

PAPERS | APRIL 01 2001

The elusive chemical potential

Ralph Baierlein



Am. J. Phys. 69, 423–434 (2001)

<https://doi.org/10.1119/1.1336839>



View
Online



Export
Citation

CrossMark

The elusive chemical potential

Ralph Baierlein

Department of Physics and Astronomy, Northern Arizona University, Flagstaff, Arizona 86011-6010

(Received 12 April 2000; accepted 10 October 2000)

This paper offers some qualitative understanding of the chemical potential, a topic that students invariably find difficult. Three “meanings” for the chemical potential are stated and then supported by analytical development. Two substantial applications—depression of the melting point and batteries—illustrate the chemical potential in action. The origin of the term “chemical potential” has its surprises, and a sketch of the history concludes the paper. © 2001 American Association of Physics Teachers.

Teachers.

[DOI: 10.1119/1.1336839]

I. INTRODUCTION

It was the semester’s end in spring 1997, and I had just finished teaching a course in thermal physics. One of my students said to me, “Now I really understand temperature, but what is the meaning of the chemical potential?” Clearly, a topic needed greater emphasis, both in class and in the draft text. I vowed to do better the next year, and—together—I spent several years looking into responses to the question.

The present article first describes three meanings of the chemical potential, next develops them analytically, and finally gives two substantial examples of how the chemical potential is used. Some observations are interleaved, and the paper concludes with a short history.

For whom is this paper intended? I wrote primarily for someone—instructor or student—who already knows about the chemical potential but would like to understand it better. Some portions of the paper are original, but much of it consists of material that is common knowledge among textbook writers. I have gathered together interpretations, insights, and examples to construct a kind of tutorial or review article on the chemical potential.

II. MEANINGS

Any response to the question, “What is the meaning of the chemical potential?” is necessarily subjective. What satisfies one person may be wholly inadequate for another. Here I offer three characterizations of the chemical potential (denoted by μ) that capture diverse aspects of its manifold nature.

- (1) *Tendency to diffuse.* As a function of position, the chemical potential measures the tendency of particles to diffuse.
- (2) *Rate of change.* When a reaction may occur, an extremum of some thermodynamic function determines equilibrium. The chemical potential measures the contribution (per particle and for an individual species) to the function’s rate of change.
- (3) *Characteristic energy.* The chemical potential provides a characteristic energy: $(\partial E/\partial N)_{S,V}$, that is, the change in energy when one particle is added to the system at constant entropy (and constant volume).

These three assertions need to be qualified by contextual conditions, as follows.

- (a) Statement (1) captures an essence, especially when the temperature T is uniform. When the temperature varies spatially, diffusion is somewhat more complex and is discussed briefly under the rubric “Further comments” in Sec. IV.
- (b) Statement (2) is valid if the temperature is uniform and fixed. If, instead, the total energy is fixed and the temperature may vary from place to place, then μ/T measures the contribution. When one looks for conditions that describe chemical *equilibrium*, one may focus on each locality separately, and then the division by temperature is inconsequential.
- (c) The system’s “external parameters” are the macroscopic environmental parameters (such as external magnetic field or container volume) that appear in the energy operator or the energy eigenvalues. All external parameters are to be held constant when the derivative in statement (3) is formed. The subscript V for volume illustrates merely the most common situation. Note that pressure does not appear in the eigenvalues, and so—in the present usage—pressure is not an external parameter.

These contextual conditions will be justified later. The next section studies diffusive equilibrium in a familiar context, “discovers” the chemical potential, and establishes the characterization in statement (1).

III. THE TENDENCY TO DIFFUSE

The density of the Earth’s atmosphere decreases with height. The concentration gradient—a greater concentration lower down—tends to make molecules diffuse upward. Gravity, however, pulls on the molecules, tending to make them diffuse downward. The two effects are in balance, canceling each other, at least on an average over short times or small volumes. Succinctly stated, the atmosphere is in equilibrium with respect to diffusion.

In general, how does thermal physics describe such a diffusive equilibrium? In this section, we consider an ideal isothermal atmosphere and calculate how gas in thermal equilibrium is distributed in height. Certain derivatives emerge

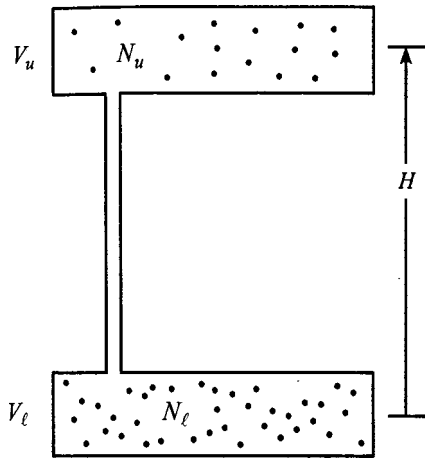


Fig. 1. The context. The narrow tube allows atoms to diffuse from one region to the other, but otherwise we may ignore it.

and play a decisive role. The underlying purpose of the section is to discover those derivatives and the method that employs them.

Figure 1 sets the scene. Two volumes, vertically thin in comparison with their horizontal extent, are separated in height by a distance H . A narrow tube connects the upper volume V_u to the lower volume V_l . A total number N_{total} of helium atoms are in thermal equilibrium at temperature T ; we treat them as forming an ideal gas. What value should we anticipate for the number N_u of atoms in the upper volume, especially in comparison with the number N_l in the lower volume?

We need the probability $P(N_l, N_u)$ that there are N_l atoms in the lower volume and N_u in the upper. The canonical probability distribution gives us that probability as a sum over the corresponding energy eigenstates Ψ_j :

$$P(N_l, N_u) = \sum_{\substack{\text{states } \Psi_j \text{ with } N_l \text{ in } V_l \\ \text{and } N_u \text{ in } V_u}} \frac{\exp(-E_j/kT)}{Z} \\ \equiv \frac{Z(N_l, N_u)}{Z}. \quad (1)$$

The symbol Z denotes the partition function for the entire system, and k is Boltzmann's constant. The second equality merely defines the symbol $Z(N_l, N_u)$ as the sum of the appropriate Boltzmann factors.

In the present context, an energy eigenvalue E_j splits naturally into two independent pieces, one for the particles in the lower volume, the other for the particles in the upper volume. Apply that split to the energy in the Boltzmann factor. Imagine holding the state of the N_u particles in the upper volume fixed and sum the Boltzmann factor $\exp(-E_j/kT)$ over the states of the N_l particles in the lower volume. That step generates the partition function $Z_l(N_l)$ for the lower volume times a Boltzmann factor with just the energy of the particles in the upper volume. Now sum over the states of the N_u particles in the upper volume. The outcome is the tidy form

$$Z(N_l, N_u) = Z_l(N_l) \times Z_u(N_u). \quad (2)$$

(More detail and intermediate steps are provided in Chap. 7 of Ref. 1.)

Common experience suggests that, given our specifically macroscopic system, the probability distribution $P(N_l, N_u)$ will have a single peak and a sharp one at that. An efficient way to find the maximum in $P(N_l, N_u)$ is to focus on the logarithm of the numerator in Eq. (1) and find its maximum. Thus one need only differentiate the logarithm of the right-hand side of Eq. (2). Note that increasing N_l by one unit entails decreasing N_u by one unit; via the chain rule, that introduces a minus sign. Thus the maximum in the probability distribution arises when N_l and N_u have values such that

$$\frac{\partial \ln Z_l}{\partial N_l} = -\frac{\partial \ln Z_u}{\partial N_u}. \quad (3)$$

The equality of two derivatives provides the criterion for the most probable situation. This is the key result. The following paragraphs reformulate the criterion, explore its implications, and generalize it.

First, why reformulate? To connect with functions that are defined in thermodynamics as well as in statistical mechanics. A system's entropy S can be expressed in terms of $\ln Z$, T , and the estimated energy $\langle E \rangle$:

$$S = \frac{\langle E \rangle}{T} + k \ln Z. \quad (4)$$

Thus the Helmholtz free energy F provides a good alternative expression for $\ln Z$:

$$F \equiv \langle E \rangle - TS = -kT \ln Z. \quad (5)$$

Equation (3) indicates that the rate of change of F with particle number is the decisive quantity; so define the *chemical potential* μ by the relation²

$$\mu(T, V, N) \equiv \left(\frac{\partial F}{\partial N} \right)_{T, V} = F(T, V, N) - F(T, V, N-1). \quad (6)$$

In the language of the chemical potential, the criterion for the most probable situation, Eq. (3), becomes

$$\mu_l(T, V_l, N_l) = \mu_u(T, V_u, N_u). \quad (7)$$

The chemical potentials for atoms in the lower and upper volumes are equal.

But what about the less probable situations? And the approach to the most probable situation? Some detail will help here. The partition function $Z_u(N_u)$ for the upper volume has the explicit form

$$Z_u(N_u) = \frac{[(V_u/\lambda_{\text{th}}^3)e^{-mgH/kT}]^{N_u}}{N_u!}, \quad (8)$$

where $\lambda_{\text{th}} \equiv h/\sqrt{2\pi mkT}$ defines the thermal de Broglie wavelength and where m denotes an atom's rest mass. The chemical potential for the atoms in the upper volume is then

$$\mu_u = mgH + kT \ln \left(\frac{N_u}{V_u} \lambda_{\text{th}}^3 \right). \quad (9)$$

For the atoms in the lower volume, μ_l has a similar structure, but the gravitational potential energy is zero.

The explicit forms for μ_l and μ_u enable one to plot the chemical potentials as functions of N_l at fixed total number of atoms. Figure 2 displays the graphs. Suppose we found the gaseous system with the number N_l significantly below its "equilibrium" or most probable value. Almost surely atoms would diffuse through the connecting tube from V_u to

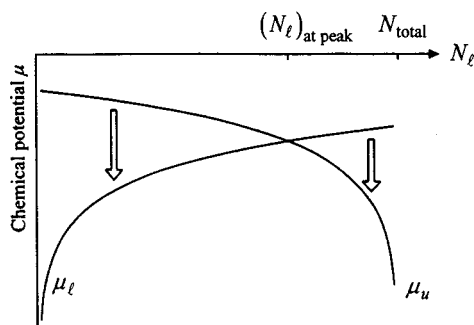


Fig. 2. Graphs of the two chemical potentials as functions of N_l and $N_u = (N_{\text{total}} - N_l)$. The arrows symbolize the direction of particle diffusion relative to the graphed values of the chemical potentials (and *not* relative to the local vertical). When N_l is less than its most probable value, particles diffuse toward the lower volume and its smaller chemical potential; when N_l is greater than its most probable value, diffusion is toward the upper volume and its smaller chemical potential.

V_l and would increase N_l toward $(N_l)_{\text{most probable}}$. Atoms would diffuse from a region where the chemical potential is μ_u to a place where it is μ_l , that is, they would diffuse toward smaller chemical potential. In Newtonian mechanics, a literal force pushes a particle in the direction of smaller potential energy. In thermal physics, diffusion “pushes” particles toward smaller chemical potential.

The details of this example justify the first characterization of the chemical potential: as a function of position, the chemical potential measures the tendency of particles to diffuse. In general, in the most probable situation, which is the only situation that thermodynamics considers, the chemical potential is uniform in space. Before the most probable situation becomes established, particles diffuse toward lower chemical potential.

Alternative derivations of these conclusions are available.^{3,4} Their great generality complements the derivation given here, whose specificity provides an opportunity to explore the “tendency to diffuse” in graphic detail. Moreover, if one uses Eq. (9) and its analog for μ_l in Eq. (7), then one finds that the number density N/V drops exponentially with height. Recovering the “isothermal atmosphere” provides students with a welcome sense of confidence in the formalism.

Just as temperature determines the diffusion of energy (by thermal conduction and by radiation), so the chemical potential determines the diffusion of particles. This parallel is profound, has been noted by many authors,⁵ and goes back at least as far as Maxwell in 1876, as Sec. IX will display.

A. Uniformity

Martin Bailyn remarks sagely, “The [chemical potentials for various particle species] ... will be uniform in equilibrium states. ... In this respect, [they] supplant density as the parameter that is uniform in material equilibrium.”⁶

That uniformity, moreover, holds good when two or more phases coexist, such as liquid and solid water, or liquid, solid, and vapor. The equality of chemical potentials across a phase boundary can serve as the basis for deriving the Clausius–Clapeyron equation (which gives the slope of the phase boundary in a pressure–temperature plane).

If a particle species is restricted to two disjoint regions, however, then the chemical potential may have different val-

ues in each region. For example, suppose the vertical tube in Fig. 1 were provided with a valve, closed initially, and that the lower volume were initially empty. If the valve were then opened, atoms would diffuse downward, but the valve could be closed again before the chemical potentials in the lower and upper regions reached equality. Equilibrium could be achieved in each region separately—but with unequal chemical potentials.

Later, when we discuss batteries, we will find such a situation. When a battery is on open circuit, the conduction electrons in the two (disjoint) metal terminals generally have different chemical potentials.

IV. EXTREMA

Now we focus on the connection between the chemical potential and extrema in certain thermodynamic functions.

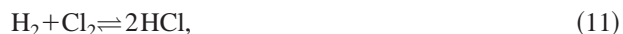
Return to the canonical probability distribution in Sec. III and its full spectrum of values for N_l and N_u . With the aid of Eqs. (5) and (2), we can form a composite Helmholtz free energy by writing

$$F(N_l, N_u) \equiv -kT \ln Z(N_l, N_u) = F_l(N_l) + F_u(N_u). \quad (10)$$

Note especially the minus sign. Where $P(N_l, N_u)$ and hence $Z(N_l, N_u)$ are relatively large and positive, $F(N_l, N_u)$ will be negative. At the maximum for $P(N_l, N_u)$ and hence the maximum for $Z(N_l, N_u)$, the function $F(N_l, N_u)$ will have its minimum. Thus we find that the composite Helmholtz free energy has a *minimum* at what thermodynamics calls the equilibrium values of N_l and N_u . This is a general property of the Helmholtz free energy (at fixed positive temperature T and fixed external parameters).⁷

A. Chemical equilibrium

A chemical reaction, such as



can come to equilibrium under conditions of constant temperature and volume. The equilibrium is characterized by a minimum in the Helmholtz free energy. How is such a condition described with the various chemical potentials?

As a preliminary step, let us generalize the chemical reaction under study. We can write the HCl reaction in the algebraic form

$$-\text{H}_2 - \text{Cl}_2 + 2\text{HCl} = 0, \quad (12)$$

which expresses—among other things—the conservation of each *atomic* species (H and Cl) during the reaction. Adopting this pattern, we write the generic form for a chemical reaction as

$$b_1 B_1 + b_2 B_2 + \cdots + b_n B_n = 0, \quad (13)$$

where each molecular species is represented by a symbol B_i and the corresponding numerical coefficient in the reaction equation is represented by the symbol b_i . For the products of a reaction, the coefficients b_i are positive; for the reactants, they are negative. Altogether, the set $\{b_i\}$ gives the number change in each molecular species when the reaction occurs once. The coefficients $\{b_i\}$ are called *stoichiometric coefficients* (from the Greek roots, *stoikheion*, meaning “element,” and *metron*, meaning “to measure”).

At equilibrium, the Helmholtz free energy will attain a minimum. Imagine that the reaction takes one step away

from equilibrium: the number N_i of molecular species B_i changes by ΔN_i , which equals the stoichiometric coefficient b_i . Then the change in the Helmholtz free energy is

$$\Delta F = \sum_i \left(\frac{\partial F}{\partial N_i} \right)_{T,V, \text{ other } N\text{'s}} \Delta N_i = \sum_i \mu_i b_i = 0. \quad (14)$$

The partial derivatives are precisely the chemical potentials, and the zero follows because the imagined step is away from the minimum. Equilibrium for the chemical reaction implies a constraint among the various chemical potentials:

$$\sum_i \mu_i b_i = 0. \quad (15)$$

The constraint provides the key equation in deriving the chemists' law of mass action.

Equation (14) provides the initial justification for the second characterization of the chemical potential: when a reaction may occur, the chemical potential measures the contribution (per particle) to the rate of change of the function whose extremum determines equilibrium.

Note that the rate of change is *not* a rate of change with time. Rather, it is a rate of change as the reaction proceeds, step by step.

B. Other extrema

In different circumstances, other thermodynamic functions attain extrema. For example, under conditions of fixed energy and fixed external parameters, the entropy attains a maximum. Does the second characterization continue to hold?

To see how the extrema in S and in the Gibbs free energy G are related to the chemical potential, we need equivalent expressions for μ , ones that correspond to the variables that are held constant while the extrema are attained. To derive those expressions, start with the change in F when the system moves from one equilibrium state to a nearby equilibrium state. The very definition of F implies

$$\Delta F = \Delta E - T\Delta S - S\Delta T. \quad (16)$$

(Here E is the thermodynamic equivalent of the estimated energy $\langle E \rangle$ in statistical mechanics.)

A formal first-order expansion implies

$$\Delta F = \left(\frac{\partial F}{\partial T} \right)_{V,N} \Delta T + \left(\frac{\partial F}{\partial V} \right)_{T,N} \Delta V + \left(\frac{\partial F}{\partial N} \right)_{T,V} \Delta N. \quad (17)$$

The last coefficient is, of course, the chemical potential μ . The first two coefficients, which are derivatives at constant N , can be evaluated by invoking energy conservation. When the system moves from an equilibrium state to another one nearby, energy conservation asserts that

$$T\Delta S = \Delta E + P\Delta V, \quad (18)$$

provided N is constant. Use Eq. (18) to eliminate $T\Delta S$ in Eq. (16) and then read off the coefficients that occur in Eq. (17) as $-S$ and $-P$, respectively. Insert these values into Eq. (17); equate the right-hand sides of Eqs. (16) and (17); and then solve for ΔE :

$$\Delta E = T\Delta S - P\Delta V + \mu\Delta N. \quad (19)$$

This equation expresses energy conservation under the *generalization* that the number N of particles may change but also under the *restriction* that all changes are from one equilibrium state to a nearby equilibrium state.⁸ The involution steps in the present derivation are required because, at the start, we knew an expression for μ only in terms of the Helmholtz free energy.

Now, however, from Eq. (19) we can read off that

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V}. \quad (20)$$

Next, divide Eq. (19) by T , solve for ΔS , and read off the relation

$$-\frac{\mu}{T} = \left(\frac{\partial S}{\partial N} \right)_{E,V}. \quad (21)$$

Finally, add $\Delta(PV)$ to both sides of Eq. (19), rearrange to have $\Delta G \equiv \Delta(E - TS + PV)$ on the left-hand side, and read off that

$$\mu = \left(\frac{\partial G}{\partial N} \right)_{T,P}. \quad (22)$$

If a reaction comes to equilibrium at fixed temperature and pressure, the Gibbs free energy attains a minimum. Equation (22) shows that the chemical potential then plays the same role that it did under conditions of fixed temperature and volume. Gibbs often considered the case of minimum energy at fixed entropy and external parameters.⁹ Now Eq. (20) shows that the chemical potential retains the same role.

If entropy, however, is the function that attains an extremum, then the pattern is broken. Equation (21) shows that the quotient μ/T takes the place of μ alone. Merely on dimensional grounds, some alteration was necessary.¹⁰

All in all, this section supports the second characterization and shows how the entropy extremum departs from the typical pattern.

C. Further comments

The sum in Eq. (14) may be split into a difference of subsums over products and reactants, respectively. In each subsum, the chemical potentials are weighted by the corresponding stoichiometric coefficients (all taken positively here). If the system has not yet reached thermodynamic equilibrium, the entire sum in Eq. (14) will be nonzero. Evolution is toward lower Helmholtz free energy. That implies evolution toward products if their weighted sum of chemical potentials is smaller than that of the reactants. Here is the analog of particle diffusion in real space toward lower chemical potential: a chemically reactive system evolves toward the side—products or reactants—that has the lower weighted sum of chemical potentials.

One may generalize the term “reaction” to apply to coexistence of phases. One need only interpret the term to mean “transfer of a particle from one phase to another.” Thereby one recaptures the property that, at thermodynamic equilibrium, the chemical potential is spatially uniform (in each connected region to which diffusion can carry the particles).

The present paragraph redeems the promise in Sec. II to return to diffusion. In an isothermal context, diffusion is determined by the gradient of the chemical potential, $\text{grad } \mu$, as

developed in Sec. III. In a more general context, one can turn to the first-order theory of relaxation toward equilibrium. Among the theory's most appropriate variables are derivatives of the entropy density with respect to the number density and the energy density. These derivatives are $-\mu/T$ [by Eq. (21)] and $1/T$. Thus particle diffusion is determined by a linear combination of $\text{grad}(\mu/T)$ and $\text{grad}(1/T)$; each term enters the sum with an appropriate (temperature-dependent) coefficient.¹¹

When a system—initially not in thermal equilibrium—evolves toward equilibrium, the process can be complex: sometimes purely diffusive (in the random-walk sense), other times hydrodynamic. James McLennan's *Introduction to Nonequilibrium Statistical Mechanics*, cited in Ref. 11, provides an excellent survey. Because of the diversity of processes, I use the word “diffusion” broadly in this paper. Its meaning ranges from a strict random walk to merely “spreads out.”

V. CHARACTERISTIC ENERGY

Now we turn to the phrase “characteristic energy” as providing a meaning for the chemical potential. Here the clearest expression for μ is the form $(\partial E/\partial N)_{S,V}$: the system's energy change when one particle is added under conditions of constant entropy (and constant external parameters). Because entropy is so often a crucial quantity, an energy change at constant entropy surely provides a characteristic energy. The question can only be this: in which contexts does this characteristic energy play a major role?

I will not attempt a comprehensive response but rather will focus on a single context, arguably the most significant context in undergraduate thermal physics. Recall that the Fermi–Dirac and Bose–Einstein distribution functions depend on the energy ε_α of a single-particle state (labeled by the index α) through the difference $\varepsilon_\alpha - \mu$. Why does the chemical potential enter in this fashion?

Derivations of the distribution functions make comparisons of entropies (or multiplicities¹²), either directly or implicitly (through the derivation of antecedent probability distributions). In entropy lies the key to why μ appears, and two routes that use the key to explain the appearance come to mind.

(1) *Additive constant.* The physical context determines the set $\{\varepsilon_\alpha\}$ of single-particle energies only up to an arbitrary additive constant. When ε_α appears in a distribution function, it must do so in a way that is independent of our choice of the additive constant. In short, ε_α must appear as a difference. With what quantity should one compare ε_α and form the difference?

The energy ε_α describes the system's energy change when one particle is added in single-particle state φ_α . Typically, such an addition induces a change in entropy. So the comparison might well be made with $(\partial E/\partial N)_{S,V}$, the system's energy change when one particle is added under conditions of constant entropy (and constant external parameters). The derivative, of course, is another expression for the chemical potential, as demonstrated in Eq. (20).

(2) *Total entropy change when a particle is added.* Some derivations of the distribution functions entail computing the total entropy change of either the system or a reservoir when a particle is added to the system of interest.¹³ To avoid irrel-

evant minus signs, focus on the system. When a particle of energy ε_α is added, the entropy change consists of two parts:

$$\begin{aligned}\Delta S &= \left(\frac{\partial S}{\partial E}\right)_{N,V} \times \varepsilon_\alpha + \left(\frac{\partial S}{\partial N}\right)_{E,V} \times 1 \\ &= \frac{1}{T} \times \varepsilon_\alpha + \left(\frac{-\mu}{T}\right) \times 1 = \frac{(\varepsilon_\alpha - \mu)}{T}.\end{aligned}\quad (23)$$

The chemical potential enters through the relationship (21). As explained in Ref. 10, derivatives of S and E with respect to N are necessarily proportional to each other (at fixed external parameters). So, the second term must be expressible in terms of μ , which appears again as a characteristic energy.

The notion of “characteristic energy” can be made broader and looser. Chemists often focus on Eq. (22) and characterize the chemical potential as the “partial molar Gibbs free energy” (provided that particle numbers are measured in moles). Some authors note that, when only one particle species is present, Eq. (22) is numerically equivalent to the relation $\mu = G/N$, and they characterize the chemical potential as the Gibbs free energy per particle.¹⁴ For typical physics students, however, the Gibbs free energy remains the most mysterious of the thermodynamic functions, and so such characterizations are of little help to them.¹⁵

VI. NUMERICAL VALUES

This section is devoted to qualitative reasoning about the numerical value that the chemical potential takes on. The aim is to develop more insight and greater familiarity. For the most part, the reasoning is based on the form for μ given in Eq. (20).

A. Absolute zero

In the limit as the temperature is reduced to absolute zero, the system settles into its ground state, and its entropy becomes zero. (For a macroscopic system, any degeneracy of the ground state, if present, would be insignificant, and so the description assumes none.) Adding a particle at constant entropy requires that the entropy remain zero. Moreover, after the addition, the system must again be in thermal equilibrium. Thus the system must be in the ground state of the new system of $(N+1)$ particles. [One could preserve the constraint “entropy=0” by using a single state (of the entire system) somewhat above the ground state, but that procedure would not meet the requirement of thermal equilibrium.]

For a system of ideal fermions, which are subject to the Pauli exclusion principle, we construct the new ground state from the old by filling a new single-particle state at the Fermi energy ε_F . Thus the system's energy increases by ε_F , and that must be the value of the chemical potential.

Consider next bosons, such as helium atoms, that obey a conservation law: the number of bosons is set initially and remains constant in time (unless we explicitly add or subtract particles). For such bosons (when treated as a quantum ideal gas), we construct the new ground state by placing another particle in the single-particle state of lowest energy. The chemical potential will equal the lowest single-particle energy, ε_1 .

B. Semi-classical ideal gas

For an ideal gas in the semi-classical domain¹⁶ (such as helium at room temperature and atmospheric pressure), the probability that any given single-particle state is occupied is quite small. An additional atom could go into any one of a great many different single-particle states. Moreover, we may use classical reasoning about multiplicity and entropy.¹⁷ Adding an atom, which may be placed virtually anywhere, surely increases the spatial part of the multiplicity and hence tends to increase the entropy. To maintain the entropy constant, as stipulated in Eq. (20), requires that the momentum part of the multiplicity decrease. In turn, that means less kinetic energy, and so the inequality $\Delta E < 0$ holds, which implies that the chemical potential is negative for an ideal gas in the semi-classical domain. (Implicit here is the stipulation that the energy E is strictly kinetic energy. Neither a potential energy due to external forces nor a rest energy, mc^2 , appears.)¹⁸

In this subsection and in the preceding one, we determined ΔE (or placed a bound on it) by assessing the change in the system's energy as the system passed from one equilibrium state to another. How the additional particle is introduced and what energy it may carry with it are secondary matters. Why? Because the system is required to come to equilibrium again and to retain—or regain—its original entropy value. To satisfy these requirements, energy may need to be extracted by cooling or added by heating. Such a process will automatically compensate for any excess or deficiency in the energy change that one imagines to accompany the literal introduction of the extra particle.

To summarize, one compares equilibrium states of N and $N+1$ particles, states that are subject to certain constraints, such as “same entropy.” This comparison is what determines ΔE . The microscopic process by which one imagines one particle to have been introduced may be informative, but it is not decisive and may be ignored.

Next, consider the explicit expression for a chemical potential in Eq. (9)—but first delete the term mgH . The derivation presumed that the thermal de Broglie wavelength is much smaller than the average interparticle separation. Consequently, the logarithm's argument is less than one, and the chemical potential itself is negative. Can such a negative value be consistent with the characterization of the chemical potential as measuring the tendency of particles to diffuse? Yes, because what matters for diffusion is how μ changes from one spatial location to another. The spatial gradient (if any) is what determines diffusion, not the size or sign of μ at any single point.

C. Photons

Photons are bosons, but they are not conserved in number. Even in a closed system at thermal equilibrium, such as a hot kitchen oven, their number fluctuates in time. There is no specific number N of photons, although—when the temperature and volume of a cavity have been given—one can compute an estimated (or mean) number of photons $\langle N \rangle$.

There are several ways to establish that the chemical potential for photons is zero (though I find none of them to be entirely satisfactory). The most straightforward route is to compare the Planck spectral distribution with the Bose–

Einstein distribution, computed for conserved bosons. The distributions match up if one sets the chemical potential in the latter to zero.

One can also use a version of Eq. (20), replacing N by $\langle N \rangle$. Everything about a (spatially uniform) photon gas in thermal equilibrium is known if one knows the temperature and volume. Specifically, the entropy can be written in terms of T and V , and so T can be expressed in terms of S and V . Therefore, the energy E (usually considered a function of T and V) can be expressed in terms of S and V . Then any derivative of E at constant S and V must be zero.

For a third method, one can examine the annihilation of an electron and a positron or their creation. In thermal equilibrium, the reaction should follow the pattern set by chemical reactions:

$$\mu_{\text{electron}} + \mu_{\text{positron}} = 2\mu_{\text{photon}} \quad (24)$$

for the two-photon process. Provided the temperature, volume, and net electrical charge have been specified, one can construct the probability that N_- electrons, N_+ positrons, and any number of photons are present. The construction follows the route outlined in Sec. III for N_l and N_u (provided one treats the electrons and positrons as uncharged semi-classical gases). The ensuing probability is a function of N_- , N_+ , T , and V . Looking for the maximum in its logarithm, one finds the criterion

$$\mu_{\text{electron}} + \mu_{\text{positron}} = 0. \quad (25)$$

Comparing Eqs. (24) and (25), one infers that the chemical potential for photons is zero.

Next, note that the spin-singlet state of positronium may decay into two or four or any higher even number of photons. Simultaneous thermal equilibrium with respect to all of these processes [in the fashion illustrated in Eq. (24) for two-photon decay] is possible only if μ_{photon} equals zero. Here one sees clearly that the absence of a conservation law for photons leads to a zero value for the chemical potential.

Finally, consider the conduction electrons in the metal wall of a hot oven. The electrons interact among themselves and with the radiation field (as well as with the metallic ions, which are ignored here). One electron can scatter off another and, in the process, emit a photon into the oven. The reaction is $e + e' \rightarrow e'' + e''' + \gamma$. The primes indicate different electronic states. (The reversed process, in which a photon is absorbed, is also possible.) In thermal equilibrium, an analogous equation should hold among chemical potentials: $2\mu_e = 2\mu_e + \mu_{\text{photon}}$. From this equality, one infers that the chemical potential for thermal radiation is zero.

If the chemical potential for photons is everywhere zero, is there any meaning to the notion of “diffusion of photons”? Yes, a spatial gradient in the temperature field determines the flow of radiant energy and hence the “diffusion of photons.”

G. Cook and R. H. Dickerson¹⁹ provide additional and alternative qualitative computations of the chemical potential.

Now the paper turns—for illustration—to two applications of the chemical potential. The first—depression of the melting point—was the subject of a Question²⁰ and four Answers²¹ in the Journal's Question and Answer section in 1997. The second application—batteries—seems to be enduringly fascinating for most physicists.

VII. DEPRESSION OF THE MELTING POINT

Consider liquid water and ice in equilibrium at $T = 273\text{ K}$ and atmospheric pressure. If alcohol or table salt is added to the liquid water and if the pressure is maintained at one atmosphere, a new equilibrium will be established at a lower temperature. Thus, when ice is in contact with the solution, its melting point is depressed. The chemical potential provides a succinct derivation of the effect, as follows.

For the sake of generality, replace the term “ice” by “solid” and the term “liquid water” by “solvent.” The “solvent” here is the liquid form of the solid but may contain a solute (such as alcohol or salt).

The solvent’s chemical potential depends on the temperature, the external pressure, and the ratio η of number densities (or concentrations), solute to solvent:

$$\eta \equiv \frac{n_{\text{solute}}}{n_{\text{solvent}}} \quad (26)$$

In the absence of solute, the solvent and solid coexist at temperature T_0 and pressure P_0 , and so their chemical potentials are equal then.

After the solute has been added, the new equilibrium occurs at temperature $T_0 + \Delta T$ such that

$$\mu_{\text{solvent}}(T_0 + \Delta T, P_0, \eta) = \mu_{\text{solid}}(T_0 + \Delta T, P_0). \quad (27)$$

Expand the left-hand side about the arguments $(T_0, P_0, 0)$ and the right-hand side about (T_0, P_0) . To evaluate the partial derivatives with respect to T , one may differentiate Eq. (22) with respect to T , interchange the order of differentiation on the right-hand side, and note that $(\partial G / \partial T)_{P,N} = -S = -Ns(T, P)$, where s denotes the entropy per molecule. [Alternatively, one may use the Gibbs–Duhem relation²² for a single species (because the derivatives are evaluated in the pure phases).] The first-order terms in the expansion of Eq. (27) yield the relation

$$-s_{\text{solvent}}\Delta T + \frac{\partial \mu_{\text{solvent}}}{\partial \eta} \eta = -s_{\text{solid}}\Delta T. \quad (28)$$

The remaining derivative is to be evaluated at zero solute concentration. Solve for ΔT :

$$\Delta T = \frac{1}{s_{\text{solvent}} - s_{\text{solid}}} \times \frac{\partial \mu_{\text{solvent}}}{\partial \eta} \eta. \quad (29)$$

Both a theoretical model, outlined in Appendix A, and experimental evidence from osmosis indicate that $\partial \mu_{\text{solvent}} / \partial \eta$ is negative. A liquid usually has an entropy per particle higher than that of the corresponding solid phase, and so the difference in entropies is usually positive. The implication then is a *depression* of the melting point that is *linear* in the solute concentration (when a first-order expansion suffices).

Qualitative reasons for a depression are given in Ref. 21. Exceptions to the positive difference of entropies are described in Ref. 23. The chemical potential provides a simple analytic approach to related phenomena: elevation of the boiling point, osmotic pressure, and various solubility problems.²⁴

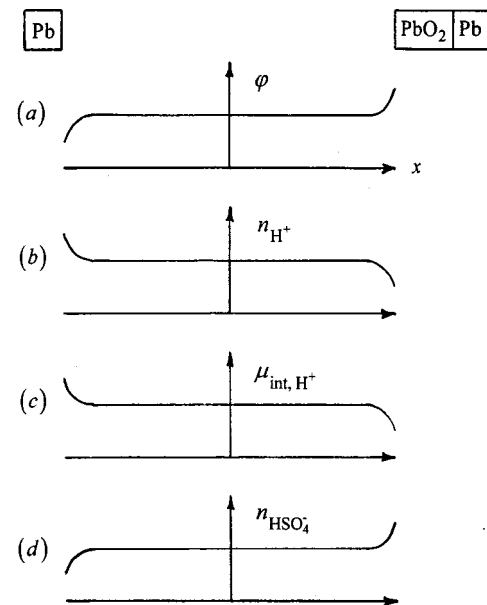


Fig. 3. For the electrolyte in a lead-acid cell, qualitative graphs of (a) the electric potential $\varphi(x)$, (b) the number density n_{H^+} of H^+ ions, (c) the intrinsic chemical potential for those ions, and (d) the number density of HSO_4^- ions. The abscissa x runs from the pure lead electrode to the terminal with lead and lead oxide. Potentials—both electric and chemical—usually contain an arbitrary additive constant. Implicitly, those constants have been chosen so that graphs (a) and (c) lie conveniently above the origin.

VIII. BATTERIES

The topic of batteries immediately prompts two questions:

- (1) Why does a potential difference arise?
- (2) How large is the potential difference between the terminals?

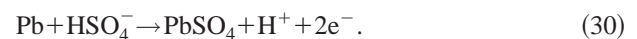
This section addresses the questions in order.

For a prelude, let me say the following. In discussing the first question, we will see how the spatial uniformity of the chemical potential helps one to infer and describe the spatial behavior of ionic concentrations and the electric potential. In addressing the second question, we will relate the potential difference to the (intrinsic) chemical potentials and stoichiometric coefficients of the particles whose reactions power the battery. An overall difference in binding energy will emerge as what sets the fundamental size of the potential difference.

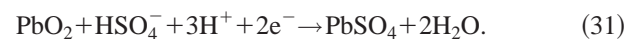
A. Why a potential difference?

To have an example before us, take a single cell of an automotive lead-acid battery. The chemical reactions are the following.

At the pure lead terminal,



At the terminal with lead and lead oxide,



Imagine commencing with two neutral electrodes and no electrolyte; then pour in a well-stirred mixture of sulfuric acid and water (and keep the cell on open circuit).

At the pure lead terminal, reaction (30) depletes the nearby solution of HSO_4^- ions and generates H^+ ions, as illustrated in Figs. 3(b) and 3(d). The two electrons contrib-

uted to the terminal make it negative. In the context of the figure, the negative charges on the terminal and the net positive charge density in the nearby electrolyte produce a leftward-directed electric field in the electrolyte. By themselves, the concentration gradients in the ion densities would produce diffusion that restores a uniform density. The electric forces oppose this tendency. Specifically, acting on the HSO_4^- ions, the field produces a rightward force and prevents the concentration gradient from eliminating the depletion. Acting on the H^+ ions, the field produces a leftward force and sustains the excess of H^+ ions.

The situation is an electrical analog of air molecules in the Earth's gravitational field: the effects of a force field and a concentration gradient cancel each other.

Over an interval of a few atomic diameters, the electrolyte has a positive charge density. Beyond that, the electrolyte is essentially neutral (when averaged over a volume that contains a hundred molecules or so). The combination of a positively charged interval and a negative surface charge on the electrode provides a region of leftward-directed electric field and a positive step in electric potential (as one's focus shifts from left to right). Figure 3(a) illustrates the step.

At the lead oxide electrode, reaction (31) depletes the nearby solution of H^+ ions [as illustrated in Fig. 3(b)] and simultaneously makes the terminal positive. The positive charges on the terminal produce a leftward-directed electric field. The electric force on the remaining H^+ ions prevents diffusion from eliminating the depletion.

The reaction depletes HSO_4^- ions also, but the 3-to-1 ratio in the reaction implies that the effect of the H^+ ions dominates. In fact, the electric force on the remaining HSO_4^- ions pulls enough of those ions into the region to produce a local excess of HSO_4^- ions, as illustrated in Fig. 3(d). (Later, a self-consistency argument will support this claim.) The electrolyte acquires a net negative charge density over an interval of a few atomic diameters. The combination of a negatively charged electrolyte and the positive surface charge on the electrode produces a leftward-directed electric field and another positive step in electric potential, as shown in Fig. 3(a).

In summary, a chemical reaction at an electrode is a sink or source of ions. Thus the reaction generates a concentration gradient and a separation of charges. The latter produces an electric field. In turn, the electric field opposes the tendency of diffusion to eliminate the concentration gradient and the charge separation. The electric field is preserved and, acting over an interval, produces a step in electric potential.

For each ion, its chemical potential may be decomposed into two terms: (1) the chemical potential that the ion would have in the absence of an electric potential and (2) the electrical potential energy of the ion due to the macroscopic electric potential $\varphi(x)$. The first term was called by Gibbs the "intrinsic potential," and it is now called the *intrinsic or internal chemical potential*. The subscript "int" may be read in either way. Thus an ion's chemical potential μ has the form

$$\mu = \mu_{\text{int}} + q_{\text{ion}}\varphi, \quad (32)$$

where q_{ion} is the ion's charge.

The intrinsic chemical potential is an increasing function of the ion's concentration (provided the physical system is stable).²⁵ Thus Fig. 3(c) displays a graph for $\mu_{\text{int}, \text{H}^+}$ that shows the same trends that the number density n_{H^+} does.

At thermal equilibrium the chemical potential for each species is uniform in space (in each connected region to which diffusion can carry the particles). For a positively charged ion, the concentration will decrease where the electric potential increases; a comparison of graphs (a) and (b) illustrates this principle. For a negatively charged ion, the opposite behavior occurs. Thus consistency among the four graphs in Fig. 3 requires the concentration variations of HSO_4^- ions to be opposite those for H^+ ions.

The gradient of Eq. (32) lends itself to a simple interpretation if the ions exist in a *dilute* aqueous solution. Then the intrinsic chemical potential for a specific ion depends spatially on merely the concentration $n(x)$ of that ion. To say that the chemical potential is uniform in space is to say that the gradient of the chemical potential is zero. Thus Eq. (32) implies

$$\left(\frac{\partial \mu_{\text{int}}}{\partial n} \right) \frac{dn}{dx} - q_{\text{ion}} E_x = 0, \quad (33)$$

where E_x denotes the x -component of the electric field. The electric force $q_{\text{ion}} E_x$ annuls the diffusive tendency of a concentration gradient, dn/dx .

B. How large is the potential difference?

To determine the potential difference when the cell has come to equilibrium on open circuit, we start with a basic principle: at chemical equilibrium, the Gibbs free energy attains a minimum (in the context of fixed temperature and external pressure). Taking into account *both* reactions (30) and (31), we have

$$\Delta G = \sum_i b_i \mu_i = 0. \quad (34)$$

The sum of stoichiometric coefficients times chemical potentials must yield zero. Note that electrons appear *twice* in this sum, once when transferred *from* an electrode and again when transferred *to* the other electrode.

Each chemical potential has the two-term structure displayed in Eq. (32). Some care is required in evaluating the term containing the electric potential. Indeed, the ions and electrons need separate treatment. We will find that the ionic electric potential terms in ΔG cancel out and that the electronic contribution introduces the potential difference.

For the ions, we may use the values of the chemical potentials at the center of the electrolyte. Although the reaction uses up ions at the electrodes, those ions are replenished from the plateau region, and that replenishment is part of the step from one equilibrium context to another.

An alternative way to justify using the central location is to note that the chemical potential for each ion is uniform throughout the electrolyte. Thus every location gives the same numerical value, but the center—because it typifies the plateau region—will ensure that, later, the ionic intrinsic chemical potentials are to be evaluated in the bulk region of the electrolyte.

The number of "conduction" electrons is the same in the reactants and the products. The electrons are merely on different electrodes. Consequently, the net charge on the ions is the same for the reactants and the products. When those net charges are multiplied by the electric potential at the center and then subtracted, the contributions cancel out. All that the

ions and neutrals contribute to the sum in Eq. (34) is their intrinsic chemical potentials, suitably weighted.

Quite the opposite is true for the conduction electrons. Two electrons are removed from the metallic lead at the positive electrode, and two are transferred to the lead at the negative electrode. The contribution of their electric potential terms to Eq. (34) is

$$-(-2e)\varphi_{\text{pos}} + (-2e)\varphi_{\text{neg}} = -(-2e)\Delta\varphi, \quad (35)$$

where $\Delta\varphi \equiv \varphi_{\text{pos}} - \varphi_{\text{neg}}$ is the positive potential difference between the terminals and where e denotes the magnitude of the electronic charge. The electron's intrinsic chemical potentials, however, have the same numerical value in the two pieces of lead,²⁶ and so they cancel out in Eq. (34).

The upshot is that the minimum property for G implies

$$\Delta G = \sum_{i \neq \text{el}} b_i \mu_{\text{int},i} - (-2e)\Delta\varphi = 0. \quad (36)$$

The summation includes the ions and neutrals but excludes the electrons.

Solving for the potential difference, one finds

$$\Delta\varphi = \frac{-1}{q_{\text{ch}}} \times \sum_{\substack{\text{ions and} \\ \text{neutrals}}} b_i \mu_{\text{int},i}, \quad (37)$$

where q_{ch} denotes the magnitude of the electrons' charges. For the ions, the intrinsic chemical potentials are to be evaluated in the bulk region of the electrolyte.

The sum obviously depends on the cell's composition. For the species in the electrolyte, each μ_{int} can be expanded to display a dependence on concentrations. Thus composition and concentrations determine the potential difference. But significantly more can be said, as follows.

For an ion in an aqueous solution, the chemical potential is difficult to calculate explicitly. If the electrolyte were a dilute gas, computation would be much simpler. The intrinsic chemical potential for each ion or neutral molecule would contain, as a term, the particle's ground-state energy $\varepsilon_{\text{g.s.}}$ (reckoned relative to the energy of some standard state, such as the energy of free neutral atoms and free electrons at rest at infinite separation).²⁷ That is to say, $\mu_{\text{int}} = \varepsilon_{\text{g.s.}} + \dots$. *Differences in binding energy* would dominate the sum in Eq. (37), would set the fundamental size of the open-circuit potential difference, and would provide the primary energy source for a current if the circuit were closed. Qualitatively the same situation prevails in an aqueous solution.²⁸

Appendix B describes a more common route to the result in Eq. (37) and reconciles the two different values for ΔG . Dana Roberts²⁹ and Wayne Saslow³⁰ provide complementary treatments of batteries. A chemist's view of a battery, couched in language that a physicist can penetrate, is given by Jerry Goodisman.³¹

Having seen—in two applications—how the chemical potential is used, we can turn to a sketch of how it arose and was named.

IX. SOME HISTORY

J. Willard Gibbs introduced the chemical potential in his great paper, "On the Equilibrium of Heterogeneous Substances," published in two parts, in 1876 and 1878. In those days, Gibbs was doing thermodynamics, not statistical mechanics, and so he differentiated the system's energy with

respect to the macroscopic mass (denoted m_i by him) of the substance denoted by the subscript i . Entropy and volume were to be held constant. Thus Gibbs introduced a macroscopic version of Eq. (20). He had denoted the system's energy and entropy by the lower case Greek letters ε and η , respectively; so, presumably, he chose the letter μ to provide a mnemonic for a derivative with respect to mass.

Early in the paper, Gibbs called his derivative merely the "potential."³² Later, he found it necessary to distinguish his derivative from the electric potential and gravitational potential. He introduced the term "intrinsic potential" for a derivative that is "entirely determined at any point in a mass by the nature and state of the mass about that point."³³ In short, the intrinsic potential is a local quantity, dependent on the local mass (or number) density, say, but not on fields created by distant charges or masses. For example, Gibbs would have called the second term on the right-hand side of Eq. (9) the intrinsic potential for the atoms in the upper volume.

Nowhere in his major paper does Gibbs use the term "chemical potential." That coinage seems to have been introduced by Wilder Dwight Bancroft.³⁴ A Ph.D. student of Wilhelm Ostwald, Bancroft was a physical chemist at Cornell and the founder of the *Journal of Physical Chemistry* (in 1896). Yale's Beinecke Library preserves five letters from Bancroft to Gibbs,³⁵ dated 1898–1899. In the earliest letter, Bancroft adheres to Gibbs's usage. In his second letter, dated 18 March 1899, however, Bancroft says he is trying to find time to write a book on electrochemistry and uses the phrase "the chemical potential μ ." He employs the phrase nonchalantly, as though it were a familiar or natural phrase. Most likely, Bancroft found a need to distinguish between the electric potential and Gibbs's (intrinsic) potential. The term "chemical potential" for the latter would make the distinction clear. (Altogether, Bancroft uses the new term again in two more of the later letters.)

The fourth letter, dated 4 June 1899, has some wonderful lines. Bancroft mentions that "it has taken me seven years hard work to find out how your equation should be applied in actual cases" and comments, "The chemical potential is still a mere phrase to everyone although Ostwald uses it with a certain specious glibness." Then he launches into his peroration:

If we can once get used to writing and thinking in terms of the chemical potential for the comparatively simple case of electromotive forces, it will not be so difficult to take the next step and to think in terms of the chemical potentials when we are dealing with systems which cannot be transposed to form a voltaic cell. So far as I can see at present, our only hope of converting organic chemistry, for instance, into a rational science lies in the development and application of the idea of the chemical potential.

Widespread understanding of how to use the chemical potential was slow in coming.

In a reply to Bancroft, Gibbs acceded to the coinage, writing about the need "to evaluate the (intrinsic or chemical) potentials involved" in a working theory of galvanic cells.³⁶

Electrochemists remain true to Bancroft's usage. In referring to Eq. (32), they would call the term μ_{int} the "chemical potential" and would cite the entire right-hand side as the "electrochemical potential."³⁷

Physicists sometimes split the right-hand side of Eq. (32) into “internal” and “external” chemical potentials and denote the sum as the “total” chemical potential.³⁸

Most often, however, physicists accept whatever emerges when one forms the derivative indicated in Eqs. (6), (20), and (22) and call the result the chemical potential. In this paper, I adopted that usage.

Nonetheless, the splitting of the chemical potential into various terms warrants further comment. Gibbs noted that the zeroes of energy and entropy are arbitrary; consequently, the chemical potential may be shifted arbitrarily in value by a term of the form³⁹ $(\text{constant}) - T \times (\text{another constant})$. This result is seen most easily by regarding the chemical potential as a derivative of the Helmholtz free energy, as in Eq. (6).

Today, the Third Law of Thermodynamics gives us a natural zero for entropy. The zero of energy, however, may seem to remain arbitrary. So long as the number of particles of species i remains constant in whatever process one considers, the zero for the energy of those particles remains irrelevant. If, however, particles of species i are created or annihilated, then one must include the rest energy, $m_i c^2$, where m_i denotes the rest mass, in the explicit expression for the chemical potential. The rest energy term is absolutely essential in a description of the early universe and many other aspects of high-temperature astrophysics.⁴⁰ In short, special relativity theory provides a natural zero for the energy of (free) particles.

Now this section returns to history. James Clerk Maxwell was fascinated by Thomas Andrews’ experiments on carbon dioxide: the coexistence of liquid and vapor, the critical point, trajectories in the pressure-temperature plane, and mixtures of CO_2 with other gases. Moreover, Maxwell had been impressed by the geometric methods that Gibbs outlined in his first two papers on thermodynamics. In this context, Maxwell developed—independently of Gibbs—his own version of the chemical potential and some associated relationships. When Gibbs’s comprehensive paper appeared, Maxwell dropped his own formalism and enthusiastically recommended Gibbs’s methods.⁴¹

Speaking to a conference of British chemists in 1876, Maxwell distinguished between what we would today call “extensive” and “intensive” thermodynamic properties. The former scale with the size of the system. The latter, in Maxwell’s words, “denote the intensity of certain physical properties of the substance.” Then Maxwell went on, explaining that “the pressure is the intensity of the tendency of the body to expand, the temperature is the intensity of its tendency to part with heat; and the [chemical] potential of any component is the intensity with which it tends to expel that substance from its mass.”⁴² The idea that the chemical potential measures the tendency of particles to diffuse is indeed an old one.

Maxwell drew an analogy between temperature and the chemical potential. The parallel was noted already near the end of Sec. III, and it is developed further in Appendix C, which explores the question, why does this paper offer three characterizations of the chemical potential, rather than just a single comprehensive characterization?

ACKNOWLEDGMENTS

Professor Martin J. Klein provided incisive help in tracking down Bancroft’s contribution. Professor Martin Baily and Professor Harvey Leff read preliminary versions and offered

comments and suggestions. Finally, Professor Yue Hu posed several good questions, and our ensuing discussion produced a correction and improvements in the typescript. To all of these generous people, I express my thanks.

APPENDIX A: HOW THE SOLVENT’S CHEMICAL POTENTIAL VARIES

This appendix focuses on how μ_{solvent} varies with the (relative) concentration of solute, denoted η . A verbal argument comes first; then a derivation with equations will support it.

As a prelude, however, let us note that a liquid is largely incompressible. To a good approximation, its volume is determined by the number of molecules and the temperature. The volume need not be considered an external parameter (unlike the situation with a gas). I will adopt this “incompressible” approximation here.

To begin the verbal argument, recall that one may think of μ_{solvent} as the change in F_{total} , the Helmholtz free energy for the system of solvent and solute, when one solvent molecule is added to the liquid (at constant temperature and constant number of solute molecules).

Addition of a solvent molecule increases the volume of the liquid. Thus *solute* molecules have more space in which to move. Their entropy increases, and so F_{total} undergoes a *supplemental* change, which is a decrease. Therefore $\mu_{\text{solvent}}(T, P, \eta)$ is less than $\mu_{\text{solvent}}(T, P, 0)$.

The supplemental decrease scales approximately linearly with N_{solute} (because S_{solute} scales approximately linearly with N_{solute} when the solute concentration is relatively small). Consequently, $\partial \mu_{\text{solvent}} / \partial \eta$ is negative and approximately constant (for small η).

A simple theoretical model enables one to confirm this line of reasoning. To construct the partition function for the solution, modify the forms in Eqs. (2) and (8) to read as follows:

$$Z(T, N_1, N_2) = \frac{(N_1 v_1 + N_2 v_2)^{N_1 + N_2}}{\lambda_{\text{th},1}^{3N_1} N_1! \times \lambda_{\text{th},2}^{3N_2} N_2!} \times \exp[(N_1 \varepsilon_1 + N_2 \varepsilon_2)/kT]. \quad (\text{A1})$$

The subscripts 1 and 2 refer to solvent and solute, respectively. Because a liquid is largely incompressible and determines its own volume, one may replace a container volume V by $N_1 v_1 + N_2 v_2$, where the constants $\{v_1, v_2\}$ denote volumes of molecular size. An attractive force (of short range) holds together the molecules of a liquid and establishes a barrier to escape. Model the effect of that force by a potential well of depth $-\varepsilon$, saying that a molecule in the liquid has potential energy $-\varepsilon$ relative to a molecule in the vapor phase.

In short, the model treats the liquid as a mixture of two ideal gases in a volume determined by the molecules themselves (as they jostle about in virtually constant contact with each other). The energy ε provides the dominant contribution to the liquid’s latent heat of vaporization.

The solvent’s chemical potential now follows by differentiation as

$$\mu_{\text{solvent}} = -\varepsilon_1 + kT \ln \left(\frac{\lambda_{\text{th},1}^3}{e v_1} \right) - kT \eta + O(\eta^2). \quad (\text{A2})$$

As expected, μ_{solvent} decreases approximately linearly with the (relative) concentration of solute. Moreover, if one keeps track of where terms originate, one finds that the term $-kT\eta$ arises from the solute's entropy.

Note also that the term linear in η is *independent* of the ratio v_2/v_1 . The terms of higher order in η do depend on v_2/v_1 . The distinction should be borne in mind if one asks whether purported qualitative reasons for the depression of the melting point are actually valid.

APPENDIX B: BATTERIES BY AN ALTERNATIVE ROUTE

The potential difference across a battery's terminals is often calculated by assessing the electrical work done when the chemical reaction proceeds by one step. This appendix outlines that alternative route and reconciles it with the route presented in the main text.

To determine the potential difference, we start with energy conservation in the form

$$T\Delta S = \Delta E + P\Delta V + w_{\text{el}}, \quad (\text{B1})$$

where w_{el} denotes the (external) electrical work done by the system. In this alternative route, "the system" consists of the ions and neutrals *but not the conduction electrons* in the terminals.

In the context of fixed temperature and pressure, Eq. (B1) may be rearranged as

$$-\Delta(E - TS + PV) \equiv -\Delta G = w_{\text{el}}. \quad (\text{B2})$$

The electrical work equals the energy that would be required to transport the electrons literally across the potential difference, and so Eq. (B2) becomes

$$-\Delta G = q_{\text{ch}}\Delta\varphi, \quad (\text{B3})$$

where (as before) q_{ch} denotes the magnitude of the electrons' charges and where the potential difference $\Delta\varphi$ is positive.

The change in G is given by

$$\Delta G = \sum_{\substack{\text{ions and} \\ \text{neutrals}}} b_i \mu_i, \quad (\text{B4})$$

the sum of stoichiometric coefficients times chemical potentials. For the ions, we may use the values of the chemical potentials at the cell's center. (Justification for this step was given in the main text.) Because the reaction preserves electric charge, that is, because the products have the same net charge as do the reactants, the value of the electric potential on the plateau, $\varphi(0)$, cancels out in the sum. All that remains is the weighted sum of *intrinsic* chemical potentials for ions and neutrals.

Upon combining these observations with Eqs. (B3) and (B4), we find the relationship that determines the potential difference:

$$\Delta\varphi = \frac{-1}{q_{\text{ch}}} \times \sum_{\substack{\text{ions and} \\ \text{neutrals}}} b_i \mu_{\text{int},i}. \quad (\text{B5})$$

For the ions, the intrinsic chemical potentials are to be evaluated in the bulk electrolyte.

The result for $\Delta\varphi$ is, of course, the same here as in the main text. The two routes differ in what they take to be "the system," and so their intermediate steps necessarily differ also. The derivation in the main text takes a comprehensive

view of the cell and sees it in chemical equilibrium (when on open circuit). Relations internal to the system determine the electric potential difference. In this appendix, a portion of the cell (taken to be the thermodynamic system) does work on another portion (namely, the conduction electrons).

APPENDIX C: WHY SO MANY?

This appendix addresses two questions.

(A) Why does this paper offer three characterizations of the chemical potential, rather than just a single comprehensive characterization?

(B) Why does the chemical potential have so many equivalent expressions, which seem to make it exceptional?

Response to Question A. Rather than immediately confront the chemical potential, let us consider temperature and ask how many characterizations it requires. One may start with either a common-sense understanding of (absolute) temperature or the general expression $1/T = (\partial S/\partial E)_{N,V}$.

In characterizing temperature or listing its "meanings," I would start with the statement

(1) *Tendency of energy to diffuse.* As a function of position, temperature measures the tendency of energy to diffuse (by thermal conduction and by radiation).

But I could not stop there. I would have to add the statement

(2) *Characteristic energy.* Temperature provides a characteristic energy: kT .

To be sure, I would have to qualify the second statement. Only if classical physics and the equipartition theorem apply does kT give the energy per particle or per mode or per "degree of freedom" (within factors of $\frac{1}{2}$ or $\frac{3}{2}$ or so). But even for a degenerate quantum system like the conduction electrons in a metal, kT provides a characteristic energy that one compares with the Fermi energy ϵ_F .

Thus temperature has at least two essential characterizations.

If only one species of particle is present, then the characterizations of the chemical potential can be reduced to two items: (1) measures the tendency of particles to diffuse and (2) provides a characteristic energy. So one could see the chemical potential as no more complicated than temperature, and both require at least two characterizations.

But, of course, the chemical potential really comes into its own when more than one species of particle is present and reactions are possible. Then, it seems to me, I am stuck with needing to assign three characterizations.

Response to Question B. Energy, entropy, Helmholtz free energy, Gibbs free energy, and enthalpy commonly appear in undergraduate thermal physics. Each of these functions is the optimal function to use in some set of physical circumstances. In each case, one may contemplate adding a particle to the system, and so the chemical potential can be expressed as a derivative of each of the five functions.

Given the same five functions, in how many straightforward ways can one express temperature as a derivative? Three: as the derivative of entropy with respect to energy, and as the derivatives of energy and enthalpy with respect to entropy. The two free energies already have temperature as a natural independent variable, so there is no straightforward way to express temperature as a derivative.

We tend, however, to view the expression $1/T = (\partial S/\partial E)_{N,V}$ and its reciprocal as a single relationship, not as two distinct relationships. Moreover, expressing tempera-

ture in terms of an enthalpy derivative lies far outside the typical physicist's working sphere.

All in all, both the chemical potential and temperature have several equivalent expressions. In undergraduate physics, the chemical potential has two more than temperature does, and those of temperature tend to be reduced to a singleton in practice. But—as far as the number of equivalent expressions goes—the chemical potential is *not* exceptional in any qualitative fashion.

¹Ralph Baierlein, *Thermal Physics* (Cambridge U. P., New York, 1999), pp. 148–155.

²The partition function Z is the sum of Boltzmann factors taken over a complete, orthogonal set of energy eigenstates. Consequently, both Z and F depend on temperature, the number of particles of each species that is present, and the system's external parameters. If more than one species is present, the chemical potential μ_i for species i is the derivative of F with respect to N_i , computed while the numbers of all other particles are kept fixed and while temperature and all external parameters are held fixed.

Equation (6) indicates that the chemical potential may be computed by applying calculus to a function of N or by forming the finite difference associated with adding one particle. When N is large, these two methods yield results that are the same for all practical purposes. Convenience alone determines the choice of method.

³Charles Kittel and Herbert Kroemer, *Thermal Physics*, 2nd ed. (Freeman, New York, 1980), pp. 118–125.

⁴F. Mandl, *Statistical Physics* (Wiley, New York, 1971), pp. 222–224.

⁵For examples, see Ref. 1, p. 154, and Ref. 3, pp. 118–120.

⁶Martin Bailyn, *A Survey of Thermodynamics* (AIP, New York, 1994), p. 213.

⁷Reference 1, pp. 233–234. Nuclear spin systems may be in thermal equilibrium at negative absolute temperatures. In such a case, the Helmholtz free energy attains a maximum. For more about negative absolute temperatures, see Ref. 1, pp. 343–347 and the references on pp. 352–353.

⁸Reference 6, pp. 212–213; Ref. 1, p. 228.

⁹J. Willard Gibbs, *The Scientific Papers of J. Willard Gibbs: Vol. I, Thermodynamics* (Ox Bow, Woodbridge, CT, 1993), pp. 56, 64, and 65, for examples.

¹⁰Relation (21) can be seen as a consequence of a remarkable mathematical identity, most often met in thermodynamics: if the variables $\{x, y, z\}$ are mutually dependent, then $(\partial x / \partial y)_z (\partial y / \partial z)_x (\partial z / \partial x)_y = -1$. To understand why the minus sign arises, note that not all of the variables can be increasing functions of the others. For a derivation, see Herbert B. Callen, *Thermodynamics* (Wiley, New York, 1960), pp. 312–313. The correspondence $\{S, N, E\} = \{x, y, z\}$ yields Eq. (21), given Eq. (20) and the expression $(\partial S / \partial E)_{N, V} = 1/T$ for absolute temperature.

¹¹A classic reference is S. R. De Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962), Chaps. 3, 4 and 11. James A. McLennan provides a succinct development in his *Introduction to Nonequilibrium Statistical Mechanics* (Prentice-Hall, Englewood Cliffs, NJ, 1989), pp. 17–25. H. B. G. Casimir develops an illuminating example in his article, “On Onsager's Principle of Microscopic Reversibility,” *Rev. Mod. Phys.* **17**, 343–350 (1945)—but beware of typographical errors. A corrected (but less lucid) version is given by H. J. Kreuzer, *Nonequilibrium Thermodynamics and its Statistical Foundations* (Oxford U. P., New York, 1981), pp. 60–62.

¹²As used here, the term “multiplicity” denotes the number of microstates associated with a given macrostate. Entropy is then Boltzmann's constant times the logarithm of the multiplicity.

¹³Reference 3, pp. 134–137. Francis W. Sears and Gerhard L. Salinger, *Thermodynamics, Kinetic Theory, and Statistical Thermodynamics*, 3rd ed. (Addison-Wesley, Reading, MA, 1975), pp. 327–331.

¹⁴C. B. P. Finn, *Thermal Physics* (Routledge and K. Paul, Boston, 1986), pp. 190–193. Also Ref. 4, pp. 224–225.

¹⁵Here are two more characterizations that offer little help. Some authors describe the chemical potential as a “driving force” (e.g., Ref. 3, p. 120). Others call μ or a difference in μ a “generalized force” [e.g., Herbert B. Callen, *Thermodynamics* (Wiley, New York, 1960), pp. 45–46]. These strike me as unfortunate uses of the word “force.” In mechanics, students learn that a “force” is a push or a pull. But no such push or pull drives the statistical process of diffusion. The terms “driving force” and “generalized force” are more likely to confuse than to enlighten.

¹⁶As used in this paper, the term “semi-classical” means that the thermal de

Brogie wavelength is much smaller than the average interparticle separation. Hence quantum mechanics has no substantial direct effect on dynamics. Nonetheless, vestiges of the indistinguishability of identical particles persist (as in division by $N!$ in the partition function), and Planck's constant remains in the most explicit expressions for entropy, partition function, and chemical potential. A distinction between fermions and bosons, however, has become irrelevant.

The analysis in Sec. III treated the helium atoms as a semi-classical ideal gas.

¹⁷See Ref. 12.

¹⁸David L. Goodstein provides similar (and complementary) reasoning in his *States of Matter* (Prentice-Hall, Englewood Cliffs, NJ, 1975), p. 18.

¹⁹G. Cook and R. H. Dickerson, “Understanding the chemical potential,” *Am. J. Phys.* **63**, 737–742 (1995).

²⁰Sherwood R. Paul, “Question #56. Ice cream making,” *Am. J. Phys.* **65**, 11 (1997).

²¹The most germane of the answers is that by F. Herrmann, “Answer to Question #56,” *Am. J. Phys.* **65**, 1135–1136 (1997). The other answers are Allen Kropf, *ibid.* **65**, 463 (1997); M. A. van Dijk, *ibid.* **65**, 463–464 (1997); and Jonathan Mitschele, *ibid.* **65**, 1136–1137 (1997).

²²Reference 1, pp. 279–280.

²³Although the entropy per particle is usually higher in a liquid than in the corresponding solid, exceptions occur for the helium isotopes ^3He and ^4He on an interval along their melting curves. See J. Wilks and D. S. Betts, *An Introduction to Liquid Helium*, 2nd ed. (Oxford U. P., New York, 1987), pp. 15–16.

²⁴Frank C. Andrews, *Thermodynamics: Principles and Applications* (Wiley, New York, 1971), pp. 211–223. Also, Frank C. Andrews, “Colligative Properties of Simple Solutions,” *Science* **194**, 567–571 (1976). For a more recent presentation, see Daniel V. Schroeder, *An Introduction to Thermal Physics* (Addison-Wesley, Reading, MA, 2000), pp. 200–208.

²⁵Recall that particles diffuse toward *lower* chemical potential. If the concentration increases locally, only an increase in μ_{int} will tend to restore the original state and hence provide stability. For a detailed derivation, see Ref. 6, pp. 232, 239, and 240.

²⁶One might wonder, does charging the electrodes significantly alter the electrons' intrinsic chemical potentials? To produce a potential difference of 2 V, say, on macroscopic electrodes whose separation is of millimeter size requires only a relatively tiny change in the number of conduction electrons present. Thus the answer is “no.”

²⁷Reference 1, pp. 246–257.

²⁸Some of the energy for dc operation may come from the environment (as an energy flow by thermal conduction that maintains the battery at ambient temperature). Such energy shows up formally as a term $T\Delta S$ in ΔG [seen most easily in Eq. (B2)]. Measurements of $d\Delta\phi/dT$ show that this contribution is typically 10% of the energy budget (in order of magnitude). Gibbs was the first to point out the role of entropy changes in determining cell voltage (Ref. 9, pp. 339–349).

²⁹Dana Roberts, “How batteries work: A gravitational analog,” *Am. J. Phys.* **51**, 829–831 (1983).

³⁰Wayne Saslow, “Voltaic cells for physicists: Two surface pumps and an internal resistance,” *Am. J. Phys.* **67**, 574–583 (1999).

³¹Jerry Goodisman, *Electrochemistry: Theoretical Foundations* (Wiley, New York, 1987).

³²Reference 9, pp. 63–65 and 93–95.

³³Reference 9, pp. 146 and 332.

³⁴A. Ya. Kipnis, “J. W. Gibbs and Chemical Thermodynamics,” in *Thermodynamics: History and Philosophy*, edited by K. Martínás, L. Ropolyi, and P. Szegedi (World Scientific, Singapore, 1991), p. 499.

³⁵The letters are listed in L. P. Wheeler, *Josiah Willard Gibbs: The History of a Great Mind*, 2nd ed. (Yale U. P., New Haven, 1962), pp. 230–231.

³⁶Reference 9, p. 425.

³⁷E. A. Guggenheim introduced the phrase “electrochemical potential” and made the distinction in “The conceptions of electrical potential difference between two phases and the individual activities of ions,” *J. Phys. Chem.* **33**, 842–849 (1929).

³⁸Reference 3, pp. 124–125.

³⁹Reference 9, p. 95–96.

⁴⁰Reference 1, pp. 262–264.

⁴¹Elizabeth Garber, Stephen G. Brush, and C. W. F. Everitt, *Maxwell on Heat and Statistical Mechanics: On “Avoiding All Personal Enquiries” of Molecules* (Lehigh U. P., Bethlehem, PA, 1995), pp. 50 and 250–265.

⁴²Reference 41, p. 259.