device that absorbs heat and converts part of that energy into work. An important example is the steam turbine, used to generate electricity in most of today's power plants. The familiar internal combustion engine used in automobiles does not actually absorb heat, but we can pretend that the thermal energy comes from outside rather than inside and treat it, also, as a heat engine.

Unfortunately, only *part* of the energy absorbed as heat can be converted to work by a heat engine. The reason is that the heat, as it flows in, brings along entropy, which must somehow be disposed of before the cycle can start over. To get rid of the entropy, every heat engine must dump some waste heat into its environment. The work produced by the engine is the difference between the heat absorbed and the waste heat expelled.

My goal in this section is to make these ideas precise, and to determine exactly how much of the heat absorbed by an engine can be converted into work. Amazingly, we can say a great deal without knowing *anything* about how the engine actually works.

Figure 4.1 shows the flow of energy into and out of a heat engine. The heat absorbed by the engine comes from a place called the **hot reservoir**, while the waste heat is dumped into the **cold reservoir**. The temperatures of these reservoirs, T_h and T_c, are assumed fixed. (In general, a **reservoir** in thermodynamics is anything that's so large that its temperature doesn't change noticeably when heat enters or leaves. For a steam engine, the hot reservoir is the place where the fuel is burned and the cold reservoir is the surrounding environment.) I'll use the symbol Q_h for the heat absorbed from the hot reservoir in some given time period, and Q_c for the heat expelled to the cold reservoir. The net work done by the engine during this time will be W. All three of these symbols will represent positive quantities; in this chapter I'm departing from my earlier sign conventions for heat and work.

Figure 4.1. Energy-flow diagram for a heat engine. Energy enters as heat from the hot reservoir, and leaves both as work and as waste heat expelled to the cold reservoir.
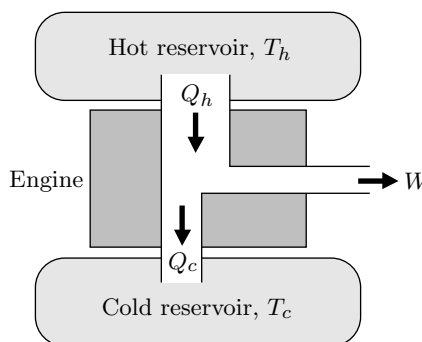
The benefit of operating a heat engine is the work produced, W. The cost of operation is the heat absorbed, Q_h. Let me therefore define the **efficiency** of an engine, e, as the benefit/cost ratio:

$$e \equiv \frac{\text{benefit}}{\text{cost}} = \frac{W}{Q_h}. \tag{4.1}$$

The question I would like to ask is this: For given values of T_h and T_c, what is the maximum possible efficiency? To answer this question, all we need are the first and second laws of thermodynamics, plus the assumption that the engine operates in cycles, returning to its original state at the end of each cycle of operation.

The first law of thermodynamics tells us that energy is conserved. Since the state of the engine must be unchanged at the end of a cycle, the energy it absorbs must be precisely equal to the energy it expels. In our notation,

$$Q_h = Q_c + W. \tag{4.2}$$

If we use this equation to eliminate W in equation 4.1, we have for the efficiency

$$e = \frac{Q_h - Q_c}{Q_h} = 1 - \frac{Q_c}{Q_h}. \tag{4.3}$$

Thus the efficiency cannot be greater than 1, and can equal 1 only if $Q_c = 0$.

To proceed further we must also invoke the second law, which tells us that the total entropy of the engine plus its surroundings can increase but not decrease. Since the state of the engine must be unchanged at the end of a cycle, the entropy it expels must be at least as much as the entropy it absorbs. (In this context, as in Section 3.2, I like to imagine entropy as a fluid that can be created but never destroyed.) Now the entropy extracted from the hot reservoir is just Q_h/T_h, while the entropy expelled to the cold reservoir is Q_c/T_c. So the second law tells us

$$\frac{Q_c}{T_c} \geq \frac{Q_h}{T_h}, \qquad \text{or} \qquad \frac{Q_c}{Q_h} \geq \frac{T_c}{T_h}. \tag{4.4}$$

Plugging this result into equation 4.3, we conclude

$$e \leq 1 - \frac{T_c}{T_h}. \tag{4.5}$$

This is our desired result. So, for instance, if $T_h = 500$ K and $T_c = 300$ K, the max-

imum possible efficiency is 40%. In general, for the greatest maximum efficiency, you should make the cold reservoir *very* cold, or the hot reservoir *very* hot, or both. The smaller the ratio T_c/T_h, the more efficient your engine can be.

It's easy to make an engine that's *less* efficient than the limit $1 - T_c/T_h$, simply by producing additional entropy during the operation. Then to dispose of this entropy you must dump extra heat into the cold reservoir, leaving less energy to convert to work. The most obvious way of producing new entropy is in the heat transfer processes themselves. For instance, when heat Q_h leaves the hot reservoir, the entropy lost by that reservoir is Q_h/T_h; but if the engine temperature at this time is *less* than T_h, then as the heat enters the engine its associated entropy will be *greater* than Q_h/T_h.

In deriving the limit (4.5) on the efficiency of an engine, we used both the first and second laws of thermodynamics. The first law told us that the efficiency can't be any greater than 1, that is, we can't get more work out than the amount of heat put in. In this context, the first law is often paraphrased, "You can't win." The second law, however, made matters worse. It told us that we can't even achieve $e = 1$ unless $T_c = 0$ or $T_h = \infty$, both of which are impossible in practice. In this context, the second law is often paraphrased, "You can't even break even."

Problem 4.1. Recall Problem 1.34, which concerned an ideal diatomic gas taken around a rectangular cycle on a PV diagram. Suppose now that this system is used as a heat engine, to convert the heat added into mechanical work.

   (a) Evaluate the efficiency of this engine for the case $V_2 = 3V_1$, $P_2 = 2P_1$.

   (b) Calculate the efficiency of an "ideal" engine operating between the same temperature extremes.

Problem 4.2. At a power plant that produces 1 GW (10^9 watts) of electricity, the steam turbines take in steam at a temperature of $500°$C, and the waste heat is expelled into the environment at $20°$C.

   (a) What is the maximum possible efficiency of this plant?

   (b) Suppose you develop a new material for making pipes and turbines, which allows the maximum steam temperature to be raised to $600°$C. Roughly how much money can you make in a year by installing your improved hardware, if you sell the additional electricity for 5 cents per kilowatt-hour? (Assume that the amount of fuel consumed at the plant is unchanged.)

Problem 4.3. A power plant produces 1 GW of electricity, at an efficiency of 40% (typical of today's coal-fired plants).

   (a) At what rate does this plant expel waste heat into its environment?

   (b) Assume first that the cold reservoir for this plant is a river whose flow rate is 100 m^3/s. By how much will the temperature of the river increase?

   (c) To avoid this "thermal pollution" of the river, the plant could instead be cooled by evaporation of river water. (This is more expensive, but in some areas it is environmentally preferable.) At what rate must the water evaporate? What fraction of the river must be evaporated?

Problem 4.4. It has been proposed to use the thermal gradient of the ocean to drive a heat engine. Suppose that at a certain location the water temperature is $22°$C at the ocean surface and $4°$C at the ocean floor.

(a) What is the maximum possible efficiency of an engine operating between these two temperatures?

(b) If the engine is to produce 1 GW of electrical power, what minimum volume of water must be processed (to suck out the heat) in every second?

The Carnot Cycle

Let me now explain how to make an engine that *does* achieve the maximum possible efficiency for a given T_h and T_c.

Every engine has a so-called "working substance," which is the material that actually absorbs heat, expels waste heat, and does work. In many heat engines the working substance is a gas. Imagine, then, that we first want the gas to absorb some heat Q_h from the hot reservoir. In the process, the entropy of the reservoir decreases by Q_h/T_h, while the entropy of the gas increases by Q_h/T_{gas}. To avoid making any *new* entropy, we would need to make $T_{\mathrm{gas}} = T_h$. This isn't quite possible, because heat won't flow between objects at the same temperature. So let's make T_{gas} just slightly less than T_h, and keep the gas at this temperature (by letting it expand) as it absorbs the heat. This step of the cycle, then, requires that the gas expand isothermally.

Similarly, during the portion of the cycle when the gas is dumping the waste heat into the cold reservoir, we want its temperature to be only infinitesimally greater than T_c, to avoid creating any new entropy. And as the heat leaves the gas, we need to compress it isothermally to keep it at this temperature.

So we have an isothermal expansion at a temperature just less than T_h, and an isothermal compression at a temperature just greater than T_c. The only remaining question is how we *get* the gas from one temperature to the other and back. We don't want any heat to flow in or out when the gas is at intermediate temperatures, so these intermediate steps must be adiabatic. The entire cycle consists of four steps, illustrated in Figures 4.2 and 4.3: isothermal expansion at T_h, adiabatic expansion from T_h to T_c, isothermal compression at T_c, and adiabatic compression from T_c back up to T_h. The theoretical importance of this cycle was first pointed out by Sadi Carnot in 1824, so the cycle is now known as the **Carnot cycle**.

It is possible to prove directly, from the formulas for isothermal and adiabatic processes in Section 1.5, that an ideal gas taken around a Carnot cycle realizes the maximum possible efficiency $1 - T_c/T_h$. But while the proof makes an interesting exercise (see Problem 4.5), it is not really necessary once one understands entropy and the second law. As long as we know that no new entropy was created during the cycle, the strict equality must hold in equation 4.4, and therefore the efficiency must be the maximum allowed by equation 4.5. This conclusion holds even if the gas isn't ideal, and, for that matter, even if the working substance isn't a gas at all.

Although a Carnot cycle is very *efficient*, it's also horribly *impractical*. The heat flows so slowly during the isothermal steps that it takes forever to get a significant amount of work out of the engine. So don't bother installing a Carnot engine in your car; while it would increase your gas mileage, you'd be passed on the highway by pedestrians.
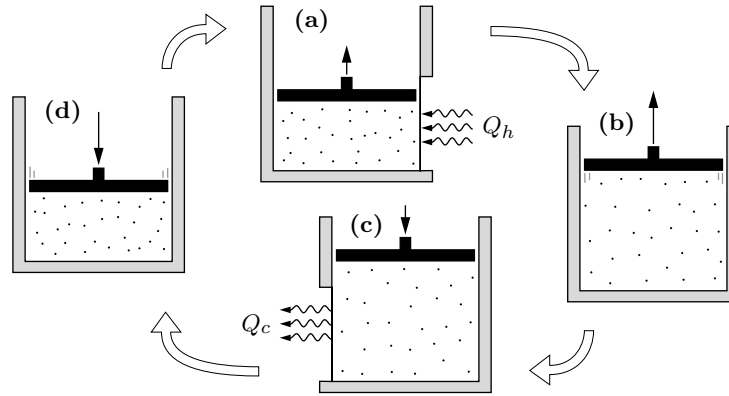
Figure 4.2. The four steps of a Carnot cycle: (a) isothermal expansion at T_h while absorbing heat; (b) adiabatic expansion to T_c; (c) isothermal compression at T_c while expelling heat; and (d) adiabatic compression back to T_h. The system must be put in thermal contact with the hot reservoir during step (a) and with the cold reservoir during step (c).
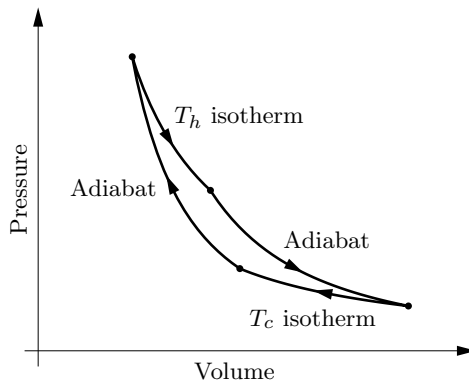

Figure 4.3. PV diagram for an ideal monatomic gas undergoing a Carnot cycle.

Problem 4.5. Prove directly (by calculating the heat taken in and the heat expelled) that a Carnot engine using an ideal gas as the working substance has an efficiency of $1 - T_c/T_h$.

Problem 4.6. To get more than an infinitesimal amount of power out of a Carnot engine, we would have to keep the temperature of its working substance below that of the hot reservoir and above that of the cold reservoir by non-infinitesimal amounts. Consider, then, a Carnot cycle in which the working substance is at temperature T_{hw} as it absorbs heat from the hot reservoir, and at temperature T_{cw} as it expels heat to the cold reservoir. Under most circumstances the rates of heat transfer will be directly proportional to the temperature differences:

$$\frac{Q_h}{\Delta t} = K(T_h - T_{hw}) \qquad \text{and} \qquad \frac{Q_c}{\Delta t} = K(T_{cw} - T_c).$$

I've assumed here for simplicity that the constants of proportionality (K) are the same for both of these processes. Let us also assume that both processes take the

same amount of time, so the Δt's are the same in both of these equations.[*]

(a) Assuming that no new entropy is created during the cycle except during the two heat transfer processes, derive an equation that relates the four temperatures T_h, T_c, T_{hw}, and T_{cw}.

(b) Assuming that the time required for the two adiabatic steps is negligible, write down an expression for the power (work per unit time) output of this engine. Use the first and second laws to write the power entirely in terms of the four temperatures (and the constant K), then eliminate T_{cw} using the result of part (a).

(c) When the cost of building an engine is much greater than the cost of fuel (as is often the case), it is desirable to optimize the engine for maximum power output, not maximum efficiency. Show that, for fixed T_h and T_c, the expression you found in part (b) has a maximum value at $T_{hw} = \frac{1}{2}(T_h + \sqrt{T_h T_c})$. (Hint: You'll have to solve a quadratic equation.) Find the corresponding expression for T_{cw}.

(d) Show that the efficiency of this engine is $1 - \sqrt{T_c/T_h}$. Evaluate this efficiency numerically for a typical coal-fired steam turbine with $T_h = 600°$C and $T_c = 25°$C, and compare to the ideal Carnot efficiency for this temperature range. Which value is closer to the actual efficiency, about 40%, of a real coal-burning power plant?

4.2 Refrigerators

A **refrigerator** is a heat engine operated in reverse, more or less. In practice, it may work in a completely different way, but if you only care about what it *does*, not how it works, you can just reverse the arrows in Figure 4.1 to obtain a generalized diagram of a refrigerator, shown in Figure 4.4. Again I'm defining all symbols to stand for positive quantities. The heat sucked out of the cold reservoir (the inside of the fridge) is Q_c, while the electrical energy supplied from the wall outlet is W.

Figure 4.4. Energy-flow diagram for a refrigerator or air conditioner. For a kitchen refrigerator, the space inside it is the cold reservoir and the space outside it is the hot reservoir. An electrically powered compressor supplies the work.

---
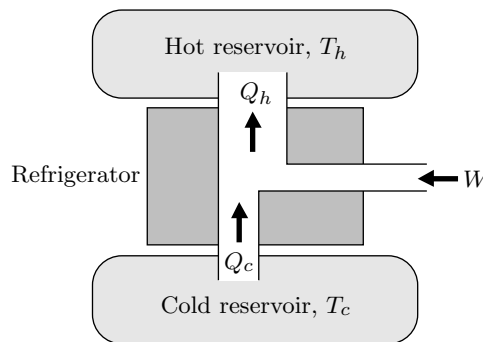
[*]Neither of these assumptions is necessary in order to obtain the final result for the efficiency in part (d). See the article on which this problem is based: F. L. Curzon and B. Ahlborn, "Efficiency of a Carnot engine at maximum power output," *American Journal of Physics* **41**, 22–24 (1975).

There's also some waste heat, Q_h, dumped into your kitchen. By the way, the same diagram could apply to an air conditioner; then the cold reservoir is the inside of your house while the hot reservoir is outside.*

How should we define the "efficiency" of a refrigerator? Again the relevant number is the benefit/cost ratio, but this time the benefit is Q_c while the cost is W. To avoid confusion with equation 4.1, this ratio is called the **coefficient of performance**:

$$\text{COP} = \frac{\text{benefit}}{\text{cost}} = \frac{Q_c}{W}. \tag{4.6}$$

Just as for a heat engine, we can now use the first and second laws to derive a limit on the COP in terms of the temperatures T_h and T_c. The first law tells us $Q_h = Q_c + W$, so

$$\text{COP} = \frac{Q_c}{Q_h - Q_c} = \frac{1}{Q_h/Q_c - 1}. \tag{4.7}$$

Notice that there's no obvious upper limit on this quantity yet; in particular, the first law allows the COP to be greater than 1.

Meanwhile, the second law says that the entropy dumped into the hot reservoir must be at least as much as the entropy absorbed from the cold reservoir:

$$\frac{Q_h}{T_h} \geq \frac{Q_c}{T_c}, \qquad \text{or} \qquad \frac{Q_h}{Q_c} \geq \frac{T_h}{T_c}. \tag{4.8}$$

(This relation is the reverse of relation 4.4 because the entropy is flowing in the opposite direction.) Plugging this inequality into equation 4.7 gives

$$\text{COP} \leq \frac{1}{T_h/T_c - 1} = \frac{T_c}{T_h - T_c}. \tag{4.9}$$

For a typical kitchen refrigerator (with freezer), T_h might be 298 K while T_c might be 255 K. In this case the coefficient of performance can be as high as 5.9. In other words, for each joule of electrical energy drawn from the wall, the coolant can suck as much as 5.9 J of heat from the inside of the refrigerator/freezer. In this ideal case, the waste heat dumped into the kitchen would be 6.9 J. As you can see from the formula, the COP is largest when T_h and T_c aren't very different. A refrigerator that cools something down to liquid helium temperature (4 K) would have to be *much* less efficient.

To make an ideal refrigerator with the maximum possible COP, one can again use a Carnot cycle, this time operated in reverse. In order to make the heat flow in the opposite direction, the working substance must be slightly *hotter* than T_h while heat is being expelled, and slightly *colder* than T_c while heat is being absorbed. Once again, this is a lousy way to do things in practice, because the heat transfer is much too slow. A more practical refrigerator is described in Section 4.4.

---

*An air conditioner usually also has a fan, which blows air around inside your house to speed up the heat flow on that side. Don't confuse the air (which never leaves the cold reservoir) with the heat (which would flow outward, though more slowly, even without the fan).

Historically, heat engines and refrigerators played a crucial role in the formulation of the second law and the identification of entropy as a quantity of interest. Early versions of the second law, derived from experience, included the statements that all heat engines must produce some waste heat, and that all refrigerators require some work input. Carnot and others invented ingenious arguments to show that these laws could be violated if you could make an engine or a refrigerator whose efficiency exceeded that of a Carnot cycle (see Problems 4.16 and 4.17). Carnot also recognized that for an ideal engine there must be a quantity, associated with heat, that flows in from the hot reservoir and out to the cold reservoir in equal amounts. But Carnot's 1824 memoir did not distinguish carefully enough between this quantity and what we now call simply "heat." At that time the relation between heat and other forms of energy was still controversial, and the simple formula Q/T eluded scientists who had not yet adopted a temperature scale measured from absolute zero. It wasn't until 1865, after these other issues were fully resolved, that Rudolf Clausius brought Carnot's quantity to the full attention of the scientific community and put it on a firm mathematical basis. He coined the term "entropy" for this quantity, after a Greek word meaning "transformation" (and because the word resembles "energy"). Clausius did not explain what entropy actually *is*, however. Ludwig Boltzmann took up that question during the following years, and had it figured out by 1877.

Problem 4.7. Why must you put an air conditioner in the window of a building, rather than in the middle of a room?

Problem 4.8. Can you cool off your kitchen by leaving the refrigerator door open? Explain.

Problem 4.9. Estimate the maximum possible COP of a household air conditioner. Use any reasonable values for the reservoir temperatures.

Problem 4.10. Suppose that heat leaks into your kitchen refrigerator at an average rate of 300 watts. Assuming ideal operation, how much power must it draw from the wall?

Problem 4.11. What is the maximum possible COP for a cyclic refrigerator operating between a high-temperature reservoir at 1 K and a low-temperature reservoir at 0.01 K?

Problem 4.12. Explain why an ideal gas taken around a rectangular PV cycle, as considered in Problems 1.34 and 4.1, cannot be used (in reverse) for refrigeration.

Problem 4.13. Under many conditions, the rate at which heat enters an air conditioned building on a hot summer day is proportional to the difference in temperature between inside and outside, $T_h - T_c$. (If the heat enters entirely by conduction, this statement will certainly be true. Radiation from direct sunlight would be an exception.) Show that, under these conditions, the cost of air conditioning should be roughly proportional to the *square* of the temperature difference. Discuss the implications, giving a numerical example.

Problem 4.14. A **heat pump** is an electrical device that heats a building by pumping heat in from the cold outside. In other words, it's the same as a refrigerator, but its purpose is to warm the hot reservoir rather than to cool the cold reservoir (even though it does both). Let us define the following standard symbols, all taken to be positive by convention:

$$T_h = \text{temperature inside building}$$
$$T_c = \text{temperature outside}$$
$$Q_h = \text{heat pumped into building in 1 day}$$
$$Q_c = \text{heat taken from outdoors in 1 day}$$
$$W = \text{electrical energy used by heat pump in 1 day}$$

(a) Explain why the "coefficient of performance" (COP) for a heat pump should be defined as Q_h/W.

(b) What relation among Q_h, Q_c, and W is implied by energy conservation alone? Will energy conservation permit the COP to be greater than 1?

(c) Use the second law of thermodynamics to derive an upper limit on the COP, in terms of the temperatures T_h and T_c alone.

(d) Explain why a heat pump is better than an electric furnace, which simply converts electrical work directly into heat. (Include some numerical estimates.)

Problem 4.15. In an **absorption refrigerator**, the energy driving the process is supplied not as work, but as heat from a gas flame. (Such refrigerators commonly use propane as fuel, and are used in locations where electricity is unavailable.[*]) Let us define the following symbols, all taken to be positive by definition:

$$Q_f = \text{heat input from flame}$$
$$Q_c = \text{heat extracted from inside refrigerator}$$
$$Q_r = \text{waste heat expelled to room}$$
$$T_f = \text{temperature of flame}$$
$$T_c = \text{temperature inside refrigerator}$$
$$T_r = \text{room temperature}$$

(a) Explain why the "coefficient of performance" (COP) for an absorption refrigerator should be defined as Q_c/Q_f.

(b) What relation among Q_f, Q_c, and Q_r is implied by energy conservation alone? Will energy conservation permit the COP to be greater than 1?

(c) Use the second law of thermodynamics to derive an upper limit on the COP, in terms of the temperatures T_f, T_c, and T_r alone.

Problem 4.16. Prove that *if* you had a heat engine whose efficiency was better than the ideal value (4.5), you could hook it up to an ordinary Carnot refrigerator to make a refrigerator that requires no work input.

---

[*]For an explanation of how an absorption refrigerator actually works, see an engineering thermodynamics textbook such as Moran and Shapiro (1995).

Problem 4.17. Prove that *if* you had a refrigerator whose COP was better than the ideal value (4.9), you could hook it up to an ordinary Carnot engine to make an engine that produces no waste heat.

4.3 Real Heat Engines

The previous sections treated heat engines and refrigerators in an idealized way, arriving at theoretical limits on their performance. These theoretical limits are extremely useful, because they tell us generally how the efficiency of an engine or refrigerator tends to depend on its operating temperatures. The limits also serve as benchmarks for judging the efficiency of any real engine or refrigerator. For instance, if you have an engine that operates between $T_c = 300$ K and $T_h = 600$ K, and its efficiency is 45%, you know there's not much point in trying to improve the design further since the highest possible efficiency is only 50%.

You may be wondering, however, how real engines and refrigerators are actually built. This is a vast subject, but in this section and the next I'll describe a few examples of real engines and refrigerators, to alleviate the abstraction of the preceding sections.

Internal Combustion Engines

Let's start with the familiar gasoline engine found in most automobiles. The working substance is a gas, initially a mixture of air and vaporized gasoline. This mixture is first injected into a cylinder and compressed, adiabatically, by a piston. A spark plug then ignites the mixture, raising its temperature and pressure while the volume doesn't change. Next the high-pressure gas pushes the piston outward, expanding adiabatically and producing mechanical work. Finally, the hot exhaust gases are expelled and replaced by a new mixture at lower temperature and pressure. The whole cycle is shown in Figure 4.5, where I've represented the exhaust/replacement step as if it were a simple lowering of pressure due to the extraction of heat. What actually happens is the piston pushes the old mixture out through a valve and pulls a new mixture in through another valve, expelling heat but doing no net work. This cycle is called the **Otto cycle**, after the German inventor Nikolaus August Otto.

Figure 4.5. The idealized Otto cycle, an approximation of what happens in a gasoline engine. In real engines the compression ratio V_1/V_2 is larger than shown here, typically 8 or 10.
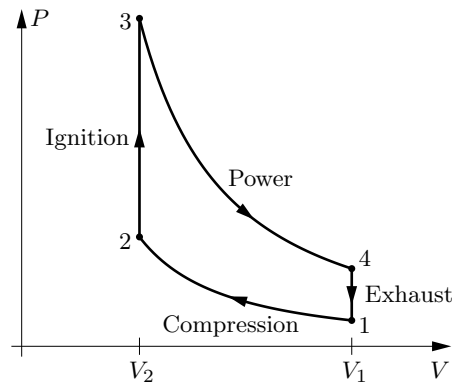
Notice that there is no "hot reservoir" connected to this engine. Instead, thermal energy is produced internally by burning the fuel. The result of this burning, however, is a gas at high temperature and pressure, exactly as if it had absorbed heat from an external source.

The efficiency of a gasoline engine is the net work produced during the cycle divided by the "heat" absorbed during the ignition step. Assuming that the gas is ideal, it's not particularly hard to express these quantities in terms of the various temperatures and volumes (see Problem 4.18). The result is fairly simple:

$$e = 1 - \left(\frac{V_2}{V_1}\right)^{\gamma-1}, \tag{4.10}$$

where V_1/V_2 is the **compression ratio** and γ is the adiabatic exponent introduced in Section 1.5. For air, $\gamma = 7/5$, while a typical compression ratio might be 8, yielding a theoretical efficiency of $1 - (1/8)^{2/5} = 0.56$. This is good, but not as good as a Carnot engine operating between the same extreme temperatures. To compare the two, recall that during an adiabatic process, $TV^{\gamma-1}$ is constant. We can therefore eliminate the volumes in equation 4.10 in favor of the temperatures at the ends of either adiabatic step:

$$e = 1 - \frac{T_1}{T_2} = 1 - \frac{T_4}{T_3}. \tag{4.11}$$

Either of these temperature ratios is greater than the ratio of the extreme temperatures, T_1/T_3, that appears in the Carnot formula. The Otto engine is therefore less efficient than the Carnot engine. (In practice, a real gasoline engine is still less efficient, because of friction, conductive heat loss, and incomplete combustion of the fuel. Today's automobile engines typically achieve efficiencies of about 20–30%.)

The obvious way to make a gasoline engine more efficient would be to use a higher compression ratio. Unfortunately, if the fuel mixture becomes too hot it will "preignite" spontaneously before the compression step is complete, causing the pressure to jump upward before point 2 in the cycle is reached. Preignition is avoided in the **Diesel engine** by compressing only air, then spraying fuel into the cylinder after the air is hot enough to ignite the fuel. The spraying/ignition is done as the piston begins to move outward, at a rate that is adjusted to maintain approximately constant pressure. An idealized version of the Diesel cycle is shown in Figure 4.6. One can derive a rather complicated formula for the efficiency of the Diesel cycle in terms of the compression ratio V_1/V_2 and the **cutoff ratio**, V_3/V_2. For a given compression ratio the efficiency is actually less than that of the Otto cycle, but Diesel engines generally have higher compression ratios (typically around 20) and hence higher efficiencies (up to about 40% in practice). As far as I know, the only limit on the compression ratio of a Diesel engine comes from the strength and melting point of the material from which it is made. A ceramic engine could in principle withstand higher temperatures and therefore achieve higher efficiency.

Figure 4.6. *PV* diagram for the Diesel cycle.
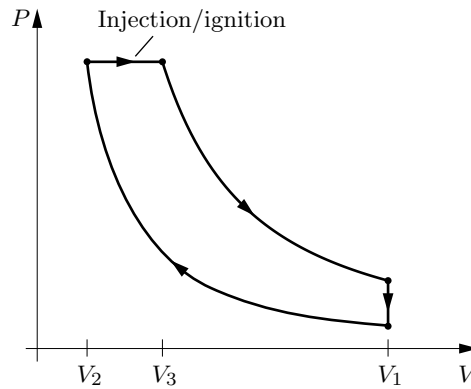
Problem 4.18. Derive equation 4.10 for the efficiency of the Otto cycle.

Problem 4.19. The amount of work done by each stroke of an automobile engine is controlled by the amount of fuel injected into the cylinder: the more fuel, the higher the temperature and pressure at points 3 and 4 in the cycle. But according to equation 4.10, the *efficiency* of the cycle depends only on the compression ratio (which is always the same for any particular engine), not on the amount of fuel consumed. Do you think this conclusion still holds when various other effects such as friction are taken into account? Would you expect a real engine to be most efficient when operating at high power or at low power? Explain.

Problem 4.20. Derive a formula for the efficiency of the Diesel cycle, in terms of the compression ratio V_1/V_2 and the cutoff ratio V_3/V_2. Show that for a given compression ratio, the Diesel cycle is less efficient than the Otto cycle. Evaluate the theoretical efficiency of a Diesel engine with a compression ratio of 18 and a cutoff ratio of 2.

Problem 4.21. The ingenious **Stirling engine** is a true heat engine that absorbs heat from an external source. The working substance can be air or any other gas. The engine consists of *two* cylinders with pistons, one in thermal contact with each reservoir (see Figure 4.7). The pistons are connected to a crankshaft in a complicated way that we'll ignore and let the engineers worry about. Between the two cylinders is a passageway where the gas flows past a **regenerator**: a temporary heat reservoir, typically made of wire mesh, whose temperature varies
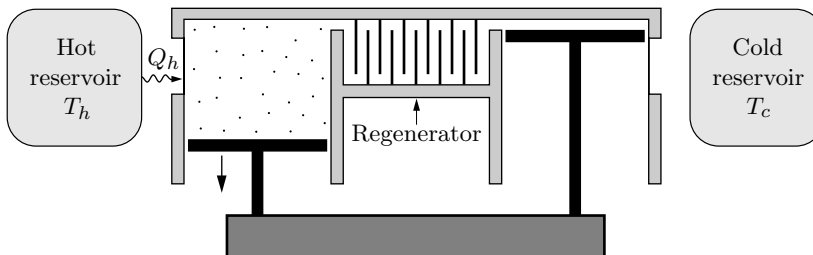

Figure 4.7. A Stirling engine, shown during the power stroke when the hot piston is moving outward and the cold piston is at rest. (For simplicity, the linkages between the two pistons are not shown.)

gradually from the hot side to the cold side. The heat capacity of the regenerator is very large, so its temperature is affected very little by the gas flowing past. The four steps of the engine's (idealized) cycle are as follows:

  i. Power stroke. While in the hot cylinder at temperature T_h, the gas absorbs heat and expands isothermally, pushing the hot piston outward. The piston in the cold cylinder remains at rest, all the way inward as shown in the figure.

 ii. Transfer to the cold cylinder. The hot piston moves in while the cold piston moves out, transferring the gas to the cold cylinder at constant volume. While on its way, the gas flows past the regenerator, giving up heat and cooling to T_c.

iii. Compression stroke. The cold piston moves in, isothermally compressing the gas back to its original volume as the gas gives up heat to the cold reservoir. The hot piston remains at rest, all the way in.

 iv. Transfer to hot cylinder. The cold piston moves the rest of the way in while the hot piston moves out, transferring the gas back to the hot cylinder at constant volume. While on its way, the gas flows past the regenerator, absorbing heat until it is again at T_h.

   (a) Draw a PV diagram for this idealized Stirling cycle.

   (b) Forget about the regenerator for the moment. Then, during step 2, the gas will give up heat to the cold reservoir instead of to the regenerator; during step 4, the gas will absorb heat from the hot reservoir. Calculate the efficiency of the engine in this case, assuming that the gas is ideal. Express your answer in terms of the temperature ratio T_c/T_h and the compression ratio (the ratio of the maximum and minimum volumes). Show that the efficiency is less than that of a Carnot engine operating between the same temperatures. Work out a numerical example.

   (c) Now put the regenerator back. Argue that, if it works perfectly, the efficiency of a Stirling engine is the same as that of a Carnot engine.

   (d) Discuss, in some detail, the various advantages and disadvantages of a Stirling engine, compared to other engines.

The Steam Engine

A very different type of engine is the **steam engine**, ubiquitous in the 19th century and still used today in large power plants. The steam does work by pushing a piston or a turbine, while the heat is provided by burning a fossil fuel or fissioning uranium. A schematic diagram of the cycle is shown in Figure 4.8, along with an idealized PV diagram for the cycle (called the **Rankine cycle**). Starting at point 1, water is pumped to high pressure (2) and then flows into a boiler, where heat is added at constant pressure. At point 3 the steam hits the turbine, where it expands adiabatically, cools, and ends up at the original low pressure (4). Finally the partially condensed fluid is cooled further in a "condenser" (a network of pipes that are in good thermal contact with the low-temperature reservoir).

The working substance in a steam engine is most definitely *not* an ideal gas—it condenses into a liquid during the cycle! Because of this complicated behavior, there's no way to calculate the efficiency of the cycle straight from the PV diagram. However, if you know the pressures everywhere and the temperature at point 3 you can look up the data needed to compute the efficiency in what are called "steam tables."
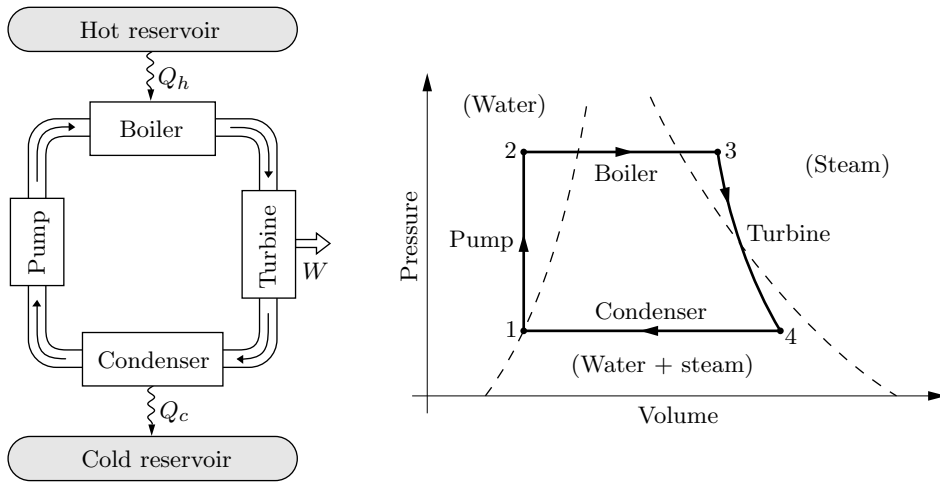
Figure 4.8. Schematic diagram of a steam engine and the associated PV cycle (not to scale), called the **Rankine cycle**. The dashed lines show where the fluid is liquid water, where it is steam, and where it is part water and part steam.

Recall from Section 1.6 that the **enthalpy** of a substance is $H = U + PV$, its energy plus the work needed to make room for it in a constant-pressure environment. Therefore the change in enthalpy is equal to the heat absorbed under constant-pressure conditions (assuming that no "other" work is done). In the Rankine cycle, heat is absorbed at constant pressure in the boiler and expelled at constant pressure in the condenser, so we can write the efficiency as

$$ e = 1 - \frac{Q_c}{Q_h} = 1 - \frac{H_4 - H_1}{H_3 - H_2} \approx 1 - \frac{H_4 - H_1}{H_3 - H_1}. \tag{4.12} $$

The last approximation, $H_2 \approx H_1$, is pretty good because the pump adds very little energy to the water, while the PV term in H is very small for liquids in comparison to gases.

Two different tables are generally used to look up the needed values of H. The first (see Table 4.1) is for "saturated" water and steam, covering any point between the dashed lines on the PV diagram, where the temperature is determined by the pressure. This table lists the enthalpy and entropy for pure water and for pure steam at the boiling point; for mixtures of water and steam one can interpolate between these two values.

The other table needed (see Table 4.2) is for "superheated" steam, in the right-most region of the PV diagram where the pressure and temperature must be specified separately. Again, the table lists the enthalpy and entropy at each point.

To compute the efficiency of a Rankine cycle, we need the enthalpies at points 1, 3, and 4. The enthalpy at point 1 can be looked up in Table 4.1, while the enthalpy at point 3 can be looked up in Table 4.2. To locate point 4 we use the fact that the expansion of the steam in the turbine is approximately adiabatic ($Q = 0$), so that ideally its entropy does not change during this step. We can look up the entropy

T (°C)	P (bar)	H_{water} (kJ)	H_{steam} (kJ)	S_{water} (kJ/K)	S_{steam} (kJ/K)
0	0.006	0	2501	0	9.156
10	0.012	42	2520	0.151	8.901
20	0.023	84	2538	0.297	8.667
30	0.042	126	2556	0.437	8.453
50	0.123	209	2592	0.704	8.076
100	1.013	419	2676	1.307	7.355

Table 4.1. Properties of saturated water/steam. Pressures are given in bars, where 1 bar = 10^5 Pa $\approx$ 1 atm. All values are for 1 kg of fluid, and are measured relative to liquid water at the triple point (0.01°C and 0.006 bar). Excerpted from Keenan et al. (1978).

P (bar)		Temperature (°C) 200	300	400	500	600
1.0	H (kJ)	2875	3074	3278	3488	3705
	S (kJ/K)	7.834	8.216	8.544	8.834	9.098
3.0	H (kJ)	2866	3069	3275	3486	3703
	S (kJ/K)	7.312	7.702	8.033	8.325	8.589
10	H (kJ)	2828	3051	3264	3479	3698
	S (kJ/K)	6.694	7.123	7.465	7.762	8.029
30	H (kJ)		2994	3231	3457	3682
	S (kJ/K)		6.539	6.921	7.234	7.509
100	H (kJ)			3097	3374	3625
	S (kJ/K)			6.212	6.597	6.903
300	H (kJ)			2151	3081	3444
	S (kJ/K)			4.473	5.791	6.233

Table 4.2. Properties of superheated steam. All values are for 1 kg of fluid, and are measured relative to liquid water at the triple point. Excerpted from Keenan et al. (1978).

at point 3 in Table 4.2, then interpolate in Table 4.1 to find what mixture of liquid and gas has the same entropy at the lower pressure.

For example, suppose that the cycle operates between a minimum pressure of 0.023 bar (where the boiling temperature is 20°C) and a maximum pressure of 300 bars, with a maximum superheated steam temperature of 600°C. Then for each kilogram of water/steam, $H_1 = 84$ kJ and $H_3 = 3444$ kJ. The entropy at point 3 is 6.233 kJ/K, and to obtain this same entropy at point 4 we need a mixture of 29% water and 71% steam. This same mixture has an enthalpy of $H_4 = 1824$ kJ, so the efficiency of the cycle is approximately

$$e \approx 1 - \frac{1824 - 84}{3444 - 84} = 48\%. \tag{4.13}$$

For comparison, an ideal Carnot engine operating over the same temperature range would have an efficiency of 66%. While the temperatures and pressures assumed here are typical of modern fossil-fuel power plants, these plants achieve an actual efficiency of only about 40%, due to a variety of complications that I've neglected. Nuclear power plants operate at lower temperatures for safety reasons, and therefore achieve efficiencies of only about 34%.

Problem 4.22. A small-scale steam engine might operate between the temperatures $20°C$ and $300°C$, with a maximum steam pressure of 10 bars. Calculate the efficiency of a Rankine cycle with these parameters.

Problem 4.23. Use the definition of enthalpy to calculate the change in enthalpy between points 1 and 2 of the Rankine cycle, for the same numerical parameters as used in the text. Recalculate the efficiency using your corrected value of H_2, and comment on the accuracy of the approximation $H_2 \approx H_1$.

Problem 4.24. Calculate the efficiency of a Rankine cycle that is modified from the parameters used in the text in each of the following three ways (one at a time), and comment briefly on the results: (a) reduce the maximum temperature to $500°C$; (b) reduce the maximum pressure to 100 bars; (c) reduce the minimum temperature to $10°C$.

Problem 4.25. In a real turbine, the entropy of the steam will increase somewhat. How will this affect the percentages of liquid and gas at point 4 in the cycle? How will the efficiency be affected?

Problem 4.26. A coal-fired power plant, with parameters similar to those used in the text above, is to deliver 1 GW (10^9 watts) of power. Estimate the amount of steam (in kilograms) that must pass through the turbine(s) each second.

Problem 4.27. In Table 4.1, why does the entropy of water increase with increasing temperature, while the entropy of steam decreases with increasing temperature?

Problem 4.28. Imagine that your dog has eaten the portion of Table 4.1 that gives entropy data; only the enthalpy data remains. Explain how you could reconstruct the missing portion of the table. Use your method to explicitly check a few of the entries for consistency. How much of Table 4.2 could you reconstruct if it were missing? Explain.

4.4 Real Refrigerators

The operation of an ordinary refrigerator or air conditioner is almost the reverse of the Rankine cycle just discussed. Again the working substance changes back and forth from a gas to a liquid, but here the fluid must have a much lower boiling temperature.

Dozens of fluids have been used as refrigerants, including carbon dioxide (which requires rather high pressures) and ammonia (which is still used in large industrial systems, despite its toxicity). Around 1930, General Motors and du Pont developed and produced the first of the nontoxic chlorofluorocarbon (CFC) refrigerants, giving them the trade name *Freon*. Of these the most familiar is Freon-12 (CCl_2F_2), used

in domestic refrigerators and automobile air conditioners. We now know, however, that CFC's that have escaped into the atmosphere are causing the breakdown of the ozone layer. The most damaging CFC's are therefore being replaced with chlorine-free compounds; the usual substitute for Freon-12 is a hydrofluorocarbon, $F_3C_2FH_2$, known by the catchy name *HFC-134a*.

A schematic sketch and *PV* diagram of the standard refrigeration cycle are shown in Figure 4.9. Beginning at point 1, the fluid (here a gas) is first compressed adiabatically, raising its pressure and temperature. It then gives up heat and gradually liquefies in the condenser (a network of pipes in thermal contact with the hot reservoir). Next it passes through a "throttling valve"—a narrow opening or porous plug—emerging on the other side at much lower pressure and temperature. Finally it absorbs heat and turns back into a gas in the evaporator (a network of pipes in thermal contact with the cold reservoir).

It's easy to express the coefficient of performance of a standard refrigerator in terms of the enthalpies of the fluid at various points around the cycle. Since the pressure is constant in the evaporator, the heat absorbed is $Q_c = H_1 - H_4$, the change in enthalpy. Similarly, the heat expelled in the condenser is $Q_h = H_2 - H_3$. So the coefficient of performance is

$$\text{COP} = \frac{Q_c}{Q_h - Q_c} = \frac{H_1 - H_4}{H_2 - H_3 - H_1 + H_4}. \tag{4.14}$$

The enthalpies at points 1, 2, and 3 can be looked up in tables, with point 2 located by assuming that the entropy is constant during the compression stage. To locate point 4 we must analyze the throttling valve in a bit more detail.
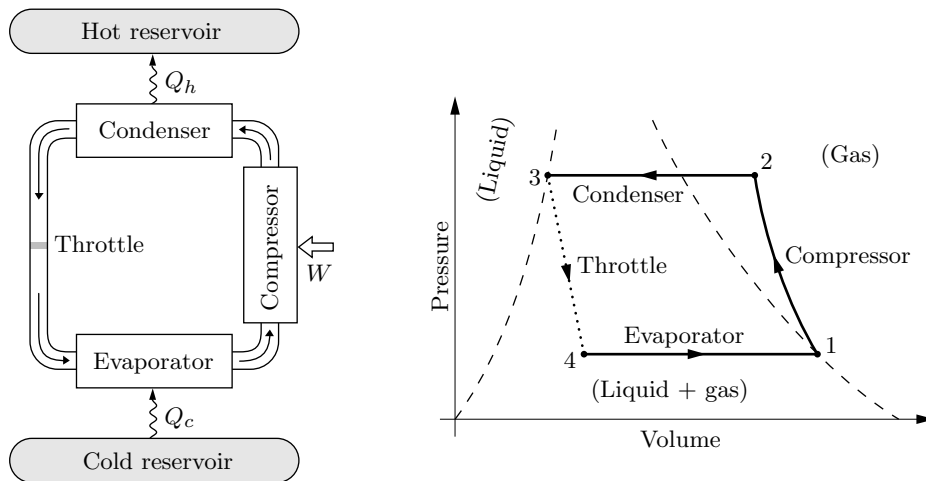

Figure 4.9. A schematic drawing and *PV* diagram (not to scale) of the standard refrigeration cycle. The dashed lines indicate where the refrigerant is liquid, gas, and a combination of the two.

The Throttling Process

The **throttling** process (also known as the **Joule-Thomson process**) is shown in Figure 4.10. I find it helpful to pretend that the fluid is being pushed through the plug by a piston, exerting pressure P_i, while a second piston, exerting pressure P_f, moves backward on the other side to make room. For a particular chunk of fluid, let the initial volume (before going through the plug) be V_i, while the final volume (on the other side) is V_f. Since there is no heat flow during this process, the change in the energy of the fluid is

$$U_f - U_i = Q + W = 0 + W_{\text{left}} + W_{\text{right}}, \tag{4.15}$$

where W_{left} is the (positive) work done by the piston on the left, and W_{right} is the (negative) work done by the piston on the right. (Ultimately, the net work is actually performed by the compressor, way over on the other side of the cycle. Here, however, we're concerned only with what's happening locally.) But the work done from the left in pushing the entire volume V_i through the plug is P_iV_i, while the work done from the right is $-P_fV_f$—negative because the piston is moving backwards. Therefore the change in energy is

$$U_f - U_i = P_iV_i - P_fV_f. \tag{4.16}$$

Putting the f's on the left and the i's on the right, this equation becomes

$$U_f + P_fV_f = U_i + P_iV_i, \qquad \text{or} \qquad H_f = H_i. \tag{4.17}$$

The enthalpy is constant during the throttling process.

The *purpose* of the throttling valve is to *cool* the fluid to below the temperature of the cold reservoir, so it can absorb heat as required. If the fluid were an ideal gas, this wouldn't work at all, since

$$H = U + PV = \frac{f}{2}NkT + NkT = \frac{f+2}{2}NkT \qquad \text{(ideal gas).} \tag{4.18}$$

Constant enthalpy would imply constant temperature! But in a dense gas or liquid, the energy U also contains a potential energy term due to the forces between the molecules:

$$H = U_{\text{potential}} + U_{\text{kinetic}} + PV. \tag{4.19}$$
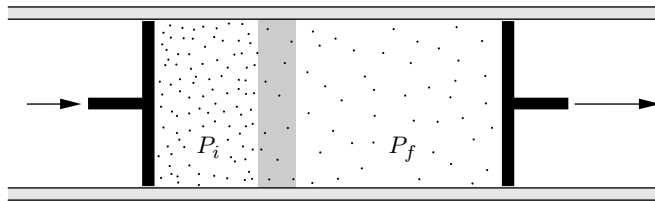

Figure 4.10. The throttling process, in which a fluid is pushed through a porous plug and then expands into a region of lower pressure.

The force between any two molecules is weakly attractive at long distances and strongly repulsive at short distances. Under most (though not all) conditions the attraction dominates; then $U_{\text{potential}}$ is negative, but becomes less negative as the pressure drops and the distance between molecules increases. To compensate for the increase in potential energy, the kinetic energy generally drops, and the fluid cools as desired.

If we use the fact that $H_4 = H_3$ in the refrigeration cycle, the coefficient of performance (4.14) simplifies to

$$\text{COP} = \frac{H_1 - H_3}{H_2 - H_1}. \tag{4.20}$$

Now one only has to look up three enthalpies. Tables 4.3 and 4.4 give enthalpy and entropy values for the refrigerant HFC-134a.

Problem 4.29. Liquid HFC-134a at its boiling point at 12 bars pressure is throttled to 1 bar pressure. What is the final temperature? What fraction of the liquid vaporizes?

P (bar)	T (°C)	H_{liquid} (kJ)	H_{gas} (kJ)	S_{liquid} (kJ/K)	S_{gas} (kJ/K)
1.0	−26.4	16	231	0.068	0.940
1.4	−18.8	26	236	0.106	0.932
2.0	−10.1	37	241	0.148	0.925
4.0	8.9	62	252	0.240	0.915
6.0	21.6	79	259	0.300	0.910
8.0	31.3	93	264	0.346	0.907
10.0	39.4	105	268	0.384	0.904
12.0	46.3	116	271	0.416	0.902

Table 4.3. Properties of the refrigerant HFC-134a under saturated conditions (at its boiling point for each pressure). All values are for 1 kg of fluid, and are measured relative to an arbitrarily chosen reference state, the saturated liquid at −40°C. Excerpted from Moran and Shapiro (1995).

P (bar)		Temperature (°C) 40	50	60
8.0	H (kJ)	274	284	295
	S (kJ/K)	0.937	0.971	1.003
10.0	H (kJ)	269	280	291
	S (kJ/K)	0.907	0.943	0.977
12.0	H (kJ)		276	287
	S (kJ/K)		0.916	0.953

Table 4.4. Properties of superheated (gaseous) refrigerant HFC-134a. All values are for 1 kg of fluid, and are measured relative to the same reference state as in Table 4.3. Excerpted from Moran and Shapiro (1995).

Problem 4.30. Consider a household refrigerator that uses HFC-134a as the refrigerant, operating between the pressures of 1.0 bar and 10 bars.

(a) The compression stage of the cycle begins with saturated vapor at 1 bar and ends at 10 bars. Assuming that the entropy is constant during compression, find the approximate temperature of the vapor after it is compressed. (You'll have to do an interpolation between the values given in Table 4.4.)

(b) Determine the enthalpy at each of the points 1, 2, 3, and 4, and calculate the coefficient of performance. Compare to the COP of a Carnot refrigerator operating between the same reservoir temperatures. Does this temperature range seem reasonable for a household refrigerator? Explain briefly.

(c) What fraction of the liquid vaporizes during the throttling step?

Problem 4.31. Suppose that the throttling valve in the refrigerator of the previous problem is replaced with a small turbine-generator in which the fluid expands adiabatically, doing work that contributes to powering the compressor. Will this change affect the COP of the refrigerator? If so, by how much? Why do you suppose real refrigerators use a throttle instead of a turbine?

Problem 4.32. Suppose you are told to design a household air conditioner using HFC-134a as its working substance. Over what range of pressures would you have it operate? Explain your reasoning. Calculate the COP for your design, and compare to the COP of an ideal Carnot refrigerator operating between the same reservoir temperatures.

Liquefaction of Gases

If you want to make something *really* cold, you normally don't just stick it into a refrigerator—instead you put it on dry ice (195 K at atmospheric pressure), or immerse it in liquid nitrogen (77 K) or even liquid helium (4.2 K). But how are gases like nitrogen and helium liquefied (or in the case of CO_2, solidified) in the first place? The most common methods all involve the throttling process.

You can liquefy carbon dioxide at room temperature, simply by compressing it isothermally to about 60 atmospheres. Then throttle it back to low pressure and it cools and partially evaporates just as in the refrigeration cycle discussed above. At pressures below 5.1 atm, however, liquid CO_2 cannot exist; instead the condensed phase is a solid, dry ice. So to make dry ice, all you have to do is hook up a tank of liquid CO_2 to a throttling valve and watch the frost form around the nozzle as the gas rushes out.

Liquefying nitrogen (or air) isn't so simple. Compress it all you want at room temperature and it never undergoes a sudden phase transformation to a liquid—it just gets denser and denser in a continuous way. (This behavior is discussed in more detail in Section 5.3.) If you start with nitrogen at, say, 300 K and 100 atm and throttle it down to 1 atm it does cool, but only to about 280 K. To get any liquid, starting from this pressure, the initial temperature must be below about 160 K. At higher initial pressures the initial temperature can be somewhat higher, but must still be well below room temperature. The first liquefaction of oxygen and nitrogen was achieved in 1877 by Louis Cailletet, who used an initial pressure of 300 atm and precooled the gases using other cold liquids. A more convenient method, though, is to use the throttled gas itself to precool the incoming gas. A device that does this
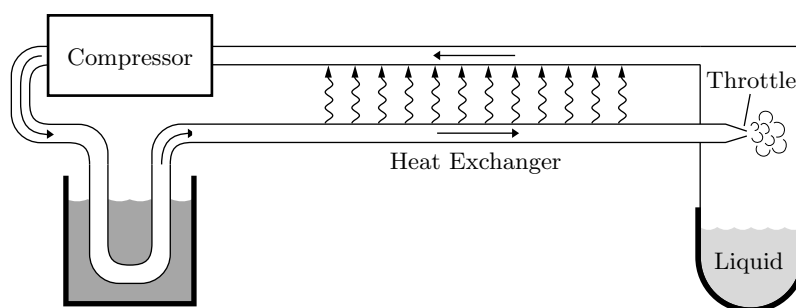
Figure 4.11. Schematic diagram of the Hampson-Linde cycle for gas liquefaction. Compressed gas is first cooled (to room temperature is sufficient if it is nitrogen or oxygen) and then passed through a heat exchanger on its way to a throttling valve. The gas cools upon throttling and returns through the heat exchanger to further cool the incoming gas. Eventually the incoming gas becomes cold enough to partially liquefy upon throttling. From then on, new gas must be added at the compressor to replace what is liquefied.

was invented by William Hampson and Carl von Linde (independently) in 1895; it is shown schematically in Figure 4.11. Instead of being discarded, the throttled gas is sent through a **heat exchanger** where it cools the incoming gas. When that gas passes through the throttle it cools even more, and thus the system gradually gets colder and colder until the throttled gas begins to liquefy.

Starting from room temperature, the Hampson-Linde cycle can be used to liquefy any gas except hydrogen or helium. These gases, when throttled starting at room temperature and any pressure, actually become *hotter*. This happens because the attractive interactions between the molecules are very weak; at high temperatures the molecules are moving too fast to notice much attraction, but they still suffer hard collisions during which there is a large *positive* potential energy. When the gas expands the collisions occur less frequently, so the average potential energy decreases and the average kinetic energy increases.

To liquefy hydrogen or helium, it is therefore necessary to first cool the gas well below room temperature, slowing down the molecules until attraction becomes more important than repulsion. Figure 4.12 shows the range of temperatures and pressures under which hydrogen will cool upon throttling. The temperature below which cooling will occur is called the **inversion temperature**; for hydrogen the maximum inversion temperature is 204 K, while for helium it is only 43 K. Hydrogen was first liquefied in 1898 by James Dewar, using liquid air for precooling. Helium was first liquefied in 1908 by Heike Kamerlingh Onnes, using liquid hydrogen for precooling. Today, however, the precooling of helium is normally accomplished without liquid hydrogen (and sometimes even without liquid nitrogen), by allowing the helium to expand adiabatically as it pushes a piston. This technique is a great advance in safety but is mechanically more challenging. The piston must operate at temperatures as low as 8 K, at which any ordinary lubricant would freeze. The helium itself is therefore used as a lubricant, with extremely small clearances to prevent a significant amount from escaping.

Figure 4.12. Lines of constant enthalpy (approximately horizontal, at intervals of 400 J/mol) and inversion curve (dashed) for hydrogen. In a throttling process the enthalpy is constant, so cooling occurs only to the left of the inversion curve, where the enthalpy lines have positive slopes. The heavy solid line at lower-left is the liquid-gas phase boundary. Data from Vargaftik (1975) and Woolley et al. (1948).
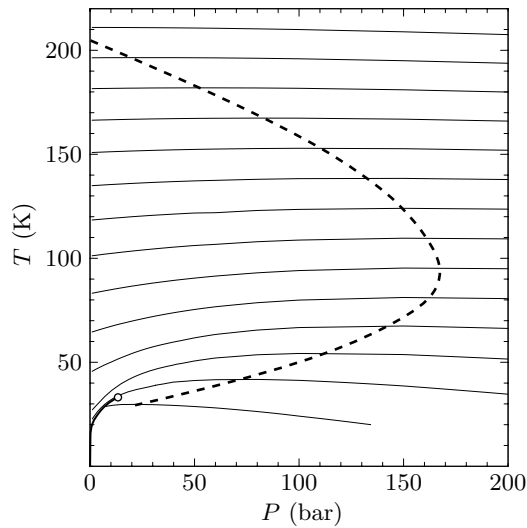
	Temperature (K)							
	77 (liq.)	77 (gas)	100	200	300	400	500	600
1 bar	−3407	2161	2856	5800	8717	11,635	14,573	17,554
100 bar			−1946	4442	8174	11,392	14,492	17,575

Table 4.5. Molar enthalpy of nitrogen (in joules) at 1 bar and 100 bars. Excerpted from Lide (1994).

Problem 4.33. Table 4.5 gives experimental values of the molar enthalpy of nitrogen at 1 bar and 100 bars. Use this data to answer the following questions about a nitrogen throttling process operating between these two pressures.

(a) If the initial temperature is 300 K, what is the final temperature? (Hint: You'll have to do an interpolation between the tabulated values.)

(b) If the initial temperature is 200 K, what is the final temperature?

(c) If the initial temperature is 100 K, what is the final temperature? What fraction of the nitrogen ends up as a liquid in this case?

(d) What is the highest initial temperature at which some liquefaction takes place?

(e) What would happen if the initial temperature were 600 K? Explain.

Problem 4.34. Consider an ideal Hampson-Linde cycle in which no heat is lost to the environment.

(a) Argue that the combination of the throttling valve and the heat exchanger is a constant-enthalpy device, so that the total enthalpy of the fluid coming out of this combination is the same as the enthalpy of the fluid going in.

(b) Let x be the fraction of the fluid that liquefies on each pass through the cycle. Show that

$$x = \frac{H_{out} - H_{in}}{H_{out} - H_{liq}},$$

where H_{in} is the enthalpy of each mole of compressed gas that goes into the heat exchanger, H_{out} is the enthalpy of each mole of low-pressure gas

that comes out of the heat exchanger, and H_{liq} is the enthalpy of each mole of liquid produced.

(c) Use the data in Table 4.5 to calculate the fraction of nitrogen liquefied on each pass through a Hampson-Linde cycle operating between 1 bar and 100 bars, with an input temperature of 300 K. Assume that the heat exchanger works perfectly, so the temperature of the low-pressure gas coming out of it is the same as the temperature of the high-pressure gas going in. Repeat the calculation for an input temperature of 200 K.

Toward Absolute Zero

At atmospheric pressure, liquid helium boils at 4.2 K. The boiling point decreases as the pressure is reduced, so it's not hard to lower the temperature of liquid helium still further, by pumping away the vapor to reduce the pressure; the helium cools through evaporation. Below about 1 K, however, this procedure becomes impractical: Even the smallest heat leak raises the temperature of the helium significantly, and the best vacuum pumps cannot remove the vapor fast enough to compensate. The rare isotope helium-3, whose normal boiling point is only 3.2 K, can be cooled to about 0.3 K by pumping to low pressure.

But isn't 1 K cold enough? Why bother trying to attain still lower temperatures? Perhaps surprisingly, there are a variety of fascinating phenomena that occur only in the millikelvin, microkelvin, and even nanokelvin ranges, including transformations of helium itself, magnetic behavior of atoms and nuclei, and "Bose-Einstein condensation" of dilute gases. To investigate these phenomena, experimenters have developed an equally fascinating array of techniques for reaching extremely low temperatures.*

To get from 1 K to a few millikelvins, the method of choice is usually the **helium dilution refrigerator**, shown schematically in Figure 4.13. The cooling occurs by "evaporation" of liquid ^{3}He, but instead of evaporating into a vacuum, it dissolves into a liquid bath of the more common isotope, ^{4}He. At subkelvin temperatures the two isotopes are relatively immiscible, like oil and water. Below about 0.1 K essentially no ^{4}He will dissolve in pure ^{3}He, while a small amount of ^{3}He, about 6%, will dissolve into otherwise pure ^{4}He. This is what happens in the "mixing chamber," where ^{3}He continuously dissolves ("evaporates") into the ^{4}He, absorbing heat in the process. The ^{3}He then diffuses upward through a heat exchanger to a "still" at 0.7 K, where heat is supplied to make it evaporate (in the conventional sense). The ^{4}He is essentially inert during the whole process: It is a "superfluid" in this temperature range, offering negligible resistance to the diffusion of ^{3}He atoms; and it is less volatile than ^{3}He, so it does not evaporate at a significant rate in the still. After evaporating in the still, the gasous ^{3}He is compressed, recooled to a liquid (by a separate bath of ^{4}He), and finally sent back through the heat exchanger to the mixing chamber.

---

*For a good overview of methods of reaching subkelvin temperatures, see Olli V. Lounasmaa, "Towards the Absolute Zero," *Physics Today* **32**, 32–41 (December, 1979). For more details on helium dilution refrigerators, see John C. Wheatley, *American Journal of Physics* **36**, 181–210 (1968).

Figure 4.13. Schematic diagram of a helium dilution refrigerator. The working substance is ^{3}He (light gray), which circulates counter-clockwise. The ^{4}He (dark gray) does not circulate.
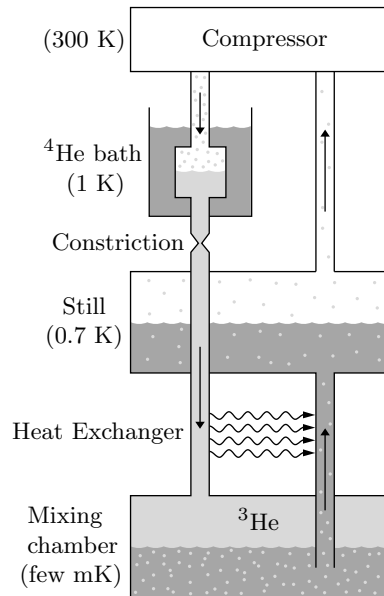
An alternative method of reaching millikelvin temperatures is **magnetic cooling**, based on the properties of paramagnetic materials. Recall from Section 3.3 that the total magnetization of an ideal two-state paramagnet is a function of the *ratio* of the magnetic field strength B to the temperature:

$$M = \mu(N_\uparrow - N_\downarrow) = N\mu \tanh\left(\frac{\mu B}{kT}\right). \tag{4.21}$$

(For an ideal paramagnet with more than two states per particle, the formula is more complicated but has the same qualitative behavior.) For an *electronic* paramagnet, whose elementary dipoles are electrons, the value of μ is such that a magnetic field of 1 T and a temperature of 1 K yield $M/N\mu = 0.59$: A significant majority of the dipoles are pointing up. Suppose, now, that we start with the system in such a state and then reduce the magnetic field strength without allowing any heat to enter. The populations of the up and down states will not change during this process, so the total magnetization is fixed and therefore the temperature must decrease in proportion to the field strength. If B decreases by a factor of 1000, so does T.

A good way to visualize this process is shown in Figure 4.14, where I've plotted the entropy of the system as a function of temperature for two different values of the magnetic field strength. For any nonzero field strength the entropy goes to zero as $T \to 0$ (as all the dipoles line up) and goes to a nonzero constant value at sufficiently high temperatures (as the alignments of the dipoles become random). The higher the field strength, the more gradually the entropy rises as a function of temperature (due to the greater tendency of the dipoles to remain aligned with the field). In a magnetic cooling process, the sample is first put in good thermal contact with a constant-temperature "reservoir" such as a liquid helium bath. The
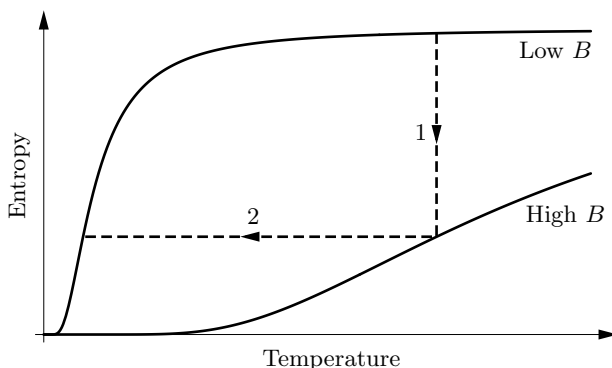
Figure 4.14. Entropy as a function of temperature for an ideal two-state paramagnet, at two different values of the magnetic field strength. (These curves were plotted from the formula derived in Problem 3.23.) The magnetic cooling process consists of an isothermal increase in the field strength (step 1), followed by an adiabatic decrease (step 2).

magnetic field is then increased, causing the entropy of the sample to drop as its temperature remains fixed. The sample is then insulated from the reservoir and the field strength reduced, resulting in a drop in temperature at constant entropy. The process is analogous to the cooling of an ideal gas by adiabatic expansion, following isothermal compression.

But why should we merely *reduce* the strength of the magnetic field—why not *eliminate* it entirely? Then, according to equation 4.21, the temperature of the paramagnet would have to go to *absolute zero* in order to maintain constant M. As you might guess, attaining absolute zero isn't so easy. The problem in this case is that no paramagnet is truly *ideal* at very low temperatures: The elementary dipoles interact with *each other*, effectively producing a magnetic field that is present even when the *applied* field is zero. Depending on the details of these interactions, the dipoles may align parallel or antiparallel to their nearest neighbors. Either way, their entropy drops almost to zero, as if there were an external magnetic field. To reach the lowest possible final temperatures, the paramagnetic material should be one in which the interactions between neighboring dipoles are extremely weak. For electronic paramagnets, the lowest temperature that can be reached by magnetic cooling is about 1 mK (see Problem 4.35).

In a *nuclear* paramagnet the dipole-dipole interactions are much weaker, so much lower temperatures can be attained. The only catch is that you also have to *start* at a lower temperature, in order to have a significant excess of one spin alignment over the other. The first nuclear magnetic cooling experiments produced temperatures of about 1 μK, and it seems that every few years someone improves the technique to achieve still lower temperatures. In 1993, researchers at Helsinki University used nuclear magnetic cooling of rhodium to produce temperatures as low as 280 *pico*kelvins, that is, 2.8×10^{-10} K.[*]

---

[*]Pertti Hakonen and Olli V. Lounasmaa, *Science* **265**, 1821–1825 (23 September, 1994).

Meanwhile, other experimenters have reached extremely low temperatures using a completely different technique: **laser cooling**. Here the system is not a liquid or solid but rather a dilute gas—a small cloud of atoms, prevented from condensing into a solid by its very low density.

Imagine that you hit an atom with laser light, tuned to just the right frequency to excite the atom into a higher-energy state. The atom will absorb a photon as it gains energy, then spontaneously emit a photon of the same frequency as it loses the energy a split second later. Photons carry momentum as well as energy, so the atom recoils each time it absorbs or emits a photon. But whereas the absorbed photons all come from the same direction (the laser), the emitted photons exit in all directions (see Figure 4.15). On average, therefore, the atom feels a *force* from the direction of the laser.

Now suppose that we tune the laser to a slightly lower frequency (longer wavelength). An atom at rest will rarely absorb photons of this frequency, so it feels hardly any force. But an atom moving *toward* the laser will see the light Doppler-shifted back to a higher frequency. It therefore absorbs plenty of photons and feels a backward force, opposing its motion. An atom moving away from the laser feels even less force than an atom at rest, but if we aim an identical laser beam at it from the opposite side, then it too feels a backward force. With laser beams coming from all six directions we can exert backward forces that tend to oppose motion in *any* direction. Put thousands or millions of atoms into the region and they'll all slow down, cooling to a very low temperature.

Even at very low speeds, though, the atoms would quickly hit the hot walls of the container (or fall to the bottom) without an additional *trapping* force that pushes them toward the center. Such a force can be created using nonuniform magnetic fields to shift the atomic energy levels and thus vary their tendency to absorb photons depending on where they are. The combination of laser cooling and trapping can readily cool a cloud of atoms to about 1 mK, without any of the hassle of liquid helium or conventional cryogenic equipment. Elaborations of the technique have recently been used to reach temperatures in the microkelvin and even nanokelvin ranges.*
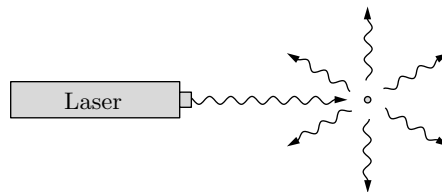

Figure 4.15. An atom that continually absorbs and reemits laser light feels a force from the direction of the laser, because the absorbed photons all come from the same direction while the emitted photons come out in all directions.

---

*For an elementary review of laser cooling and its applications, see Steven Chu, "Laser Trapping of Neutral Particles," *Scientific American* **266**, 71–76 (February, 1992). A bibliography of articles on trapping of neutral atoms has been compiled by N. R. Newbury and C. Wieman, *American Journal of Physics* **64**, 18–20 (1996).

Problem 4.35. The magnetic field created *by* a dipole has a strength of approximately $(\mu_0/4\pi)(\mu/r^3)$, where r is the distance from the dipole and μ_0 is the "permeability of free space," equal to exactly $4\pi \times 10^{-7}$ in SI units. (In the formula I'm neglecting the variation of field strength with angle, which is at most a factor of 2.) Consider a paramagnetic salt like iron ammonium alum, in which the magnetic moment μ of each dipole is approximately one Bohr magneton $(9 \times 10^{-24}$ J/T), with the dipoles separated by a distance of 1 nm. Assume that the dipoles interact only via ordinary magnetic forces.

(a) Estimate the strength of the magnetic field at the location of a dipole, due to its neighboring dipoles. This is the effective field strength even when there is no externally applied field.

(b) If a magnetic cooling experiment using this material begins with an external field strength of 1 T, by about what factor will the temperature decrease when the external field is turned off?

(c) Estimate the temperature at which the entropy of this material rises most steeply as a function of temperature, in the absence of an externally applied field.

(d) If the final temperature in a cooling experiment is significantly less than the temperature you found in part (c), the material ends up in a state where $\partial S/\partial T$ is very small and therefore its heat capacity is very small. Explain why it would be impractical to try to reach such a low temperature with this material.

Problem 4.36. An apparent limit on the temperature achievable by laser cooling is reached when an atom's recoil energy from absorbing or emitting a single photon is comparable to its total kinetic energy. Make a rough estimate of this limiting temperature for rubidium atoms that are cooled using laser light with a wavelength of 780 nm.

Problem 4.37. A common (but imprecise) way of stating the third law of thermodynamics is "You can't reach absolute zero." Discuss how the third law, as stated in Section 3.2, puts limits on how low a temperature can be attained by various refrigeration techniques.

According to this principle, the production of heat alone is not sufficient to give birth to the impelling power: it is necessary that there should also be cold; without it, the heat would be useless.

—Sadi Carnot, *Reflections on the Motive Power of Fire*, trans. R. H. Thurston (Macmillan, New York, 1890).