

Reading assignment. Schroeder, section 3.4.

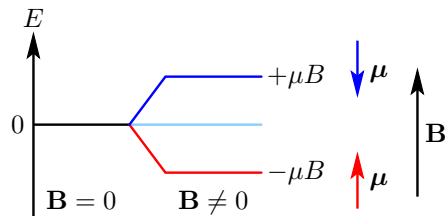
0.1 The paramagnet revisited

Now we'll explore in more detail the properties of the two-state paramagnet, one of our toy systems. It's actually much more than a toy, though, since it is physically realizable to a good approximation. One of the prime examples consists of a crystal lattice composed of atoms whose nuclei have net magnetic moments that are weakly coupled to each other and very weakly coupled to the lattice itself. This makes it possible for equilibrium to be established among the magnetic moments much faster than between the moment system and the lattice, and that makes it possible to explore the rather curious properties of this system experimentally.

Recall that the magnetic moments in a two-state paramagnet arise from the spins of particles, possibly electrons or particles in a nucleus. These spin-1/2 particles have just two possible energy states in an external magnetic field, either aligned (“up”) with the magnetic field or antialigned (“down”) with the field. The energy of a single moment is

$$E = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad (1)$$

and the quantization leads to just two possible energies $E = \pm\mu B$:



The aligned (up) moments have lower energy than the antialigned moments. Notice that the energies of the two alignments become equal as the strength of the external magnetic field goes to zero, so we've chosen the zero of energy to be that state, with the quantized energies in a nonvanishing field lying symmetrically above and below that energy.

The total energy of a system of identical moments in an external field is then

$$U = \sum_i (\pm)_i \mu B = \mu B (N_\downarrow - N_\uparrow). \quad (2)$$

If the numbers of aligned and antialigned moments are not equal, the systems as a whole has a net magnetic moment and is said to be magnetized. The net magnetic moment, or *magnetization*, is

$$M = \sum_i (\pm)_i \mu = \mu (N_\uparrow - N_\downarrow) = n\mu, \quad (3)$$

where $n = N_{\uparrow} - N_{\downarrow}$ is the net number of aligned moments. As you can see, the energy can be written in terms of the magnetization:

$$U = -MB = -n\mu B. \quad (4)$$

[EOC, Mon. 2/13/2006, #15; HW05 closed, due Mon. 2/20/2006]

The macrostates can be characterized by any of U , n , N_{\uparrow} , or N_{\downarrow} , and as we've seen before, the number of corresponding microstates (multiplicity) is given by the binomial coefficient

$$\Omega(N, N_{\uparrow}) = \binom{N}{N_{\uparrow}}, \quad (5)$$

just the number of ways of choosing N_{\uparrow} up moments out of a total of N . It's just like flipping coins. Written as a function of the net number of aligned moments, which is proportional to the energy, the multiplicity looks like:

$$\Omega(N, n) = \frac{N!}{\left(\frac{N+n}{2}\right)! \left(\frac{N-n}{2}\right)!}. \quad (6)$$

To find the entropy, $S = k \ln \omega$, we'll need the logarithm, but the manipulations will be a little easier if we first apply Stirling's approximation, then take the log. We'll use the cruder form of Stirling's approximation,

$$N! \approx N^N e^{-N}, \quad (7)$$

omitting the factor $\sqrt{2\pi N}$, which limits us to very large values of N . With that approximation, we have

$$\begin{aligned} \Omega(N, n) &\approx \frac{N^N e^{-N}}{\left(\frac{N+n}{2}\right)^{(N+n)/2} e^{-(N+n)/2} \left(\frac{N-n}{2}\right)^{(N-n)/2} e^{-(N-n)/2}} \\ &= \frac{2^N}{\left(1 + \frac{n}{N}\right)^{(N+n)/2} \left(1 - \frac{n}{N}\right)^{(N-n)/2}}. \end{aligned} \quad (8)$$

The logarithm is then, with no further approximations,

$$\begin{aligned} \ln \Omega(N, n) &\approx N \ln 2 \\ &\quad - \frac{N}{2} \left[\left(1 + \frac{n}{N}\right) \ln \left(1 + \frac{n}{N}\right) + \left(1 - \frac{n}{N}\right) \ln \left(1 - \frac{n}{N}\right) \right]. \end{aligned} \quad (9)$$

The entropy is simply $k \ln \Omega$, but it will be convenient to express it as a function of $U = -n\mu B$, rather than n :

$$\begin{aligned} S(U, N) &= k \ln \Omega(N, U) \\ &\approx kN \ln 2 - \frac{kN}{2} \left[\left(1 - \frac{U}{N\mu B}\right) \ln \left(1 - \frac{U}{N\mu B}\right) \right. \\ &\quad \left. + \left(1 + \frac{U}{N\mu B}\right) \ln \left(1 + \frac{U}{N\mu B}\right) \right]. \end{aligned} \quad (10)$$

Now, let's see what the entropy looks like as a function of U . There is a low-energy limit, in which all the magnetic moments are aligned with the external field. In this limit $n \rightarrow N$ and $U \rightarrow U_{\min} = -N\mu B$, which makes the factor

$$\left(\frac{N-n}{2}\right)! = 0! \quad (11)$$

in the exact expression for the entropy. In view of the presence of this vanishing argument to one of the factorials, we should be cautious about using the approximate forms we've derived via Stirling's approximation in this limit. In fact, our approximate forms do give the correct limiting values of the entropy in the low-energy and high-energy limits, but we'll have more confidence in the result if we use the exact expression to find those limiting values:

$$\Omega(N, N) = \frac{N!}{\left(\frac{N+N}{2}\right)! \left(\frac{N-N}{2}\right)!} = 1. \quad (12)$$

This means the entropy $S = k \ln \Omega$ vanishes at the low-energy limit $U = -N\mu B$. It's easy to see that the same is true at the upper limit $U = U_{\max} = N\mu B$. Notice that the energy limits for the system are symmetric, $U_{\min} = -U_{\max}$, since we chose the zero of energy for a single moment to be centered between the energies of the two states.

As we've seen in the homework, the binomial coefficient reaches a maximum at its midpoint. This means the entropy does as well, and the exact value at the midpoint is just

$$S_{\max} = k \ln \frac{N!}{[(N/2)!]^2}. \quad (13)$$

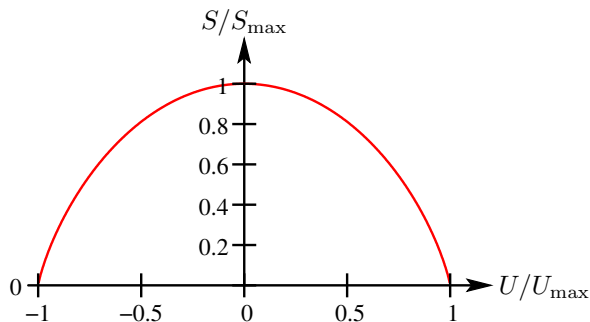
The value at the maximum, based on our approximate expression for $S(U, N)$, valid for large N , is easily seen to be

$$S_{\max} \approx kN \ln 2, \quad (14)$$

when $U = 0$. It can also be obtained from the exact value of S_{\max} by applying Stirling's approximation. We'll see explicitly in a little while that the partial derivative of S with respect to U really does vanish at $U = 0$.

With these characterizations of $S(U, N)$ in mind, we should find it very

plausible that its graph, for fixed N , looks like:

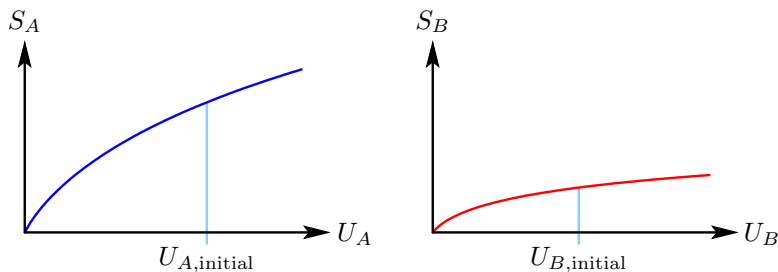


You may recall that I described that curve once before, when discussing the concept of temperature as the inverse of the derivative of S with respect to U :

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{V,N}, \quad (15)$$

though there's no volume dependence in the ideal two-state paramagnet problem. We'll find that this system exhibits some very unusual thermal behavior, but before we try to understand that, it would be useful to undertake an exercise that will help us to review the connection between the derivative $\partial S/\partial U$ and thermal behavior in normal systems.

Exercise. [Not collected] This is Schroeder's Problem 3.3, on p. 90, reworded somewhat. Suppose we have two "normal" systems, where by "normal" I mean they have entropy functions $S(U)$ that grow without bound as the energy increases, the energy itself being capable of growing without bound. The Sackur-Tetrode entropy expression for the ideal gas is an example. The entropy curves for the two systems are shown here, drawn to the same scale:



The initial energies of the systems, prior to establishment of thermal contact between them, are indicated. They are then brought into thermal contact, but remain isolated from anything else. Explain what happens subsequently and why, *without* using the word "temperature."

Answer: The slopes of the $S(U)$ curves give a measure of the rate at which the entropy changes for a small change in energy:

$$\Delta S = \frac{\partial S}{\partial U} \Delta U. \quad (16)$$

If we imagine a small amount of energy being transferred from system A to system B , the fact that the slope of $S_B(U_B)$ at $U_{B,\text{initial}}$ is smaller than that of $S_A(U_A)$ at $U_{A,\text{initial}}$ means that the increase of the energy of B results in a smaller increase in its entropy than the decrease in entropy of A resulting from its loss of energy. That would make the total entropy of the combined system, $\Delta S = \Delta S_A + \Delta S_B$, decrease, contrary to the second law of thermodynamics—it would cause a decrease in the multiplicity of the combined system.

If the energy were transferred from B to A instead, then the overall entropy would increase, since the entropy of A would go up faster with energy than that of B would go down. Thus, increasing overall entropy requires energy to be transferred from the system with the smaller slope to the system with the greater slope. At the point where the energies are such that the slopes are equal, there is no further entropic benefit to transfer of energy in either direction, so the combined system settles into equilibrium when that condition is reached.

Recalling the definition of temperature:

$$\frac{1}{T} = \frac{\partial S}{\partial U}, \quad (17)$$

you'll see that the system with the smaller slope of $S(U)$ has the higher temperature. So energy is transferred until the slopes, and therefore the temperatures, are equal.

Now let's think a bit about the meaning of the curve $S(U)$ for the two-state paramagnet. At negative energies, the majority of the moments are aligned with the external field, and the slope of $S(U)$ is positive. The slope decreases with increasing energy, indicating an increase in temperature, or a decrease in coldness ($1/T$). Eventually, $U = 0$ is reached, and the slope of $S(U)$ vanishes, corresponding to infinite temperature or zero coldness. If any additional energy is added to the system, the slope and the temperature become negative, as does the coldness. But notice that the temperature then increases from negative infinity as the energy increases above $U = 0$, whereas the coldness merely passes continuously through zero to become negative. Thus coldness, or $1/T$ is a more mathematically palatable measure of the tendency of the system to accept or give up energy than is the temperature, which is discontinuous.

If we put one of these systems with $U < 0$ in thermal contact with a normal system, the two will behave normally, in that the one with the greater slope of $S(U)$, that is, greater coldness, will absorb energy from the one with the smaller slope. But it's clear that there is no way to heat

the paramagnet to energy $U > 0$ by placing it into thermal contact with any normal system. There are no normal systems with temperature greater than infinity or coldness less than zero! The only ways to add more energy to a paramagnet with $U = 0$ are nonthermal, and a very clever approach is to do it through experimental trickery. The system is equilibrated with its environment at finite temperature, so that an excess of the moments are aligned with the external field. Then the field is quickly reversed, and the system is tricked into being in a state of positive energy and negative temperature.

Then what happens? Well, the system is in thermal contact with a normal system (its environment), and the one with the larger derivative $\partial S/\partial U$ is colder, so it absorbs energy from the other one. The paramagnet has negative coldness (negative slope) in this high-energy state, so all normal systems are colder and will absorb energy from it. Thus, energy flows out of the paramagnetic system to its surroundings. This happens because the entropy of the combined system is maximized in thermal equilibrium. Reduction of the energy of the paramagnetic system when $U > 0$ increases its entropy (look at the curve), and an increase in the energy of a normal system increases its entropy. Thus, the total entropy always increases whenever energy is transferred from a negative-temperature system to a positive-temperature system. This means negative temperatures are hotter than positive ones.

HW Problem. Schroeder problem 3.21, p. 107.

HW Problem. Schroeder problem 3.23, p. 107.

HW Problem. Schroeder problem 3.25, p. 108.

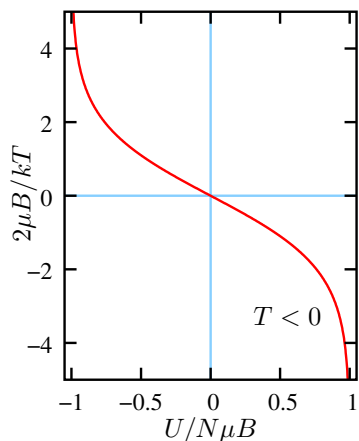
[EOC, Wed. 2/15/2006, #16]

Now let's calculate the inverse temperature explicitly from our expression for the entropy:

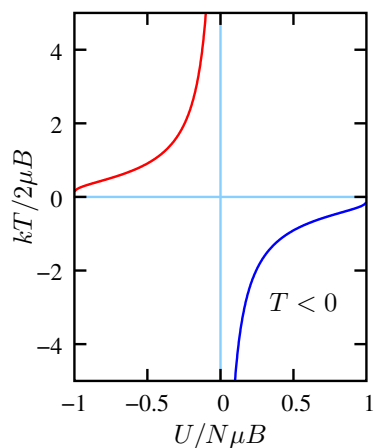
$$\begin{aligned}
 \frac{1}{T} &= \left(\frac{\partial S}{\partial U} \right)_N \\
 &= -\frac{kN}{2} \left[-\frac{1}{N\mu B} \ln \left(1 - \frac{U}{N\mu B} \right) - \frac{1}{N\mu B} \right. \\
 &\quad \left. + \frac{1}{N\mu B} \ln \left(1 + \frac{U}{N\mu B} \right) + \frac{1}{N\mu B} \right] \\
 &= \frac{k}{2\mu B} \ln \frac{1 - \frac{U}{N\mu B}}{1 + \frac{U}{N\mu B}}.
 \end{aligned} \tag{18}$$

It's easy to see that the coldness vanishes at $U = 0$, goes to positive infinity at $U \rightarrow -N\mu B$, and goes to negative infinity as $U \rightarrow N\mu B$. Here's the

plot:



Notice that it's nice and continuous away from the infinities at the energy extrema. By contrast, its inverse, the temperature has a nasty discontinuity right at zero energy:

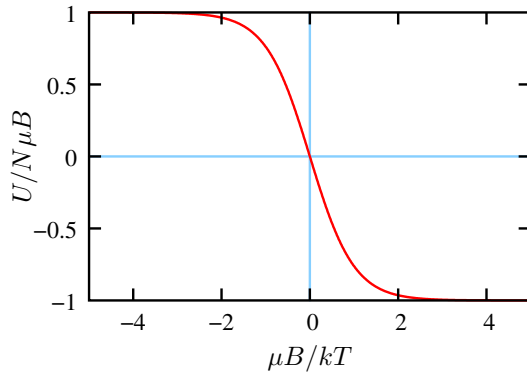


This is why I claimed above that coldness, or $1/T$, is a more useful quantity than temperature, at least in these systems that can have entropy that decreases with increasing energy.

Now, it's often useful to know the dependence of U on T , especially if we want to find the heat capacity. To solve for U , just exponentiate $2\mu B/kT$, solve for $U/N\mu B$, and rearrange slightly, obtaining

$$\begin{aligned}
 U &= N\mu B \frac{1 - e^{2\mu B/kT}}{1 + e^{2\mu B/kT}} \\
 &= -N\mu B \tanh \frac{\mu B}{kT}.
 \end{aligned}
 \tag{19}$$

That last step can be obtained by factoring out $e^{\mu B/kT}$ from both the numerator and denominator, then making use of the definitions of $\sinh x$ and $\cosh x$ in terms of exponentials with real arguments and the definition of $\tanh x$ as $\sinh x / \cosh x$. Of course, a plot of $U(1/T)$ just looks like the plot of $1/T(U)$ turned on its side:



Since the energy and the magnetization are proportional, this also gives us an expression for the magnetization as a function of T :

$$M = -\frac{U}{B} = N\mu \tanh \frac{\mu B}{kT}. \quad (20)$$

The heat capacity can now be calculated from $U(T)$:

$$\begin{aligned} C_B &= \left(\frac{\partial U}{\partial T} \right)_{B,N} \\ &= Nk \left(\frac{\mu B}{kT} \right)^2 \left(1 - \tanh^2 \frac{\mu B}{kT} \right) \\ &= Nk \frac{(\mu B/kT)^2}{\cosh^2(\mu B/kT)}. \end{aligned} \quad (21)$$

If we think about what was special about this system, that is, what endows it with the strange capability of having negative temperature, we'll see that the key is the fact that the macrostate multiplicity, and therefore the entropy, has a maximum between the extremes of energy. That arises because the energy is bounded from above, as well as from below, and there are fewer ways to choose aligned moments when their number is close to the total number of moments than when their number is half that total. That is, there are fewer ways to distribute the energy among the constituents, when most have high energy or most have low energy, than when their numbers are more evenly balanced. Thus, it's the existence of a maximum energy per constituent, hence overall, that gives this system its peculiar properties. No system having constituents with kinetic energy, which is unbounded, can have negative temperature.

It's worth noting that this system is mathematically identical to a number of other systems, among them the random walk in one dimension and the binary alloy, though in the latter, the interactions between the constituents generally play an important role, just as is the case in ferromagnetic systems.

Reading assignment. Schroeder, sections **3.5** and **3.6**.

0.2 Mechanical contact

As we've seen, the partial derivative of the entropy $(\partial S/\partial U)_{V,N}$, which we have defined to be the inverse temperature, plays the central role in determining the direction of energy flow when two systems are placed in thermal contact. The system with the larger value of that derivative is colder than the other one, and the entropy increases when energy flows from the hotter system to the colder one. Energy flow in the opposite direction is not absolutely forbidden, but it simply doesn't happen because that would require the combined system to change to a macrostate with fewer microstates. Thus, probability, and therefore the second law of thermodynamics, requires energy flow from the hotter system to the colder one.

But the entropy is generally also a function of other macroscopic variables besides the energy; for simple systems,

$$S = S(U, V, N), \quad (22)$$

the Sackur-Tetrode expression for the entropy of the ideal gas being a prime example. It is natural to suppose, indeed it should be obvious, that the other partial derivatives of S must also play roles similar to that of $(\partial S/\partial U)_{V,N}$ in the establishment of equilibrium between systems that can exchange those other variables, V , and N upon which S depends. In this section, we'll explore the role of $(\partial S/\partial V)_{U,N}$ for systems in mechanical contact, and in the next section we'll look at the role of $(\partial S/\partial N)_{U,V}$ for systems in diffusive contact.

Consider a pair of systems separated by a movable, diathermal partition, so they can exchange both volume and energy. As usual, they are to be isolated from everything else, so the total energy is fixed:

$$U_1 + U_2 = U = \text{constant}, \quad (23)$$

as is the total volume:

$$V_1 + V_2 = V = \text{constant}. \quad (24)$$

The particle numbers N_1 and N_2 are individually fixed, since particles cannot penetrate the partition.

The condition for equilibrium, the maximization of the entropy of the combined system with respect to exchanges of both volume and energy, requires the partial derivatives of S with respect to parameters characterizing both energy exchange and volume exchange to vanish. We'll take the energy U_1 of the first system to characterize the energy distribution between the systems, so in equilibrium we must have

$$\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1, N_1} = \left(\frac{\partial S_2}{\partial U_2}\right)_{V_1, N_1} \quad \text{or} \quad \frac{1}{T_1} = \frac{1}{T_2}. \quad (25)$$

To characterize the distribution of volume between the two subsystems, we'll make the analogous choice of V_1 as the parameter. In this case, the total entropy can be expressed in terms of that parameter as:

$$\begin{aligned} S(U_1, V_1) &= S_1(U_1, V_1) + S_2(U_2, V_2) \\ &= S_1(U_1, V_1) + S_2(U - U_1, V - V_1), \end{aligned} \quad (26)$$

where the particle numbers have been suppressed. In equilibrium, the partial derivative of S with respect to V_1 must vanish:

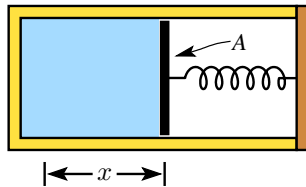
$$\begin{aligned} \left(\frac{\partial S}{\partial V_1}\right)_{U_1} &= \left(\frac{\partial S_1}{\partial V_1}\right)_{U_1} + \left(\frac{\partial S_2}{\partial V_1}\right)_{U_1} \\ &= \left(\frac{\partial S_1}{\partial V_1}\right)_{U_1} + \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2} \underbrace{\frac{dV_2}{dV_1}}_{-1} \\ &= 0. \end{aligned} \quad (27)$$

This gives us a way to express the condition for equilibrium in terms of the partial derivatives of the subsystem entropies with respect to their individual entropies:

$$\left(\frac{\partial S_1}{\partial V_1}\right)_{U_1} = \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2}. \quad (28)$$

This should all look very familiar, as it is exactly analogous to the calculation we did originally for energy exchange by systems in thermal contact.

We can discover the significance of those derivatives by considering a system consisting of a fluid (real gas or liquid) behind a piston with force applied by a compressed spring:



The entire system is isolated, so its total energy U_{tot} is fixed. The energy of the fluid system is

$$U = U(x) = U_{\text{tot}} - \frac{1}{2}Kx^2, \quad (29)$$

where K is the spring constant, and x is the distance by which the spring is compressed from its equilibrium length. The volume occupied by the fluid is

$$V = V(x) = V_0 + Ax, \quad (30)$$

where A is the cross-sectional area of the piston, and $V_0 = V - Ax$ is the volume that would be occupied by the fluid if the spring were in its equilibrium position.

The entropy of the system is then a function of the compression of the spring:

$$S = S(U, V, N) = S[U(x), V(x), N], \quad (31)$$

and the condition for equilibrium of the system is

$$\begin{aligned} \frac{dS}{dx} &= \left(\frac{\partial S}{\partial U}\right)_{V,N} \frac{dU}{dx} + \left(\frac{\partial S}{\partial V}\right)_{U,N} \frac{dV}{dx} \\ &= \frac{1}{T}(-Kx) + \left(\frac{\partial S}{\partial V}\right)_{U,N} A \\ &= 0. \end{aligned} \quad (32)$$

Thus, in equilibrium the partial derivative has the value

$$\left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{Kx}{AT} = \frac{F}{AT} = \frac{P}{T}, \quad (33)$$

where $F = Kx$ is the magnitude of the force exerted on the piston by the spring. So the slope of $S(V)$ when U and N are held fixed is proportional to the pressure.

Let's try plugging that into our equilibrium condition:

$$\left(\frac{\partial S_1}{\partial V_1}\right)_{U_1, N_1} = \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2, N_2} \Rightarrow \frac{P_1}{T_1} = \frac{P_2}{T_2}. \quad (34)$$

Since we know the temperatures must be equal in equilibrium, this implies the pressures must also be equal:

$$P_1 = P_2. \quad (35)$$

That shouldn't be too surprising—if they weren't equal, there would be a net force on the partition, and it would shift until the force vanished.

With that result from the ideal gas in mind, we'll now simply take

$$\left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{P}{T} \quad (36)$$

to define P , once T is defined by $\partial S/\partial U = 1/T$.

Now we have macroscopic parameters T and P that characterize the state of a system and can be used to see how entropy changes when the corresponding variables U and V are changed:

$$dS = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV \quad (\text{N fixed}) \quad (37)$$

which implies

$$\boxed{dS = \frac{1}{T} dU + \frac{P}{T} dV} \quad (38)$$

This should be remembered, along with $S = S(U, V, N)$.

Alternatively, we can rearrange this to get

$$\boxed{dU = T dS - P dV} \quad (39)$$

which is the more common form. This also follows from $U = U(S, V, N)$:

$$dU = \underbrace{\left(\frac{\partial U}{\partial S}\right)_{V,N}}_T dS + \underbrace{\left(\frac{\partial U}{\partial V}\right)_{S,N}}_{-P} dV \quad (\text{N fixed}) \quad (40)$$

To see the identification of $(\partial U/\partial V)_{S,N}$ with pressure, set the total differential dS of the entropy to zero, thus establishing the constant-entropy condition. This gives

$$\frac{1}{T} dU + \frac{P}{T} dV = 0, \quad (41)$$

which can be rearranged to find

$$\left(\frac{\partial U}{\partial V}\right)_{S,N} = \frac{-P/T}{1/T} = -P. \quad (42)$$

HW Problem. Schroeder problem 3.31, p. 114.

HW Problem. Schroeder problem 3.32, p. 114.

HW Problem. Schroeder problem 3.34, pp. 114–115.