THE REFRIGERATOR AND THE UNIVERSE

UNDERSTANDING THE LAWS OF ENERGY

MARTIN GOLDSTEIN AND INGE F. GOLDSTEIN

5
Engines and Refrigerators: The Second Law

A heat engine is a device for converting thermal energy into mechanical energy. Most commonly the source of the thermal energy is the combustion of a fuel, though nuclear, solar, and geothermal energy have been used. Heat engines require at least two temperatures for operation; a steam engine, for example, needs a hot boiler and a cold condenser. A considerable portion of the thermal energy supplied by the fuel is discharged to the condenser and eventually wasted on heating the surroundings. This was true at the beginning of the Industrial Revolution, and it is true today, for all forms of heat engine, including the internal combustion engine of the automobile. If an engine could be designed to operate without this waste of thermal energy, the consequences would be tremendous: we would no longer need fuel to operate our engines, and there would never be an "energy shortage"—more precisely, a shortage of inexpensive fossil fuels.

In the nineteenth century Sadi Carnot combined the old idea that perpetual-motion machines are impossible (an early and incomplete version of the first law) with the caloric theory (heat is a substance) to prove that in a heat engine a discharge of heat to the cold surroundings was inevitable. From this starting point, he and others deduced some remarkable conclusions about the efficiency of heat engines and about the properties of matter, conclusions that were repeatedly confirmed in the laboratory.

By the time the first law had been discovered and the caloric theory discarded, scientists were faced with a paradox: Carnot's conclusions about the properties of matter were almost certainly true, but the caloric theory he derived them from was certainly false.

The paradox was resolved by the conjecture that there must be a hitherto unknown law of nature, now known as the second law of thermodynamics, that states as an axiom that a *cyclically operating* heat engine cannot convert thermal energy to mechanical energy unless the device uses at least two temperatures and wastes some of the thermal energy at the lower temperature. The meaning of the phrase *cyclically operating*, and the reason for the restriction, will be given later. From this new starting point Carnot's conclusions about the efficiencies of heat engines and the properties of matter still follow. There really was no paradox: it is logically possible for a correct conclusion to follow from false premises.

In this chapter we will analyze the operation of heat engines using the first and second laws and derive Carnot's conclusions. The analysis, although subtle, relies more on verbal reasoning than on mathematics, and it is worth going through for the sake of its scientific importance.

Previously we showed how the first law implied the existence of a property of matter called *thermal energy*, which could be determined in the laboratory. The second law can be shown to imply the existence of an additional property of matter, called *entropy*, which can also be determined in the laboratory. The first law states that the energy of the universe is conserved: the total amount of it remains the same for all time. The second law states that the entropy of the Universe is *not* conserved but that the total amount of it can never decrease.

The second law provides a criterion of *possibility:* only processes that cause a *net increase* in entropy are possible, though not everything that is possible need happen. It follows that if there is an upper limit to the entropy that the universe can have, then when that limit is reached nothing further can happen. The universe, having exhausted its capacity for change, would have run down like a clock.

Heat Engines, Possible and Impossible

In Chapter 3 we explained how various forms of mechanical or electrical energy could be converted to thermal energy through friction or through the electrical resistance of a conducting wire. Now we want to look in some detail at the *reverse*—the conversion of thermal energy to mechanical energy, a conversion that takes place in the engines of automobiles and in power plants that generate electricity from fossil or nuclear fuels. What, if any, are the limitations on the conversion of thermal energy to other forms?

Before we address that question, let us examine how some specific examples of heat engines actually work. The first successful cyclical steam engine was put into operation by Thomas Newcomen in England in 1712, to pump water out of coal mines. It included a cylinder with a moving piston, similar to the cylinder and piston of a modern automobile engine. The back-and-forth motion of the piston operated a lever, the "great beam," which in turn operated the pump. In one cycle of operation, the cylinder was first filled with steam from the boiler, pushing the piston outward. Then cold water was sprayed into the cylinder, condensing the steam and creating a partial vacuum. The pressure of the atmosphere then forced the piston downward to its starting point—this was the power stroke. The freshly condensed, still-warm water was returned to the boiler to be reused. The repeated return of all components of the engine, including the working substance, to the same state in which they began is what is meant by *cyclical operation*.

Although Newcomen's engine made grossly inefficient use of the energy of the fuel, it had one crucial feature shared by all other heat engines invented since: it did not operate at a single temperature. Instead, it received thermal energy at a high temperature, provided by the boiler, and gave a portion of it out at a low temperature, usually the temperature of the ambient environment or close to it, so that a portion of the thermal energy supplied by the boiler was wasted on warming the environment.

The cylinder of the Newcomen engine undergoes large changes in temperature during a cycle of operation: it is heated to 100°C by the incoming steam, then cooled down when the cold water is injected to condense the steam. James Watt recognized that the repeated cooling and reheating of the necessarily thick metal walls of the cylinder was a source of inefficiency, and in 1769 he patented an improved engine, in which the steam was condensed in a different chamber from the cylinder in which the piston moved. In his engine the cylinder could be hot all the time, and the condensing chamber always cold (see Figure 5.1). This innovation led to a dramatic improvement in the work the engine could do per bushel of coal burned for fuel, and Watt's career, and much else, was launched. One feature that Watt's engine shared with Newcomen's was the return of the water produced by condensation of steam to the boiler: the water is thus *recycled*. (In some steam locomotives, low-temperature steam was discarded in each cycle, which is why such locomotives had to stop from time to time to take on more water.)

The Internal Combustion Engine

We can compare and contrast Watt's engine with the internal combustion engine of the automobile. In the cylinders of the automobile engine the power comes from the explosive combustion of a gasoline-air mixture, which drives the piston forcefully outward to increase the rotational speed, and thus the kinetic energy, of a heavy flywheel. Valves operated by the flywheel then open to allow the gaseous products of the combustion to be driven out of the cylinder on the return stroke of the piston and discharged through the exhaust system; then the valves close and the cylinder is ready to receive new fuel and air.

The internal combustion engine differs from the steam engine in a number of significant ways. One is that the combustion products of the

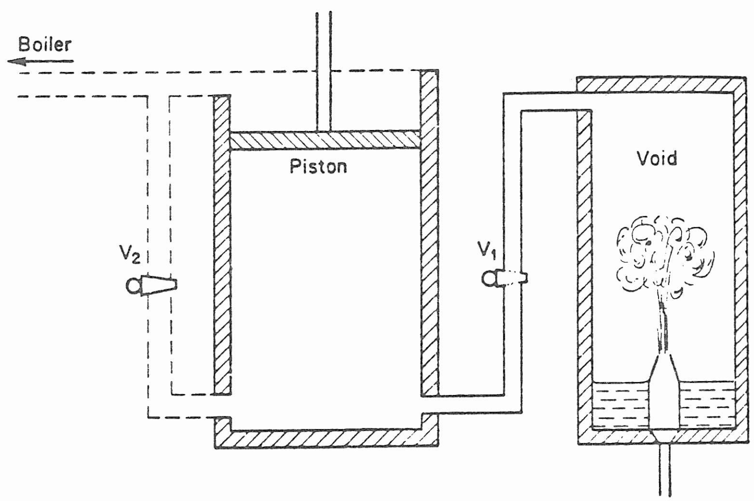

Figure 5.1 Watt's Steam Engine
In the first step of operation of the Watt engine, steam from the boiler enters the cylinder, allowing the piston to move outward. In the second step the valve to the boiler (V_2) is closed and that to the condenser (V_1) is opened. A jet of cold water sprayed into the condenser condenses the steam, thus creating a partial vacuum in the condenser and in the cylinder. The pressure of the atmosphere forcefully drives the piston down: this is the power stroke.

fuel are also the working substance of the engine, while in the steam engine the fuel (coal, usually) only supplies heat to the working substance (water), the combustion products being discharged directly to the environment. A second difference is that in the automobile the working substance is not recycled, while in Watt's engine the water is converted to steam, condensed, and reused over and over again. This cyclical operation has a useful side effect: it simplifies the bookkeeping on energy inputs and outputs of the engine. Since the working substance is repeatedly brought back to its starting state, all its properties, including its thermal energy, must repeatedly return to the same values.

The two engines do share one important feature: in both, only a part of the thermal energy supplied by the combustion of fuel is converted to the mechanical energy of motion, the rest being discharged to the cool environment. Now this waste in the operation of a heat engine sounds like a drawback. Is there no way to design a heat engine that would convert thermal to mechanical energy without waste?

A Solution to the Energy Shortage

Suppose it *were* possible to extract the thermal energy of matter and convert it to useful work without a discharge of some of that thermal energy to a second, and lower, temperature. We would never again have to think of "energy shortages." There is an enormous amount of energy around us, in the form of atomic and molecular motion. Each atom, moving with a speed v, has a kinetic energy of $\frac{1}{2}mv^2$. The mass m of any one atom is of course very small, but there are enormous numbers of them in any substantial amount of matter, and their velocities are large, about the velocity of sound. For example, at a temperature of 20°C—about room temperature—air has about 122,000,000 joules of energy for each metric ton (1,000 kilograms) in the translational motion of its molecules and almost as much additional energy in their rotational and vibrational motion. Now if we could extract this energy from the air, we would necessarily cool it. The first law requires this. While it might be unreasonable to expect to extract all the energy of motion and cool the air to absolute zero (which would solidify it), we would do very well if we could extract as much energy as would cool it by only 10°C. On this much cooling the metric ton of air would supply about 10,000,000 joules, equivalent to nearly 3 kilowatt-hours.

So we need only to invent some engine that draws in air rapidly, extracts this much energy from each metric ton, expels the cooled air, and uses the extracted energy to power a vehicle. We could then drive from New York to San Francisco in this fashion without needing, or paying for, gasoline.

What happens to the energy we have extracted from the air to operate this device? It has first been converted to the motion of our vehicle and then, by the action of friction (in the moving parts of the engine, in the tires moving over the pavement, in air resistance, in the brakes of the vehicle when we need to stop), into a heating of the surrounding environment. Eventually, as the heated bodies cool down in the ambient air, all the energy returns to the air from which it was extracted and is thus ready to be reused for the return trip. So the energy supply is inexhaustible as well as free. All we have to do is invent the energy-extracting engine, and we do not have to worry about the energy shortage any longer.

By now we begin to feel a little uneasy. It is not obvious why we cannot make a heat engine that takes thermal energy from the surrounding environment and uses it all to do work. The first law certainly does not prohibit it: we are not creating energy from nothing, only recycling it—converting it from one form (thermal energy) to another (kinetic energy of a moving vehicle) and back again. But still we feel instinctively that it must be impossible. If it isn't impossible, why hasn't it been invented?

A New Law of Nature

This statement, "Such an engine is impossible," doesn't sound like a law of science: scientific laws should be experimentally testable, and how do we test this one? Certainly if anyone succeeded in building such an engine, we would know at once that the statement is false. But suppose a hundred years go by and no one succeeds, in spite of many attempts. Would we then have enough confidence in the statement to declare it a scientific law, or should we wait another hundred years?

Fortunately, testing this statement is not as difficult as it may sound. Logical analysis often permits us to draw an extraordinary number of inferences from a simple-sounding statement: if the inferences can be experimentally tested, if there are enough of them, if at least some of them are surprising, and if they are all confirmed in the laboratory, we become very confident in the truth of that statement. The reader who remembers how many theorems, corollaries, and homework problems follow from the simple axioms and definitions of plane geometry will

have had vivid experience of the power of logical analysis. The first law itself can be stated in a few words: energy is conserved. Yet its consequences are far-reaching. Other major scientific theories can also be stated simply—sometimes in words, sometimes in a few mathematical equations—and yet generate a wide range of experimentally testable implications by logical and mathematical reasoning.

The statement that it is impossible for a cyclically operating device to convert thermal energy to mechanical energy without wasting some thermal energy on cold surroundings is only one example. It is called the *second law of thermodynamics*. We have as much confidence in it as in any scientific law yet discovered, not because no one has succeeded in building the engine since Kelvin first formulated the law in essentially these terms in 1848 ("It is impossible . . . to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects") but rather because an extraordinary number of logical consequences have been drawn from it that have been repeatedly confirmed in the laboratory. The reach of the second law is vast—it makes predictions we can test about the metabolism of a living cell, how the circulation of molten rock below the surface of the earth moves continents, and what happens when stars collapse—but its roots, as we discuss below, are embedded in a very particular problem of a very particular technology: that of the heat engine.

We will describe in this chapter why the scientists involved were led to make this statement as a law of nature; how they were able to draw conclusions from it for both the efficiencies of heat engines and the properties of matter; and what kinds of confirmatory evidence led to its acceptance as one of the greatest of scientific generalizations.

Wrong Theories and Right Answers: Carnot's Contribution

We will begin by going back in time to a period shortly before the discovery of the first law. The period is the 1820s, with the Industrial Revolution in full swing, powered both literally and figuratively by the steam engine. A question then (and now) of considerable economic importance was how to get the maximum output of work for a given quantity of fuel. The steam engine was the prevailing technology: there was constant competition to improve it, and experiments using substances other than water as the working substance were being made.

The French engineer Sadi Carnot addressed this question in a monograph published in 1824. Carnot was the son of a famous father, Lazare

Carnot, himself both a scientist and a prominent figure in the French Revolution. The elder Carnot's scientific work had a direct influence on his son's reasoning about heat engines. His goal was to discover the general principles by which machines in general operated, rather than to focus only on specific kinds such as water-powered pumps or windmills. He started from the definition of a machine as an intermediary body transmitting motion between two or more other bodies not otherwise connected, and assumed the principle of conservation of *vis viva* and its equivalency to work done. In effect, he used the laws of mechanics in the absence of friction to determine the maximum efficiency possible to any machine, and he recognized that friction reduces efficiency below that maximum. The impossibility of perpetual motion was his guiding principle.

That perpetual-motion machines are impossible is an old idea, as we have noted earlier. It is regarded today as an alternative way of stating the first law, although it was originally applied only to purely mechanical devices, as Lazare Carnot did. Sadi Carnot's achievement was to extend this idea, erroneously, to heat engines.

Sadi Carnot (named after a medieval Persian poet popular in French translation at that time) was born in 1796. He was educated by his father until he was sixteen, when he entered the Ecole Polytechnique. In 1814 he and his fellow students participated in the unsuccessful defense of Paris after Waterloo. In the 1820s, when not on military service as a captain, he studied physics and economics at the Sorbonne and at other schools, but he did not neglect the practical applications of his interests, visiting factories and studying the organization of industries and trades in France and elsewhere. In 1824 his one published work, *Reflections on the Motive Power of Fire, and on Machines Fitted to Develop that Power,* appeared. It received one favorable review but sold very few copies and had very little immediate influence on the practical aspects of heat engine design. In 1828, dissatisfied with the reactionary policies of the monarchical regime in France (he was his father's son in politics as well as science), he resigned from the army to devote full time to his studies in physics and economics. In 1832 he caught scarlet fever and suffered what was then called "brain fever" (meningitis, probably) as a sequel. Greatly weakened, he was convalescing in the country but contracted cholera when an epidemic broke out and died in one day. He was thirty-six years old when he died.

Carnot's monograph was written for a popular rather than a technically trained audience. It began:

> Every one knows that heat can produce motion. That it possesses vast motive-power no one can doubt, in these days when the steam-engine is everywhere so well known.
>
> To heat also are due the vast movements which take place on the earth. It causes the agitations of the atmosphere, the ascension of clouds, the fall of rain and of meteors, the currents of water which channel the surface of the globe, and of which man has thus far employed but a small portion. Even earthquakes and volcanic eruptions are the result of heat [Carnot was guessing here, but he guessed right].
>
> From this immense reservoir we may draw the moving force necessary for our purposes. Nature, in providing us with combustibles on all sides, has given us the power to produce, at all times and in all places, heat and the impelling power which is the result of it. To develop this power, to appropriate it to our uses, is the object of heat-engines.
>
> The study of these engines is of the greatest interest, their importance is enormous, their use is continually increasing, and they seem destined to produce a great revolution in the civilized world.

Heat Engines and Waterwheels

In his analysis Carnot used the common view of his time that heat was a substance just as water is a substance. He assumed that a heat engine was analogous to a waterwheel. As the "motive power," the work obtainable, from falling water is the product of the weight of the water and the height through which it falls, so the motive power of heat should depend on the quantity of heat, which Carnot expressed in calories, and the "height" (the difference in temperature) through which it falls. Just as no work is obtainable from water at a single level—there must be a lower level to which it can fall—so work can be done only when heat is allowed to flow from a high to a low temperature. Of course, water can fall from a high level to a low level and do no work in the process; it will take place spontaneously in the absence of a waterwheel to make use of the potential for work. Carnot assumed that heat behaves the same: it flows spontaneously from a hot body to a cold one. It will do work if we interpose an appropriate engine. If we don't, it flows anyway.

Because Carnot thought of heat as a substance that could not be created or destroyed, he took for granted that all the heat supplied by the fuel was eventually given out to the surroundings, just as all the water falling on the wheel of a water mill flows out at the bottom.

Carnot carried his analysis further. He knew that the maximum work that can be done by falling water depends only on the product of the water's weight and the height of the fall, not on the design of the water wheel, and that this is true for all liquids, not just water. If alcohol or olive oil rained from the skies and formed rivers whose flow we wanted to use for power, the maximum work obtainable would still be the product of the weight of the liquid and the height through which it falls. In other words, the determination of maximum work is independent of the particular working substance used. Carnot argued by analogy that the maximum work a heat engine can do must also be expressed in terms independent of the working substance, that it must depend only on the quantity of heat used and the difference in the two temperatures of operation. As noted by Carnot, engines had been proposed and constructed using working substances other than water, with no dramatic improvement in efficiency:

> Wherever there exists a difference of temperature, wherever it has been possible for the equilibrium of the caloric to be reestablished [by this Carnot meant the tendency of "caloric" to flow from hot to cold bodies to produce a uniform temperature], it is possible to have also the production of impelling power. Steam is a means of realizing this power, but it is not the only one. All substances in nature can be employed for this purpose, all are susceptible of changes of volume, of successive contractions and dilatations, through the alternation of heat and cold. All are capable of overcoming in their changes of volume certain resistances, and of thus developing the impelling power. A solid body—a metallic bar for example—alternately increases and diminishes in length, and can move bodies fastened to its end. A liquid alternately heated and cooled increases and diminishes in volume, and can overcome obstacles of greater or less size, opposed to its dilatation. An aeriform fluid [a gas] is susceptible of considerable change of volume by variations of temperature. If it is enclosed in an expansible space, such as a cylinder provided with a piston, it will produce movements of great extent. Vapors of all substances capable of passing into a gaseous condition, as of alcohol, of mercury, of sulphur, etc., may fulfill the same office as vapor of water. The latter, alternately heated and cooled, would produce motive power in the shape of permanent gases, that is, without ever returning to a liquid state. Most of these substances have been proposed, many even have been tried, although up to this time perhaps without remarkable success.

All Substances Give the Same Efficiency

The ability of a substance, whether air, water, or solid iron, to be used in a heat engine to produce work depends on its ability to change shape or volume as its temperature is changed and to exert forces on pistons or other restraining bodies. We know that different substances expand on heating by different amounts and exert different forces, and we would expect therefore that some would work better in heat engines than others. Carnot's reasoning implies that they do not: to the extent that it is possible to eliminate frictional losses and other sources of inefficient operation, all substances will do the same work for the same heat input and the same temperatures of operation. The importance of his conclusion is not so much that it tells us something we didn't know about heat engines, but rather that it tells us something we didn't know about substances. It tells us that those properties of *each* substance that determine its performance as the working substance in a heat engine must be interrelated in such a way as to give the same efficiency as *all* other substances.

Carnot accommodated his analysis to the accepted theory of the time, the caloric theory. He believed incorrectly that heat was an indestructible substance: the heat extracted from the high-temperature source (the boiler) must eventually all be delivered to the low-temperature condenser; the condenser is usually kept cold by contact with an outside source of cold water, such as a river, or else with the ambient air, so all the heat given out by the fuel eventually finds its way to this ultimate coolant, where it can no longer be used for doing work. Later, when the first law was discovered, it was realized that "heat" is *not* conserved, that in the operation of a heat engine some of the "heat" is converted to work.

After Carnot's death some notes he had taken on ideas he intended to pursue were found: their significance, like that of his published work, was realized much later. Aware of the challenge to the caloric theory in Rumford's work, he was becoming skeptical of the caloric theory and giving serious thought to the kinetic theory and was designing experiments similar to those done by Joule a decade later. Carnot's early death does not affect us like the premature deaths of Mozart, of Schubert, of

Keats, in which we have lost something we will never have and cannot even imagine. At most he might have discovered both the first and second laws ten or so years earlier than they were actually discovered, and instead of being merely one of a large number of scientific geniuses of the nineteenth century he would have been among its few towering figures, together with James Clerk Maxwell and Darwin. His story leaves us not with the sense of something irreplaceable lost forever, but rather with a sense of the potential of a remarkable human being unfulfilled.

The Second Law and Its Consequences

Joule's experiments showed that energy, not heat, is always conserved. It is essential to keep the difference between the two in mind, especially when even today we talk loosely about the "heat" in a body. It is always correct to use the term *thermal energy* in place of *heat*, but *heat* may properly be used for the thermal energy transferred from one body to another because of a difference in their temperatures. The thermal energy transferred between the working substances of heat engines and boilers or condensers is transferred in just this way, so we will often use the term *heat* for it. Heat engines can be simply described as devices that convert heat into work, though in more precise language they convert thermal energy to mechanical.

Once Joule showed that it is energy that is conserved, it was necessary to reexamine Carnot's conclusion, based on the now-discarded caloric theory, that the maximum work a heat engine could do does not depend on the working substance. It became apparent that the analogy on which Carnot's reasoning rested was a false one. A heat engine and a water wheel are not examples of the same kind of thing. The quantity of water is not changed by its passage through a water wheel: as many gallons leave as enter. In the operation of a heat engine, however, the first law tells us that the quantity of heat flowing out to the cold environment is *less* than that flowing in from the burning fuel. The difference is equal to the work done by the engine: energy must be conserved. If a heat engine could really operate without losing heat to a cold environment, it would merely mean that *all* the thermal energy supplied by the fuel is converted to work: energy is not being created from nothing. So it would not be a perpetual-motion machine.

That Carnot's analogy between a heat engine and a water wheel was inappropriate does not mean that his conclusion—that the maximum work a heat engine can do does not depend on the working substance—

must be false. In fact many of the logical implications of that conclusion, although surprising, had been confirmed by experiments. Kelvin, taking this convincing experimental support as equally convincing support for the caloric theory used by Carnot, was reluctant to believe Joule's results, though to his credit he recognized how revolutionary they would be if true.

It was the German physicist Rudolf Clausius who saw how to reconcile Carnot's reasoning with the first law. To do so he proposed a new law of nature, independent of the first law. His wording of it is logically equivalent to Kelvin's given earlier, which we can paraphrase as follows: *It is impossible for a heat engine operating cyclically, and gaining heat from an outside body at a single temperature, to do any net work.* As noted earlier, cyclical operation means that the engine, including the working substance, is repeatedly brought back to its starting state as the engine operates.

Why Cyclical Operation at a Single Temperature Does No Work

Let us imagine a system consisting of a flywheel, free to rotate, in close proximity to a large tank of water equipped with a sensitive thermometer, so that any change in the temperature, and hence in the thermal energy of the water, can be noted and measured. We would not be surprised to observe that if the flywheel is set in motion it will gradually slow down and the temperature of the water would rise in proportion, as the first law requires, but we would be surprised to observe the reverse of this sequence—the wheel gradually speeding up while the temperature falls—even though these events would not violate the first law. Even if we were unaware of Kelvin's statement, we would find it hard to believe that thermal energy, the random, disorganized motion of molecules, can spontaneously convert itself into the organized motion of the wheel.

Now let us imagine that some clever engineer places a device whose nature he refuses to reveal, enclosed within the traditional black box, next to the wheel and the tank, and asks if we would still be surprised to find the wheel set into motion at the expense of the thermal energy of the water. We would of course decline to answer until we know what is inside the box. Our engineer still refuses to tell us, but offers instead to guarantee that at the end of the experiment, its contents will be in exactly the same state as at the beginning. The positions of any pistons, the state of charge of any batteries, the temperature, pressure, and volume of any substances, and the energies of all components will be the same as before. Given this much information, can we now tell if the wheel will acquire kinetic energy at the expense of the thermal energy of the water?

It may seem rash to insist that under these conditions the wheel can't begin to move, but Carnot, Kelvin, and Clausius were just that rash. They could not conceive of a complete conversion of thermal energy into the mechanical energy of a macroscopic body unless *something* changed elsewhere to account for it, some transfer of energy from one body to another: some flow of heat to a lower temperature than that of the water in the tank, for example, or else a battery operating a motor to start the wheel moving and itself being discharged as a result. If nothing else in the universe had been found to have changed, the wheel could not have started moving.

Clearly, if the device in the black box has operated cyclically, therefore returning to its initial state, it could not be the site of that necessary rearrangement of energy elsewhere that would make the motion of the wheel possible. *Something* has to change, and with a cyclically operating engine the only other *something* in the universe of this experiment, nothing has.

Why Two Temperatures Enable Work to Be Done

Carnot, and later Clausius, used the properties of an ideal gas, rather than steam, to calculate the efficiency of a heat engine because its properties were better known then. We have described them in our discussion of the kinetic-molecular theory in Chapter 4. Unlike steam, an ideal gas doesn't condense, but its pressure changes when its temperature is changed. The way an ideal gas would actually operate in a heat engine shows simply why heat engines need more than one temperature to operate.

Since the engine, like all those to which Kelvin's statement applies, is operated in a cycle, the gas, after expanding and doing work, must be recompressed to its original unexpanded state. This requires work to be done on the gas, and therefore reduces the net work done in the cycle. If the gas is compressed at the same temperature it expanded, the work done on it in compression must at least equal, and in real operation exceed, the work the gas did on expansion.

There is a way to get a net performance of work by the gas in the cycle: cool the gas before compressing it, reducing its pressure and thus

the work needed to compress it. The net work done by the gas in one cycle of operation would then be the *difference* between the work it does on expanding while warm and the work we must do on it to recompress it while it is cool. That difference is clearly larger the more the gas is cooled for the compressive step, and therefore larger the greater the temperature difference between the expansive and compressive steps.

Efficiency and the Ideal of Reversibility

Efficiency is another one of those words that are used in common speech and scientific discourse with different meanings. In the science of heat engines the meaning is spelled out very precisely in quantitative terms. First of all, the definition takes into account the fact that both thermal energy and the increase in potential or kinetic energy when mechanical work is used, say, to raise a weight or accelerate a flywheel are both forms of energy and can be measured in the same units. Further, the net quantity of work done by a heat engine cannot be greater than the quantity of heat the fuel has supplied; it will, in fact, invariably be less. The efficiency (we use the symbol *Eff.*) is defined as the ratio of the work done, W, to the heat Q_H supplied at the higher temperature of operation:

$$Eff. = \frac{W}{Q_H}$$

and must necessarily have a value between 0 and 1.

We have mentioned that the first law requires that the work done by a heat engine is the difference between the heat (Q_H) flowing in at the high temperature (T_H) and the heat (Q_L) discarded at the low temperature (T_L):

$$W = Q_H - Q_L$$

which permits the formula for efficiency to be written:

$$Eff. = \frac{W}{Q_H} = \frac{Q_H - Q_L}{Q_H}$$

We will need this formula later in the chapter.

Obviously engines can vary in efficiency, depending on flaws in design and ordinary wear and tear. They may have leaky valves or rusty cylinders; the mechanical connections to the valves may be poorly timed; the spark plugs may misfire; the boiler may lose heat because of poor insulation. All sorts of things will reduce efficiency, but even with the best design possible, the efficiency would be limited by the frictional losses that inevitably occur when there is motion and by the slowness with which heat flows. We can reduce these losses, though we cannot eliminate them, by operating the engine slowly, which requires keeping the various opposing forces as close to a state of balance as possible. This will ensure that the work done is close to the maximum possible. In practice we can never eliminate frictional losses completely, but we can often estimate their influence on the work and heat inputs and outputs and calculate the efficiency of an idealized, frictionless machine.

It is easier to visualize the idealized operation of a simpler device than a heat engine. Imagine the seesaw discussed in Chapter 2 with two riders perfectly balanced. To raise either person and increase his or her potential energy, we could add a small additional weight to the other side of the seesaw. If there were no friction, a grain of sand would be more than enough—provided we did not care how much time the raising took. In this scenario of no friction and unlimited time, the gain in potential energy of one person is just matched by the loss in potential energy of the other (if we neglect the loss in potential energy of the grain of sand). A process run in this idealized fashion is said to be *reversible.*

A heat engine can be *imagined* to be run reversibly, and if an engine were run in this manner, the work done for a given input of heat, and therefore the efficiency, would be the maximum possible. As a practical matter we do not run heat engines this way because we need to get things done in a reasonable time and we are prepared to pay the price in reduced efficiency. A reversible process, because of the near balance continuously maintained between the forces driving it, may have its direction reversed at any moment with negligible effort: just shift the grain of sand from one side to the other. When the process is run backward instead of forward, what were previously gains become losses and losses gains: the person who gained potential energy when the grain of sand was on one side loses the same amount when it is on the other.

Carnot pointed out that a heat engine run backward is a *refrigerator:* in return for the work input to operate it, it removes heat from a cold region and discards it to already warm surroundings. Inputs have become outputs. If an engine is run backward reversibly, each flow of heat

changes direction, but the amounts of heat flowing are the same. Instead of work done *by* the engine, we must do the same quantity of work *on* the engine to run it backward reversibly. This is not true for real engines and real refrigerators: the engines do *less* work than the maximum possible, the refrigerators require *more* work to run them than the minimum possible.

The reversible process is one more example of the usefulness in science of an idealization that is never achieved in practice but that provides a standard to which real processes can be usefully compared. We have already encountered other examples: the idealized frictionless world imagined by Galileo, in which moving bodies never stop moving, and the ideal gas, which behaves like no real gas ever does. These idealized entities can often be approximated closely by real ones, and how the ideal would behave if it really existed inferred by an extrapolation from the real. The idealized process sets limits that the real process cannot overcome.

In the nineteenth century, it was difficult to calculate the reversible efficiency most substances would have if used as the working substance of a heat engine, because it requires detailed knowledge of a number of physical properties not usually available then. Ideal gases do follow certain simple regularities of behavior, however, so Clausius was able to calculate the reversible efficiency of a heat engine operating cyclically with an ideal gas as the working substance. His result is expressed by the equation:

$$Eff. = \frac{T_H - T_L}{T_H} \qquad \text{(reversible, ideal gas)}$$

T_H and T_L are the higher and lower temperatures of operation. The equation shows that if the higher and lower temperatures are equal ($T_H = T_L$), the efficiency is zero: no work at all is done.

Now we come to the most subtle part of the argument. We have stated as our basic hypothesis that it is impossible for a heat engine operating cyclically, and exchanging heat with an outside body at a single temperature, to do any net work. We can see that an ideal gas engine satisfies this hypothesis, for it would do no net work when $T_H = T_L$. Any substance not an ideal gas will have properties different from an ideal gas: its pressure may change with temperature and volume in a different way, the heat it absorbs on expanding may be different. We will expect it therefore to have a different reversible efficiency also. All that our basic hypothesis seems to require is that its efficiency should also become zero when $T_H = T_L$. Thus it does not seem that a substance whose efficiency is one-and-a-half times the ideal gas efficiency or two-thirds the ideal gas efficiency should violate the hypothesis.

Carnot's argument, as modified by Clausius, proceeds as follows. Suppose there were a substance X with a reversible efficiency different from that of the ideal gas. If so, it should be possible to create a new heat engine, one whose net effect would be to convert heat gained at a single temperature into work while operating in a cycle, by combining a reversible heat engine using the ideal gas with one using substance X. This, however, would contradict our basic hypothesis that such an engine is impossible. We therefore conclude that if the basic hypothesis is correct, all working substances must give engines with the *same reversible efficiency*. Since we know the efficiency given by an ideal gas (because it happened to be easy to calculate), we know it for all other substances as well.

The Paradox of the Refrigerator

A heat engine and a water wheel are both devices for doing work, and both can be run in reverse. A water pump is in a sense a water wheel run in reverse. It doesn't do work *for* us; instead, work must be done *on* it to make it pump the water uphill. A heat engine run in reverse also requires an input of work, in return for which it removes heat from a cold reservoir and gives out heat to a warm reservoir; in other words it does the job of a refrigerator or an air conditioner. The generic term for all such devices is, unsurprisingly, *heat pump*. The heat output of a refrigerator can be detected by placing one's hand on the side or back of the unit; the heat output of an air conditioner outside the cooled room is perhaps more familiar.

Heat flows into a heat engine from what is often referred to as a *heat source*, which in turn is usually heated by burning fuel, and a lesser quantity of heat flows out to the surrounding cold environment, which is called a *sink* for heat. When we consider both engines and heat pumps operating together, the "source" and the "sink" can both give heat and receive it. As a more neutral term for source and sink, we will use *heat reservoir*.

There is an apparent paradox when we compare the reversible efficiency of an ideal gas heat engine with its performance when it is run in reverse as a heat pump. From a heat engine we want the *maximum* work output for a given heat input, the heat being supplied by the

reservoir at the upper temperature of operation. The Clausius formula for the ideal gas engine tells us the efficiency, defined as W/Q_H, is greater the greater the difference of temperatures. To operate a heat pump we want to use the *minimum* work input to withdraw a given quantity of heat from the body (the inside of a refrigerator or an air-conditioned room) we are trying to keep cool, which is at the lower temperature of operation. To characterize how well the refrigerator does its job we use a "coefficient of performance" *(C.O.P.)*, which is defined as the *ratio* of the *heat* extracted from the colder reservoir, Q_L, to the *work* we must do to extract it, W. For an ideal gas heat pump operating reversibly it can be shown that:

$$C.O.P. = \frac{Q_L}{W} = \frac{T_L}{T_H - T_L} \qquad \text{(reversible, ideal gas)}$$

A comparison of the formula for the coefficient of performance of the heat pump with the formula for the efficiency of a heat engine shows that both depend on the difference between the two temperatures T_H and T_L but in opposite ways: the *greater* the difference in the operating temperatures, the *greater* the efficiency of the reversible heat engine and the *lesser* the coefficient of performance of the reversible heat pump. Put more simply, a heat engine performs better and a heat pump worse when the temperature difference between the heat reservoirs are greater. This sounds paradoxical, but on second thought it is not: it means we use more electrical energy to operate an air conditioner or refrigerator on a very hot day than on a merely warm one. The importance of this fact for the Carnot-Clausius argument is that if working substances could differ in their reversible efficiencies, the less efficient substances, although they make worse reversible heat engines, would make better reversible heat pumps.

An Impossible Engine

Figures 5.2 and 5.3 both show reversible, cyclically operating machines—a heat engine and a heat pump, respectively. In the engine, heat from the high-temperature reservoir expands the gas to drive the engine's piston; the outputs are the work done on the flywheel which the piston drives and the heat transferred to the low-temperature reservoir. Because the operation is cyclical there is an exact energy balance: the input is equal to the output ($Q_H = W_{ENG} + Q_L$). The heat

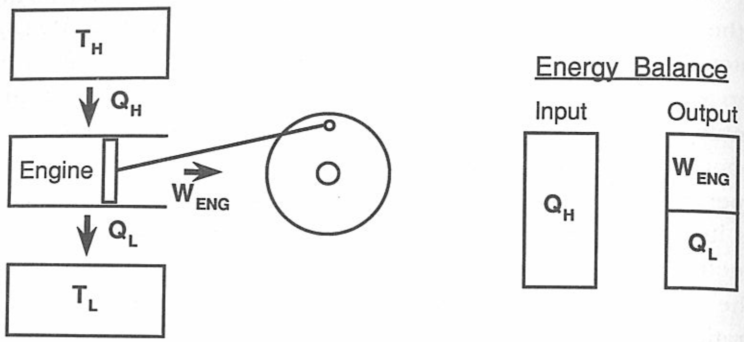

Figure 5.2 Energy Balance of a Reversible Heat Engine

In this highly schematic representation, a heat engine is shown as a cylinder with a moving piston that imparts motion to a flywheel on the *right;* it operates both cyclically and reversibly. The high-temperature reservoir (at a temperature T_H) is represented by a rectangle above the engine, the low-temperature reservoir (at a temperature T_L) by a rectangle below. The quantities of heat flowing between the reservoirs and the engine are symbolized by arrows labeled Q_H and Q_L, which show the direction of flow: *from* the reservoir at T_H *toward* the engine (an *input* to the engine), and *from* the engine *toward* the reservoir at T_L (an *output* from the engine). In addition, an arrow labeled W_{ENG} (for net work done *by* the engine) points from the engine toward the flywheel, symbolizing that the engine does work on the flywheel (an *output* from the engine).

The energy balance of the engine during one cycle of operation is shown by the two rectangles labeled *Input* and *Output.* We find it convenient to diagram *work* inputs and outputs differently from *heat* inputs and outputs. When a heat engine goes through a complete cycle, the working substance varies in temperature in different phases of the cycle, exchanges heat with external reservoirs, and is at times doing work on external bodies (during expansion of the working substance) and at others having work done on it (during compression). We will lump together work done *by* the engine in some parts of its cycle and work done *on* the engine in other parts of the cycle to obtain a single *net* work, which for an engine appears as an output. On the other hand, heat inputs and outputs will be kept separate according to the temperature of the working substance during the input or output. In the engine, Q_H represents a heat input at the higher temperature of operation, Q_L an output at the lower temperature. The diagram for the energy balance therefore shows a heat input at the higher temperature, a heat output at the lower temperature, and a net work output.

Since we only consider cyclical operation here, and the energy is restored to its original value once in each cycle, net energy inputs must equal net energy outputs. In other words, the net work and heat outputs must be combined to give a total energy output, which must equal the heat input.

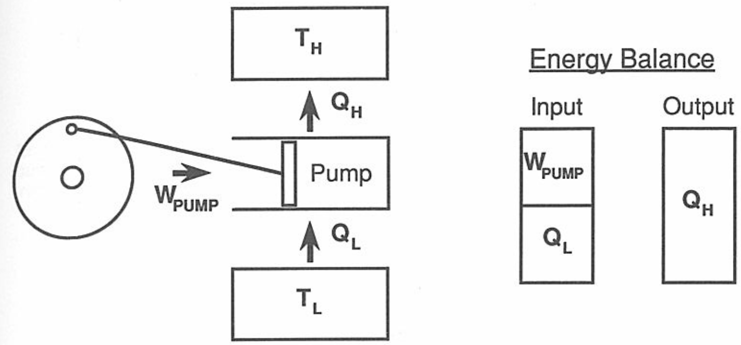

Figure 5.3 Energy Balance of a Reversible Heat Pump

In this corresponding diagram for a heat pump, the high-temperature and low-temperature reservoirs are represented as they were for the engine. The flywheel, which must do work to operate the pump, and which slows down as it does so, is shown to the *left* of the pump. The arrows showing heat flows Q_H and Q_L and work done on the pump W_{PUMP} point in directions opposite from those in the previous diagram: Q_H toward the reservoir T_H, Q_L and W_{PUMP} toward the pump.

The energy balance diagram shows a net work input W_{PUMP} and heat input Q_L, which combined give a total energy input that equals the heat output Q_H.

pump is the reverse of the heat engine. Work is done by the flywheel to drive the pump's piston, which pumps heat from the low-temperature reservoir; these make up the energy input. The pump then delivers an output of heat to the high-temperature reservoir. Again the energy input is exactly equal to the output ($W_{PUMP} + Q_L = Q_H$). Since a heat engine can do the work to set a flywheel into motion, and a moving flywheel can do the work to operate a heat pump, we could combine them into a composite engine. We will find that if the working substances of engine and pump differed in their reversible efficiencies, the second law would be violated.

We show a hypothetical "composite engine" in Figure 5.4. Our starting hypothesis is that there is a substance X with a lower reversible efficiency than the ideal gas. We imagine an ideal gas heat engine providing the work needed to operate a substance X heat pump. There is both a high-temperature reservoir at T_H, from which the engine extracts heat and to which the pump delivers it, and a low-temperature reservoir at T_L, to which the engine delivers heat and from which the

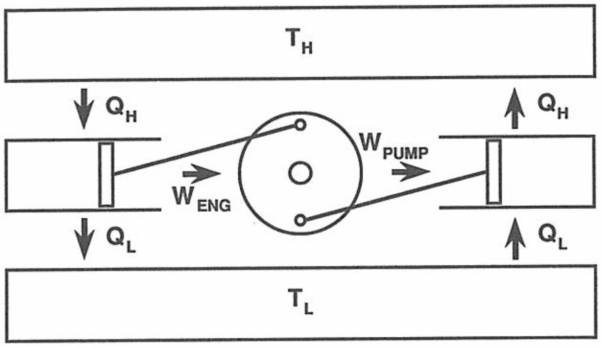

Figure 5.4 Energy Balance of a Composite Engine: Can Heat Be Converted Completely to Work?

The composite engine combines a reversible heat engine and a reversible heat pump. The energy inputs and outputs are drawn so that we may apply them to two different cases: first, we consider the operation of the engine and the pump when both use the ideal gas as the working substance; and, second, we consider what would happen if the pump used the less efficient substance *X*. The operating conditions are chosen so that in both cases the heat extracted from the high-temperature reservoir is exactly equal to the heat delivered to it by the pump.

In the first case, all the work output of the engine is needed to operate the pump. The heat input from the low-temperature reservoir to the pump is exactly equal to the heat output of the engine to the T_L reservoir. The net effect of the operation of the first composite engine is *nothing at all:* (1) all heat given out by either reservoir is restored to it; (2) no net work is done on any other bodies. It is exactly as though 100 liters of water falling 20 meters was used to operate a reversible pump that pumped the same 100 liters of water back up the 20 meters it fell. (Obviously, only idealized, reversible operation could accomplish this net cancellation of all effects.)

In the second case we come to the heart of the argument. As before, the heat pump, this time using substance *X*, is operated so as to deliver, during one cycle of operation, the *same* quantity of heat Q_H to the high-temperature reservoir as the ideal gas engine received from it in that same cycle. This ensures that the

pump extracts it. We imagine operating the engine and pump in such a way that the pump delivers the *same* quantity of heat to the high-temperature reservoir in each cycle as the engine receives from it, hence the high-temperature reservoir is restored in each cycle of operation to its initial condition. *Because it is operated cyclically we can consider it part of the cyclically operating engine.* We can imagine it as being inside a black

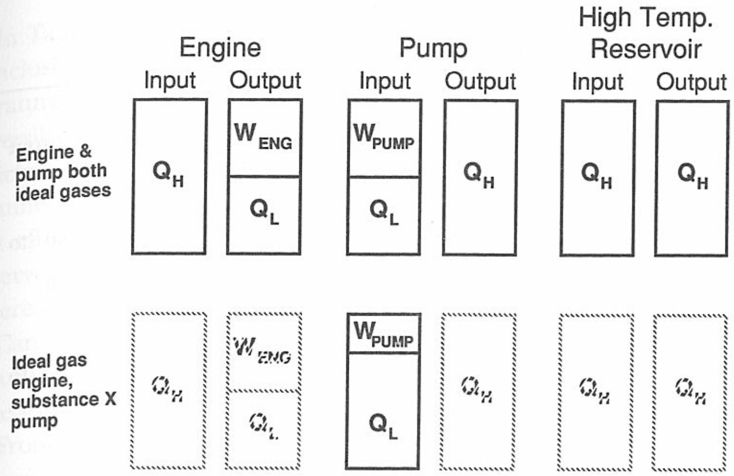

Figure 5.4 continued

high-temperature reservoir is operated cyclically. Because of the lower reversible efficiency of substance *X*, *less* work input is needed to deliver the *same* quantity of heat to the high-temperature reservoir than an ideal gas heat pump would need. (As we noted in the text, if there were a less efficient substance than an ideal gas, it would make a worse reversible engine but a better reversible heat pump.)

The energy balances for this composite engine (consisting of an ideal gas heat engine, a substance *X* heat pump, and the high-temperature reservoir, the whole operating in a cycle) are identical with those of the first case *except for the input to the heat pump*. The ideal gas engine now has a work output large enough to operate the substance *X* heat pump and *do other work besides*. Energy conservation is satisfied because a net heat equivalent to this extra work is extracted from the low-temperature reservoir. The composite engine has therefore performed work and extracted thermal energy from a reservoir at a *single* temperature, T_H. There is no lower-temperature reservoir to which heat is delivered at all. This arrangement clearly violates Kelvin's hypothesis.

box with the engine and the pump, the whole constituting the composite heat engine.

The detailed analysis shows that the composite engine does the impossible. Operating cyclically, it withdraws heat from a reservoir at a single temperature T_L and converts it all to work. No heat is discharged to a temperature lower than T_L.

Table 5.1

Energy Calculations for a Composite Engine

Heat pump operated with ideal gas	Input	Output
Engine		
Heat	100 (from T_H)	50 (to T_L)
Net work	0	50
Pump		
Heat	50 (from T_L)	100 (to T_H)
Net work	50	0
Composite		
Net heat	0	0
Net work	0	0
Heat pump operated with substance X	**Input**	**Output**
Engine		
Heat	100 (from T_H)	50 (to T_L)
Net work	0	50
Pump		
Heat	75 (from T_L)	100 (to T_H)
Net work	25	0
Composite		
Net heat	25 (from T_L)	0
Net work	0	25

Note: The composite engine consists of an ideal gas heat engine, a heat pump, and a heat reservoir at T_H. Energy calculations are shown first for the heat pump operated with an ideal gas as the working substance and then for the pump operated with substance X, whose reversible efficiency is half that of an ideal gas. Temperatures T_H = 600 K and T_L = 300 K are assumed. Heat and work are measured in joules.

In Table 5.1 we use a numerical example to demonstrate the same conclusion. To get simple numbers, we assume upper and lower temperatures at 600 K (327°C) and 300 K (27°C), giving an ideal gas reversible efficiency of 0.50. We further assume that the reversible efficiency of substance X is 0.25 for these temperatures. Finally, we assume that in one cycle of operation Q_H will always be 100 joules. The net effect is that 25 joules of thermal energy are extracted from the reservoir at 300 K and converted to work, with no other changes anywhere else.

Can substance X have a higher reversible efficiency? If it did, we would use it for our heat engine and the ideal gas for the heat pump. The logic is the same, and the result is equally impossible.

From this analysis we find that Carnot's conclusion was correct after all, even though his argument was flawed by his use of the caloric theory. All substances must have the *same* reversible efficiency in a heat engine operated cyclically, an efficiency that depends only on the temperatures at which heat exchanges occur.

The Universal Efficiency

Since we know this efficiency for the ideal gas, we know it for every other substance as well—namely:

$$\mathit{Eff.} = \frac{T_H - T_L}{T_H} \qquad \text{(reversible, any substance whatever)}$$

This is the crucial result of the argument, and a necessary step to exploring all the consequences of the second law.

In following the Carnot-Clausius argument that *if* any substance had a different reversible efficiency from an ideal gas *then* Kelvin's hypothesis would be violated, we have done a careful analysis of energy inputs and outputs for a complicated device. It is easy to get lost in the details and lose sight of the overall purpose of the argument. Let us give an example of something analogous and more familiar: the commodity prices reported for different markets. For example, we may read in the paper that the price of gold in London this morning is $465.50 per ounce, while in New York it is $467. It may occur to us that if we could buy a large quantity of gold in London and sell it in New York on the same day, we could get rich. While this ignores brokers' commissions and time delays during which price changes could wipe out potential profits, clever people do sometimes get rich on similar transactions.

In the work of Carnot and Clausius, the reversible efficiencies of different working substances in heat engines are somewhat like the price of gold, but there are no quick ways to get rich: the price is the same in all markets.

Improving Efficiency

Our analysis has focused on the efficiency of idealized reversible engines. But efficiency is a goal in the operation of real heat engines also, both for reasons of economy and for the protection of the environment.

The formula for the reversible efficiency of a steam engine operating between the boiling point of water and an ambient environment at about 25°C gives a result of about 0.2. (*Eff.* = (373 − 298 K) / 373 K = 0.2.) This is a fairly low efficiency. Do we really need to waste 80 percent of the heat provided by our fuel to operate a steam engine? The answer, according to Carnot's principle, is, no, we don't. If we can raise the upper temperature of operation we can obtain a higher efficiency, and there are many ways to do this.

First of all, water under pressure boils at a higher temperature and produces high-pressure steam. High-pressure steam engines were developed before the second law of thermodynamics was discovered, and in fact heat engine efficiencies were improved dramatically over time with very little input from the laboratory. (This fact led one historian of science to remark that science owes more to the steam engine than the steam engine owes to science.) Most electricity-generating plants running on fossil fuels use high-pressure steam at 500°C (773 K) and reach a real, as opposed to a reversible, efficiency of about 40 percent (the ideal efficiency would be 61 percent for these temperatures). While other working substances might permit higher temperatures than those possible with inexpensive and convenient water, we run into limitations on the ability of available materials that can be used for the moving parts of heat engines—such as pistons, cylinders, or turbines—to withstand higher temperatures. Proposals have been made for a process that would convert thermal energy directly to electricity from hot, electrically charged gases flowing in magnetic fields (first studied in the atmospheres of stars; the field is called *magnetohydrodynamics*). Estimates have been made that gases at temperatures of 2,000–2,500°C could be used and that real (not reversible) efficiencies of 50–60 percent could be achieved.

The conclusion that all working substances have the same *reversible* efficiency can be a misleading guide to the practical details of heat engine design. No *real* heat engine or heat pump is run reversibly: we cannot wait inordinately long periods of time to get jobs done. Real efficiencies fall below the reversible limit and depend on such properties of the working substance as the time it takes heat to flow into it (its "thermal conductivity"), its boiling point and melting point in relation to the temperatures of operation, and various properties of the materials of which the engine itself is made. It is no accident that we use water in steam engines but not in refrigerators, where ammonia and other low-boiling substances which evaporate faster and do not freeze under operating conditions perform better.

It should also be remembered that the Carnot limit applies only to the conversion of thermal energy to other forms. There is no comparable restriction on the conversion of mechanical energy to electrical, as in a hydroelectric station, or electrical to chemical, as when we charge a battery, or chemical to electrical to mechanical, as when we use the charged battery to operate the starter of an automobile. In a later chapter we will examine how living organisms do work by a direct conversion of the chemical energy of foods to motion. There is no heat engine, and no Carnot limit. None of these processes involves cyclical operation, however; if they did, the second law assures us that they would do no net work.

Entropy: A New Property of Matter

The importance of the second law is not so much what it tells us about heat engines, but rather what it tells us about the properties of matter. Any body, whether it is a single pure substance or a complicated combination of substances, can be imagined to be the working substance of a heat engine. When the body is subjected to a series of changes in temperature and in the forces between it and other systems, there will be heat and work inputs and outputs predictable from the properties of the body. We can then calculate what its reversible efficiency in some cyclical process would be from the calculated inputs and outputs. This efficiency will depend on one hand on the specific properties of the body, but on the other hand it must always be given by the general formula

$$Eff. = \frac{T_H - T_L}{T_H}$$

This in turn implies that the physical properties from which the efficiency is calculated must be interrelated in just such a way as to always give us that general formula. Thus we have been able to learn something useful and important about those physical properties, even though we may not have discovered a better heat engine. This approach was used, as we will see in Chapter 6, by Kelvin and his brother early in the history of thermodynamics to explain some baffling properties of water and ice.

There is, however, an alternative and more powerful way to explore the consequences of the second law. We have already shown that the first law implies that matter has a previously unsuspected property, thermal energy, and prescribes procedures for determining this property. The statements of the second law given earlier can be shown by some logical and mathematical analysis to imply that matter has an additional unsuspected property, called *Entropy* by its discoverer Rudolf Clausius, and to prescribe procedures for determining this new property.

The first law requires that when any process takes place, the total energy change for all bodies involved must be zero: what one body loses another must gain. The second law requires that when any process takes place, the total entropy change for all bodies involved cannot be negative. This means that the total entropy cannot ever decrease, but it may increase, and in real (as opposed to reversible) processes, where friction and other forms of inefficiency always operate, the total entropy will *always* increase. Unlike energy, entropy is not conserved. The amount of it in the universe gets greater all the time.

The reasoning that led Clausius to the concept of entropy as a property of matter involves some difficult logical and mathematical steps and requires the use of calculus. In the following summary some of these mathematical steps are omitted; we hope that even without them the reader can get some insight into the argument.

First, what do we mean when we speak of a "property of matter"? Let us think of definite samples of matter—a kilogram of water, 2,000 kilograms of a mixture of 20% oxygen and 80% nitrogen, 1 gram of platinum—and consider what properties these samples might have: density, electrical resistance, speed of sound propagation, hardness, color, and so on. Some instrument or combination of instruments will measure each of these properties on our sample of matter. We can easily verify by simple experiments that changes in the "state" of the matter—a change in its temperature or in the pressure or in any electrical fields acting on it—will produce changes in each of its properties, though the changes need not always be large. And we expect also that if, after having changed the state, we then return to the original state (that is, to the original temperature, pressure, electric field) so that the process has been a *cyclical* one, the properties will return to their original values. If they did not, we would hesitate to call them "properties."

For example, measurements performed on any sample of matter undergoing rapid and violent chemical reaction—the contents of a blast furnace, or a sample of rubber decomposing at 1,000°C—would give unreliable and rapidly changing values for anything we try to measure. Our concept of "property" therefore seems to require two things: first, that we should be able to perform a direct measurement of the property, and, second, that the value should be stable: there should be no net change during a cyclical process. This stability implies that the property has a definite value associated with each definite state of the sample of matter. If we change the state, the property takes on a new value, and we can measure it again in the new state.

Properties That Cannot Be Directly Measured

Now, although we have said energy and entropy are also properties of matter, they lack one of the crucial defining features: there are no instruments to measure them, no energy-meters or entropy-scopes we can apply to our kilogram of water or our gram of platinum that will give us numerical values for the properties. Why do we claim that they are properties, then? While we cannot measure either of them directly when our sample of matter is in a definite state, we can measure the *changes* in energy or entropy when the sample changes from one state to another. Because we can measure these changes, we can show in the laboratory that the net changes of either energy or entropy in cyclical processes always add up to zero. So although we can't measure the value of the energy or entropy of a kilogram of water at, say, 20°C, 10 atmospheres pressure, and an electric field of 50 volts per meter, we can measure their net changes when the water is brought to that state from 0°C, 1 atmosphere pressure, and zero electric field. We therefore can measure these particular properties relative to some arbitrarily selected reference state, but otherwise they are properties as definite and as stable as electrical resistance, hardness, and all the others.

This idea that the existence of a property can be proved by showing that changes in all cyclical processes are zero can be made clearer by a more easily visualized analogy. We will consider the property "height-above-sea-level" or, more simply, "elevation." There is a definite numerical value of elevation at each point on the earth's surface. Suppose we are exploring a mountain all trails on which are within sight of a landmark at sea level, and we are equipped with surveyor's instruments, including a telescope. At any point on any trail we are able to make a measurement of its elevation. For example, we can first determine with our instruments that the elevation of a base camp is 250 meters and, after climbing to the summit, determine that *its* elevation is 4,000 meters. We expect, of course, that if we return to any point after wandering anywhere else on the mountain, a repetition of the measurement of elevation will give the same answer as before. This verifies that elevation is a property (barring earthquakes, of course).

Now, however, suppose that our particular mountain is always shrouded in an impenetrable fog that debars us from using our telescope. Is there any way we could test whether elevation is determined by our position on the mountain? Suppose we were to use a meter stick and a carpenter's level to measure accurately in centimeters the *change* in elevation for each step we take along the trail. Thus, for example, we could determine by tediously adding up all the changes of each step that the summit is 3,750 meters higher in elevation than the base camp. Obviously we could also add up all the changes in elevation for any hike that returns us to some particular point on the mountain—the hike is thus cyclical. We would find that the net change on returning to any starting point is zero. We would conclude that elevation really is a property. We could therefore determine the elevation of any point on the mountain relative to some arbitrary reference point, such as the base camp, even though we do not know the elevation above sea level of either base camp or summit.

How Entropy Is Defined

Clausius was aware that the net heat inputs and outputs of a heat engine do not add up to zero. The first law forbids it, since if work is done by the engine, it must take in more heat than it gives out. But he did realize that there was a quantity related to the heat inputs and outputs that did add up to zero for any *reversible* cycle whatever. Let us follow his reasoning:

The first law requires that the work, W, must equal the difference between the heat received by the engine, Q_H, from the hot reservoir and the heat, Q_L, discharged to the cold reservoir:

$$W = Q_H - Q_L$$

This permits the equation for the reversible efficiency of a heat engine operating between two temperatures to be written:

$$Eff. = \frac{W}{Q_H} = \frac{Q_H - Q_L}{Q_H} = \frac{T_H - T_L}{T_H}$$

A little algebraic manipulation of the above equation gives us

$$\frac{Q_H}{Q_L} = \frac{T_H}{T_L}$$

In words, the ratio of the heat extracted from the hot reservoir to the heat discharged to the cold one is equal to the ratio of the absolute temperatures of the two reservoirs.

A little further algebraic manipulation leads to:

$$\frac{Q_L}{T_L} = \frac{Q_H}{T_H}$$

Q_L is the quantity of heat lost by the engine to the cold reservoir and Q_H the quantity of heat gained from the hot one. We paraphrase the above equation by saying that the loss of Q_L/T_L is equal to the gain of Q_H/T_H. This doesn't by itself make Q/T a property, as the heat engine operating reversibly with only two temperatures is a very restricted kind of cycle. But Clausius was able to prove, by using calculus, that any completely arbitrary reversible cycle, no matter at how many different temperatures it gains or loses heat, can be divided up into a large number of two-temperature reversible cycles. For each of them the gain of Q/T at the upper temperature is balanced by a loss of Q/T that just cancels it at the lower temperature. It follows that for any arbitrary reversible cycle the net changes of Q/T add up to zero. The conclusion: since the net change around any arbitrary reversible cycle is zero, there must be a property of matter associated with those changes. He chose the name *entropy* [in German, *Entropie*] from a Greek root meaning "turn" or "change"; the same root is present in other English words, such as *tropism* and *troposphere*.

Entropy Changes in Real Processes

What about heat engines operating under real conditions, which means nonreversibly, and with lower efficiencies? First of all, once we are convinced that entropy, determined by measurements performed during reversible changes, is a property of matter, then we must concede that the change in entropy of the working substance of a heat engine operating in a cycle must be zero, even if the cycle was not carried out reversibly. This may sound contradictory, but it isn't (we recognize that this point is one of the more difficult for the reader not trained in science or mathematics). Entropy is a property whose numerical value must be determined by a reversible process, but once determined it is a property and depends only on the state of the system. Hence if the working substance in the heat engine has been returned to the same state, it must have again the same entropy as before.

Let us turn instead to the two reservoirs, which are also samples of matter and which therefore also have the property entropy. The reservoirs do not operate cyclically: in reversible operation the hot reservoir gives out the heat Q_H, and thus loses an amount of entropy Q_H/T_H, while the cold reservoir gains an amount of entropy Q_L/T_L. So, like the engine, there has been no net entropy change in the reservoirs. Now let us consider the opposite extreme from reversible operation: an engine so inefficient that no work is done at all. Under such conditions we can ignore the heat engine entirely and imagine that the quantity of heat Q_H flows directly from the hot reservoir to the cold one, so that $Q_L = Q_H$. The entropy decrease of the hot reservoir is still Q_H/T_H, but now the entropy increase of the cold one is Q_H/T_L. Since the temperature T_L is smaller than T_H, the entropy increase of the cold reservoir is now greater than the decrease of the hot reservoir. There is therefore a net entropy increase of the two reservoirs.

If the real engine, although not operating reversibly, is not this ridiculously inefficient and does some work, it is easy to show that there is a net entropy increase, although this increase is less than if the heat were just to flow unimpeded from the hot to the cold reservoir.

Once this net increase of entropy in real operation of a two-temperature heat engine is established, further detailed logical analysis shows that it is also true for all processes in nature, not just those of conventional heat engines, and not just those involving heat reservoirs at just two temperatures. The conclusion: there can be no net decrease of entropy in any natural process. In reversible processes, an unattainable ideal, there is no net change in entropy. In all real processes there must be an increase.

What Is Entropy, Anyway?

Attempts to find an intuitively simple meaning for entropy without reference to a molecular picture have not been particularly successful.

The Carnot limit on the efficiency of even reversible heat engines implies that not all the thermal energy of a heat source is available for conversion to mechanical energy, and there is a relation between the entropy of the heat source and the unavailable portion of the thermal energy. Unfortunately, the amount of thermal energy that is unavailable depends also on the temperature of the ambient environment that we are using as our cold reservoir, and that in turn depends on where we happen to site the engine. So entropy is related to the concept of unavailable thermal energy, but not in a simple way.

We have come to the end of a long story. Let us summarize it before we go on in the next chapter to describe one of the first experimental tests of the second law and discuss some of its implications.

Starting with the conviction that a heat engine operating cyclically at a single temperature should be unable to do any work, Carnot and Clausius were able to derive a remarkable consequence: that the maximum (or ideal) efficiency of a heat engine should depend only on the temperatures of operation, not on the specific properties of the working substance used. Once this was established, an even more remarkable consequence followed that extended beyond the operation of heat engines to describe the properties of all matter and the course of all natural processes. This was the discovery of entropy. Entropy is a property of matter, readily measured in the laboratory. The total amount of entropy, unlike that of energy, is not conserved, but rather must increase in all real processes.

Appendix: Entropy Changes and How They Are Determined

When we described how the thermal energy in matter is determined in the laboratory, we stated that we were required by the nature of energy to determine it relative to an arbitrarily selected reference state. The energy relative to this state was measured by keeping track of the total heat flowing into the substance and the total work done on it during a process going from the reference state to the state we are studying. It doesn't matter what kind of process is used or whether it is carried out reversibly or not.

Entropy is also determined relative to a reference state, and its determination also requires us to have knowledge both of the heat absorbed and the work done. The process by which the reference state is changed to the new state we are studying, in contrast to changes in energy, must be a reversible one: it must be a process of negligible frictional loss, or anything equivalent to friction that would reduce the work done below its maximum possible value for the path of change. Such a process cannot be carried out exactly in practice, but it can be closely approximated in the laboratory, with corrections made for any frictional losses. Under reversible conditions, and if the process takes place at a single temperature, the change in entropy of the process, ΔS, is given by the heat absorbed divided by the absolute temperature at which the process is carried out:

$$\Delta S = \frac{Q}{T}$$

As an example, consider the following experiment, which approximates a reversible process. When water is boiled slowly at 100°C under normal atmospheric pressure and is thus converted to steam also at normal atmospheric pressure, the heat absorbed per kilogram of water evaporated (the latent heat) is 2,254,000 joules. The absolute temperature corresponding to 100°C is 373 K. Dividing the latent heat by this temperature, we find the entropy change on evaporation to be 6,043 joules per kelvin per kilogram. In other words, the entropy of 1 kilogram of steam at this temperature and pressure is 6,043 joules per kelvin greater than that of 1 kilogram of liquid water.

More generally, we need to determine entropy changes for processes in which the temperature changes. The formula for these cases has to be expressed in the language of calculus, but the calculation of the entropy change is still quite straightforward. In Table 5.2 we give the energy and entropy of 1 kilogram of water as measured in the laboratory at various temperatures. We have chosen the temperature 0°C as the reference state, so we arbitrarily call the energy and entropy zero in that state.

Now let us calculate the energy and entropy changes when 1 kilogram of cold water at 0°C is brought into contact with 1 kilogram of hot water at 100°C. We know, first, the energy cannot change. We note that the energy of 2 kilograms of 50°C water is just equal to the sum of the energies of 1 kilogram of 0°C water and 1 kilogram of 100°C water (418,400 joules). It follows that the final temperature is 50°C. (This is not a particularly surprising result, but accurate measurements show that it is not strictly true because the specific heat of water depends slightly on temperature between 0°C and 100°C.) The entropy at the start is the sum of the entropies of the two kilograms of water, one at 0°C and one at 100°C. Thus $S = 0 + 1{,}305 = 1{,}305$ joules per kelvin.

The entropy at the end is twice the entropy of one kilogram of water at 50°C, or $S = 2 \times 703 = 1{,}406$ joules per kelvin.

There has therefore been a net increase in entropy of 101 joules per kelvin. The hot water has *decreased* in entropy in the process, but the cold water has *increased* by a greater amount, so that the net change is an increase. The requirement of a *net* increase clearly does not mean that some of the participants in the process cannot undergo an entropy decrease (as did the hot water). This numerical example illustrates the process considered earlier: a flow of a quantity of heat Q_H from a body at a higher temperature T_H to another body at a lower temperature T_L, leading always to a net increase of the total entropy.

The entropy of liquid water, as can be seen from the table, increases when the temperature increases. So does the entropy of everything else,

Table 5.2
The Energy and Entropy of 1 Kilogram of Water

Temperature (°C)	E (joules)	S (joules per kelvin)
0	0	0
10	41,840	151
20	83,680	297
30	125,500	435
40	167,400	573
50	209,200	703
60	251,000	833
70	292,900	954
80	334,700	1,075
90	376,600	1,192
100	418,400	1,305

provided pressure and other relevant variables remain the same; it is one of the most general rules about entropy.

There are other ways to increase entropy: expanding a gas to a larger volume while holding its temperature constant always increases its entropy. Compressing liquids and solids to smaller volumes at constant temperature usually results in decreases of their entropies; but there are exceptions to this rule: very cold water (between 0° and 4°C) is one.

When soluble substances like sugar dissolve in water there is a net entropy increase. The process of dissolving something in a large volume of a liquid solvent is in some ways like expanding a gas.

When substances undergo "changes of phase"—a solid melts, a liquid vaporizes—the phase formed by heating has the higher entropy, so invariably liquids have higher entropies than the solids they are formed from, and gases have higher entropies than liquids. Since water always has a higher entropy than ice at the same temperature, one might think that the freezing of water on a cold day leads to a decrease in entropy and a violation of the second law. But it is the *net* entropy change that must be an increase. When water freezes outdoors, it gives out heat to the cold surroundings, increasing the entropy of those surroundings by more than enough to cause a net entropy increase.