Heat Engines and the Second Law of Thermodynamics

I. INTRODUCTION

The science of thermodynamics was born from the realization that microscopic energy (such as the internal kinetic energy in an ideal gas, for example) can be used to do macroscopic, useful work W. One of the first industrial applications of this science was the development of the steam engine. Today we understand that thermodynamics governs the useful extraction of microscopic energy in all cases, including the work done by an internal combustion engine or the work done by a biological system such as the human body.

We have already studied the **first law of thermodynamics**

$$\Delta U = U_f - U_i = Q - W, \tag{1}$$

which tells us that the change in internal energy U of a system is equal to the heat flow Q (*into* the system) minus the work W (done *by* the system). By considering the first law you might come to the conclusion that we could continually extract useful work W in the following manner. First, we connect our system to an external **reservoir** that is at a temperature that is higher than that of the system.[1] Because the reservoir is at a higher temperature, heat Q flows from the reservoir into the system. At the same time we construct the system so that it can continually do work W without having its internal energy U change. Thus, from the first law [Eq. (1)] we would have $W = Q$, which implies that any amount of heat that we extract from the reservoir can be used to do useful work. The conservation of energy is satisfied, and everything is hunky-dory.

Well, in fact, no one has ever figured out how to construct such a system. In reality, the heat extracted from the high-temperature reservoir can never be fully utilized to do sustained, useful work. There is always a certain amount of heat Q that must be removed from the system (to another, lower-temperature reservoir, as we shall see) in any such system that can do sustained, useful work. The combination of (1) a high-temperature reservoir, (2) low-temperature reservoir and (3) the system that does the work (on the outside world) is generically known as a **heat engine**. The heat that must be removed from the system is know as **waste heat**. In a biological system such as the human body this waste heat is partially responsible for the increase in body temperature that accompanies physical exertion.

II. HEAT ENGINES

To understand the basis of the claim that we cannot convert all heat to sustained, useful work, lets let the

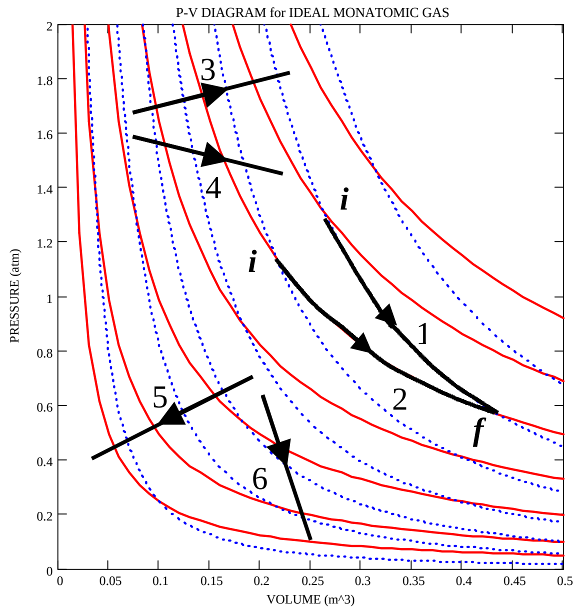

Figure 1. P-V diagram for an ideal monatomic gas. Path 1 is along an adiabat, path 2 along an isotherm. Along paths 3 and 4 the temperature is increasing and heat is flowing into the system. Along paths 5 and 6 the temperature is decreasing and heat is flowing out of the system.

system be an ideal monatomic gas (in discussing heat engines the system is also known as the **working substance** of the heat engine). In particular let's study the P-V diagram for a monatomic gas, as shown in Fig. 1. Recall, the series of solid lines in the diagram are **isotherms** while the dotted lines are **adiabats**. For example, if we let the gas expand or contract in such a way that it follows one of the solid lines, then its temperature remains constant. If it moves along one of the dotted lines then there is no heat flow into or out of the system. Conversely, if the system is allowed to expand or contract so that it crosses the solid lines then its temperature will be changing. For example, if the system expands along an adiabat (as shown by path 1) then its temperature decreases along that path since it is moving from a higher temperature isotherm to a lower temperature isotherm. Similarly, if the system expands or contracts such that it crosses the adiabats, heat will either be flowing into or out of the system. For example, along path 2 heat is flowing into the system. If the system moves (generally) from left to right or bottom to top (path 3 or 4, e.g.) the temperature increases and the heat flow is into the system. Conversely, if the system moves (generally) from right to left or top to bottom (path 5 or 6, e.g.), the temperature decreases and the heat flow is out of the

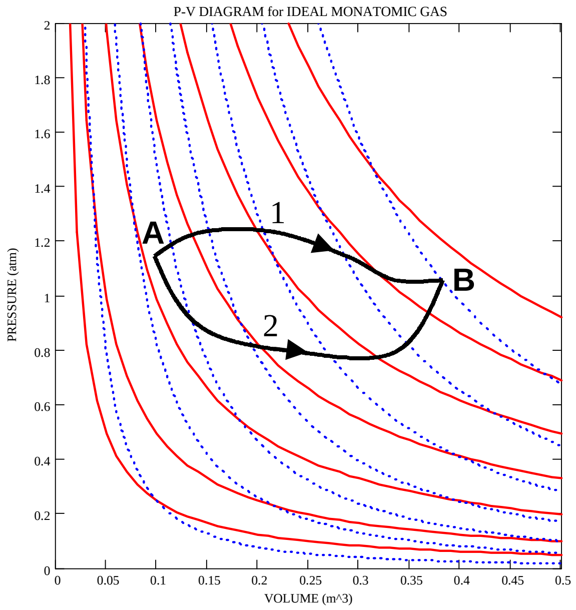

Figure 2. P-V diagram for an ideal monatomic gas. Along path 1 the work done by the system is greater than the work done along path 2. Because point **B** is at a greater volume than point **A**, the work W is positive in both cases.

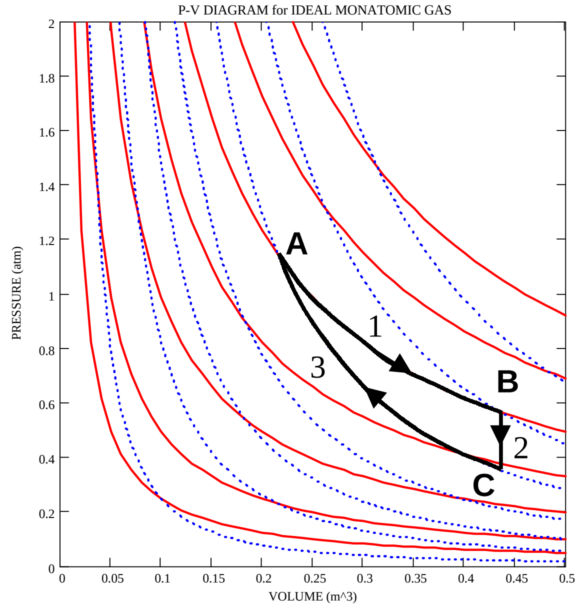

Figure 3. P-V diagram for an ideal monatomic gas. Path 1 is along an isotherm. Path 2 is along an isochor. Path 3 is along an adiabat. Heat flow in (Q_{in}) occurs along path 1. Heat flow out (Q_{out}) occurs along path 2. The net work in one thermodynamic cycle equals the area inside the closed loop.

system.

The other thing that we must keep in mind about the P-V diagram is that if we move the system between two points on the diagram then the work W done by the system is the area under the curve between the two points. The work is positive if the second point is at a larger volume V than the first point, while the work is negative if the second point is at a lower V. Looking at Fig. 2, if we move from point **A** to point **B** along path 1 the system does positive work. Note that we could also move from point **A** to point **B** along path 2. The system would still do positive work, but it would be *less* than that along path 1. This is because the area under path 2 is less than that under path 1.

Let's now think about how we can make a useful heat engine out of our ideal gas. Consider the P-V diagram in Fig. 3. We could start the system at point **A** in this diagram and let it expand along an isotherm, as indicated by path 1. In fact, along path 1 we do have total conversion of Q (let's call this Q_{in} since this is heat that is flowing into the system along the isotherm) to W since $\Delta U = 0$. However, we have not achieved very much; the gas is sitting at a larger volume. In order to get any *sustained*, *net* work out of the system we must get the system back to a smaller volume state so that it can re-expand and do more work for us. We could simply have the system move back along path 1, but the work done by the engine on this reverse path is the negative of that just done by the engine, and thus the *net* work done by the heat engine would be *zero*. However, we can move the system back to point **A** along a different path. For example we could move it back to point **A** along paths 2 and 3. In this case the net work in starting out at **A** and moving along paths 1, 2, and 3 back to **A** is positive! Why is this? Well, remember that the work done in moving between any two points is the area under the curve. The area under path 1 (when the system is doing positive work) is greater that the area under path 3 (when the system is doing negative work), so the net work done by the system is positive. We can now run the engine around this **thermodynamic cycle** as many times as we want and extract as much useful work as we need. One more point, since the net work is the area under path 1 minus the area under path 3, the *total net work in one cycle is simply the area enclose by the cycle.* This is true for any cycle. Of course, the system must move around the closed path *clockwise.* If the system moves around the cycle counter-clockwise then the outside word is doing the net positive work (on the system). (This is the basis of refrigeration, which we briefly discuss below.)

Now back to our statement about waste heat (which we now designate as Q_{out} since it is heat that leaves the system.) Notice in Fig. 3 that along path 1 heat is flowing into the system since the system is moving from left to right across the adiabats (remember, there are actually an infinite number of adiabats; only a few have been drawn). The only way that we could get back to point **A** from point **B** without throwing away any heat would be to move along an adiabat from point **B** to **A**, but this is impossible. In fact, *any* path that gets the system back to point **A** must cross the adiabats in the right to left (or top to bottom) direction, which means that some amount of heat Q_{out} must leave the system.[2] For the cycle shown, the waste heat Q_{out} is extracted along path 2 (since path 3 is along an adiabat). The other important fact about any cycle where the system does net work is that the waste heat Q_{out} leaves the system when the system is (at least on average) at a *lower temperature* than when Q_{in} flows into the system. This can be seen for the cycle in Fig. 3 by noting that Q_{in} occurs while the system is at a constant temperature (along path 1), while Q_{out} occurs as the system is cooling (along path 2). Thus, in moving from point **B** to point **C** along path 2, the system must be hooked up to a reservoir that is at a lower temperature than that when the system was moving along path 1 (or else Q_{out} would not flow out of the system).

Summarizing, a heat engine is a system that moves along a thermodynamic cycle, extracting Q_{in} from a high-temperature reservoir, expelling waste heat Q_{out} to a lower temperature reservoir, and in the process doing useful net work W.

Because $U_f = U_i$ around a complete thermodynamic cycle, the first law tells us that $W = Q_{in} - Q_{out}$. The efficiency e of the engine is defined as the ratio of W to Q_{in} around a complete cycle. That is,

$$e = \frac{W}{Q_{in}} = 1 - \frac{Q_{out}}{Q_{in}}. \tag{2}$$

Because $Q_{out} \neq 0$ the efficiency can never be equal to 1. Therefore, no heat engine is 100% efficient.

III. SECOND LAW OF THERMODYNAMICS

That the efficiency of a heat engine can never be 100% is the essence of the **second law of thermodynamics**, although there are many other different, but equivalent, statements of this law. Lets look at two of the more common statements of this law.

One statement of the law is *one cannot devise a thermodynamic cycle whereby heat is converted to work at a single temperature*. However, keeping the system at a single temperature would mean that the system would stay on the same isotherm. As we have seen in our discussion of Fig. 3, no net work can be done in that case.

Another statement of the second law is *without external work done* ***on*** *the system, heat only flows from a hot to a colder substance*. This is equivalent to the statement in the last paragraph, as the following example shows. If we could convert heat to work at a single temperature then we could devise a heat engine that did net useful work with only heat input at that temperature; the mechanical work could then be converted to internal energy (through friction) at some higher temperature location and this internal energy could then placed into a higher-temperature reservoir. The net result would be the flow of heat from the lower temperature (heat-engine-input) reservoir to a higher temperature reservoir *without* any *external* work on the system.

In fact, we *can* move heat from a lower-temperature reservoir to a higher-temperature reservoir; it just takes external work *on* the system. Refrigerators and heat pumps are examples of devices that indeed transfer heat from a cold-temperature location to a high-temperature location – but they certainly do not do this spontaneously! External work must be performed on the working substance of these devices in order to perform this heat transfer.

IV. ISOCHOR – ISOBAR HEAT ENGINE

Any particular type of heat engine is defined by the paths that compose its thermodynamic cycle. For example, a fairly simple heat engine to think about is the isochor-isobar heat engine, illustrated in Fig. 4. The net work done for this cycle is easy to calculate because it is a rectangle: $W = \Delta P\, \Delta V$, where $\Delta P = P_1 - P_2$ and $\Delta V = V_2 - V_1$ are the lengths of the vertical and horizontal sides of the rectangle, respectively.[3] By noting which direction the paths cross the adiabats we deduce that along paths 1 and 4 heat flows into the system, and along paths 2 and 3 heat flows out of the system. Although it is a bit tricky to show, the average temperature during the Q_{in} paths is larger than the average temperature during the Q_{out} paths.

Let's see what we can say about the efficiency of an isochor-isobar engine. To do this we need to know two of the following three items: W, Q_{in}, and Q_{out}. Along the isochors $W = 0$, while along the isobars the work is simply P time the change in volume, which is ΔV along path 1 and $-\Delta V$ along path 3. As we discussed in class, along an isobar $Q = (5/2)\, Nk_B\, \Delta T$. Combining this with $P\, \Delta V = Nk_B\, |\Delta T|$ yields $|\, Q\,| = (5/2)\; P\, \Delta V$. Similarly, along an isochor $Q = (3/2)\, Nk_B\, \Delta T$. Combining this with $V\, \Delta P = Nk_B\, |\Delta T|$ yields $|\, Q\,| = (3/2)\; V\, \Delta P$. The results for all four paths are summarized in the table below. Using these expressions the efficiency can now be calculated from Eq. (2) as

$$e = \frac{W}{Q_{in}} = \frac{\Delta P\, \Delta V}{\frac{5}{2} P_1\, \Delta V + \frac{3}{2} V_1\, \Delta P} \tag{3}$$

The right hand side can be rearranged to express the efficiency as

$$e = \frac{1}{\frac{5}{2}\frac{P_1}{\Delta P} + \frac{3}{2}\frac{V_1}{\Delta V}} \tag{4}$$

This equation actually tells us something very interesting. First, the efficiency can be made as large as possible

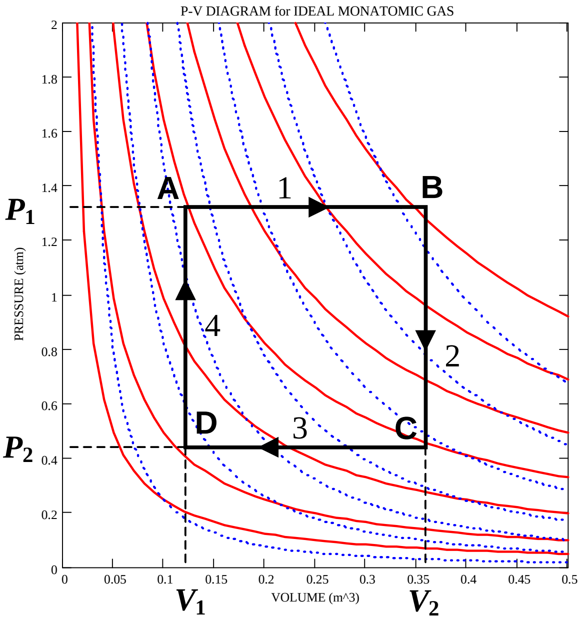

Figure 4. P-V diagram for an ideal monatomic gas, showing an isochor-isobar thermodynamic cycle. Paths 1 and 3 are along isobars. Paths 2 and 4 are along isochors. Heat flow in (Q_{in}) occurs along paths 1 and 4. Heat flow out (Q_{out}) occurs along paths 2 and 3. The net work W in one thermodynamic cycle equals the area inside the closed loop, $\Delta P \Delta V$.

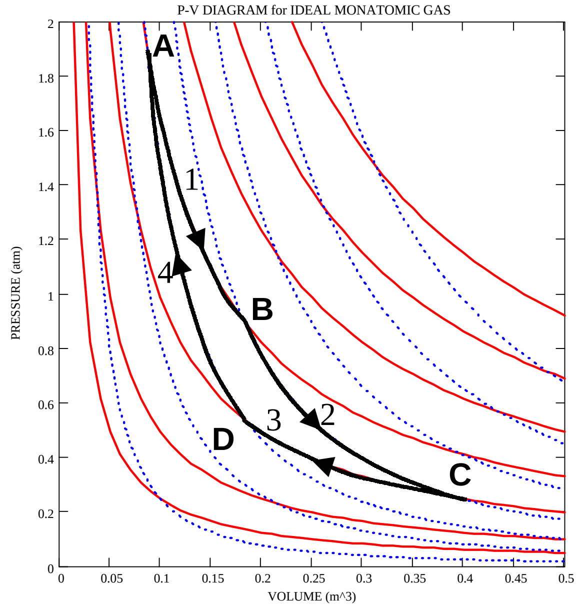

Figure 5. P-V diagram for an ideal monatomic gas, showing an isotherm-adiabat (Carnot) thermodynamic cycle. Paths 1 and 3 are along isotherms. Paths 2 and 4 are along adiabats. Heat flow in (Q_{in}) occurs along path 1. Heat flow out (Q_{out}) occurs along path 3. The net work W in one thermodynamic cycle equals the area inside the closed loop.

if (1) V_1 is as small as possible and V_2 as large as possible, so that $V_1/\Delta V \approx 0$, and (2) P_2 is made as small as possible and P_1 as large as possible, so that $P_1 \approx \Delta P$. In these limits we come close to the maximum efficiency of $2/5 = 40\%$ for an isochor-isobar heat engine. That is a lot of waste heat, at least 60% of the input heat, but we could probably live with that; the heat engine can still do a considerable amount of useful work.

Path	W	Q_{in}	Q_{out}
1	$P_1 \Delta V$	$(5/2)\ P_1 \Delta V$	
2			$(3/2)\ V_2 \Delta P$
3	$-P_2 \Delta V$		$(5/2)\ P_2 \Delta V$
4		$(3/2)\ V_1 \Delta P$	

V. ISOTHERM-ADIABAT HEAT ENGINE

Perhaps the most famous heat engine is the isotherm-adiabat heat engine, also known as the Carnot-cycle heat engine, named after the first person to write down the second law of thermodynamics, Sadi Carnot. The Carnot cycle is illustrated in Fig. 5. Starting at **A** the system is first expanded isothermally at a high temperature T_H (path 1, the Q_{in} phase) and then adiabatically (path 2). It is then compressed isothermally at a low temperature T_C (path 3, the Q_{out} phase) and then adiabatically (path 4) back to point **A**. While it is difficult to build a real Carnot engine due to its oblique P-V diagram, it is conceptually simple since there is heat flow only along the isotherms. Thus, both the high-temperature and low-temperature reservoirs can each be held at a fixed temperature, the hot reservoir just above T_H and the cold reservoir just below T_C.

Furthermore, this engine has the property that $Q_{out}/Q_{in} = T_C/T_H$, so that the efficiency can written as

$$e = 1 - \frac{T_C}{T_H} \tag{5}$$

Keep in mind that these two temperatures are absolute temperatures. In principle, if the cold reservoir were at absolute zero then the efficiency would indeed be 100%. However, one can never reach absolute zero (this is essentially the **third law of thermodynamics**) so that no heat engine is 100% efficient. However, Eq. (5) clearly shows that the smaller the ratio of T_C to T_H, the more efficient will be the engine.

VI. REFIGERATORS AND HEAT PUMPS

A refrigerator or heat pump is simply a heat engine that is run in reverse, reverse meaning that the thermodynamic cycle is traversed *counterclockwise*. Consider the cycle in Fig. 3. If we run the system around this cycle counterclockwise, then the outside word does net work on the system. Furthermore, path 2 is now the Q_{in} path and path 1 is the Q_{out} path. Remember, path 2 is at a lower average temperature than path 1, so we have heat flowing in from a low temperature reservoir and heat flowing out to a high temperature reservoir – a refrigerator or heat pump!

---

[1] We assume that the reservoir is so large that the heat flow from or to it does not change its temperature. However, we do not need to assume that the reservoir is at a constant temperature. As we will find out when we discuss entropy, to make the engine as efficient as possible we want the temperature of the reservoir to be only slightly different in temperature from the system. For heat flow in (Q_{in}) the reservoir will be at a slightly higher temperature than the working substance. For heat flow out (Q_{out}) the reservoir will be at a slightly lower temperature than the system.

[2] The standard convention for heat flow is that Q is positive for heat flow in and negative for heat flow out. However, when discussing heat engines we let both Q_{in} and Q_{out} be positive, but keep in mind that Q_{in} always flows in and Q_{out} always flows out.

[3] Keep in mind that throughout this discussion both ΔP and ΔV as defined are positive.