where we've used the definition of enthalpy in terms of energy:

$$
\begin{equation*}
H=U+P V \tag{22}
\end{equation*}
$$

Let's see what happens when we try this out on an ideal gas. Its energy is given by the equipartition theorem:

$$
\begin{equation*}
U=\frac{f}{2} N k T \tag{23}
\end{equation*}
$$

where $f$ is the number of degrees of freedom, and its $P V$ product is given by the ideal-gas law as

$$
\begin{equation*}
P V=N k T \tag{24}
\end{equation*}
$$

Therefore, its enthalpy is

$$
\begin{equation*}
H_{\text {ideal }}=\frac{f}{2} N k T+N k T=\left(\frac{f}{2}+1\right) N k T \tag{25}
\end{equation*}
$$

Since $H_{\text {ideal }}$ is proportional to $T$, the constancy of the enthalpy in a throttling process implies constancy of the temperature as well. That is, the effect vanishes for an ideal gas. Thus, if a throttling process is to change the temperature, that effect must originate in the nonideality of real gases. [EOC, Fri. 2/24/2006, \#20]

The key to understanding the process lies in the potential energy of interaction between the molecules. Typically, this looks something like:


So if you reduce the pressure, the molecules become farther apart, raising the potential energy of interaction and correspondingly lowering the kinetic energy, and with it, the temperature. But it's also easy to see that at high enough pressure or high enough temperature, the potential energy of interaction can decrease with throttling, causing the temperature to increase, instead of decrease. So, obtaining a cooling effect may require precooling or working at lower pressure.

Here's a rough graph of constant-enthalpy curves for nitrogen gas in the
pressure-temperature plane:


The inversion curve passes through the maxima of the nonmonotonic isenthalps, marking the boundary between points where a throttling process cools the gas, to the left of the inversion curve, and points where a throttling process wamrs the gas.

The parameter that describes whether the process warms or cools the gas at any pressure and temperature is the slope of the isenthalp passing through that point:

$$
\begin{align*}
\text { slope of isenthalp } & =\left(\frac{\partial T}{\partial P}\right)_{H}  \tag{26}\\
& =\text { "Joule-Thomson coefficient." }
\end{align*}
$$

HW Problem. Schroeder problem 4.29, p. 140.

## Chapter 1

## Thermodynamic Potentials

Reading assignment. Schroeder, section 5.2.

### 1.1 Thermodynamic potentials

So far we've mostly focused on the entropy as the central and most fundamental quantity in thermal physics. We've seen its crucial importance in understanding and determining equilibrium states when systems are permitted to exchange energy, volume, or particles. But it's not always the most convenient quantity to use in thermodynamic calculations, in spite of its fundamental importance. Working with the energy $U$ and potentials derived from it is more traditional in thermodynamics, but $U$ itself is also not always convenient. We've seen a few instances in which the enthalpy

$$
\begin{equation*}
H=U+P V \tag{1.1}
\end{equation*}
$$

is a convenient thermodynamic quantity to use for calculations. This is most naturally useful when the process of interest takes place under constant-pressure conditions. Then the volume is not fixed, and some mechanical work contributes to $U$, which is inconvenient to keep track of explicitly. Working with $H$ relieves us of that need.

The best way to understand all of this in a general way is first to recall the thermodynamic identity:

$$
\begin{equation*}
d U=T d S-P d V+\mu d N \tag{1.2}
\end{equation*}
$$

relating changes in $U$ to changes in $S, V$, and $N$. These are the "natural" variables upon which $U$ depends, and knowledge of the function

$$
\begin{equation*}
U=U(S, V, N) \tag{1.3}
\end{equation*}
$$

tells us everything we can know about the thermodynamics of a system. This is true in the same sense in which knowledge of

$$
\begin{equation*}
S=S(U, V, N) \tag{1.4}
\end{equation*}
$$

gives us complete knowledge of the thermodynamic behavior of a system. $U=U(S, V, N)$ contains the same information about the relation among the four variables $U, S, V$, and $N$ as $S=S(U, V, N)$.

However, if pressure, rather than volume, is controlled by experimental conditions, the enthalpy is more convenient; it satisfies the thermodynamic identity

$$
\begin{align*}
d H & =d U+P d V+V d P \\
& =T d S+V d P+\mu d N . \tag{1.5}
\end{align*}
$$

This shows that the natural variables upon which $H$ depends are $S, P$, and $N$ :

$$
\begin{equation*}
H=H(S, P, N) \tag{1.6}
\end{equation*}
$$

and that if the pressure is fixed $(d P=0)$, only changes in entropy and particle number can make $d H \neq 0$. In effect, we have replaced the independent variable $V$ by the independent variable $P$, which is a derivative quantity:

$$
\begin{equation*}
P=-\left(\frac{\partial U}{\partial V}\right)_{S, N}=T\left(\frac{\partial S}{\partial V}\right)_{U, N} . \tag{1.7}
\end{equation*}
$$

### 1.1.1 Legendre transformations

It's worthwhile to take a look at the mathematical formulation of this transformation, which is known as a Legendre transformation, in the simple case of a function of a single variable. Suppose we have some known function

$$
\begin{equation*}
y=y(x) . \tag{1.8}
\end{equation*}
$$

Knowledge of that function is equivalent to complete tabulation of pairs $x$, $y$ :

$$
\begin{array}{cc}
x & y \\
\hline x_{1} & y_{1}=y\left(x_{1}\right) \\
x_{1} & y_{2}=y\left(x_{2}\right) \\
\vdots & \vdots
\end{array}
$$

We'd like a way of expressing that same information in terms of the derivative quantity

$$
\begin{equation*}
m(x)=\frac{d y}{d x}, \tag{1.9}
\end{equation*}
$$

the slope of the curve. Can we do that by simply rewriting $y$ as a function of $m$ :

$$
\begin{equation*}
y=y(m) ? \tag{1.10}
\end{equation*}
$$

Consider the graph:


The "impostors" have the same slope at each point of the curve, hence the same $y(m)$, but they have different dependences on $x$, and they do not contain the same information as $y(x)$. So, if we merely rewrite $y$ in terms of $m$, we have lost information-we no longer know which is the correct curve $y(x)$. But, we can characterize the same curve by coordinate pairs if those pairs are the slope and the $y$ intercept of the tangent to each point on the curve:


The tangent to the impostor has the same slope as the correct curve, but the $y$ intercept of the tangent line is different. Thus, we can express the correct curve as the alternative set of coordinate pairs

$$
\begin{equation*}
b(m)=\left\{\left(m_{1}, b_{1}\right),\left(m_{2}, b_{2}\right), \ldots\right\} \tag{1.11}
\end{equation*}
$$

Let's see how this function is related to $y(x)$. First recall the connection between $y, x, m$, and $b$ :

$$
\begin{equation*}
m=\frac{y-b}{x-0} \quad \text { or } \quad y=m x+b \tag{1.12}
\end{equation*}
$$

But $m$ can be expressed as a function of $x$ :

$$
\begin{equation*}
m=\frac{d y(x)}{d x}=m(x) \tag{1.13}
\end{equation*}
$$

so if we can invert this to get $x(m)$, then we can write $b$ as a function of $m$ as the independent variable:

$$
\begin{equation*}
b(m)=y[x(m)]-m x(m) . \tag{1.14}
\end{equation*}
$$

This is a Legendre transformation of the original functional relationship $y=y(x)$.

Example. Let

$$
\begin{equation*}
y(x)=\left(x-x_{0}\right)^{2}+y_{0} . \tag{1.15}
\end{equation*}
$$

The slope is

$$
\begin{equation*}
m=\frac{d y}{d x}=2\left(x-x_{0}\right)=m(x) \tag{1.16}
\end{equation*}
$$

This can be inverted to get $x(m)$ :

$$
\begin{equation*}
x=\frac{m}{2}+x_{0}=x(m) \tag{1.17}
\end{equation*}
$$

In terms of $m, y$ is

$$
\begin{align*}
y & =\left(x-x_{0}\right)^{2}+y_{0} \\
& =\left(\frac{m}{2}\right)^{2}+y_{0}  \tag{1.18}\\
& =y(m),
\end{align*}
$$

and

$$
\begin{align*}
b & =y-m x \\
& =\left(\frac{m}{2}\right)^{2}+y_{0}-m\left(\frac{m}{2}+x_{0}\right)  \tag{1.19}\\
& =-\frac{m^{2}}{4}-m x_{0}+y_{0} \\
& =b(m)
\end{align*}
$$

By construction, the points on this curve $b(m)$ are exactly those on the curve $y(x)$. Let's demonstrate that for two of the points.

- For $m=0$ the $y$ intercept is $b(0)=y_{0}$. The corresponding value of $x$ from (1.17) is

$$
\begin{equation*}
x(0)=x_{0} . \tag{1.20}
\end{equation*}
$$

We can find the corresponding value of $y$ on the original curve $y(x)$ by plugging this value of $x$ into the original expression for the curve (1.15):

$$
\begin{equation*}
y\left(x_{0}\right)=\left(x_{0}-x_{0}\right)^{2}+y_{0}=y_{0} \tag{1.21}
\end{equation*}
$$

And, we can find the corresponding value of $y$ on the straight line $y=m x+b$ by plugging it into that expression:

$$
\begin{equation*}
y\left(x_{0}\right)=m x_{0}+b=y_{0} \tag{1.22}
\end{equation*}
$$

So, both the original curve (1.15) and the straight line $y(x)=0 x+y_{0}$ obtained from $m$ and $b(m)$ for $m=0$ give the same mapping between $x$ and $y$ at that point. That is, the straight line intersects the original curve at that point-actually it is tangent to the curve, since $m=$ $d y / d x$ at that point, but it's the correspondence between $x$ and $y$ that we're interested in here.
If, instead, we simply calculate $y$ from the expression for $y(m)$ in (1.18), we obtain

$$
\begin{equation*}
y(m)=y(0)=\left(\frac{0}{2}\right)^{2}+y_{0}=y_{0} \tag{1.23}
\end{equation*}
$$

That is the correct value of $y$ for that slope, but that expression alone gives us no information on the corresponding value of $x$-any point on the line $y=y_{0}$ could correspond to that value of $y$.

- For $m=1$ the $y$ intercept is $b(1)=-\frac{1}{4}-x_{0}+y_{0}$. The corresponding value of $x$ is:

$$
\begin{equation*}
x(1)=\frac{1}{2}+x_{0} \tag{1.24}
\end{equation*}
$$

The value of $y$ that the original function maps this value of $x$ to is

$$
\begin{equation*}
y\left(\frac{1}{2}+x_{0}\right)=\left(\frac{1}{2}+x_{0}-x_{0}\right)^{2}+y_{0}=\frac{1}{4}+y_{0} \tag{1.25}
\end{equation*}
$$

and the value of $y$ that the tangent line maps this $x$ to is

$$
\begin{equation*}
y=m x+b=1\left(\frac{1}{2}+x_{0}\right)-\frac{1}{4}-x_{0}+y_{0}=\frac{1}{4}+y_{0} \tag{1.26}
\end{equation*}
$$

Clearly, the curve and the line coincide at that point, too, but this line is not the same straight line we had for $m=0$. As $m$ changes, the tangent line changes as well; but for every pair $(m, b)$, the corresponding straight line $y=m x+b$ coincides with the original curve $y=\left(x-x_{0}\right)^{2}+y_{0}$ at one point $(x, y)$.

### 1.1.2 Thermodynamic potentials

In thermodynamics, the situation is much the same, except that there are generally several variables tied up in some functional relationship, and there are strong physical motivations for making Legendre transformations. Any of the independent variables upon which the entropy $S$ or the energy $U$ depend can be "traded" for the variable that is the corresponding derivative quantity. Our catalog of derivative quantities for simple systems includes

$$
\begin{align*}
\frac{1}{T} & =\left(\frac{\partial S}{\partial U}\right)_{V, N} & \frac{P}{T} & =\left(\frac{\partial S}{\partial V}\right)_{U, N}
\end{align*}-\frac{\mu}{T}=\left(\frac{\partial S}{\partial N}\right)_{U, V} .
$$

We have already encountered one of the Legendre transformationss of the energy, in which the volume was exchanged for the negative of the pressure as an independent variable, the enthalpy. But it's easy to see that there are several other possibilities, including transformations on more than one variable. The most commonly encountered Legendre transformations of $U(S, V, N)$ are:

$$
\begin{align*}
H(S, P, N) & =U-(-P) V & & \text { (enthalpy) } \\
F(T, V, N) & =U-T S & & \text { (Helmholtz free energy) }  \tag{1.28}\\
G(T, P, N) & =U-T S-(-P) V & & \text { (Gibbs free energy) }
\end{align*}
$$

These, together with the energy itself, are called thermodynamic potentials. Each of them is most naturally useful when the independent variables upon which it depends are experimentally controlled parameters. This is perhaps easiest to see when one considers the thermodynamic identities they satisfy:

$$
\begin{align*}
d U & =T d S-P d V+\mu d N \\
& =\left(\frac{\partial U}{\partial S}\right)_{V, N} d S+\left(\frac{\partial U}{\partial V}\right)_{S, N} d V+\left(\frac{\partial U}{\partial N}\right)_{S, V} d N \\
d H & =d U+P d V+V d P \\
& =T d S+V d P+\mu d N \\
& =\left(\frac{\partial H}{\partial S}\right)_{P, N} d S+\left(\frac{\partial H}{\partial P}\right)_{S, N} d P+\left(\frac{\partial H}{\partial N}\right)_{S, P} d N \\
d F & =d U-T d S-S d T  \tag{1.29}\\
& =-S d T-P d V+\mu d N \\
& =\left(\frac{\partial F}{\partial T}\right)_{V, N} d T+\left(\frac{\partial F}{\partial V}\right)_{T, N} d V+\left(\frac{\partial F}{\partial N}\right)_{T, V} d N \\
d G & =d U-T d S-S d T+P d V+V d P \\
& =-S d T+V d P+\mu d N \\
& =\left(\frac{\partial G}{\partial T}\right)_{P, N} d T+\left(\frac{\partial G}{\partial P}\right)_{T, N} d P+\left(\frac{\partial G}{\partial N}\right)_{T, P} d N
\end{align*}
$$

So, for example, the Helmholtz free energy $F$ is convenient when the temperature is constant, so that $d T=0$ and changes in $F$ depend only on changes in volume and particle number. The Gibbs free energy $G$ is convenient when both pressure and temperature are fixed, as is commonly the case in chemical reactions performed in a vessel that is open to the air.

One other important result follows from the information in (1.29): each of the derivative quantities, $T, P$, and $\mu$, as well as the variables $S, V$, and $N$, can be expressed in terms of derivatives of some of the potentials. The
correspondence between the terms in the thermodynamic identity satisfied by any one of the potentials and the total differential of that potential when viewed as a function of its natural variables gives the definitions of the three derivative quantities of that potential. The meanings of the derivatives of the energy are familiar, but we can also easily extract from (1.29) things like

$$
\begin{equation*}
T=\left(\frac{\partial H}{\partial S}\right)_{P, N} \quad \text { and } \quad V=\left(\frac{\partial G}{\partial P}\right)_{T, N} \tag{1.30}
\end{equation*}
$$

These can be very handy bits of information to have available when one is manipulating thermodynamic expressions.

One can, of course, construct several additional thermodynamic potentials through other combinations of Legendre transformations of the energy. In addition, each of the potentials can be expressed in terms of Legendre transformations of other potentials. For example,

$$
\begin{equation*}
G=F+P V=H-T S \quad \text { and } \quad F=G-P V=H-T S-P V \tag{1.31}
\end{equation*}
$$

You can see that there are many possibilities.
There is also an analogous family of Legendre transformations of the entropy, those functions often being called Massieu functions.

One important thing to keep in mind is that you need not memorize this entire blizzard of formulas. Nor need you spend time looking through several books for just the right one when you need one. If you only remember the thermodynamic identity for the energy, $d U=T d S-P d V+\mu d N$, and either the natural variables upon which each potential depends or the definition of each potential in terms of a Legendre transformation of the energy, together with the list of variables upon which $U$ depends, it is extremely easy to construct all the other formulas in a matter of seconds. Even the definitions of the derivative quantities $T, P$, and $\mu$ in terms of $U$ need not be memorized.

HW Problem. Schroeder problem 5.12, pp. 158-159.
HW Problem. Schroeder problem 5.14, p. 159.

