

Electrode and corrosion physics

Preface

These notes have arisen from a series of Journal Clubs held at Imperial College and King’s College London, beginning in 2018. The object was for a number of physicists to try and understand the puzzling language of electrochemistry. I first whimsically ventured to call these “Electrophysics Notes”, but now I think “Electrode and Corrosion Physics” is a better title.

The approach is rather unorthodox because we assumed familiarity with electrostatics and density functional theory, but not with even elementary electrochemistry. The Appendix on thermodynamics, and in particular chemical potential, is adapted from my lecture notes to the MSE 307 class at Imperial College.

The notes are typeset in Plain \TeX , and the date is entered automatically. This means that readers can keep track of corrections and the addition of further sections; we don’t know how far these notes will evolve, or into what eventual form they will evolve.

I am very grateful for the insights of Sasha Lozovoi, Andrew Horsfield and Mike Finnis.

Motivation was provided by the UKRI / EPSRC grant, EP/R005230/1, “A tool for atomic scale simulation of corrosion: applications to Mg and Ti alloys.”

Tony Paxton, King’s College London, 28 May 2021

Contents

1.	The lemon lamp	1
2.	Simple electrochemical cells	3
2.1	The simplest cell	3
2.2	The next simplest cell	5
3.	Inner potential, work function and contact potential	6
3.1	Density functional theory	7
3.2	Inner potential difference	9
3.3	Contact potential	11
3.4	Work function	12
3.5	Contact potential again	15
4.	Electrochemical potential; real potential	17
4.1	Electrochemical potential	17
4.2	Real Potential	18
5.	Outer electric potential	18
6.	The Bockris point	24
7.	Electron work function of an electrolyte	25
8.	Work function and real potential of a species in an electrolyte	26
8.1	Electrochemical potential of an ion	26
8.2	Ion work function of an electrolyte	28
8.3	Confused and confusing remarks	30
9.	Equilibrium and reversible work of the electrochemical cell	31
9.1	Standard reaction free enthalpy	31
9.2	Inner potential difference across a metal–solution interface	33
9.3	Open circuit voltage—electromotive force	34
9.4	The electrochemical series	36
9.5	Nernst equation	39
10.	Relative and absolute ion work function and real potential	40
10.1	Measurement of the proton work function and real potential in water	40
10.2	Absolute electrode potential	42
11.	Electrode capacitance and electrocapillarity	45
11.1	Thermodynamics of the interphase	45
11.2	The electrocapillarity and Lippmann equations	54
11.3	Measurement	59
12.	Atomistic models of the interphase and interphase capacitance	61
12.1	Gouy–Chapman theory	62

13.	Kinetics	64
13.1	Corrosion	64
13.2	First order rate equation	65
13.3	Rate constant according to Bockris	66
13.4	Velocity at the saddle point	69
13.5	Equilibrium constant in terms of partition function	70
13.6	Rate constant in terms of partition function	74
13.7	Formulation from first principles	75
14.	Single electrode in equilibrium and at an overpotential	81
14.1	Detailed balancing	81
14.2	Electronation	82
14.2	De-electronation	83
14.4	Further interpretations of the symmetry factor—overpotential	85
14.5	Equilibrium exchange current density	87
14.6	Interpretation of the equilibrium potential	88
14.7	Nernst equation from the point of view of kinetics	89
14.8	Measurement of the exchange current density	91
14.9	Interpretation of the potential of zero charge	93
14.10	Final remarks on the baseness and nobility of metals	94
15.	Overpotential and the ButlerVolmer equation	95
16.	The Evans diagram and the corrosion potential	99
16.1	Corrosion	99
16.2	Redox equilibrium	100
16.3	Corrosion potential	103
16.4	Evans diagram	105
17.	Surface film and Pourbaix diagram	108
17.1	Passivation	108
17.2	Pourbaix diagram	109
18.	The evils of chloride	113
	Further reading	114

	Appendix—outline of thermodynamics	115
1	Closed systems	115
1.1	State functions	115
1.2	Conditions for equilibrium in closed systems	117
2.	Open systems, chemical potential	119
2.1	Conditions for equilibrium in open systems	123
2.2	Chemical potential in a closed system	124
3.	Activity	125
3.1	Ideal mixture	128
3.2	Non ideal mixture	129
3.3	Unimolar standard state	132
3.4	Chemical potential and activity of ionic species	133
3.5	Electrolyte chemical potential and activity	135
3.6	Measurement of mean ion activity coefficient	137
	Further reading	139

List of symbols

* denotes not measurable

I have tried as hard as possible to avoid multiple definitions of the same symbol, but in some cases such as h , V , n and γ this has been unavoidable. The table and the context should serve to avoid any ambiguity.

symbol	description	defining equation
ϵ_0	permittivity of free space, $1/4\pi\epsilon_0 = 9 \times 10^9 \text{ [N m}^2 \text{ C}^{-2}\text{]}$	
ϵ	permittivity, dielectric constant	
e	elementary charge, $+1.602 \times 10^{-19}$ Coulomb	
k	Boltzmann constant	
L	Avogadro constant	
R	gas constant = Boltzmann constant \times Avogadro constant	$R = 8.3 \text{ J/mole K}$
F	Faraday constant = elementary charge \times Avogadro constant	$F = 96500 \text{ C/mol}$
K	equilibrium constant	(9.1.3)
n	total number of moles, $\sum_i n_i$	
n	number of electrons involved in an electrochemical reaction	(9.3.6)
\mathbf{p}	dipole moment per unit area	(3.2.4) <i>et seq.</i>
ϕ, ϕ_m, ϕ_s	inner, or Galvani (electric) potential (in metal or solution)*	(3.2.3)
ψ, ψ_m, ψ_s	outer, or Volta (electric) potential (in metal or solution)	(5.1)
χ, χ_m, χ_s	dipole (electric) potential (in metal or solution)*	$\phi = \psi + \chi$ (5.1)
${}^A\Delta^B X$	$X_A - X_B$ for a potential X	(3.2.3) and footnote
\mathcal{E}	electromotance, e.m.f., $\mathcal{E} = {}^m\Delta^m \phi$	(9.3.5) and above
\mathcal{E}°	standard e.m.f. of an electrochemical cell	(9.5.3)
$\Delta_c \phi$	compensation potential (e.m.f.) of cell with air gap	(10.1.1) and page 12
q_m, q_s	charge on metal, electrolyte	section 5
μ	electron chemical potential relative to DFT zero	(3.1.1)
$\mu_e, \mu_e^{\text{Pt}} \dots$	Fermi level* ($\mu_e = \mu - \bar{V}$)	(3.1.2)
$\tilde{\mu}_e, \tilde{\mu}_e^{\text{Pt}} \dots$	electron electrochemical potential, $\tilde{\mu}_e = \mu_e - e\phi_m$	(3.4.4)
W	electron work function of a metal, $W = -\tilde{\mu}_e$ for an uncharged metal	(3.4.2)
W_s	electron work function of a solution	(6.1)
$W_i^{s,\circ}$	ion work function of species i in solution in standard state	(8.2.1)
G	free enthalpy (Gibbs function)	Appendix, section 1
ΔG°	standard free enthalpy change in a chemical reaction	(9.1.2)
$\Delta_{\text{sol}} G_i^\circ$	standard free enthalpy of solvation of species i	(8.2.5)
$\Delta_{\text{sol}}^r G_i^\circ$	standard real free enthalpy of solvation of species i	(8.2.7)
μ^\bullet	free enthalpy per mole of pure substance	(A3.4)
μ^∞	infinitely dilute standard chemical potential	(A3.7c) and page 63
μ^\square	unimolar standard chemical potential	(A3.12) and page 63
μ^\ominus	standard chemical potential of gas at one bar pressure	(A3.5)

symbol	description	defining equation
$\tilde{\mu}_i^{\text{g},\circ}$	electrochemical potential of charged species i in standard gas phase*	(8.2.1)
$\tilde{\mu}_i^{\text{s},\circ}$	electrochemical potential of charged species i in standard solution*	(8.2.1)
$\alpha_i^{\text{s},\circ}$	standard real potential of species i in solution	(8.2.3)
μ_i	chemical potential of species i	(A2.7)
μ_+, μ_-	ion chemical potential*	after (A3.22)
$\tilde{\mu}_i$	electrochemical potential,* $\tilde{\mu}_i = \mu_i + z_i F \phi_s$	(4.1.1)
α	real potential, $\alpha_i = \mu_i + z_i F \chi_s$	(4.2.1)
z_i	charge number of species i	(4.1.1)
z_+, z_-	charge number of cation and anion	above (A3.13)
a	activity, generally $a_i = \gamma_i(T, p, x_1, x_2, \dots, x_N) \times x_i$	(A3.1) and figure A2
a_i^\bullet	Raoultian activity of species i , $a_i^\bullet = \gamma_i^\bullet x_i$	(A3.6) and above
n_i	number of moles of species i (a mole is an Avogadro number of objects)	
x_i	concentration of species i as mole or atom fraction	(A3.9)
c_i	concentration of species i in moles per kg solvent	(A3.11)
m_i	<i>molality</i> : normalised concentration of species i (moles per kg solvent)	(A3.11)
m_\pm	geometric mean ion molality	(A3.20)
h^∞	infinitely dilute Henrian activity $h^\infty = x$	(A3.7d)
h^r	rational Henrian activity (referred to mole fraction, x)	(A3.8b)
h	unimolal, practical Henrian activity (referred to molality, m)	(A3.12a)
h	partial molar enthalpy (also v, f, g)	Appendix, section 2
h_+, h_-	'thermodynamic' practical activity of an ion*	(A3.23)
h_+, h_-	'working' practical activity of an ion	(A3.26)
h_\pm	geometric mean ion activity	(A3.24)
h_{el}	practical Henrian activity of electrolyte	(A3.22)
γ^\bullet	Raoultian activity coefficient (not constant, but $\gamma^\bullet \rightarrow 1$ as $x \rightarrow 1$)	(A3.7a)
γ^∞	Henry's law constant, $a = \gamma^\infty x$ for $x \rightarrow 0$	after (A3.7c)
γ^r	rational Henrian activity coefficient (not constant, but $\gamma^r \rightarrow \gamma^\infty$ as $x \rightarrow 0$)	(A3.8a)
γ	practical Henrian activity coefficient (not constant)	(A3.12b)
γ_+, γ_-	single ion activity coefficient*	(A3.16)
γ_{el}	practical Henrian activity coefficient of electrolyte	(A3.22)
γ_\pm	mean ion activity coefficient	(A3.19)
γ	surface tension, interfacial excess free enthalpy	(11.1.2)
Γ_i	interfacial excess of component i	(11.1.7a)
u, v, s, n	layer extensive quantities per unit area	page 47
\bar{V}	average Hartree plus external potential (energy) relative to DFT zero	figure 3–1
V	voltage, volume	
σ	surface charge density	(3.3.1), (11.2.4)
ρ	volume charge density	section 12.1
C	electrical capacitance	(3.3.1)
\mathcal{C}	(<i>differential</i>) capacity, capacitance per unit area, $\mathcal{C} = d\sigma/dV$	(11.2.5)

symbol	description	defining equation
ΔG^\ddagger	free enthalpy of activation	page 68
ΔH^\ddagger	enthalpy of activation	page 68
ΔS^\ddagger	entropy of activation	page 68
$\Delta G^{\circ\ddagger}$	standard free enthalpy of activation	(13.1.1)
$\Delta H^{\circ\ddagger}$	standard enthalpy of activation	(13.1.1)
$\Delta S^{\circ\ddagger}$	standard entropy of activation	(13.1.1)
ΔG_e^\ddagger	electronation free enthalpy of activation	page 82
ΔG_d^\ddagger	de-electronation free enthalpy of activation	page 83
$\Delta\phi_{eq}$	equilibrium interphase potential difference	(14.1.2)
$\Delta\phi_{eq}^\circ$	standard equilibrium interphase potential difference	(14.7.2)
k_1	first order rate coefficient	(13.1.1)
k_e	electronation rate coefficient	(14.7.1)
k_d	de-electronation rate coefficient	(14.7.1)
\bar{c}	areal concentration (moles per unit area)	page 82
\bar{a}_{kink}	activity of a metal surface kink site (mole m ⁻²)	page 83
r_e^0	zero-field rate of electronation (mole m ⁻² s ⁻¹)	page 82
r_d^0	zero-field rate of de-electronation (mole m ⁻² s ⁻¹)	page 84
r_e	rate of electronation (mole m ⁻² s ⁻¹)	(14.3.1)
r_d	rate of de-electronation (mole m ⁻² s ⁻¹)	(14.2.1)
ν_e	electronation frequency prefactor (s ⁻¹)	page 82
ν_d	de-electronation frequency prefactor (s ⁻¹)	page 83
i_e	electronation current density (amp m ⁻²)	page 82
i_d	de-electronation current density (amp m ⁻²)	page 88
i_e^{eq}	equilibrium electronation current density (amp m ⁻²)	(14.2.2)
i_d^{eq}	equilibrium de-electronation current density (amp m ⁻²)	(14.3.2)
i_0	= $i_e^{eq} = i_d^{eq}$: equilibrium exchange current density	page 88
i	Bulter–Volmer current density (amp m ⁻²)	(15.10.1)
\mathcal{E}_{cor}	corrosion potential (V)	page 104
i_{cor}	corrosion current density (amp m ⁻²)	page 105

1. The lemon lamp

We shall begin with the lemon lamp. I cut open a lemon and drive a zinc nail and a copper nail into the fruit without allowing them to touch. If I then use crocodile clips and wire to connect the two heads through a light bulb, the bulb will illuminate. See figure 1–1. I suggest you try this at home; you could use a galvanised nail and a brass screw.

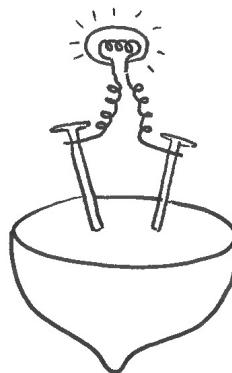


FIGURE 1–1

This is an electrochemical cell, as illustrated in figure 1–2. Zinc and copper metal “electrodes” are immersed in an acid solution; that is, an aqueous solution containing hydrogen ions (protons). A chemical reaction occurs spontaneously on the left hand electrode,



In this reaction a Zn atom at the surface of the electrode detaches itself into the solution, *but it leaves behind two electrons*. In this way it becomes a Zn^{++} “cation” free to wander in the liquid. Of course, this is corrosion of the zinc. The two electrons serve to reduce the charge carried by the electrode by $2e^-$ where, by convention, e is the fundamental constant—the charge on the proton, 1.602×10^{-19} Coulomb. Because of this spontaneous reaction this electrode is called the *electron sink*. Faraday called it the “anode”. In the terminology of Bockris, the zinc suffers “de-electronation”, also called “oxidation”. It is called that because its *oxidation state*—the number of plus signs in the superscript—is increased (from zero to two).

A second chemical reaction occurs spontaneously on the right hand electrode,



Two hydrogen ions in the acid solution swipe two electrons from the electron gas in the copper and combine to form hydrogen gas which bubbles off out of the liquid and into the air above. In this way this electrode, called the “cathode” by Faraday, is an *electron source* and the reaction is “electronation”, or “reduction” since the oxidation state of the H^+ ions are reduced from plus one to zero.[†]

[†] Note that you can easily tell an oxidation from a reduction reaction because in oxidation the electrons appear on the right of the arrow, while in reduction they appear on the left.

Because the two electrodes are connected by wire the electrons which are piling up on the zinc will repel each other because of the Coulomb force and will flow down the wire to the copper where they can serve the electronation reaction. In the process they can light up the bulb. Note that even if the wires are at open circuit so that electrons cannot flow, each reaction will proceed to some limited extent until the charge accumulating on the metals prevents any further reaction. A general conclusion is that if a metal is immersed in water containing some ions, the metal may become slightly charged.

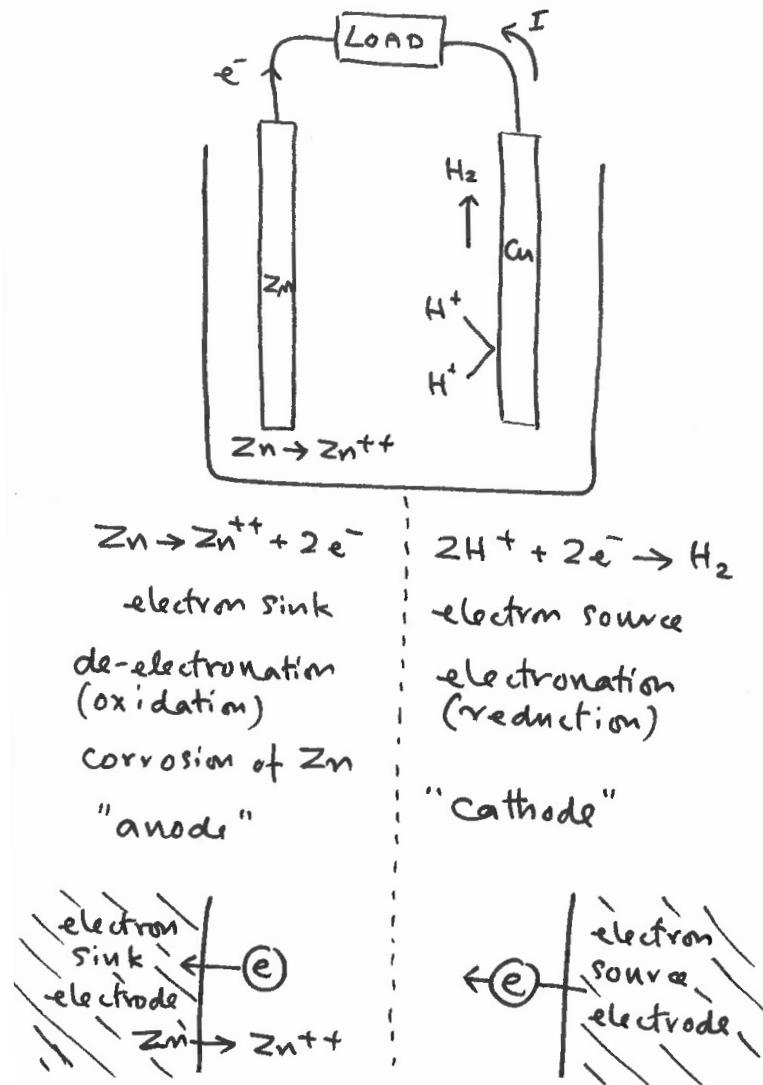


FIGURE 1-2

2. Simple electrochemical cells

2.1 The simplest cell

The simplest electrochemical cell is illustrated in figure 2-1.

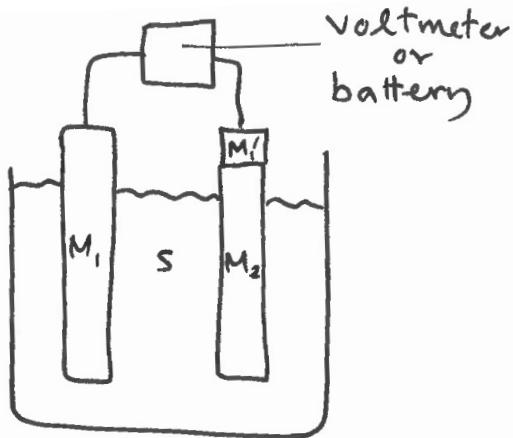


FIGURE 2-1

Two electrodes made of metal M_1 and metal M_2 are immersed in a solution, or *electrolyte*, S . They are connected so as to close the circuit through a battery or voltmeter. In the lemon lamp I glossed over the question of what metal the wires are made of. This is important all the same since the junction between dissimilar metals sets up an electric potential difference (see section 3.4) and this will contribute to the voltage at the voltmeter. In this simplest cell we take it that the wire is made of the same metal as one of the electrodes, say, M_1 . Because the wires on either side of the voltmeter (or battery) may be at a different electric potentials we indicate this in the cartoon by including a contact between wire of metal M'_1 and electrode of metal M_2 .

This cell can be illustrated more symbolically as in figure 2-2. This places the elements of the cell in line from left to right.

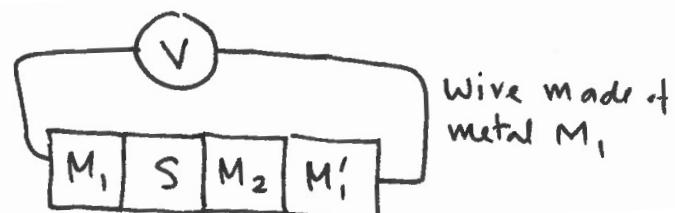
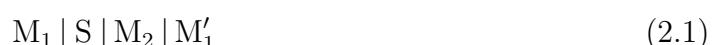


FIGURE 2-2

It is conventional to denote such cell in a formulaic way as follows, sometimes called the “cell diagram,”



The vertical bars indicate interfaces between the phases in contact. We assume that the electric potential is uniform within each phase but there may be an electric potential

difference across each interface. The convention is to put the anode to the left where the reaction is de-electronation and the cathode on the right where the reaction is electronation (reduction).

We anticipate by our experience with the lemon lamp that an electric potential difference (PD) will be set up between the terminals of the voltmeter. Since we assume that the electric potential is uniform within each phase then there may be an electric potential drop across each interface and the sum of these will make up the total PD across the voltmeter. This effect can be summarised in a *third* cartoon representation of the cell, namely the ladder diagram of figure 2-3.

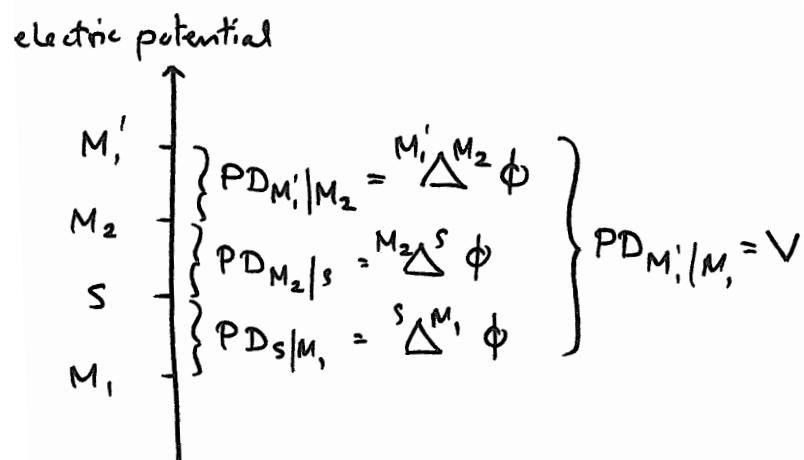


FIGURE 2-3

We concentrate for the present on the case where the circuit is closed through a high impedance voltmeter as we are interested in the electric potential difference that is set up by creating the sequence of interfaces. But keep in the back of your mind the thought that a battery in place of the voltmeter can be used to modify the PD. For example if the zinc electrode is made more positive it will drive the reaction (1.1) to the right and increase the tendency of the zinc to corrode. Hence the metallurgical dictum: “anodes are positive and they corrode.”

It is very important to appreciate what can and what cannot be measured. Clearly we can measure the total PD across the voltmeter; but thus far we have not established that we can measure any one of the electric potential drops across an interface. Generally we cannot because any such measurement necessarily introduces new electrodes and probes and new PDs that conspire to add to the one we’re trying to measure. As an exact statement we may assert that *it is not possible to measure a single electrode potential.*

2.2 The next simplest electrochemical cell

Immediately I have a counter response to that assertion. Figure 2–4 shows the *next* simplest electrochemical cell.

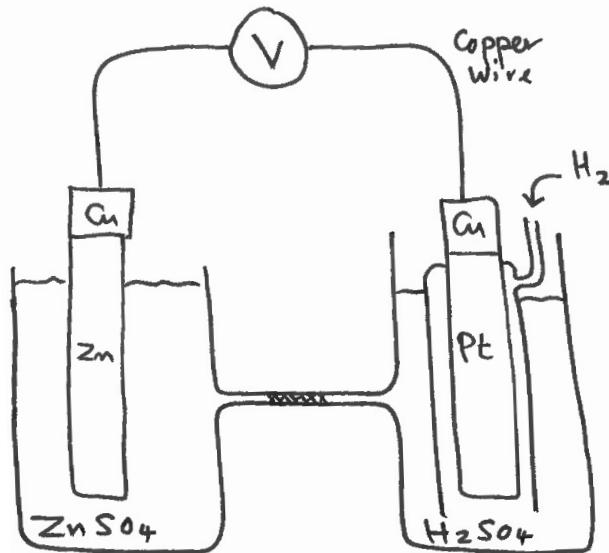
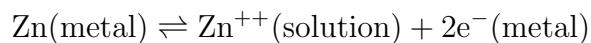


FIGURE 2–4

In this figure the single electrodes are separated by a narrow tube stuffed with some salt that allows ions to flow and complete the circuit but enables each electrode to be immersed in a different electrolyte. On the left is a Zn electrode immersed in a solution of zinc sulphate and so there can be an equilibrium of the reaction



The other cell is the famous “standard hydrogen electrode” (SHE) which consists of a platinum wire surrounded by hydrogen gas at exactly one bar pressure immersed in a solution of sulphuric acid at such a strength that the hydrogen ion activity (Appendix) is exactly one. The zinc and platinum electrodes are connected through copper wires to a voltmeter or battery. Figure 2–5 shows the cell in the in-line cartoon form.

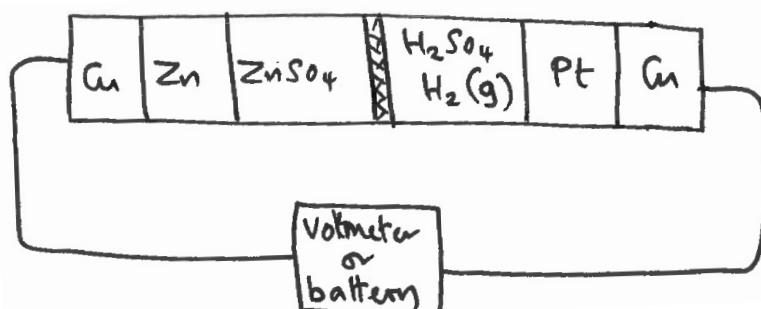


FIGURE 2–5

This set up does permit the measurement of a single electrode potential because the SHE is a so-called “non polarisable”, *reversible* single electrode. This means that its

single electrode potential is both known and unchanging as a result of it being connected to a second single electrode to form an electrochemical cell. A single electrode can be modelled in a first approximation as a capacitor and resistor connected in parallel, see figure 2–6. If the resistance is zero then the electrode is said to be “non polarisable” so that if we try to change the electric potential on the electrode, current will leak out into the electrolyte and we find that the electric potential cannot be altered. Conversely, if the resistance is infinite then the interface acts as a capacitor and is “ideally polarisable” (see section 12) and changes in electric potential are accommodated by charge accumulation across the interface. Therefore since all the electric potential drops are known except one, that at the Zn|S interface; then by varying the concentration of zinc sulphate and observing the voltmeter the single electrode potential may be inferred. This is only possible if the Zn|S interface is ideally polarisable.

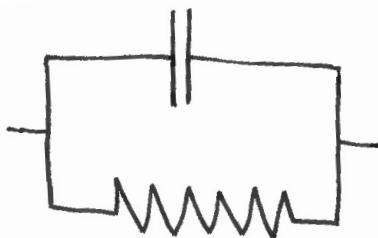


FIGURE 2–6

3. Inner potential, work function and contact potential

The surface of a metal has features absent in the bulk; in particular, since the electrons spill out into the vacuum there will exist an electric dipole pointing normal to the surface from the vacuum into the metal. Similarly for pure water: it has been established experimentally that water molecules at the surface are oriented with their oxygen atoms pointing into the vacuum; hence as for a metal surface, there is an electric dipole that points from the vacuum towards the water; the dipole potential at the surface of water is estimated from experiment to be 80mV (see section 10.1). If a metal is immersed in an electrolyte it may become charged. The liquid close to the metal surface is bound to have a molecular structure that is different from the bulk. Not only are the water molecules differently oriented near the surface anyway, the additional charge will further modify the structure into what is called, loosely, the *double layer*. Pictures of the double layer are associated with the authors Helmholtz, Gouy and Chapman, and Stern (see Bockris, and sections 11 and 12 below). A common feature of the surfaces of both the metal and the electrolyte are the appearance of surface *electric dipoles*. After the interface between a metal and an electrolyte is established these dipoles will persist and the narrow region spanning the interface on either side is often referred to in the literature as the “interphase.” The purpose of the following few sections is to dissect the interphases such as indicated by a vertical bar in (2.1) in a semi-quantitative manner.

Now there needs to be some remark about the word “potential” which has two meanings in all the books and literature. I have carefully used “electric potential” where convenient to describe just that: a quantity calibrated in Volts or Joules per Coulomb in SI units.

Authors persistently use “potential” when they mean potential energy whose SI units are Joules. It is not really possible to write either “electric potential” or “potential energy” all the time without becoming prolix. And indeed quantities called chemical potential, or inner potential are so common that we must use them—although the former is in Joules and latter in Volts. The most egregious villains are in density functional theory: very confusing is Lang and Kohn on the work function, whose ϕ is called a potential and has the usual symbol of electric potential, but actually is a potential energy. At best I can warn you to be careful and hope that ambiguity doesn’t creep into these notes. Where possible, we will use lower case Greek, ϕ , ψ and χ for electric potentials and U for potential energy. But we will use V for voltage. For example if the electric potential somewhere is ϕ , then the potential energy of a charged object at the same place is $q\phi$ where q is the charge of the object. If the object is an electron then the potential energy is $-e\phi$ (measured in electron volts, eV). Finally recall that *all* instances of either electric potential or potential energy *must* be referenced to some arbitrary and agreed zero (see section 10.2): it is always safest to mention potential differences rather than absolute potentials. For example in electrostatics we often take the electric potential as relative to a point at infinity out of range of all charges. In solid state physics if we treat only a bulk solid with no surfaces but periodic boundary conditions, then we choose a potential energy zero with reference to some mean energy within the solid (for example the average Hartree plus external potential energies).

3.1 Density functional theory

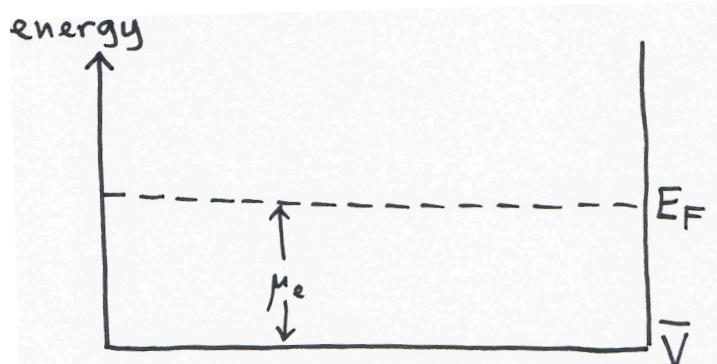


FIGURE 3–1

Figure 3–1 shows a sketch of the energy levels in a metal. The lower horizontal line indicates the average potential energy, $\bar{V} = -e\bar{\phi}$, of all the electrons in the metal. Because electrons are Fermions they occupy successively higher energy levels such that the energy of the highest occupied state is the Fermi energy, E_F . The bulk of the metal, far from its surfaces, must have its electronic structure exactly as described using periodic boundary conditions and as long as the metal is not charged we may describe the electronic levels using bulk density functional theory (DFT). A postulate of the DFT is that the total energy is formed of three terms, each a functional only of the ground state electron density, $n(\mathbf{r})$ (number of electrons per unit volume),

$$E_{\text{tot}}[n] = T[n] + V_{\text{en}}[n] + V_{\text{ee}}[n]$$

which are the expectation values of the the kinetic energy, electron–nucleus and electron–

electron operators. The number of electrons is

$$N = \int n(\mathbf{r}) d\mathbf{r}$$

Often we pull out the universal functional

$$F[n] = \left\langle \Psi \left| \hat{T} + \hat{V}_{ee} \right| \Psi \right\rangle$$

(Ψ is the ground state many-electron wavefunction) and write

$$E_{\text{tot}} = F + V_{\text{en}} = F + \int n(\mathbf{r}) v_{\text{ext}} d\mathbf{r}$$

where v_{ext} is the external potential (*i.e.*, that due to the nuclei). Note that v_{ext} is a potential energy, not an electric potential. In addition, we pull out from F the classical Coulomb self energy of the inhomogeneous electron gas, namely the Hartree energy,

$$F = E_{\text{H}} + G$$

which defines the universal functional, G , which contains all the contributions to the total energy from exchange and correlation, and the interacting part of the kinetic energy. The second Hohenberg–Kohn principle states that the ground state density is that which minimises the total energy: this requires a constrained minimisation,

$$\frac{\delta}{\delta n} \left\{ E_{\text{tot}}[n] - \mu \left(\int n(\mathbf{r}) d\mathbf{r} - N \right) \right\} = 0$$

in which μ is a Lagrange multiplier. This leads to the Euler–Lagrange equation

$$\begin{aligned} \frac{\delta E_{\text{tot}}}{\delta n} &= v_{\text{ext}} + V_{\text{H}} + \frac{\delta G}{\delta n} \\ &= \mu \end{aligned}$$

where, the $\delta/\delta n$ being a functional derivative with respect to the function $n(\mathbf{r})$,

$$V_{\text{H}} = \frac{\delta E_{\text{H}}}{\delta n}$$

is the “Hartree potential” (again, actually a potential energy) and

$$\begin{aligned} \mu &= v_{\text{ext}} + V_{\text{H}} + \frac{\delta G}{\delta n} \\ &= \left(\frac{\partial E_{\text{tot}}}{\partial N} \right)_{v_{\text{ext}}} \end{aligned} \tag{3.1.1}$$

is the zero temperature chemical potential, or variation of the total energy with respect to an infinitesimal change in electron number at fixed external potential.

In DFT the zero of energy is the energy of all the electrons and nuclei infinitely separated (the nuclei are not disassembled) so E_{tot} and μ are both referred to the DFT energy zero. On the other hand, if we make the choice that the zero of energy is the average potential energy of the electrons in the metal then we have

$$\bar{V} = \bar{v}_{\text{ext}} + \bar{V}_H = 0$$

and we may define

$$\mu_e = \mu - \bar{V} = \frac{\delta G}{\delta n} \quad (3.1.2)$$

as the “chemical potential of the electron in the metal.” The overbar denotes a spatial average, say, over a unit cell of the metal crystal. We shall adhere to the convention that we call μ_e the “chemical potential” and that it is always measured from the “mean inner potential”, \bar{V} .

3.2 Inner potential difference

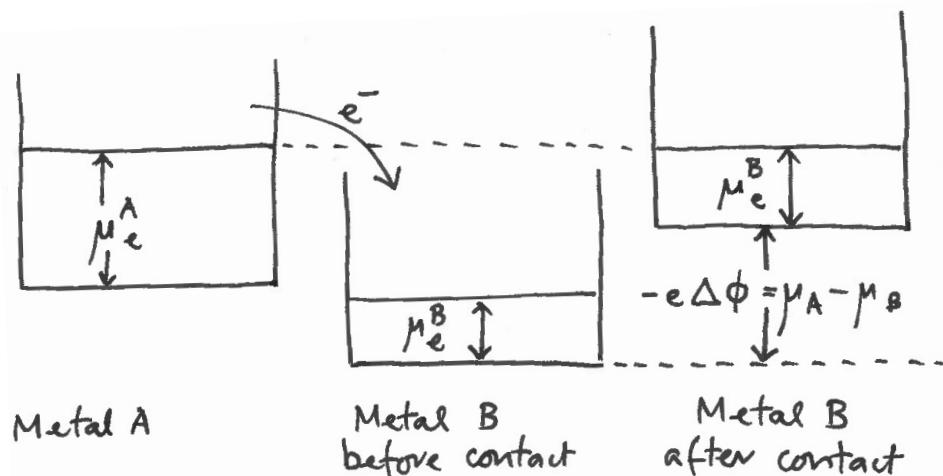


FIGURE 3-2

Consider two metals, A and B (figure 3-2). The Fermi energies of electrons in metals A and B will not be the same; therefore if I put specimens of the two metals into electrical contact, as for example at the $M_2|M'_1$ interface in (2.1) (or if I connect them with a wire) then if, say, $\mu_A > \mu_B$, relative to the DFT zero of energy, electrons will tend to flow from metal A to metal B. However in doing so, each metal will become charged, resulting in a shift in the potential energy, \bar{V} , of the electrons, and this will introduce an electric potential difference, $\Delta\phi = \phi_B - \phi_A$ between interior points in the two metals. As the electric potential in metal A becomes increasingly positive and in metal B negative then transfer will be resisted until the Fermi levels align by virtue of the shift in the potential energies in each metal. This is illustrated in figure 3-2. It is clear by construction that energy levels are shifted so that $-e\Delta\phi = \mu_e^A - \mu_e^B$. We therefore have, in equilibrium,[†]

$$\mu_e^A - e\phi_A = \mu_e^B - e\phi_B \quad (3.2.1)$$

[†] Suppose N electrons are transferred from metal A to metal B. If we refer the total energy to the average potential energy, \bar{V} , in each metal and hence define $\bar{E}(N) = E_{\text{tot}} - \bar{V}$ as

where $-e$ is the charge on the electron. The electric potential across the metal junction is sketched in figure 3–3. Because a metal is a conductor, all the excess charge resides at the interface.

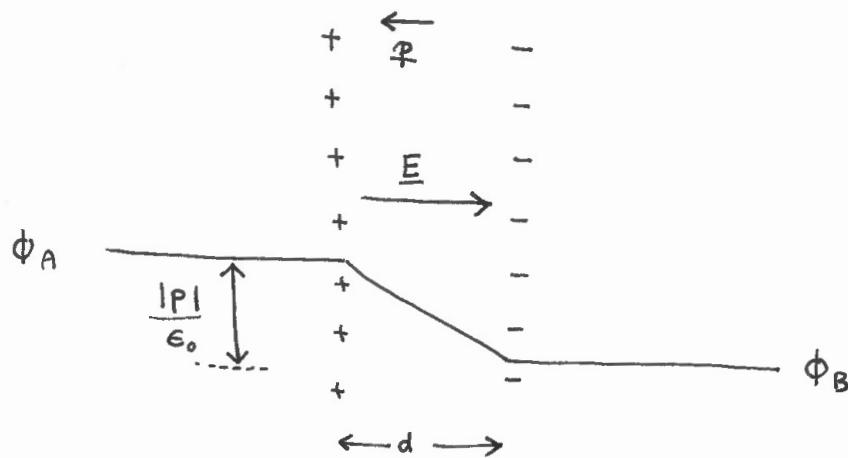


FIGURE 3–3

The charge averaged over the area of the interface forms a one dimensional distribution of charge density, $\delta\rho(x)$, if x is the cartesian component normal to the interface. This generates a dipole moment per unit area, \mathbf{p} . If we treat the double layer as a capacitor, as illustrated in figure 3–3, then the electric field is zero to either side and is constant and equal to σ/ϵ_0 in the gap of width d where σ is the charge per unit area. Therefore the electric potential drop acrosss the interface is

$$\phi_B - \phi_A = -\frac{1}{\epsilon_0} \sigma d = \frac{1}{\epsilon_0} p \quad (3.2.2)$$

in SI units. (Note that p is negative since \mathbf{p} points in the negative x -direction.) This is the “inner potential difference”, or “Galvani potential difference”. Following conventional, if cumbersome, notation we must write this as[†]

$$\phi_B - \phi_A = {}^B\Delta^A\phi \quad (3.2.3)$$

the total energy once N electrons have been added to the metal, then the total energy of the assembly is $\bar{E}_A(-N) + \bar{E}_B(N)$ plus the energy stored in the capacitor that is created by the charge dipole layer at the interface—figure 3–3. Then if C is the capacitance,

$$\begin{aligned} \bar{E}(N) &= \bar{E}_A(-N) + \bar{E}_B(N) + \frac{1}{2} \frac{e^2 N^2}{C} \\ &= \bar{E}_A(0) - N \frac{\partial \bar{E}_A}{\partial N} + \bar{E}_B(0) + N \frac{\partial \bar{E}_B}{\partial N} + \frac{1}{2} \frac{e^2 N^2}{C} \\ &= \bar{E}_A(0) - N \mu_e^A + \bar{E}_B(0) + N \mu_e^B + \frac{1}{2} \frac{e^2 N^2}{C} \end{aligned}$$

If this is minimised with respect to N we obtain (3.2.1) after noting that by comparison with a flat plate capacitor, $-eN/C = \phi_B - \phi_A$.

[†] Not only is it ugly, it is awkward and confusing; but it can be useful to shorten formulas

The result (3.2.2) is much more general than it seems. As long as the one dimensional charge density, $\delta\rho(x)$, is zero on both sides far from the interface and that it integrates to zero between these limits so that there is a dipole but the interface is not charged, then the work done in carrying an electron from left to right across the dipole layer is

$$W_{\text{dipole}} = -e(\phi_B - \phi_A) = -\frac{1}{\epsilon_0} ep \quad (3.2.4)$$

Briefly, the reasoning is this. The perturbation in charge density, $\delta\rho$, that creates the dipole layer is zero far to the left and far to the right of the interface and so is zero for all $x < a$ and $x > b$ where $x = a$ and $x = b$ are points in the interior of metals A and B respectively. The difference in potential $\phi_B - \phi_A$ is the work done in taking a unit test charge from $x = a$ to $x = b$ and this is minus the integral of the electric field, $E(x)$, between $x = a$ and $x = b$,

$$\begin{aligned} \phi(b) - \phi(a) &= - \int_a^b E dx = - \int_a^b E \left(\frac{d}{dx} x \right) dx \\ &= \int_a^b x \frac{dE}{dx} dx + [Ex]_a^b \\ &= \frac{1}{\epsilon_0} \int_a^b x \delta\rho(x) dx \\ &= \frac{1}{\epsilon_0} p \end{aligned}$$

We inserted a “one” in the first line and integrated by parts with the boundary term vanishing since there is no electric field in the interior of the metal. The third line follows from Poisson’s equation,

$$\frac{dE}{dx} = -\frac{d^2\phi}{dx^2} = \frac{1}{\epsilon_0} \delta\rho(x)$$

and the integral in the third line is by definition the dipole moment of the charge distribution.

It must be recalled that $\phi_B - \phi_A$ *cannot be measured*. It may have been sensible to call this the “contact potential”, and indeed Bockris does so, but this phrase is reserved for a subtly different electric potential difference between two metals in contact which *is* measurable and is the subject of the next section.

3.3 Contact potential

There is no doubt that the electric potential at a point “just outside” the surface of a metal, say 10^{-5} cm distant, is not the same as the electric potential at infinity (out of

when we add lots of PDs together. I am following Bockris and John West in putting the superscripts in the same order on both sides of (3.2.3). Fawcett, and Cheng and Sprik place them the other way around. This notation is probably best avoided: Schmickler and Santos don’t use it.

reach of all charges and fields). The reason for this is that as electrons spill out of the surface of the metal into the vacuum they create a charge dipole layer exactly as envisaged in figure 3–3 if metal B is replaced by vacuum. The details of the rearrangement of charge, both at the atomic level and as averaged over thousands of surface atoms depends on the indices of the crystal face and at the macroscopic level on the state of the surface—its roughness, contamination by adsorbates, oxidation products and so on. So each actual surface of a metal specimen will have a different distribution of polarisation (bound) charge which will give rise to stray electric fields outside the metal. These particularly arise at the corners of a single crystal or as “patch fields” which are external electric field lines connecting the surfaces of different grains in a polycrystalline sample. The existence of this electric potential is proved beyond doubt by experiments first made by Kelvin. The simplest realisation is to place the faces of two metals, or different faces of two single crystals of the same metal, into the form of a capacitor so that the two plane surfaces are parallel and separated by a distance d . If there exists an electric potential difference $\Delta\psi$ between points close to each surface then there must be an electric field, E , and surface charge densities $\pm\sigma$ such that

$$\Delta\psi = Ed = \frac{\epsilon_0\sigma}{d} \quad ; \quad C = \frac{A\epsilon_0}{d} \quad (3.3.1)$$

by analogy with a parallel plate capacitor of area A and capacitance C . If the surfaces are moved further apart or closer together such that d changes but $\Delta\psi$ is to remain the same then the surface charge density must change and this will lead to a measurable current if the two metals are connected through an ammeter. This is indeed observed; and if in addition to the ammeter a battery and potentiostat are introduced into the circuit then a bias potential may be imposed. When this bias potential is exactly equal and opposite to $\Delta\psi$ then no current will flow when d is altered and this measured bias potential provides a *measurement* of $\Delta\psi$. This is the Kelvin probe. The bias potential that exactly opposes the contact potential is called the *compensation potential*. For two metals, A and B, the measured difference in outer electric potential, $\psi_B - \psi_A$ is called the *contact potential* difference. It is not the same as the inner potential difference (3.2.3) which is not measurable but which does contribute to the total PD measured across an electrochemical cell due to the use of different metal electrodes and electrical connectors.

3.4 Work function

The Kelvin probe can also be used to measure the work function (see Fawcett, section 8.7 for details of how this is done). It is easy to get confused by what is meant by work function and so here I will attempt to stay well grounded in either experiment or the density functional theory. However the two approaches actually give slightly different definitions of work function. Loosely stated the work function of a metal is the amount of work needed to remove an electron from the metal (as observed for example in thermionic emission or the photoelectric effect). Naturally an electron that has the highest (kinetic) energy will be the easiest to remove so the electron will originally be at the Fermi surface.

- Kittel defines the work function as “...the difference in potential energy of an electron between the vacuum level and the Fermi level”. The vacuum level, according to

Kittel is “...the energy of the electron at rest at a point sufficiently far outside the surface so that the electrostatic image force on the electron may be neglected—more than 100Å [10⁻⁶ cm] from the surface.” Different crystal surfaces have different work functions. For example tungsten has these work functions.

surface	work function (eV)
(100)	4.63
(110)	5.25
(111)	4.47

This table reflects the fact stated earlier that the electric potential is not the same just outside different faces of the same metal.

2. The definition of work function in DFT is,

$$W = [V(\infty) + E_{\text{tot}}(N - 1)] - E_{\text{tot}}(N) \quad (3.4.1)$$

In words: this is the difference in total energy between a piece of metal containing N electrons, and the same piece containing $N - 1$ electrons plus the potential energy of an electron at infinity. This definition can be made consistent with the first by replacing $V(\infty)$ with $-e\psi$ where ψ is the “outer electric potential” associated with the metal surface.

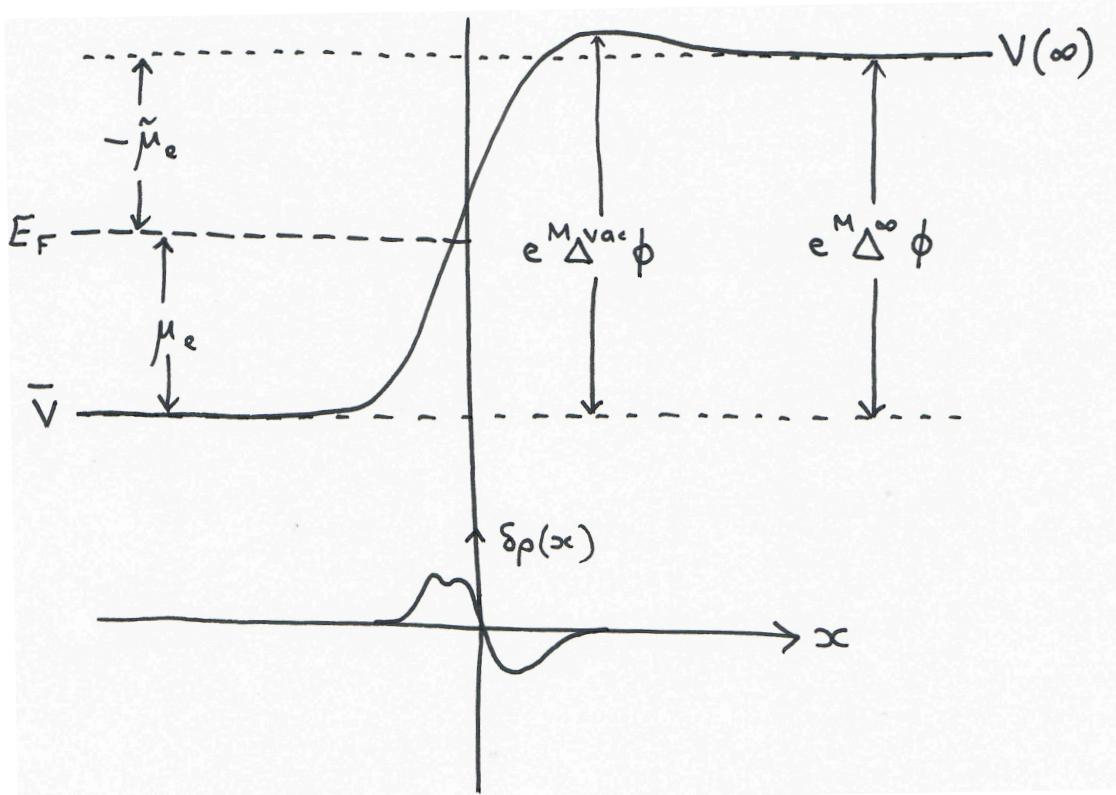


FIGURE 3-4

A way to picture the description embodied in (3.4.1) is to modify figure 3-1 to include the presence of a surface and vacuum.[†] Figure 3-4 shows the metal to the left with the

[†] By “vacuum” we really mean gas phase. In fact in the vacuum outside a metal there is

quantities \bar{V} , E_F and μ_e as in figure 3–1. To the right is the vacuum, and the potential energy of an electron is now sketched as increasing from left to right up to $V(\infty)$ as $x \rightarrow \infty$. Lang and Kohn show that without approximation in DFT, (3.4.1) is formally and exactly equal to

$$W = V(\infty) - \mu \quad (3.4.2)$$

in which μ is given by (3.1.1). Lang and Kohn then add and subtract \bar{V} and arrive at the formula

$$W = \underbrace{(V(\infty) - \bar{V})}_{\text{dipole potential}} - \underbrace{(\mu - \bar{V})}_{\mu_e} = e^m \Delta^\infty \phi - \mu_e \quad (3.4.3)$$

where ${}^m \Delta^\infty \phi = \phi_m - \phi_\infty$ is the difference in electric potential between the interior of the metal and infinity, and μ_e is the chemical potential as defined in (3.1.2). Note that ${}^m \Delta^\infty \phi$ is *positive* because the electric potential is more positive in the metal than outside (that's why the electrons are trapped inside). The first term in (3.4.3) is $-e {}^\infty \Delta^m \phi = e^m \Delta^\infty \phi > 0$.

To make contact with experimental measurement and to allow for variations of work function with crystal orientation of the surface we may replace ${}^m \Delta^\infty \phi$ with

$${}^m \Delta^{\text{vac}} \phi = \phi - \psi = \chi > 0$$

where ψ is the electric potential at the vacuum level and ϕ is the electric potential in the interior of the metal. Both ϕ and ψ are relative to the electric potential at infinity. The “dipole (electric) potential” is (3.2.2)

$$\chi = -\frac{p}{\epsilon_0} \quad (p < 0)$$

and $q\chi$ is the work done in taking a test charge, q , from the vacuum level to the interior of the metal through the dipole layer of strength p per unit area which arises from the perturbation of the charge density from its bulk distribution due to the creation of the surface (figure 3–3). This difference in density is sketched in figure 3–4.

Finally, note that if $\phi \equiv {}^m \Delta^\infty \phi$ is the electric potential in the metal relative to infinity, then it is clear from figure 3–4 that we may define a quantity by construction

$$-\tilde{\mu}_e + \mu_e = e\phi$$

from which we define the “electrochemical potential” of the electron as

$$\tilde{\mu}_e = \mu_e - e\phi \quad (3.4.4)$$

It is then clear that since from (3.4.3) $W = e\phi - \mu_e$, the work function is equal to minus the electrochemical potential of the electron,

$$W = -\tilde{\mu}_e$$

a finite probability of finding an electron according to the Fermi–Dirac distribution. If there were not then there would be no possibility of equilibrium between the condensed and the gas phase.

It is also evident that (3.2.1) states that in equilibrium the electrochemical potentials of electrons in each metal are equal, which is the usual statement of equilibrium. It is very important to note that μ , μ_e and $\tilde{\mu}_e$ are *not* different quantities. They are the same quantity but referred to different zeros of energy. This is analogous to chemical potentials in chemical thermodynamics which are only defined with respect to some standard state.

3.5 Contact potential again

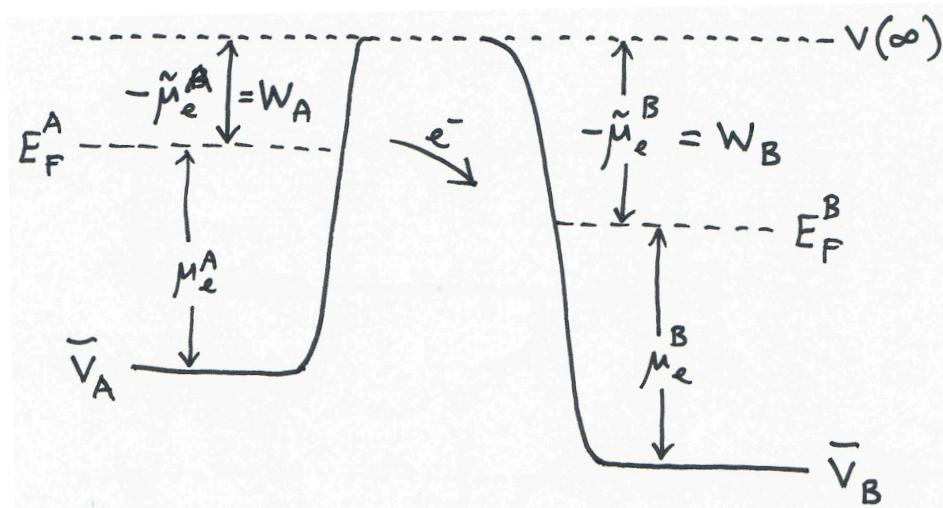


FIGURE 3-5

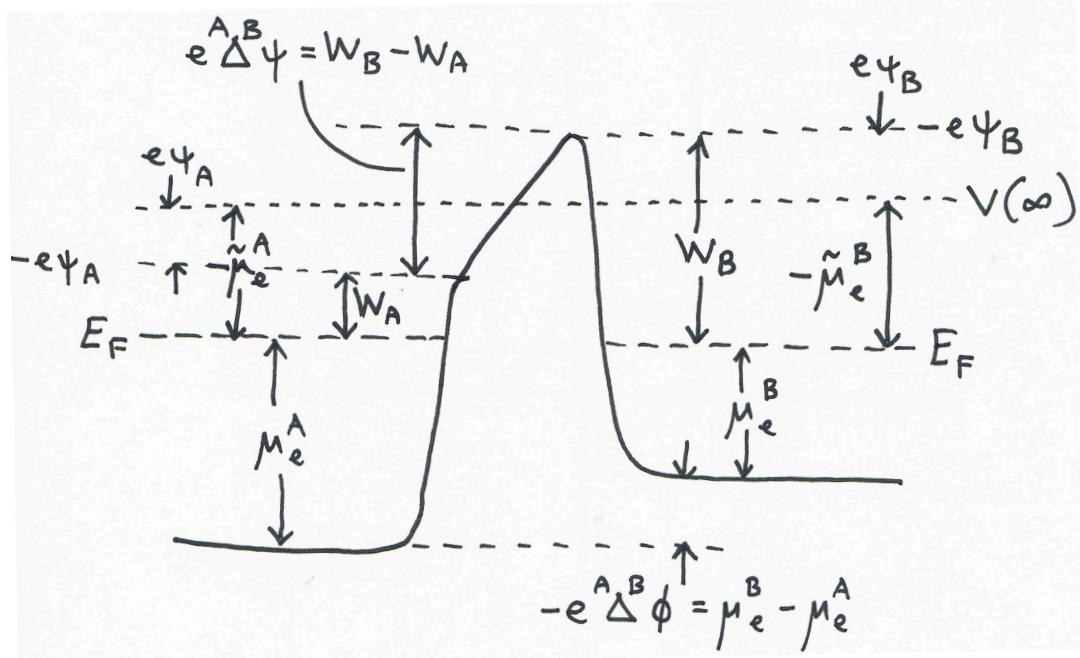


FIGURE 3-6

We can now return to the experiment in which two metals are electrically connected. Figure 3-5 shows energy diagrams for metals A and B before they are connected. These are two diagrams as figure 3-4 with energy levels, $E_F = \mu$ and \bar{V} for each metal referred to the DFT zero of energy. Because the Fermi energy of metal A is higher than that

of metal B we expect that when they are connected electrically by a wire, electrons will flow so as to equalise the Fermi energies. This was already demonstrated to be the equilibrium condition in section 3.2. The condition for equilibrium is $\tilde{\mu}_e^A = \tilde{\mu}_e^B$ and this is demonstrated in Figure 3–6, adapted from Cusack, which is a cartoon of the energy levels once the metals are electrically in contact. The solid line shows $-e$ times the electric potential in the two metals and in the vacuum separating them. The slope in the electric potential is exactly the electric field that exists in the gap between two metal surfaces in Kelvin’s experiment. Because of the transfer of electrons both metals are charged, and so the electric potential in each metal is uniformly shifted. Metal A is positively charged and so its electric potential is raised and consequently its electronic energy levels are lowered. (It’s a real nuisance that electrons are negatively charged—should we blame Faraday? On the other hand it does make you think—figures 3–4 and 3–5 would be “upside down” if we were plotting electric potential, or if you like, electrons flow uphill. The same issue crops up in semiconductor band diagrams in textbooks.)

It is clear now that the work function is *not* minus the electrochemical potential if the metal is charged. Instead as seen from figure 3–6, by construction, we have the more general formula

$$W = -e\psi - \tilde{\mu}_e \quad (3.4.5)$$

where ψ is the electric potential “just outside” the metal.

The difference between the electric potentials just outside the two metals is $\psi_B - \psi_A$ and is called the contact potential; either by construction or by subtracting (3.4.5) for two metals and noting that their electrochemical potentials are equal when in contact, we see that

$$W_B - W_A = -e(\psi_B - \psi_A)$$

In words: *the difference in “outer potential” of two electrically connected metals is equal to $-1/e$ times the difference in work function*. We also see that the difference in inner potentials is $+e$ times the difference in chemical potentials, which is consistent with (3.2.1).

The following table shows some Fermi energies, μ_e , (taken from Moruzzi, Janak and Williams) and work functions (taken from Ashcroft and Mermin). All are in eV; the inner potential difference in Volts is just the difference in Fermi levels, if two of the metals are in contact. The difference in outer potential in Volts is the difference in work functions.

	Li	Mg	Al	Ag	Cu	Zn	Fe	Ti
μ_e/eV	4.4	5.9	8.4	6.5	8.6	8.3	11.3	8.7
W/eV	2.4	3.6	4.3	4.3	4.4	4.3	4.3	4.3

Interestingly, as long as we know μ_e and W for two metals, we can infer both potential differences, ${}^A\Delta^B\phi$ and ${}^A\Delta^B\psi$, which appears in contradiction of the observation that only ${}^A\Delta^B\psi$ is *measurable*. The point is, I suppose, that μ_e is calculated, not measured.

4. Electrochemical potential; real potential

4.1 Electrochemical potential

The meaning of chemical potential is explained in some detail in the Appendix to these notes. At constant temperature and pressure, it is the infinitesimal change in free enthalpy with respect to the mole number of species i keeping the mole numbers of all other species unchanged,

$$\mu_i = \left(\frac{\partial G}{\partial n_i} \right)_{T,p,n_j \neq i} \quad (\text{A2.7})$$

At constant temperature and volume, replace G with F , the free energy. Since G and F are always expressed with respect to some arbitrary zero of energy, the chemical potential is similarly undetermined to within a constant. This is usually expressed in terms of a standard chemical potential. For example, if you imagine taking a single atom or molecule from some reservoir and adding it to the body in question, then you may take the reservoir to be gas phase at standard temperature and pressure or pure substance; these two choices involve different amounts of reversible work and so the chemical potentials referred to these two standard states will not be the same. Just as in electrostatics, thermodynamics is about *changes* in energy; it is only necessary for consistency to keep to the same zero of energy when comparing energies. In electrolytes and solutions the chemical potential refers to a complicated process. Adding a molecule or atom to a solution involves much complex rearrangement of its environment as the host solvent molecules and preexisting solute molecules rearrange themselves to accommodate the new species. Indeed the work done will depend on the concentration of the species in question already present if the solute molecules are interacting. For example, we speak of the *free enthalpy of solvation* which includes the reversible work done by the solvent as its molecules create the so-called solvation shell (at constant T and p). Again, one defines a standard free enthalpy of solvation, for example at standard temperature and pressure and at unit activity (see the Appendix, section 3) of the solute.

If the species in question is *charged* then a polar solvent such as water will reorganise its molecular dipole moments the best to screen the electric field introduced by the charged species. Equation (A2.7) still defines the chemical potential but it is conventional to separate out from the chemical potential that part which is a result of the chemistry alone—solvation, effects of the concentration, temperature and pressure—and that part of the work done due to having to place the charge of the species i (namely $q_i = ez_i$ where z_i is called the “charge number” of the species) into the interior of a phase whose inner potential, ϕ , is different to that of the reservoir. If we agree that the species is at zero potential with respect to infinity in the reservoir, then we define the electrochemical potential as

$$\tilde{\mu}_i = \mu_i + Fz_i\phi \quad (4.1.1)$$

This introduces the *Faraday constant*, F , which is the charge on a mole of protons. It is $+e$ times the Avogadro number. You may ask, how can I redefine the chemical potential (A2.7) so as to somehow exclude all effects due to the charge on the species i ? After all the structure of the solvation shell will depend on the charge. This means that the new μ_i in (4.1.1) appears to be ill-determined. The only recourse is to recognise

that $\tilde{\mu}_i$ in (4.1.1) is the same quantity as defined rigorously in (A2.7) and so the new μ_i in (4.1.1) is then uniquely determined by the operation of subtracting $Fz_i\phi$ from the chemical potential. I'm sorry—this is a horrible clash of notation: I have used μ_i to mean two different things in the same section. It shan't happen again... What I should have done is either written that (A2.7) is the chemical potential for an *uncharged* species, or to have used $\tilde{\mu}_i$ in the left hand side. To my mind, as in the case of the electron, μ and $\tilde{\mu}$ are anyway the same quantity, but referred to different zeros of energy.

4.2 Real potential

In the electrochemical literature it is common to introduce an additional potential energy, which is related to the work function and is useful when work functions are defined for chemical species. We start by rewriting (4.1.1)

$$\tilde{\mu}_i = \mu_i + Fz_i(\psi + \chi)$$

(see (5.1) in the following section). We have to ignore the subtle differences in potential between infinity, and the vacuum level where the electric potential is ψ (figure 3–4). Then if the phase is uncharged, we assert that $\psi = 0$ and give a new name to the electrochemical potential in those circumstances: the *real potential* of species i ,

$$\alpha_i = \mu_i + Fz_i\chi \tag{4.2.1}$$

The point is that, according to Sprik, this is a measurable quantity; although as I understand it only measurable after making an extra-thermodynamic assumption (see ahead, section 10). In our definition number 1 of the work function in section 3.4, it should be clear that the work function of a metal is minus the real potential of an electron in the metal.

5. Outer electric potential

In electrochemistry the electrode and the solution are treated in their continuum electrostatics quite similarly. Certainly both are assumed perfect conductors in which the electric potential is uniform up to the surface or interphase and there is no electric field unless a current is flowing. For example we later will see that a work function can also be associated with the electrolyte; but this is a more involved quantity, firstly because the object withdrawn to the vacuum level is not usually an electron—it's an ion—and secondly because the work done depends on the state of solvation and the concentration of the ion in solution.

The inner electric potential difference (3.2.3) is called the difference in “Galvani potential”, always given the symbol ϕ . Whereas under some limiting circumstances as we have seen the Galvani potential difference can be measured (for example, for a polarisable single electrode connected to a reversible reference electrode) in principle the absolute Galvani potential within the bulk of an electrode or electrolyte cannot be measured. What *can* be measured is the so called outer electric potential, or “Volta potential” as

we have seen in the Kelvin probe experiment. And so it is common to split up the Galvani potential into two terms, the Volta potential, ψ , and the remainder called the dipole potential, χ . The latter is obviously not measurable. We have

$$\phi = \psi + \chi \quad (5.1)$$

Or, at an interface, A|B, (see equation (3.2.3) and footnote on page 10)

$${}^B\Delta^A\phi = {}^B\Delta^A\psi + {}^B\Delta^A\chi \quad (5.2)$$

The Volta potential difference, $\psi_B - \psi_A$, according to Bockris is “the contribution to the potential difference across an electrified interface arising from the charges on the two phases.” Imagine the process of taking a positive test charge from infinity where $\phi = 0$ and carrying it through the surface to the inside of an uncharged phase (electrode or electrolyte). We are asked to imagine first carrying the test charge to a certain point, P_B , just outside the surface where it is very close but still outside the range of the image potential. Now the image potential falls off like $1/r$ so rigorously speaking there is no such point; but this is also implicit in the definition of work function, above. For now, believe that this is possible and conclude that the electric potential at this point is still zero since the test charge has not encountered any electric field in its journey. Now suppose that the phase is charged. Again we imagine taking our test charge to P_B and now there will be positive or negative work done and the electric potential at P_B will not be zero. Because the point P_B is “outside the range of the image force” the work done is entirely due to moving the charge through the electric field of the charged phase. The electric potential thus defined at point P_B is the outer electric potential, or *Volta potential*, ψ . By construction, the Volta potential of an uncharged phase is zero. This procedure is illustrated both for solution and the electrode in figure 5–1.

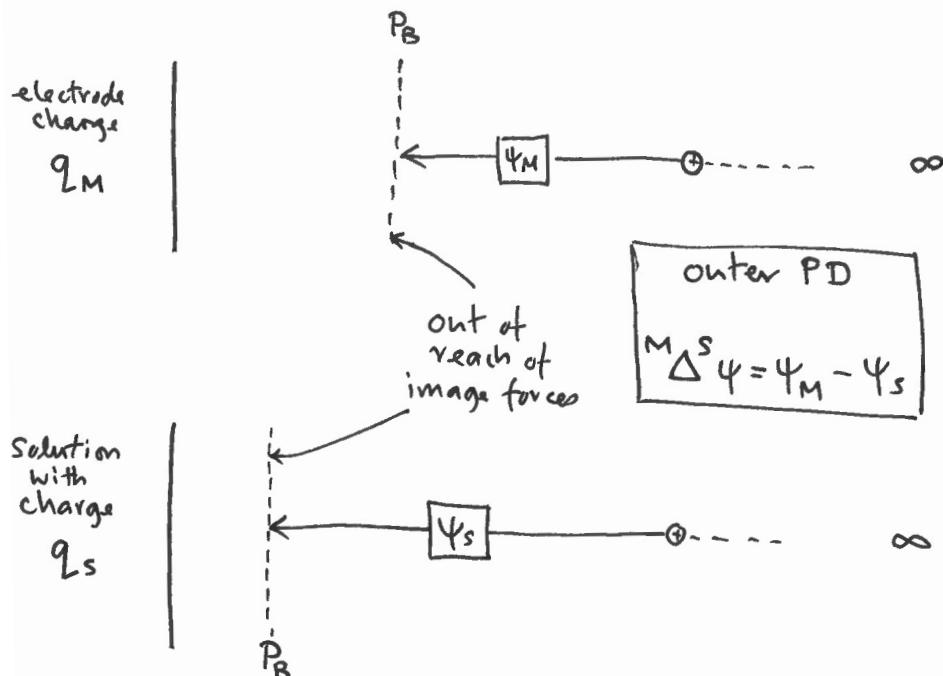


FIGURE 5–1

Bockris devised a thought experiment to clarify what is meant by the Volta potential difference across the interphase separating bulk metal (electrode) and bulk solution (electrolyte). This is illustrated in figure 5–2. The entire assembly at the top of the figure shows the charged electrode on the left, carrying a charge q_m and possessing a dipole layer at its surface where it meets the solution carrying a charge q_s . This charge is distributed near the surface of the solution since we assume that the solution is a conductor and will not support an electric field. The possibly charged layer of solution abutting the interface may be quite wide before giving way to electrolyte with bulk properties; this layer is the Gouy–Chapman or Stern double (or triple) layer (see section 12, below) and as in the case of the metal we expect a complex rearrangement of molecules giving rise to dipole (and presumably higher multipole) moments. The simplest assumption as in the metal is to allow a one dimensional distribution, $\delta\rho(x)$, which integrates to zero so that the work done in taking a test charge across the double layer is equal to ϵ_0^{-1} times the dipole moment per unit area. Now Bockris imagines separating the interphase so that all the metal is exposed to vacuum as is all the solution. He then makes both halves neutral so that any dipole layer due to the charging or due to contact between the two phases vanishes. He then recharges the two phases to the original charges, q_m and q_s ; and uses a test charge as in figure 5–1 to find the outer potentials, ψ_m and ψ_s . Then the Volta potential difference across the interphase is ${}^m\Delta^s\psi$.

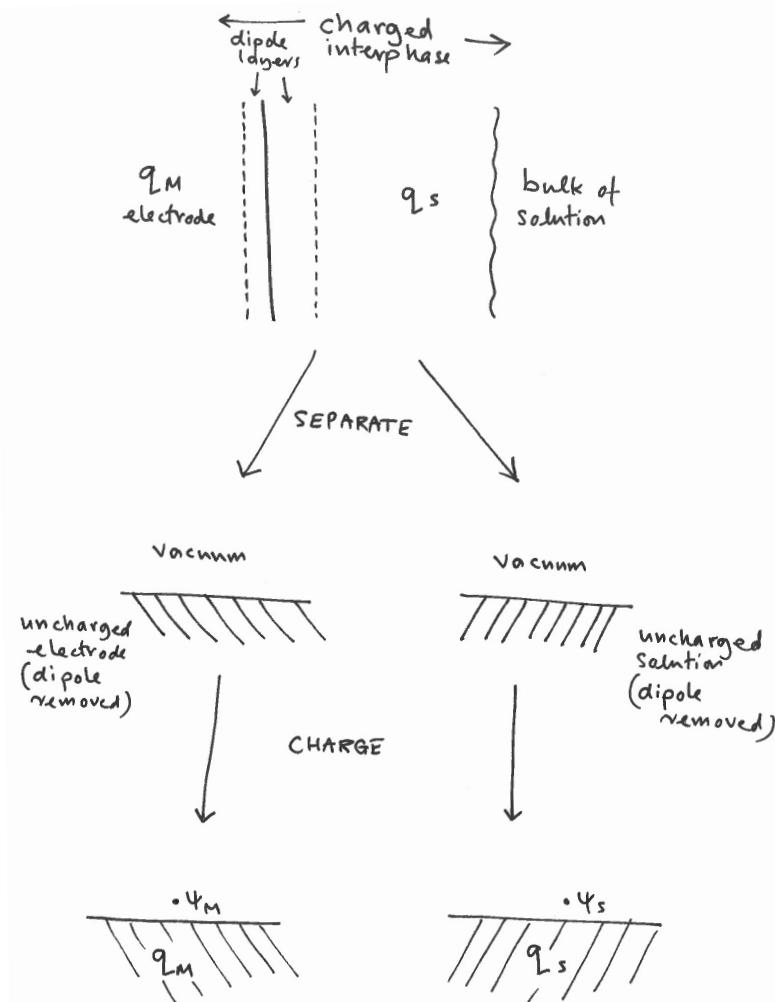


FIGURE 5–2

In order to clarify the dipole potential difference across the same interphase as in figure 5–2, Bockris devised a second thought experiment as illustrated in figure 5–3. The interphase is again dismantled and discharged; but this time the arrangement of molecules that made up the dipole, or “double”, layer in the solution is reassembled on the surface of the solution. Now the separated phases exposed to vacuum are uncharged but have the dipole layers reinstated.

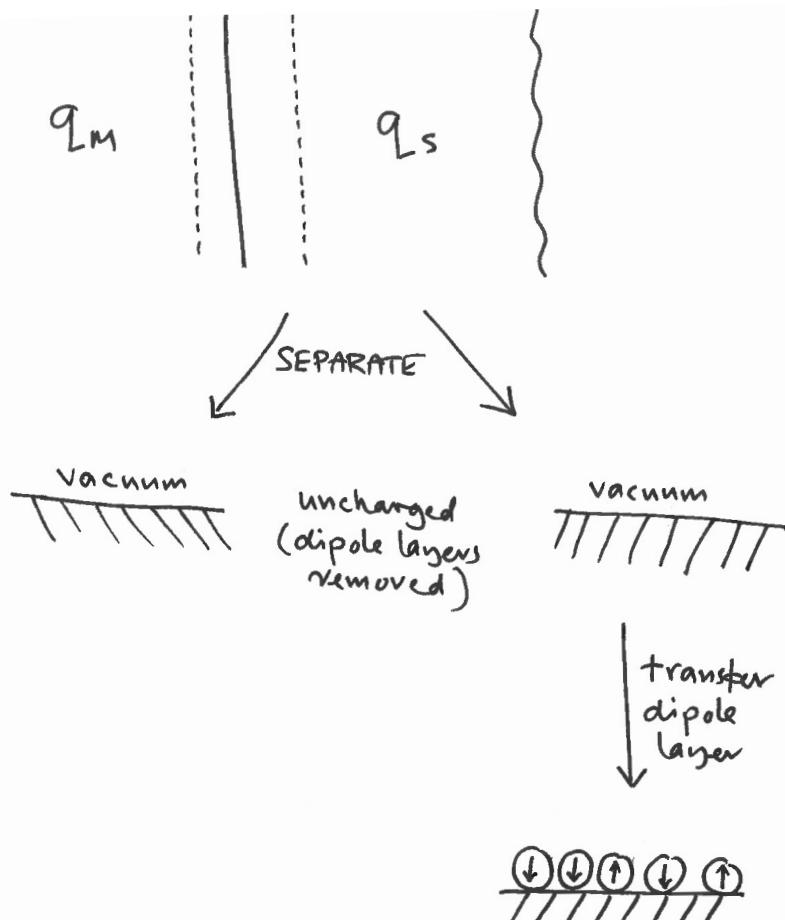


FIGURE 5-3

A positive test charge is now carried from infinity to the interior of the phase where it has bulk properties, see figure 5–4. This defines the χ -potential difference as ${}^m\Delta^s\chi$. In the case of the metal in figure 5–4 it is worthwhile to compare χ_m with the work function. There are two differences. Firstly, the test charge is positive and not negative and it travels in the other direction. A subtle point we'll come back to is that the test charge is not an electron so there will be no exchange contribution to the image force. Secondly in the second thought experiment we bring in the test charge from infinity and not from the vacuum level. The reason for the difference in electric potential between infinity and the vacuum level of a neutral metal described in section 3 is that a crystal naturally has several faces possibly with dissimilar dipole layers. Then the work required to move a test charge from infinity to the vacuum level just outside each face is affected by the electric fields produced by the changes in charge distribution at the corners of the crystal and as a result of differences in surface crystal orientation. *In all of Bockris's thought experiments and in electrochemical textbooks this subtle point is avoided. It is implicitly*

assumed that the extent of the surfaces and interfaces involved is infinite.[†] The process envisaged in Bockris's thought experiment number two is illustrated in figure 5–5. In the upper figure which illustrates the outer, Volta potential associated with the solution, ψ_s is only non zero because the solution carries a net charge $q_s \neq 0$. If the solution were uncharged then the implication is that the potential at infinity is the same as at the point P_B (which we may as well call the Bockris point); indicating that the surface is of infinite extent so there is no electric field.

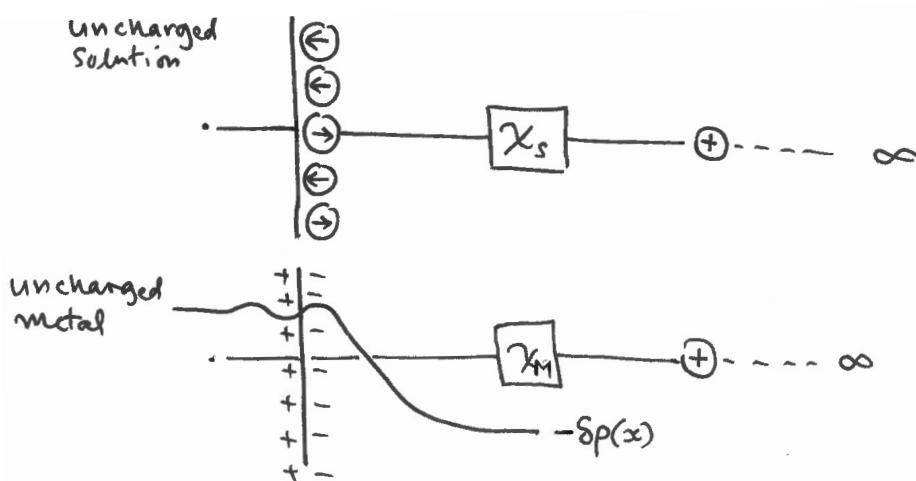


FIGURE 5-4

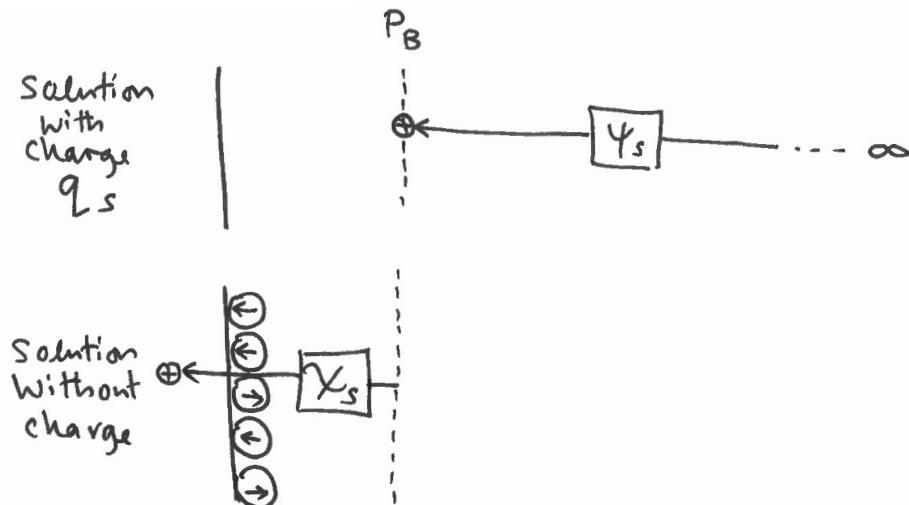


FIGURE 5-5

[†] Imagine two thin sheets of metal joined along their flat surfaces. If these really are infinite in the y and z directions in the sheet then to take a test charge from $-\infty$ to ∞ requires a journey through the three dipole layers and the associated work must be done: hence $\phi(-\infty) \neq \phi(\infty)$. On the other hand if the area of metal normal to x is large but not infinite then the test charge can be taken around the edge keeping essentially at infinity for the whole trip, in which case $\phi(-\infty) = \phi(\infty)$.

Figure 5–5 illustrates the construction of the inner potential from the outer potential and the dipole potential (5.1),

$$\phi_s = \psi_s + \chi_s \quad (5.3)$$

A test charge is taken from infinity to the Bockris point; the solution is discharged while retaining its dipole double layer and the test charge is carried across the double layer into the bulk of the solution where the electric potential is ϕ_s relative to infinity. In conclusion, for any of the interfaces indicated by vertical bars in, say, (2.1) the electric potential difference between two contiguous phases, figure 2–3, were it measurable, would be the difference in inner, or Galvani, potentials which is a sum of the differences of Volta and dipole potentials associated with each bulk phase.

Bockris's thought experiments can be summarised in the diagram shown as figure 5–6 and taken from Fawcett. The difference in Galvani, or inner, potential, ${}^\alpha\Delta^\beta\phi$, between phases α and β , be they each either electrode or solution, can only be made rigorously by reference to a surface of each phase exposed to vacuum. For each phase a test charge is taken from infinity to the Bockris point and the work done per unit charge is the outer, or Volta, potential of each phase. If either phase is uncharged then this is zero for that phase. To avoid the complications of Bockris's second thought experiment we have to assume that all of the work against the electric field due to the charged phase has now been done so that as the test charge is now moved from P_B to the interior where the properties are uniform and bulk-like, the only work done is that against the dipole layer—this can be any charge distribution, $\delta\rho(x)$, as long as it integrates to zero between P_B and the interior point. The dipole potential is then just ϵ_0^{-1} times the dipole moment per unit area. The vertical arrows in the figure show the electric potential differences between the two Bockris points which is ${}^\alpha\Delta^\beta\psi$, and between the two interior points which is ${}^\alpha\Delta^\beta\phi$. This electric potential difference is the result of taking a test charge from the interior of the β phase to its Bockris point, on to infinity and then back to the Bockris point of the α phase and finally into the interior of that phase. This results in

$${}^\alpha\Delta^\beta\phi = -\chi_\beta - \psi_\beta + \psi_\alpha + \chi_\alpha$$

This confirms equation (5.2).

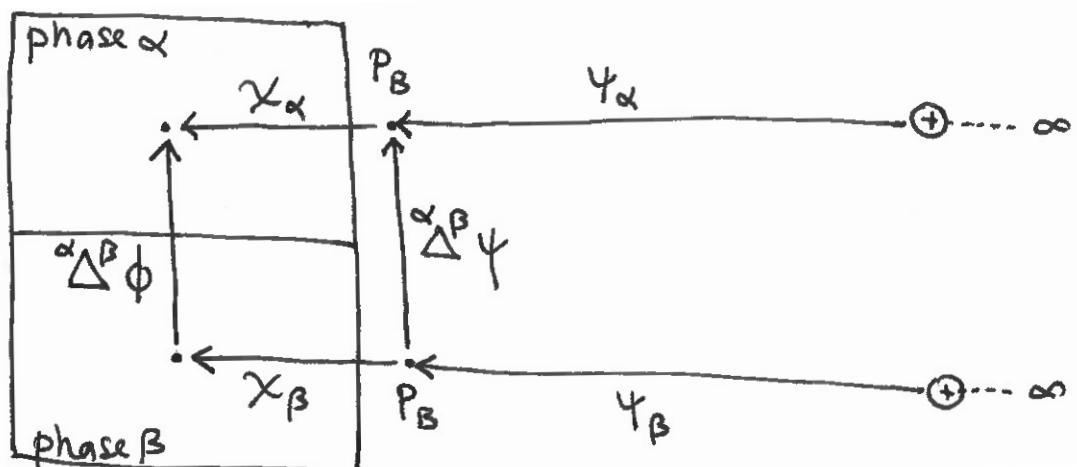


FIGURE 5–6

6. The Bockris point

How far is the Bockris point from the surface of a charged metal or electrolyte? Neglecting for the moment the image potential, suppose that the phase is in the shape of a sphere of radius r and carries a total charge Q . Then the outer electric potential with respect to a point at infinity at a distance s from the surface is

$$\psi_Q = \frac{1}{4\pi\epsilon_0} \frac{Q}{r+s}$$

Both Bockris and Fawcett assert that since $s \ll r$ we may write this as

$$\psi_Q = \frac{1}{4\pi\epsilon_0} \frac{Q}{r}$$

indicating that the outer potential is constant. Bockris argues that an appropriate distance is 10^{-5} cm; Fawcett suggests it is 10^{-4} cm. Both these are beyond the position of the vacuum level for a metal which Kittel states is greater than 10^{-6} cm. It seems unsatisfactory to neglect s altogether since an elementary Taylor expansion results in

$$\frac{1}{r+s} = \frac{1}{r} \left(1 - \frac{s}{r}\right)$$

to first order and this is not constant, but is linear in s . The image potential seen by a test charge q at s , with respect to infinity, if the sphere is conducting and grounded is (see Griffiths)

$$\psi_i = -\frac{1}{4\pi\epsilon_0} \frac{1}{2} \frac{qr}{s(s+2r)}$$

The total Volta potential, Coulomb plus image, is hence

$$\begin{aligned} \psi &= \frac{1}{4\pi\epsilon_0} \left(\frac{Q}{r+s} - \frac{1}{2} \frac{qr}{s(s+2r)} \right) \\ &= \frac{1}{4\pi\epsilon_0} \frac{Q}{r} \left(1 - \frac{s}{r} + \frac{q}{Q} \left(\frac{1}{8} - \frac{1}{4} \frac{r}{s} \right) \right) + \mathcal{O}\left(\frac{s}{r}\right)^2 \end{aligned}$$

It's not obvious that this function is constant for any $s \ll r$ and when plotted as in figure 6–1, it clearly is not.

However this plots the case when the test charge is as much as one tenth of the charge on the sphere. If we take a more realistic case of $q = e$, and $Q = 1.25 \times 10^7 e$ which will charge a sphere of $r = 1$ cm to about a volt, then we are rescued by making a plot against $\log s$. This is done in figure 6–2. Now there is quite a wide plateau, and indeed these calculations are consistent with similar plots shown by Bockris. In conclusion, the Bockris point, as asserted by Bockris, is no closer than 10^{-5} cm from the surface. Infinity is seen to be at about 10–100 cm. On the other hand the image potential is very short ranged and as Kittel states is negligible at about 10^{-6} cm from the surface.

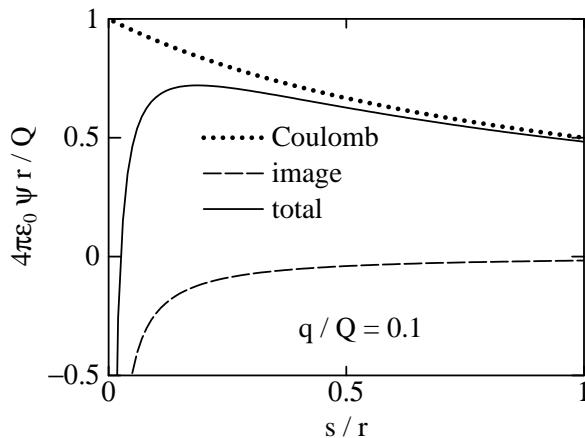


FIGURE 6-1

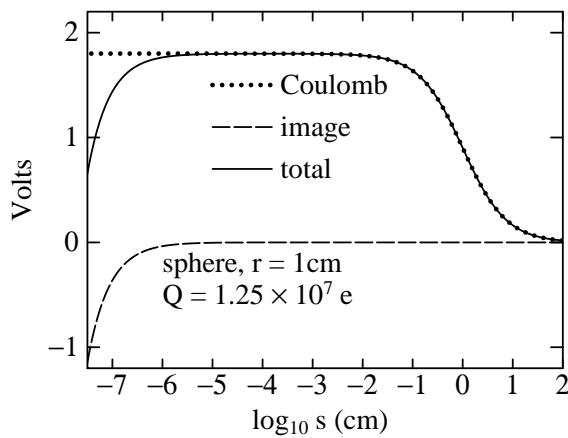


FIGURE 6-2

Note that in figure 6-2 the test charge is taken to be a proton (or positron or anti-muon). This is to acknowledge that this a classical electrostatic calculation and the exchange contribution to the image potential is not included as it should be for an electron.

7. Electron work function of an electrolyte

In order to analyse the potential difference across an electrochemical cell, as well as asking about the work function of the metals involved we may ask about the solutions. Sometimes the work function for an ionic species is required, and usually that is expressed in terms of the real potential, say, of species i in phase α (Cheng and Sprik). In addition we may ask for the work done in withdrawing an electron from an electrolyte. An elegant thought experiment by Schmickler and Santos illustrates this procedure and comes up with an operational definition of work function for an electrolyte. These authors imagine a metal-electrolyte interface with both phases also being exposed to vacuum. The idea is to try and transfer an electron from the solution to the metal via the vacuum. The solution contains equal numbers of moles of Fe^{++} and Fe^{+++} ions. The electron transfer is accomplished in the following five steps.

1. Take an Fe^{++} ion out of the solution and into the vacuum just above the surface of the solution. The amount of work required is minus $\Delta_{\text{sol}}^{\text{r}} G(\text{Fe}^{++})$, which is the *real* free enthalpy of solvation of the Fe^{++} ion. This is a measurable real potential, and it includes the work done against the surface electric dipole belonging to the solution.
2. Remove an electron from the Fe^{++} ion,



The work done is the third ionisation potential of iron, I_3 .

3. Return the Fe^{+++} ion to the solution. This gains us the real free enthalpy of solvation of the Fe^{+++} ion.
4. Transfer the electron from the Bockris point of the electrolyte to the Bockris point of the metal. The work done is

$$-e(\psi_m - \psi_s)$$

5. Lastly put the electron into the metal and gain its work function, W_m .

The total work done should be zero if the electrons are in equilibrium in the assembly, so the energy budget is

$$0 = -\Delta_{\text{sol}}^{\text{r}} G(\text{Fe}^{++}) + I_3 + \Delta_{\text{sol}}^{\text{r}} G(\text{Fe}^{+++}) - e(\psi_m - \psi_s) - W_m$$

or

$$-e(\psi_m - \psi_s) = W_m - [\Delta_{\text{sol}}^{\text{r}} G(\text{Fe}^{+++}) - \Delta_{\text{sol}}^{\text{r}} G(\text{Fe}^{++}) + I_3]$$

Recall that for two metals, A and B the difference in work functions is $-e$ times the difference in outer potentials. This implies that the “work function” of the electrolyte is the sum of terms in brackets,

$$W_s = \Delta_{\text{sol}}^{\text{r}} G(\text{Fe}^{+++}) - \Delta_{\text{sol}}^{\text{r}} G(\text{Fe}^{++}) + I_3 \quad (6.1)$$

all three quantities on the right being measurable. However this work function depends on many things. It is the work function of the $\text{Fe}^{++}/\text{Fe}^{+++}$ “redox couple” since we have used iron ions as the vehicle for carrying the electron away from the solution. The work function would of course be different if we used other cations. It also depends on temperature and on the composition. It is commonplace to construct *standard* real potentials and work functions based in unit activity and standard temperature and pressure.

8. Work function and real potential of a species in an electrolyte

8.1 Electrochemical potential of an ion

The chemical potential of a pure substance as described in the Appendix is nothing other than the partial molar free enthalpy, reflecting the fact that free enthalpy is an extensive state function which is homogeneous and of first order. This allows us to derive equation (A2.10) for the case of a substance made up of more than one component. Whereas G is only determined to within an arbitrary zero of energy, one might think that μ is independent of the choice of zero since the derivative implies a *difference* in free enthalpy which ought to cancel out the dependence on that choice. This is true for a pure substance. On the other hand the chemical potential of a species i in a multi-component substance as expressed in equation (A3.1) is still undefined because we need to refer it to a “standard” chemical potential which depends only on temperature and pressure. This is because the differentiation in (A2.7) still concerns a thought experiment whereby the number of moles of component i is varied while keeping T , p and the mole numbers of all other components fixed. We are reversibly adding an infinitesimal number of moles, dn , and so the total change in free enthalpy is

$$G(n_1 \dots, n_i + dn_i, \dots, n_N) - (G(n_1 \dots, n_i \dots, n_N) + \mu_i^\circ dn_i \text{ (moles of } i \text{ in a reservoir)})$$

So the chemical potential is not independent of the zero of energy because it depends on the nature of the reservoir and the molar free enthalpy, μ_i° , of species i in the reservoir—this is the standard chemical potential. It may be the free enthalpy per mole in the pure substance solid state, or the gas phase, at standard T and p ; or it may be the free enthalpy per mole in a dilute solution. In solution chemistry, we talk of Henrian

and Raoultian standard states, and molar and molal standard states meaning that the species is in unit concentration either in moles per litre or moles per kg of solvent. Nonetheless in the thought experiment, although we need to describe the nature of the reservoir, we do not need to consider the journey that the fraction of a mole of species takes when it leaves the reservoir and arrives in the system under study.

Everything changes when we have the complication that the species i and the system (electrode or electrolyte) may be charged. (In normal circumstances in electrochemistry, electrolytes are neutral however). In that case the reversible work done in carrying a charged component from infinity to the interior of a phase, as we have seen, depends on the details of the surface dipole layer and upon the Galvani potential within the bulk of the phase. It is completely unreasonable to assert that the system is in the shape of a thin specimen infinitely extended in the plane. If that object is charged then the electric field is constant and the electric potential is linear in the distance from the sheet, so it diverges to $\phi = \pm\infty$ at infinity so there can exist no field-free infinity. As Cheng and Sprik emphasise, the electrochemical potential is only defined for a system of finite spatial extent. What, as far as I have found, no textbook acknowledges, is that the electrochemical potential *depends on the shape of the phase*. Moreover, many authors omit to acknowledge the difference between the electric potential at infinity and the Volta potential of an uncharged phase. This is clearly a problem with the definition of the *real potential* (section 4.2) in Fawcett, which author also confuses himself in whether the electron work function is measured with respect to the potential energy of the electron at infinity or at the Bockris point (Fawcett, section 6.6).

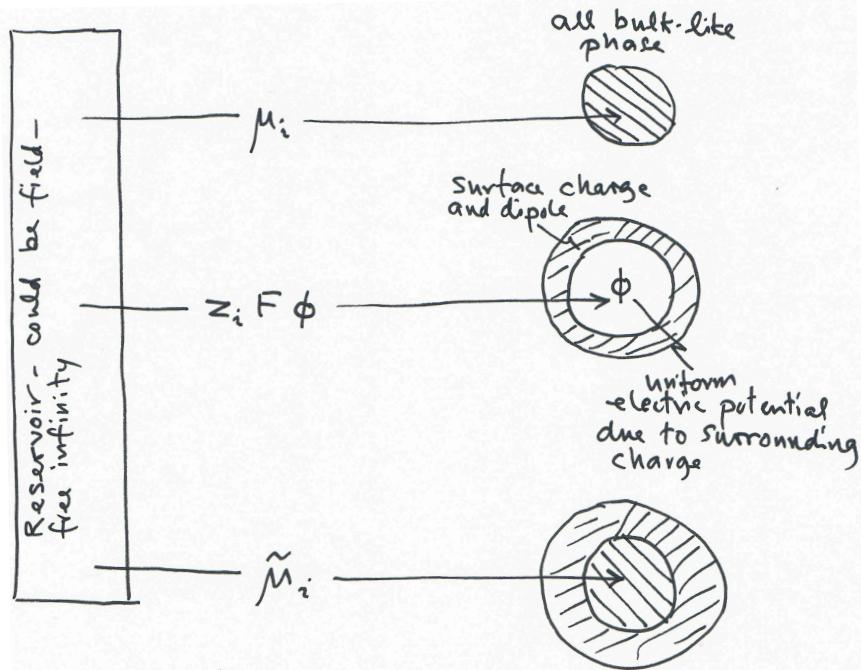


FIGURE 8-1

Figure 8-1 is adapted from Fawcett (and attributed to Parsons) to illustrate some of the points in this section. The upper cartoon is to illustrate the chemical potential. We imagine the system as being bulk-like and homogeneous up to its surface which is taken

as spherical. Since in the this case neither the phase nor the component is charged the shape is irrelevant. If however the phase *is* charged then the charge will reside at the surface (since we take electrolytes to be conducting) and there will be a charged surface dipole layer. The process shown in the lowest cartoon is the equivalent of the Bockris thought experiment in figure 5–5. Now instead of a test charge, we bring in the charged species i from infinity, past the Bockris point where the Volta electric potential is ψ , through the double layer and into the interior where the Galvani potential is ϕ . This serves to *define* the electrochemical potential, but it must be emphasised that this depends on, (i) the *shape* of the phase since the Volta potential depends on the distance from any corners and the presence of patch fields; and (ii) the details of the surface through which the component enters the phase if the surface is not isotropic. In the case of a liquid electrolyte the surface is almost certainly isotropic however. The centre cartoon is to illustrate how the electrochemical potential is divided, as in equation (4.1.1) into the electric work and the remainder. We imagine that the bulk has been excavated from the phase leaving the charged dipole layer surrounding empty space. The electric potential within the space is different from that at infinity due to the surface charge. Hence the reversible work per mole is $z_i F \phi$ where z_i is the charge number of component i and F is the Faraday constant. Again, this must depend on the shape: the electric potential inside a hollow charged conductor depends on its shape.

Consider the case of one-to-one electrolyte, for example NaCl dissolved in water. The electrochemical potentials of the cation and anion, respectively are,

$$\tilde{\mu}_{\text{Na}^+} = \mu_{\text{Na}} + F\phi \quad ; \quad \tilde{\mu}_{\text{Cl}^-} = \mu_{\text{Cl}} - F\phi$$

Now the Galvani potential is in principle not measurable hence the *individual electrochemical potentials of ions are not measurable*; hence they are not proper thermodynamic quantities. (See the Appendix to these notes, sections 3.4–3.6.) This is a most important observation in electrochemistry, not to be ignored. What is measurable is the electrochemical potential of NaCl in the solution, since by addition the charge term cancels and we have

$$\tilde{\mu}_{\text{NaCl}} = \tilde{\mu}_{\text{Na}^+} + \tilde{\mu}_{\text{Cl}^-} = \mu_{\text{Na}} + \mu_{\text{Cl}} = \mu_{\text{NaCl}}$$

8.2 Ion work function of an electrolyte

The work function of an ion in an electrolyte is defined as the reversible work to remove one mole of ions from an *uncharged* solution, in which they are in a standard state of unit concentration and standard T and p , and taken to field-free infinity. The symbol for this normally is W_i^s for the work function of component i in a solution (superscript “s”) phase (Fawcett, Cheng and Sprik) but this fails to acknowledge that this is a standard state quantity as are the chemical potentials in its definition (8.2.1), so I will add a superscript to make this explicit,[†]

$$W_i^{s,\circ} = \tilde{\mu}_i^{\text{g},\circ} - \tilde{\mu}_i^{\text{s},\circ} \tag{8.2.1}$$

[†] Instead of the superscript “s” meaning solution, Cheng and Sprik substitute a superscript α to distinguish possible different phases in a multiphase system. This won’t be necessary here and anyway I recoil from the symbol α_i^α which is hardly aesthetic.

Recall that the electrochemical potential of a component i in a phase depends on its concentration, T and p . The superscript \circ on $\tilde{\mu}_i^{\text{s},\circ}$ is to indicate that this is the electrochemical potential of component i when it finds itself in the solution phase at unit concentration and standard T and p . In all that follows we assume that the solution is uncharged (although the species i is not) then $\phi_s = \psi_s + \chi_s = \chi_s$ (5.3); and since the electric potential at infinity where the ion is in the gas phase is zero, we have,

$$W_i^{\text{s},\circ} = \mu_i^{\text{g},\circ} - \mu_i^{\text{s},\circ} - z_i F \chi_s = \mu_i^{\text{g},\circ} - \alpha_i^{\text{s},\circ} \quad (8.2.2)$$

where $\alpha_i^{\text{s},\circ}$ is the standard real potential (energy) of species i in the solution phase (4.2.1),

$$\begin{aligned} \alpha_i^{\text{s},\circ} &= \mu_i^{\text{s},\circ} + z_i F \chi_s = \tilde{\mu}_i^{\text{s},\circ} - z_i F \psi_s \\ &\equiv \tilde{\mu}_i^{\text{s},\circ} \end{aligned} \quad (8.2.3)$$

The first line follows because $\tilde{\mu} = \mu + q\phi$; the second because in the uncharged solution the Volta potential ψ_s is zero. Combining (8.2.2) and (8.2.3), we have

$$W_i^{\text{s},\circ} = \mu_i^{\text{g},\circ} - \tilde{\mu}_i^{\text{s},\circ} \quad (8.2.4)$$

If we write this as $W_i^{\text{s},\circ} = -\tilde{\mu}_i^{\text{s},\circ} + \mu_i^{\text{g},\circ}$, then an interpretation in words is that the ion work function in its standard state in an electrolyte is the work to remove the ion to the vacuum level plus the subsequent work to transfer it from vacuum into gas at standard T and p . I think that once it is in gas phase we must assume that it is accompanied by an equal number of ions of opposite charge so that the gas is neutral.

For an uncharged solute species, i , and an uncharged solvent, we define the standard free enthalpy of solvation as the reversible work at constant T and p needed to take a mole of solute molecules from the gas phase at standard T and p , and insert them into the solvent to achieve unit concentration at standard T and p ,

$$\Delta_{\text{sol}}G_i^\circ = \mu_i^{\text{s},\circ} - \mu_i^{\text{g},\circ} \quad (8.2.5)$$

Cheng and Sprik apply the same formula to define a standard free enthalpy of solvation for a charged species.[†] In that case we could write, using (4.1.1) and (8.2.4),

$$\begin{aligned} -\Delta_{\text{sol}}G_i^\circ &= \mu_i^{\text{g},\circ} - \mu_i^{\text{s},\circ} \\ &= \mu_i^{\text{g},\circ} - (\tilde{\mu}_i^{\text{s},\circ} - z_i F \phi_s) \\ &= W_i^{\text{s},\circ} + z_i F \phi_s \\ &= W_i^{\text{s},\circ} + z_i F \chi_s \end{aligned}$$

Therefore,

$$\Delta_{\text{sol}}G_i^\circ = -(W_i^{\text{s},\circ} + z_i F \chi_s) \quad (8.2.6)$$

[†] I have difficulty with this, because to me μ_i for a charged species is not physically motivated as I pointed out in comments after equation (4.1.1), above. To me, $\tilde{\mu}_i$ is well defined (although not directly measurable if the component is charged) and μ_i is just the *remainder* after the electric work has been subtracted.

which is equation (8) in Cheng and Sprik. We have already encountered the free enthalpy of solvation in section 7, above. There we were concerned with the *electron* work function of an electrolyte, but in order to remove an electron we used an iron ion as the carrier. However in that case since the ion is charged the correct quantity is the *real* solvation free enthalpy, which in contrast to (8.2.5) is,

$$\Delta_{\text{sol}}^{\text{r}} G_i^{\circ} = \tilde{\mu}_i^{\text{s},\circ} - \mu_i^{\text{g},\circ} \quad (8.2.7)$$

Then it follows from (8.2.3) and (8.2.2) that

$$\begin{aligned} \Delta_{\text{sol}}^{\text{r}} G_i^{\circ} &= \alpha_i^{\text{s},\circ} - \mu_i^{\text{g},\circ} \\ &\equiv -W_i^{\text{s},\circ} \end{aligned} \quad (8.2.8)$$

Finally since the quantities defined in this section are all standard free enthalpies, it is unsurprising that in the case that the solution is not at unit concentration, we use (A3.1),

$$\begin{aligned} \alpha_i^{\text{s}} &= \alpha_i^{\text{s},\circ} + RT \ln a_i \\ \Delta_{\text{sol}}^{\text{r}} G_i &= \Delta_{\text{sol}}^{\text{r}} G_i^{\circ} + RT \ln a_i \end{aligned} \quad (8.2.9)$$

where the activity is defined with respect to unit concentration. I don't think that people ever write $W_i^{\text{s},\circ} = W_i^{\text{s}} + RT \ln a_i$ which is probably why they don't put a superscript \circ on the ion work function.

8.3 Confused and confusing remarks

The object in taking the notes in this direction and to this stage is to put us in a position to follow the authoritative paper by Cheng and Sprik which takes central importance since it provides a quantitative link between experimental electrochemistry and electronic structure theory. There remain for me at least two outstanding questions which have led me into confusion.

1. We insist in accord with electrochemical practice that solutions and electrolytes are globally uncharged. This is implicit in the development from equation (5) to (8) in Cheng and Sprik although they do not state explicitly that $\psi_s = 0$. In that case (8.2.3) and (8.2.8) are *identities*. This means that there is *absolutely no point* in inventing either the real potential or the ion work function. The real potential of charged species i is identical to its electrochemical potential when it's in an uncharged solution. The ion work function is minus the (real) standard free enthalpy of solution. In my view we needn't call this "real" because for a charged species, the free enthalpy of solution (8.2.5) is unphysical so we can use the term "free enthalpy of solution" for the measurable quantity (8.2.9) and (8.2.7).
2. Cheng and Sprik state, "The electrochemical potential, $\tilde{\mu}_i$, is the negative of the reversible work for transferring a particle from solution all the way to infinity. A separate potential $[\alpha_i]$ is introduced for the work of bringing the particle to just outside the surface of a phase. This is the real potential..." This seems to me to be unnecessarily complicated because it is then assumed that the solution is uncharged and hence $\psi_s = 0$ and $\alpha_i = \tilde{\mu}_i$. For even an uncharged metal the electric potential

at the vacuum level (Bockris point) is not the same as at infinity. In an extreme case like tungsten the difference in outer potentials at different crystal faces is as much as 0.75 V. In the case of an electrolyte this difference is apparently ignored in electrochemistry. I think I can see the justification for this. Firstly a liquid has isotropic surfaces; as long as it is uncharged then the outer potential is unlikely to depend on its shape. However if the electrolyte is in physical contact with a metal electrode at an interphase then surely the molecular structure and dipole moment will be very different to that at the surface of the electrolyte. In that case can we continue to imagine that the electrolyte surface is isotropic? It is interesting that the difference in inner potential between two metals in contact can be found from bulk properties, namely the Fermi level, μ_e , *independent of reference to the vacuum* as we've seen in section 3.2. On the other hand in the Bockris thought experiments and the developments in later sections, we have exploited the decomposition of the inner potential into outer and surface (or dipole) potentials (5.3). In that sense the inner potential difference is accessible by reference to the vacuum level. Perhaps this rescues us from the issue that with an interphase the electrolyte boundaries are not isotropic: by constructing the inner potential difference by reference to the surface that faces vacuum (or presumably that part of the electrolyte exposed to air in an electrochemical cell) we avoid the need to worry about the molecular structure of the interphase (figure 5–6).

9. Equilibrium and reversible work of the electrochemical cell

Possibly the most pressing questions in electrochemistry are, which metals are more likely to corrode; and how much energy can I get out of an electrochemical cell?

9.1 Standard reaction free enthalpy

Suppose we are interested in a chemical reaction



in which n_A moles of a substance A and n_B moles of a substance B react to form n_C moles of a substance C and n_D moles of a substance D. On the left are *reactants*, and on the right are *products*. If this reaction occurs spontaneously then the change in free enthalpy is, in view of (A2.10),

$$\Delta G = (n_C \mu_C + n_D \mu_D) - (n_A \mu_A + n_B \mu_B) < 0$$

If we use (A3.1) to expand the chemical potentials we get

$$\begin{aligned} \Delta G &= (n_C \mu_C^\circ + n_C RT \ln a_C + n_D \mu_D^\circ + n_D RT \ln a_D) \\ &\quad - (n_A \mu_A^\circ + n_A RT \ln a_A + n_B \mu_B^\circ + n_B RT \ln a_B) \end{aligned}$$

Gathering terms, we have

$$\begin{aligned} \Delta G &= (n_C \mu_C^\circ + n_D \mu_D^\circ) - (n_A \mu_A^\circ + n_B \mu_B^\circ) \\ &\quad + (n_C RT \ln a_C + n_D RT \ln a_D) - (n_A RT \ln a_A + n_B RT \ln a_B) \\ &= \Delta G^\circ + RT \ln \frac{a_C^{n_C} a_D^{n_D}}{a_A^{n_A} a_B^{n_B}} \\ &= \Delta G^\circ + RT \ln K \end{aligned}$$

This serves to define for us the *standard free enthalpy change* of the reaction (9.1.1)

$$\Delta G^\circ = (n_C \mu_C^\circ + n_D \mu_D^\circ) - (n_A \mu_A^\circ + n_B \mu_B^\circ) \quad (9.1.2)$$

and the *equilibrium constant*,

$$K = \frac{a_C^{n_C} a_D^{n_D}}{a_A^{n_A} a_B^{n_B}} \quad (9.1.3)$$

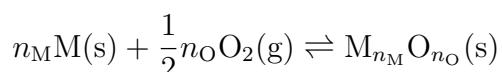
It is easy enough to generalise (9.1.2) and (9.1.3) to the case of any number of reactants and products using summation and product signs. If the reaction (9.1.1) is in equilibrium, that is the rate of reaction to the right is equal to the rate of reaction to the left, then $\Delta G = 0$ and we have

$$K = e^{-\Delta G^\circ / RT}$$

This is called the “van’t Hoff isotherm” and given the standard free enthalpy of the reaction allows us to calculate the equilibrium activities of the component species. Observe that the equilibrium constant is not constant at all—it depends exponentially on the temperature.

I have used the generic superscript \circ to indicate the standard state. But once we specify a reaction I can use for example \bullet if the component is in the pure substance standard state, or $^\ominus$ if the substance is in the one bar gaseous state; and I can mix these in the formula-pair (9.1.2) and (9.1.3) as long as I use the same standard for μ° and the activity for a particular species.

In electrochemistry, and in corrosion in particular, we are interested in the oxidation of metals. For a metal M, the reaction of interest is



where “s” stands for solid phase and “g” stands for gas phase. The standard free enthalpy of formation of the oxide is (9.1.2)

$$\Delta_f G^\circ = \mu_{M_{n_M} O_{n_O}}^\bullet - n_M \mu_M^\bullet - \frac{1}{2} n_O \mu_{O_2}^\ominus$$

We are free to choose zeros of energy such that the free enthalpy of a pure substance in its standard state is zero. Some measured standard free enthalpies of formation (in kilojoule per mole) with respect to the pure reactants are in the following table.

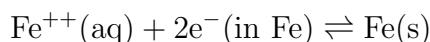
MgO	TiO ₂	NiO	Al ₂ O ₃	FeO	Fe ₂ O ₃	Fe ₃ O ₄	ZnO	H ₂ O
-300	-90	-245	-1700	-270	-820	-1120	-350	-285
Au ₂ O ₃	MnO	Mn ₂ O	HgO	SiO ₂	AgO	V ₂ O ₃	RuO ₂	
-2	-360	-460	-60	-860	-10	-65	-250	

All these are negative, which reflects Humankind’s struggle through the copper, bronze and iron ages to win metals from their ores and the fight to prevent these returning to

the oxide—that is, corroding and oxidising. The only metal found in its noble state in the earth is gold (apart from some meteoric iron) which is why it is highly prized. At the end of civilisation when Gaia returns to her natural owners there will be no trace of the metals that were won from their ores during the disastrous evolutionary experiment of combining a large fore-brain with opposable thumbs (with acknowledgment to Amory Lovins).

9.2 Inner potential difference across a metal–solution interface

In section 3.2 we calculated the inner potential difference across a metal–metal interface: it's the difference in Fermi levels. Now we are in a position to do the same at an electrode–electrolyte interface. Consider the “half-cell” in which Fe metal is immersed in a solution of Fe^{++} ions. The following chemical reaction will occur,



“aq” means in aqueous solution. The free enthalpy change per mole is

$$\begin{aligned}\Delta G &= \mu_{\text{Fe}} - \tilde{\mu}_{\text{Fe}^{++}} - 2\tilde{\mu}_{\text{e}} \\ &= (\mu_{\text{Fe}}^\bullet + RT \ln a_{\text{Fe}}) - (\mu_{\text{Fe}^{++}}^\circ + RT \ln h_{\text{Fe}^{++}} + 2F\phi_s) - 2(2\mu_{\text{e}}^{\text{Fe}} - F\phi_m)\end{aligned}$$

$h_{\text{Fe}^{++}}$ is the practical Henrian “working” activity of iron ions (Appendix section 3.6 equation (A3.26a)) and the superscript $^\circ$ indicates the unimolar (one mole per kg solvent) standard state.[†] $\mu_{\text{e}}^{\text{Fe}}$ is the (standard) chemical potential of the electron in the metal, $\tilde{\mu}_{\text{e}} = \mu_{\text{e}} - F\phi_m$.[‡] Having immersed the metal in the electrolyte, when the half-cell comes to equilibrium $\Delta G = 0$ and we have,

$$\begin{aligned}2F(\phi_m - \phi_s) &= (\mu_{\text{Fe}^{++}}^\circ - \mu_{\text{Fe}}^\bullet + 2\mu_{\text{e}}^{\text{Fe}}) + RT \ln \frac{h_{\text{Fe}^{++}}}{a_{\text{Fe}}} \\ &= \mu_{\text{Fe}^{++}}^\circ + 2\mu_{\text{e}}^{\text{Fe}} + RT \ln h_{\text{Fe}^{++}}\end{aligned}\quad (9.2.1)$$

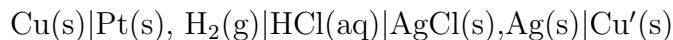
The second line follows because the standard free enthalpy of the pure solid iron is zero and its activity is one. Equation (9.2.1) should allow us to deduce ${}^m\Delta^s\phi$ from the concentration of iron ions in the electrolyte. However this can only be known relative to some other half cell: the absolute value of $\mu_{\text{Fe}^{++}}^\circ$ can only be measured relative to another ion as the electrochemical potential of a charged ion is not measurable (section 8.1). Moreover the activity of an individual ion is also not measurable; we know the concentration of ions in moles per kg of water, but we don't know the activity coefficient—we have to assume a mean ion activity coefficient (A3.19). Relative to the hydrogen ion, the standard (unimolar) chemical potential of the Fe^{++} ion is -85 kJ/mol . We could complicate the problem by imagining an iron alloy; in which case we'd need to know the standard free energy of formation of the alloy, its iron content and the activity of Fe in the alloy.

[†] If you can't be bothered with the details of how the “activity” is worked out, then at a first reading of this and what follows in sections 9 and 10 it is quite adequate to take the activity, h , of some species as its concentration in moles per kg solvent, and its standard chemical potential, μ° , as its value at 1 mol/kg.

[‡] Compare this with (3.4.4). I will not distinguish in the notation between free enthalpy *per mole* and free enthalpy *per particle*. It will be obvious from the context.

9.3 Open circuit voltage—electromotive force

Consider the electrochemical cell shown in figure 9–1. On the left is a hydrogen electrode; the electrolyte is hydrochloric acid, HCl; the right hand electrode is silver on which silver chloride has been deposited in lumps by the action of the acid. The “cell diagram” is



The Pt and Ag electrodes are connected via copper wire through a high impedance voltmeter, so that no current flows but the electric potential difference, $\phi_{\text{Cu}'} - \phi_{\text{Cu}}$ can be measured. We wish to calculate this “open circuit” voltage and determine how much useful work can be done by the cell if a current is allowed to flow. By convention the oxidation reaction is placed on the left, the reduction reaction on the right, and if current is allowed to flow then positive charge will flow from left to right through the electrolyte and negative charge (electrons) will flow from left to right through the external circuit.

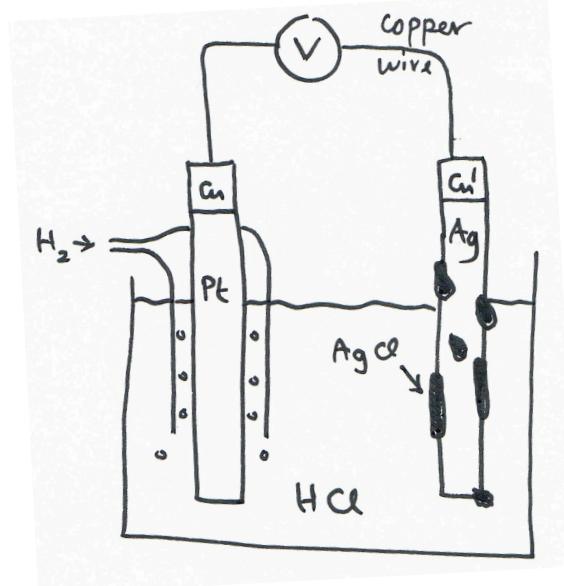
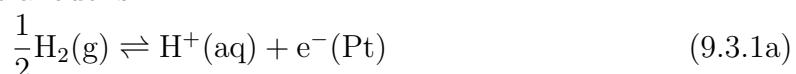


FIGURE 9–1

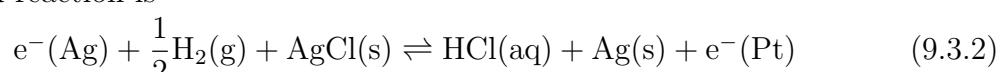
The oxidation reaction at the anode is



and the reduction reaction at the cathode is



The combined reaction is



The change in free enthalpy is

$$\Delta G = \mu_{\text{HCl}} + \mu_{\text{Ag}} + \tilde{\mu}_{\text{e}}^{\text{Pt}} - \left(\frac{1}{2}\mu_{\text{H}_2} + \mu_{\text{AgCl}} + \tilde{\mu}_{\text{e}}^{\text{Ag}} \right) \quad (9.3.3\text{a})$$

$$= \mu_{\text{HCl}}^\circ + \mu_{\text{Ag}}^\bullet - \frac{1}{2}\mu_{\text{H}_2}^\ominus - \mu_{\text{AgCl}}^\circ + RT \ln \frac{h_{\text{HCl}}}{p^{1/2}} + (\tilde{\mu}_{\text{e}}^{\text{Pt}} - \tilde{\mu}_{\text{e}}^{\text{Ag}}) \quad (9.3.3\text{a})$$

$$= \Delta G^\circ + RT \ln h_{\text{HCl}} + (\tilde{\mu}_{\text{e}}^{\text{Pt}} - \tilde{\mu}_{\text{e}}^{\text{Ag}}) \quad (9.3.3\text{b})$$

in which h_{HCl} is the electrolyte activity on the practical Henrian scale of moles per kg of water, (see Appendix (A3.12a) and (A3.22a)) and p is the pressure of hydrogen as measured in bar. The molar free enthalpies of the pure hydrogen and silver can be taken as zero and the pressure of hydrogen is adjusted to one bar. AgCl is not a pure solid, but we assume for simplicity that it is in its standard state. In this way we have the *standard free enthalpy* for the reaction (9.3.2),

$$\Delta G^\circ = \mu_{\text{HCl}}^\circ - \mu_{\text{AgCl}}^\circ$$

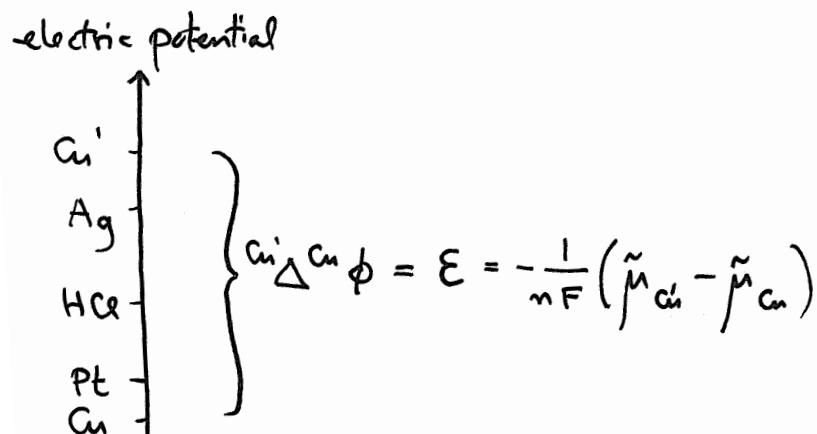


FIGURE 9–2

When both the reduction and oxidation reactions (9.3.1) come to equilibrium then at open circuit the free enthalpy change is zero, $\Delta G = 0$, even though the electrons in the electrodes are not at equilibrium. However since each electrode is in contact with its copper wire the electrons across these junctions will have the same electrochemical potential and while

$$\tilde{\mu}_e^{\text{Pt}} = \tilde{\mu}_e^{\text{Cu}} ; \quad \tilde{\mu}_e^{\text{Ag}} = \tilde{\mu}_e^{\text{Cu}'} \quad (9.3.4)$$

the electrochemical potentials in the two wires connected across the voltmeter are not the same. At this stage it's helpful to draw a ladder diagram as in figure 2–3. This is shown in figure 9–2. The electric potential of the anode is negative compared to the electric potential of the cathode. That is, $\phi_{\text{Cu}'} > \phi_{\text{Cu}}$. Conventionally represented as the electrode on the left, the anode is the electron sink (figure 1–2) so a metal atom is losing an electron to become a cation which will migrate to the right through the electrolyte ("cations" migrate to the "cathode"). The electrons left behind being negatively charged serve to *lower* the electric potential on the anode.[†] The right hand electrode, the cathode, is conversely more positive than the left hand electrode. The

[†] This is confusing because we learn, "anodes are positive and they corrode". True. If I turn the voltmeter into a battery and connect the positive terminal to the left hand electrode, this will remove electrons from the electrode, make its electric potential more positive and drive the oxidation (corrosion) of the electrode metal. In this case, the platinum is not corroding. The de-electronation reaction is the oxidation of hydrogen gas to hydrogen ions (protons).

total potential difference across the electrochemical cell, as indicated in figure 9–2, is $\mathcal{E} = \phi_{\text{Cu}'} - \phi_{\text{Cu}} > 0$ and is called the “electromotive force”, or electromotance, of the cell. It is measured in Volts. The existence of this non zero e.m.f. is due entirely to the difference in electrochemical potentials on the two electrodes. It must be true therefore that

$$\mathcal{E} = -\frac{1}{F}(\tilde{\mu}_{\text{e}}^{\text{Ag}} - \tilde{\mu}_{\text{e}}^{\text{Pt}}) = -\frac{1}{F}(\tilde{\mu}_{\text{e}}^{\text{Cu}'} - \tilde{\mu}_{\text{e}}^{\text{Cu}})$$

Putting this fact into (9.3.3) with $\Delta G = 0$ at open circuit, and using (9.3.4) we arrive at the formula we are looking for, namely,

$$-F\mathcal{E} = \Delta G^\circ + RT \ln h_{\text{HCl}} \quad (9.3.5)$$

All the quantities in (9.3.5) can be measured: the concentration of HCl is ours to control and the standard molar free enthalpies can be looked up in tables of data to obtain the standard free enthalpy change, ΔG° , for the cell reaction.

As to the amount of work that the cell might do, we note that at open circuit the cell reactions (9.3.1) are in equilibrium even though the electrochemical potentials of the electrons in the Pt and Ag electrodes are not. If we were to allow some current to flow through the external circuit, but *slowly*: then we may allow the cell to do some reversible work. If one mole of reactants is permitted to become one mole of products, then the change in free enthalpy is ΔG . If in the process n moles of positive charge are transferred from left to right through the electrolyte then an equal and opposite electronic charge, $-nF$, is transferred from left to right through the copper wire across an electric potential difference of $\Delta\phi = \mathcal{E}$, which will do an amount of work equal to $-nF\mathcal{E}$. *This statement lies at the heart of electrochemistry, batteries, fuel cells...* This is exactly how *chemical energy* is converted into *electrical energy*; and of course the reverse is also true: if the charge on $-nF$ moles of electrons is carried from right to left by a power source then one mole of product will be converted to one mole of reactant in the cell. This is how electricity is stored as chemical energy in a battery. Perhaps the most important equation in electrochemistry is the one that expresses the equivalence of chemical and electrical work—from the above argument it is clear that

$$\Delta G = -nF\mathcal{E} \quad (9.3.6)$$

The meaning of the number, n , can be a bit confusing. It is the number of moles of electrons involved in the combined “redox” reaction. So in the case of (9.3.2), $n = 1$.

9.4 The electrochemical series

We've seen that we cannot measure a single electrode potential. Nevertheless it is very useful to rank them by comparing different metal single electrodes to a reference electrode under standard conditions. The reference usually used is the standard hydrogen electrode (SHE). It is evident from figures 2–3 and 9–2 that at least formally we can divide the electric potential drop across the cell into that due to the silver half cell and that due to the SHE.

$$(\phi_{\text{Ag}} - \phi_{\text{Pt}}) = (\phi_{\text{Ag}} - \phi_{\text{HCl}}) + (\phi_{\text{HCl}} - \phi_{\text{Pt}})$$

or

$${}^{\text{Ag}}\Delta^{\text{Pt}}\phi = {}^{\text{Ag}}\Delta^{\text{HCl}}\phi - {}^{\text{Pt}}\Delta^{\text{HCl}}\phi \quad (9.4.1)$$

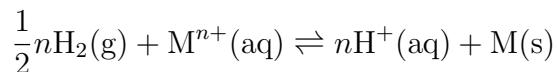
Suppose I am interested in a particular metal, M, which may be oxidised in an electrolyte, S, containing M^{n+} metal ions having a charge of ne . I can construct an electrochemical cell just as shown in figure 2–4, but with the SHE on the left,[†]



(the ‘||’ is a salt bridge as in figure 2–4) where the activities are practical “working” ones (see the Appendix equation (A3.26a) and section 3.6). For this cell I will have that the e.m.f. according to (9.4.1) is

$$\mathcal{E} = {}^{\text{m}}\Delta^{\text{s}}\phi - {}^{\text{Pt}}\Delta^{\text{s}}\phi = \mathcal{E}_{\text{M}}$$

essentially the difference between two half-cell potentials. By definition, \mathcal{E}_{M} is the *single potential* of the $\text{M}^{n+}(\text{aq})|\text{M(s)}$ electrode *on the SHE scale*. The cell reaction is



for which the free enthalpy change is

$$\begin{aligned}\Delta G &= \mu_{\text{M}}^\bullet + n\mu_{\text{H}^+}^\circ - \frac{1}{2}n\mu_{\text{H}_2}^\bullet - (\mu_{\text{M}^{n+}}^\circ + RT \ln h_{\text{M}^{n+}}) \\ &= -(\mu_{\text{M}^{n+}}^\circ + RT \ln h_{\text{M}^{n+}})\end{aligned}$$

(pure substance standard chemical potentials being zero). Since we arbitrarily set the e.m.f. of the SHE to zero, we have, using (9.3.6)

$$\begin{aligned}\mathcal{E}_{\text{M}} &= -\frac{\Delta G}{nF} \\ &= \frac{\mu_{\text{M}^{n+}}^\circ}{nF} + \frac{RT}{nF} \ln h_{\text{M}^{n+}} \quad (9.4.2a)\end{aligned}$$

$$= \mathcal{E}_{\text{M}}^\circ + \frac{RT}{nF} \ln h_{\text{M}^{n+}} \quad (9.4.2b)$$

We have followed the usual convention and set $\mu_{\text{H}^+}^\circ = 0$, which means that the standard unimolar chemical potential of the metal ion is implicitly referred to that of the hydrogen ion. Below is a table of μ° for some cations in kilojoule per mole.

[†] The SHE is defined as having a unit activity of hydrogen ions. This is problematic for me, because even if the activity of an individual ion does exist in thermodynamics, it is certainly not measurable, see the Appendix, section 3.5. The best authority I can find is “Corrosion”, edited by L. L. Sheir *et al.*, (Butterworth Heinemann, 3rd edition 1994). There the SHE by definition has unit activity, on an unspecified scale, of H^+ ions which is identified as being the same as a molality of $m_{\text{H}^+} = 1.2$ moles HCl per kg water. See the Appendix, section 3.6. Moore gives the mean ion activity (A3.20) as one.

H ⁺	Ti ⁺⁺	Ni ⁺	Al ⁺⁺⁺	Fe ⁺⁺	Fe ⁺⁺⁺	Zn ⁺⁺
0	-160	-50	-480	-85	-11	-150
Au ⁺⁺⁺	Mg ⁺⁺	Hg ⁺	Sn ⁺⁺	Ag ⁺	Cu ⁺	Cu ⁺⁺
+410	-450	+160	-25	+75	+50	+65

The more negative is the standard chemical potential, the more work can the cation do at the anode of an electrochemical cell.

Au ³⁺ + 3e ⁻ → Au.....	1.50
Pd ²⁺ + 2e ⁻ → Pd	0.987
Hg ²⁺ + 2e ⁻ → Hg.....	0.854
Ag ⁺ + e ⁻ → Ag.....	0.800
Hg ₂ ²⁺ + 2e ⁻ → 2Hg.....	0.789
Cu ⁺ + e ⁻ → Cu.....	0.521
Cu ²⁺ + 2e ⁻ → Cu.....	0.337
2H ⁺ + 2e ⁻ → H ₂	(Reference) 0.000
Pb ²⁺ + 2e ⁻ → Pb	-0.126
Sn ₂ ⁻ + 2e ⁻ → Sn.....	-0.136
Ni ²⁺ + 2e ⁻ → Ni.....	-0.250
Co ²⁺ + 2e ⁻ → Ni	-0.277
Tl ⁺ + e ⁻ → Tl	-0.336
In ³⁺ + 3e ⁻ → In	-0.342
Cd ²⁺ + 2e ⁻ → Cd	-0.403
Fe ²⁺ + 2e ⁻ → Fe	-0.440
Ga ³⁺ + 3e ⁻ → Ga	-0.53
Cr ³⁺ + 3e ⁻ → Cr.....	-0.74
Cr ²⁺ + 2e ⁻ → Cr.....	-0.91
Zn ²⁺ + 2e ⁻ → Zn	-0.763
Mn ²⁺ + 2e ⁻ → Mn	-1.18
Zr ⁴⁺ + 4e ⁻ → Zr	-1.53
Ti ²⁺ + 2e ⁻ → Ti	-1.63
Al ³⁺ + 3e ⁻ → Al	-1.66
Hf ⁴⁺ + 4e ⁻ → Hf	-1.70
U ³⁺ + 3e ⁻ → U	-1.80
Be ²⁺ + 2e ⁻ → Be	-1.85
Mg ²⁺ + 2e ⁻ → Mg	-2.37
Na ⁺ + e ⁻ → Na	-2.71
Ca ²⁺ + 2e ⁻ → Ca	-2.87
K ⁺ + e ⁻ → K	-2.93
Li ⁺ + e ⁻ → Li	-3.05

Electrochemical, or *electromotive force*, series.

\mathcal{E}_M° is the *standard single cell potential on the SHE scale*. These may be compiled into a ranked table called the *electrochemical series* or the *electromotive force series*. If at the first page of these notes, the lemon lamp, you asked how is it that the zinc corrodes

and not the copper; the answer is in the electrochemical series. When the two half cells $\text{Zn}^{++}(\text{aq})|\text{Zn(s)}$ and $\text{Cu}^{++}(\text{aq})|\text{Cu(s)}$ are connected electrically, and they share the same electrolyte (lemon juice which is acidic) then the $\text{Zn}^{++}(\text{aq})|\text{Zn(s)}$ electrode having a more negative standard electrode potential (-0.76V , compared to $+0.34\text{V}$) becomes the anode and de-electronation occurs at the Zn electrode and electronation at the Cu electrode. As expressed in (9.4.2) $\mathcal{E}_{\text{Zn}}^{\circ} < \mathcal{E}_{\text{Cu}}^{\circ}$ because $\mu_{\text{Zn}^{n+}}^{\circ} < \mu_{\text{Cu}^{n+}}^{\circ}$, that is, the Zn ion has the more negative standard chemical potential in solution in water. Evidently the system can gain more free enthalpy if the Zn dissolves, rather than the Cu. This to my mind does not really *explain* why the Zn chooses to corrode. You may gaze at the electrochemical series and wonder how Nature has organised metals in this way. To a physicist, the noble metals are Cu, Ag and Au. To the chemist the noble metals are those at the top of the electrochemical series, Au, Pt, Pd; they call Cu, Ag, Au the *coinage metals*. (The $\text{Pt}^{++}(\text{aq})|\text{Pt(s)}$ single electrochemical potential is $+1.2\text{V}$ —it's missing from the table I found.) The metals with a large negative $\mathcal{E}_{\text{M}}^{\circ}$ are called *base metals*. To a corrosion scientist the fact that a metal of interest, for example Mg, is a base metal can be depressing. In the design of Mg alloys poor corrosion resistance is one of the principal reasons why this least dense of engineering alloys is less ubiquitous than you may expect. But there is almost nothing that can be done to combat the fundamental fact that $\mathcal{E}_{\text{Mg}}^{\circ} = -2.37\text{V}$. I have no deep insight as to why the metals are ordered as they are. Fawcett gives three factors that determine the position of a metal in the series: (i) the free enthalpy of solvation of the metal ion in water; (ii) the free enthalpy of the metal, which depends on its crystal and electronic structure; (iii) the ionisation energy of the metal atom. I'll return to this in section 14.4.

9.5 Nernst equation

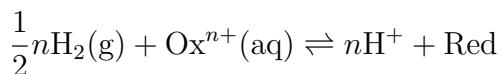
The half cell reaction can be generalised to consider a substance “Ox”, which is reduced by acquiring an electron from the electrode to become the reduced substance, “Red”,



The half cell single potential can be measured as in section 9.4 by connecting to a SHE



The cell reaction is



The e.m.f. is

$$\mathcal{E} = \mathcal{E}^{\circ} + \frac{RT}{nF} \ln \frac{a_{\text{Ox}}}{a_{\text{Red}}} \quad (9.5.2)$$

The standard e.m.f. is

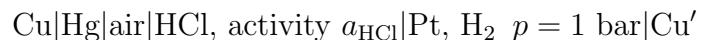
$$\mathcal{E}^{\circ} = \frac{1}{nF} (\mu_{\text{Ox}}^{\circ} - \mu_{\text{Red}}^{\circ}) \quad (9.5.3)$$

(9.5.2) is called the *Nernst equation*. The numerical value of the standard e.m.f. will depend on the scale and convention that are used for the activity (see the Appendix, section 3). We return to the Nernst equation in section 14.4, page 89.

10. Relative and absolute ion work function and real potential

10.1 Measurement of the proton work function and real potential in water

The electrochemical potential of a charged species is *in principle* not measurable. However measurements can be made if assumptions are made that extend beyond the scope of equilibrium thermodynamics. Many measurements are made using a device called “Kenrick’s apparatus.” The idea is to create an air gap between a metal electrode and an electrolyte and create conditions in which the Volta potential difference between the two surfaces is zero. This can be done as already described in section 3.4 using the Kelvin probe. There, a bias potential is applied of an amount just to prevent any current flowing if the surfaces are moved closer to each other or further away. That so called *compensation potential*, denoted $\Delta_c\phi$, is equal to minus the difference in Volta potential, permitting the measurement of the Volta potential difference between two surfaces (see Fawcett, section 8.7 for details). In Kenrick’s apparatus, rather than applying a bias potential, it is arranged that the Volta potential difference is zero by design. This is achieved using a liquid mercury electrode that is arranged to flow rapidly past a stream of electrolyte flowing in the opposite direction, with an air gap in between. Because the two opposing surfaces are in rapid flow it is assumed that charge cannot distribute in time so that the Volta potential difference across the air gap is zero (see Fawcett, section 8.7). I have to say I do not find this convincing: the time taken for charge to distribute in a metal must be roughly the inverse plasma frequency; I don’t know what it is in an electrolyte. But if the system is in steady state, or worse, then I don’t see how we can apply thermodynamic principles. However, it’s been around a long time and it works. (You really can abuse thermodynamics if you don’t mind offending the ghost of Josiah Willard Gibbs—when I was a metallurgy undergraduate doing problems, they occasionally required the student to “assume equilibrium in the blast furnace.”) Using this apparatus, and making one very reasonable and not at all serious non-thermodynamic assumption it is possible to measure the *work function of the proton in water*. The assumption is necessary because the activity of a proton cannot be measured. So in hydrochloric acid, HCl, it is assumed that the activity of the H^+ ion is equal to the mean ion activity. The Kenrick apparatus for this experiment uses streams of mercury and HCl, the electrolyte being connected to a standard hydrogen electrode. The cell diagram is



The potential measured across the copper leads is a compensation potential by virtue of

$$\psi_s - \psi_{\text{Hg}} = 0$$

and is given by

$$\begin{aligned} F\Delta_c\phi &= \tilde{\mu}_{\text{e}}^{\text{Cu}} - \tilde{\mu}_{\text{e}}^{\text{Cu}'} \\ &= \tilde{\mu}_{\text{e}}^{\text{Hg}} - \tilde{\mu}_{\text{e}}^{\text{Pt}} \\ &= \mu_{\text{e}}^{\text{Hg}} - \mu_{\text{e}}^{\text{Pt}} - F(\phi_{\text{Pt}} - \phi_{\text{Hg}}) \end{aligned} \tag{10.1.1}$$

By considering the equilibrium of the SHE half cell, using procedures just as in section 9,

one finds for ${}^{\text{HCl}}\Delta^{\text{Pt}}\phi$,

$$F(\phi_s - \phi_{\text{Pt}}) = \frac{1}{2}\mu_{\text{H}_2}^{\text{g},\circ} - \mu_{\text{H}^+}^{\text{s},\circ} - \mu_{\text{e}}^{\text{Pt}} - RT \ln a_{\text{H}^+}$$

where a_{H^+} is the activity of the hydrogen ion (we are not specifying any particular reference standard state and are using the generic $^\circ$ superscript.) Combining this with (10.1.1) to eliminate Pt potentials in favour of Hg potentials results in

$$F\Delta_c\phi = F(\chi_s - \chi_{\text{Hg}}) + RT \ln a_{\text{H}^+} - \Delta G^\circ$$

where

$$\Delta G^\circ = \frac{1}{2}\mu_{\text{H}_2}^{\text{g},\circ} - \mu_{\text{H}^+}^{\text{s},\circ} - \mu_{\text{e}}^{\text{Hg}}$$

Here the dipole potential difference enters since in view of the air gap and the flowing electrode and electrolyte the Volta potential difference is zero and ${}^s\Delta^{\text{Hg}}\phi = {}^s\Delta^{\text{Hg}}\chi$. The experimenter measures the voltage across the copper leads with a very high impedance voltmeter as a function of the strength of the acid. Assuming that the activity of H^+ is equal to the mean ion activity which can be measured (this is the “extrathermodynamic” assumption: $a_{\text{H}^+} = a_\pm = \sqrt{a_{\text{HCl}}}$, Appendix, section 3.4, (A3.24a)) then plotting $\Delta_c\phi - RT \ln m_\pm$, where $m_\pm = m_{\text{HCl}}$ is the molal concentration of the acid (assuming in the dilute limit that the mean ion activity coefficient is one) against the square root of the HCl concentration and extrapolating to zero concentration yields a value of **-56.3mV** in the infinitely dilute limit (see Fawcett, section 8.7). Because it’s constant we don’t yet need to know what ΔG° is.

Next, employ the fact that in this limit the dipole potential of the infinitely dilute acid is the same as that of water, $\chi_{\text{H}_2\text{O}}$. Then,

$$\begin{aligned} F\Delta_c\phi - RT \ln m_\pm (m_\pm \rightarrow 0) &= -56.3\text{mV} \times F \\ &= \mu_{\text{H}^+}^{\text{s},\circ} + F\chi_{\text{H}_2\text{O}} + \mu_{\text{e}}^{\text{Hg}} - F\chi_{\text{Hg}} - \frac{1}{2}\mu_{\text{H}_2}^{\text{g},\circ} \\ &= \alpha_{\text{H}^+}^{\text{s},\circ} + \alpha_{\text{e,Hg}}^{\text{s},\circ} - \frac{1}{2}\mu_{\text{H}_2}^{\text{g},\circ} \end{aligned} \quad (10.1.2)$$

using, in the second line, the definition of the real potential (4.2.1) and remembering that the charge on the electron is negative. Now, the real potential of the electron in a metal is the same as minus its work function, which for mercury is 4.5eV. The last term in (10.1.2) is the standard chemical potential of hydrogen in the gas phase which is zero for the pure substance. Hence (see second footnote, page 33),

$$\alpha_{\text{H}^+}^{\text{s},\circ} = -0.056 + 4.5 = 4.44\text{eV} \quad (10.1.3)$$

The proton work function is (8.2.8)

$$W_{\text{H}^+}^{\text{s},\circ} = \mu_{\text{H}^+}^{\text{g},\circ} - \alpha_{\text{H}^+}^{\text{s},\circ}$$

The first term, $\mu_{\text{H}^+}^{\text{g},\circ}$, is the standard free enthalpy of formation of a proton in the gas phase. Relative to the standard free enthalpy of the hydrogen molecule, this is half the

dissociation (bond) energy (4.52eV) plus the ionisation energy (13.6eV), namely 15.9eV. The work function of the proton in water then comes out as 11.4eV (1103 kJ/mol). Once that is known, then using (8.2.6) and the measured standard free enthalpy of solvation of hydrogen ions (again not strictly measurable, but using sensible extrathermodynamic assumptions) it is possible to deduce from experiment that the electric dipole potential at the surface of water is 80mV. This presents a challenge to DFT to try and calculate it.

10.2 Absolute electrode potential

The electrochemical series ranks electrode potentials and provides a single electrode potential *relative to the standard hydrogen single electrode potential*. This is inevitable because thermodynamics does not furnish us with an absolute zero of energy. Hence we choose the one most convenient to us. So far we have encountered at least five choices. (i) In the DFT using periodic boundary conditions and hence with no reference to the surfaces of the substance or the world beyond it, we use the average potential energy seen by an electron, \bar{V} , as a convenient reference (section 3.1). (ii) The DFT total energy, E_{tot} , is calculated with reference to the situation where all the electrons and nuclei are scattered, each to its own field-free infinity (the nuclei are not disassembled into nucleons or quarks however). (iii) We have used the electric potential at infinity or at the vacuum level (“Bockris point”) as convenient zeros of potential. (iv) The reader may be bemused and sceptical of our conveniently setting certain standard chemical potentials to zero. We can do this if we allow that the zero of energy is somewhat similar to the DFT zero, but instead we envisage a number of reservoirs of pure substance, solid or ideal gas at one bar and 298K, infinitely far from each other and from which we are drawing our components to make up the substance under study. (v) We have used the single SHE potential as a zero of electric potential. Electrochemists also use other standards such as the standard calomel electrode, but they can easily be calibrated to the SHE by constructing a suitable electrochemical cell.

The question now is, can we find an *absolute* single electrode potential for the SHE relative to the thermodynamic reference (iv)? This cannot be done without using assumptions beyond the equilibrium thermodynamics, but in fact there is an argument first put forward by Sergio Trassati in the 1980s.

We start by looking at section 9.3 in a different way. There we used the fact that the e.m.f. of the cell in figure 9–1 was the difference in electrochemical potentials of the electrons in the right and left hand leads. Instead, pursue the following line of argument. The equilibrium in the left hand cell (9.3.1a) implies equality of the electrochemical potentials,

$$\begin{aligned}\frac{1}{2}\mu_{\text{H}_2} &= \tilde{\mu}_{\text{H}^+} + \tilde{\mu}_{\text{e}}^{\text{Pt}} \\ &= \mu_{\text{H}^+} + F\phi_s + \mu_{\text{e}}^{\text{Pt}} - F\phi_{\text{Pt}} \\ &= \mu_{\text{H}^+}^\square + RT \ln h_{\text{H}^+} + F\phi_s + \mu_{\text{e}}^{\text{Pt}} - F\phi_{\text{Pt}}\end{aligned}$$

This is the *standard* hydrogen electrode, so we'll have unit activity of H^+ and 1 bar

pressure of hydrogen gas. Then after rearranging,

$$F(\phi_{\text{Pt}} - \phi_s) = -\frac{1}{2}\mu_{\text{H}_2}^\ominus + \mu_{\text{H}^+}^\square + \mu_e^{\text{Pt}} \quad (10.2.1)$$

Equilibrium of the right hand cell (9.3.1b) similarly implies the equality of electrochemical potentials on either side of the reaction,

$$\mu_{\text{AgCl}} + \tilde{\mu}_e^{\text{Ag}} = \mu_{\text{Ag}} + \tilde{\mu}_{\text{Cl}^-}$$

and after expanding the electrochemical potentials and rearranging exactly as above,

$$\begin{aligned} F(\phi_s - \phi_{\text{Ag}}) &= \mu_{\text{Ag}}^\bullet - \mu_{\text{AgCl}}^\bullet + \mu_{\text{Cl}^-}^\square + RT \ln h_{\text{Cl}^-} - \mu_e^{\text{Ag}} \\ &= -\mu_{\text{AgCl}}^\bullet + \mu_{\text{Cl}^-}^\square + RT \ln h_{\text{Cl}^-} - \mu_e^{\text{Ag}} \end{aligned} \quad (10.2.2)$$

since Ag is in its pure substance standard state. The open circuit e.m.f. is

$$\begin{aligned} \mathcal{E} &= \phi_{\text{Cu}'} - \phi_{\text{Cu}} \\ &= (\phi_{\text{Pt}} - \phi_{\text{Cu}}) + (\phi_s - \phi_{\text{Pt}}) + (\phi_{\text{Ag}} - \phi_s) + (\phi_{\text{Cu}'} - \phi_{\text{Ag}}) \end{aligned} \quad (10.2.3)$$

In view of (3.2.1) and second footnote, page 33, we have,

$$\begin{aligned} \phi_{\text{Pt}} - \phi_{\text{Cu}} &= \frac{1}{F} (\mu_e^{\text{Pt}} - \mu_e^{\text{Cu}}) \\ \phi_{\text{Cu}'} - \phi_{\text{Ag}} &= \frac{1}{F} (\mu_e^{\text{Cu}'} - \mu_e^{\text{Ag}}) \end{aligned}$$

A key point is that the Fermi levels in both copper leads are the same even though their electrons are at different electrochemical potentials. This is because the Fermi level is the same as the bandwidth (figure 3–1). So $\mu_e^{\text{Cu}'} = \mu_e^{\text{Cu}}$ and (10.2.3) becomes

$$\begin{aligned} \mathcal{E} &= \left((\phi_{\text{Ag}} - \phi_s) - \frac{1}{F} \mu_e^{\text{Ag}} \right) - \left((\phi_{\text{Pt}} - \phi_s) - \frac{1}{F} \mu_e^{\text{Pt}} \right) \\ &= \mathcal{E}_{\text{Ag}|\text{AgCl}} - \mathcal{E}_{\text{H}^+|\text{Pt}} \end{aligned} \quad (10.2.4)$$

Two points are worth noting.

- (i) Comparing (10.2.4) with (10.2.1) and (10.2.2) it is clear that the Fermi levels in each metal cancel. This will always happen in a formula for the full cell e.m.f. but not in the formula for a half cell such as (9.2.1). However because of this cancellation you will see in textbooks such as John West's a statement like, “we assume that electrons in metals are in their standard state so their standard chemical potentials can be set to zero.”
- (ii) In spite of this we leave (10.2.4) in the form that it is because it has separated the e.m.f. into single electrode potentials, each depending only on the properties of its own electrode.

Because of point (ii) we could imagine adding and subtracting a constant, K_{abs} , to (10.2.4) as an attempt to identify an *absolute* electrode potential,

$$\begin{aligned} \mathcal{E}_{\text{Ag}|\text{AgCl}}^{\text{abs}} &= (\phi_{\text{Ag}} - \phi_s) - \frac{1}{F} \mu_e^{\text{Ag}} + K_{\text{abs}} \\ \mathcal{E}_{\text{H}^+|\text{Pt}}^{\text{abs}} &= \mathcal{E}_{\text{SHE}}^{\text{abs}} = (\phi_{\text{Pt}} - \phi_s) - \frac{1}{F} \mu_e^{\text{Pt}} + K_{\text{abs}} \end{aligned}$$

Now, Trassatti's idea is to modify the cell in section 9.3 so as to introduce an air gap into the electrolyte,



This is a thought experiment so maybe it doesn't matter how to realise this in practice, but it is imagined that the Volta potential difference across the air gap can be made to be zero, presumably by some streaming of the electrolyte as in Kenrick's apparatus (section 10.1). In that case the e.m.f. at open circuit is a compensation potential (section 10.1) and in the manner of (10.2.3) this can be written as a sum of potential differences across the phase boundaries indicated by ']' in the cell diagram,

$$\begin{aligned}\Delta_c\phi &= (\phi_{\text{Cu}'} - \phi_{\text{Cu}}) \\ &= (\phi_{\text{Pt}} - \phi_{\text{Cu}}) + (\phi_s - \phi_{\text{Pt}}) + (\phi_{\text{air}} - \phi_s) + (\phi_s - \phi_{\text{air}}) + (\phi_{\text{Ag}} - \phi_s) + (\phi_{\text{Cu}'} - \phi_{\text{Ag}})\end{aligned}$$

Because of zero Volta potential difference across the air gap, we have

$$(\psi_{\text{air}} - \psi_s) = (\psi_s - \psi_{\text{air}}) = 0$$

and so the Galvani potential potential difference is just the dipole potential difference between the HCl electrolyte and air (we know that the surface dipole potential of water is about 80mV, section 10.1, but we don't need to use this number in what follows),

$$(\phi_{\text{air}} - \phi_s) = -(\phi_s - \phi_{\text{air}}) = \chi_s$$

Going through the same steps leading to equations (10.2.1), (10.2.2) and (10.2.4), we arrive at

$$\begin{aligned}\mathcal{E} &= \left((\phi_{\text{Ag}} - \phi_s) - \frac{1}{F}\mu_e^{\text{Ag}} + \chi_s \right) - \left((\phi_{\text{Pt}} - \phi_s) - \frac{1}{F}\mu_e^{\text{Pt}} + \chi_s \right) \\ &= \mathcal{E}_{\text{Ag}|\text{AgCl}}^{\text{abs}} - \mathcal{E}_{\text{H}^+|\text{Pt}}^{\text{abs}}\end{aligned}$$

We can assert that we have now found the absolute single SHE potential because the air gap and the high impedance across the voltmeter have isolated the two half cells. In this case using $K_{\text{abs}} = \chi_s$ identifies the constant needed to refer the SHE to the "thermodynamic" zero of energy. Now using (10.2.1) we get

$$\begin{aligned}\mathcal{E}_{\text{SHE}}^{\text{abs}} &= \phi_{\text{Pt}} - \phi_s - \frac{1}{F}\mu_e^{\text{Pt}} + \chi_s \\ &= \frac{1}{F} \left(\mu_{\text{H}^+}^\ominus - \frac{1}{2}\mu_{\text{H}_2}^\ominus + \mu_e^{\text{Pt}} \right) - \frac{1}{F}\mu_e^{\text{Pt}} + \chi_s \\ &= \mu_{\text{H}^+}^\ominus + \frac{1}{F}\chi_s \\ &= \frac{1}{F}\alpha_{\text{H}^+}^{s,o}\end{aligned}$$

In the second line the Fermi energy of Pt has cancelled as promised. We also take the chemical potential of hydrogen gas in its standard state of one bar pressure and 298K as

zero. The last line follows from (4.2.1) and leaves us with the remarkably simple result that the absolute e.m.f. of the standard hydrogen electrode is the real potential of the proton in the electrolyte. Assuming that this is the same as for water, $\chi_s = \chi_{\text{H}_2\text{O}}$ (and I don't see how we can assume this—1.2 molal HCl is not dilute) then we have already calculated the answer in (10.1.3),

$$\mathcal{E}_{\text{SHE}}^{\text{abs}} = 4.44V$$

This should be a great help in computer simulations of half cells, when comparing to measurements against the SHE. An alternative derivation and a refinement of the answer to 4.42V can be found in Cheng and Sprik.

11. Electrode capacitance and electrocapillarity

One of the most striking observations in electrochemistry is the parallel between the electrical behaviour of the electrode–electrolyte interphase and an equivalent electrical circuit. We are reminded of the mechanical behaviour of polymers which can be mapped onto “circuits” made up of combinations of Maxwell and Voigt elements in series and parallel. The equivalent electrical circuits in electrochemistry are usually combinations of capacitors, resistors and mass transfer impedances; rather similar to the analogy between a mechanical damped oscillator and an LCR electrical circuit. To understand this, we need to look into detail into the structure and thermodynamics of the interphase.

11.1 Thermodynamics of the interphase

Interfacial thermodynamics is a rather difficult subject. In fact it has led some authors and even textbooks into confusion. Here I will follow Cahn and Finnis. The subject was of course rigorously addressed by Gibbs and it is not often that you come across a work that improves on Gibbs. But Cahn is a case in point and although his treatment is complicated and the notation cumbersome, it exposes neatly where error has arisen.

In order to avoid curvature terms in the free enthalpy (see Butt, Graf and Kappl), we will consider only a flat, planar interface between two phases α and β . These may be two electrolytes or an electrolyte and a metal electrode. A key point is that when the two phases are connected to create an interface or “interphase” the chemical potentials of each component throughout the assembly must remain the same as in the bulk separated phases; hence the mole fractions of each component as well as the other extensive quantities will need to adjust. This gives rise to the notion of *excess* extensive properties compared to the amounts in the unconnected phases.[†]

[†] In thermodynamics there are two meanings of “excess”. One is as an excess of some quantity compared to a fictitious system which is ideal, and is not universally regarded as at all useful. The second interpretation is the one we use here and the “excess over the ideal” is not a concept I will use.

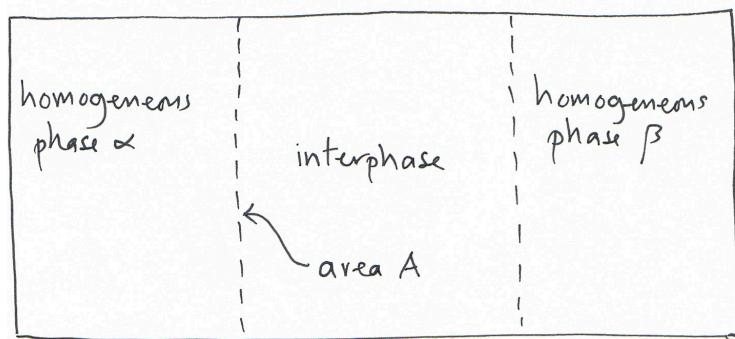


FIGURE 11-1

The solid line in figure 11-1 outlines a piece taken from the assembly of the two phases, α and β , and the interface, and is intended to represent a rectangular or cylindrical portion of the assembly, having cross sectional area A , whose axis is perpendicular to the planar interface which is somewhere between the dotted lines. The central part contains the interface (not shown). Our *system*[†] is the piece of matter enclosed by the boundary. We follow Cahn and Finnis and treat this as a *closed* system (see Appendix, section A2.2). We extend equation (A.25) as follows. If the interface has an area A the combined first and second laws read,

$$dU = TdS - pdV + \mu_i dn_i + \gamma dA \quad (11.1.1)$$

The third term contains the chemical potentials, μ_i (Appendix, (A2.4)) and n_i is the total number of moles of component (species) i in the system outlined by the solid line. The sum over i will be left implicit as in the Einstein summation convention. In the fourth term the interfacial, or surface tension, γ , appears in a work term: if the area is increased by dA then an amount of work γdA is done on the system. It is vital to appreciate what is meant by increasing the area. This may not be done by stretching; new surface must be created. For example, think of two immiscible liquids in an upright test tube. If the tube is tilted then the interfacial area increases; and although you don't realise it, in tilting the tube you have performed the work needed to increase the interfacial area. This is impossible if one or both of the phases is solid. This does introduce a difficulty: we wish to treat the system as closed, but we can't think of even a thought experiment in which the interfacial area is increased at constant volume. Finnis circumvents this issue by allowing that a perfectly legitimate process is simply to redraw the boundary in figure 11-1 by making it narrower and taller so as to keep the total volume fixed but increasing the area, A .

The differentials in (11.1.1) are all extensive quantities and in view of Euler's theorem for homogeneous functions of first order (Appendix, section 2) can be integrated to give,

$$U = TS - pV + \mu_i n_i + \gamma A$$

so that the interfacial tension is seen to be

$$\gamma = \frac{1}{A} (U - TS + pV - \mu_i n_i) = \frac{1}{A} (G - \mu_i n_i) \quad (11.1.2)$$

[†] Recall Bockris's words: "All thermodynamic thinking begins with a definition of the portion of the universe under study, i.e., the system."

This has a simple interpretation. G is the free enthalpy of the system with interphase enclosed by the solid boundary in figure 11–1 and $\mu_i n_i$ is the free enthalpy of the separated parts before assembly in view of the Appendix (A2.10) and in view of the point that before and after assembly the chemical potentials are those of the components in the bulk homogeneous phases. This exposes the interfacial tension as an *excess interfacial free enthalpy*.

Although we have allowed that the choice of boundary in figure 11–1 is arbitrary, for all that follows it is necessary to insist that boundary encloses enough material to either side so that material in the left and right portions has exactly the properties of the homogeneous phases α and β before they were joined. The central piece must be wide enough so that the (dotted) boundary between it and the left and right portions lies in material that has these bulk, homogeneous properties. Under those restrictions only, the width of the system and the width of the central region is arbitrary. You can see this from equation (11.1.2). Suppose I decide to extend the left hand boundary by adding more moles of homogeneous phase α to the system. *This doesn't change the interfacial tension* because an equal amount, say $\mu_i \Delta n_i^\alpha$ is added to both terms in (11.1.2) which cancel out. An equivalent way of seeing this is to ask, what if I have a surface and I add a further layer of atoms? This doesn't change the surface energy because the atoms that *were* on the surface are *now* subsurface and nothing has changed as regards the properties of the surface. A mathematical way to see this is that since in the homogeneous phases we must have $G = n_i \mu_i$ we must have in each phase that

$$\begin{aligned} U^\alpha - TS^\alpha + pV^\alpha - n_i^\alpha \mu_i &= 0 \\ U^\beta - TS^\beta + pV^\beta - n_i^\beta \mu_i &= 0 \end{aligned}$$

This means that we may as well remove as much of the left and right hand slabs of the system as we wish and just leave the central part—its width, of course, is still arbitrary but it must be wide enough so that the matter to the left and right (at the dotted lines in figure 11–1) has the properties of the homogeneous phases. In the appendix, we defined molar extensive quantities by using a lower case symbol, for example v is the molar volume. We now similarly define *layer quantities* by dividing the internal energy, the volume, the entropy and the number of moles of component i in the central, interphase region (*the layer*) by the area A . I will deviate from Cahn's and Finnis's notation and denote these u , v , s and n . Individually, these are meaningless because of the arbitrary choice of width of the layer (the separation of the dotted lines in figure 11–1); however as long as they appear in certain *proper* combinations they can contribute to measurable thermodynamic excess quantities. For example the interfacial tension, using (11.1.2), is

$$\gamma = u - Ts + pv - n_i \mu_i$$

from which, by differentiating and combining with (11.1.1) divided through by A , can be obtained a version of the “Gibbs adsorption equation,”

$$d\gamma = -sdT + vdp - n_i d\mu_i \quad (11.1.3a)$$

Because the layer quantities appear in a *proper* combination this is independent of the size of the layer (interphase) on account of the argument just given which carries through from (11.1.2).

Cahn points out how equation (11.1.3a) can lead one into error. The reader may be tempted to assert that at constant T and p (11.1.3a) becomes the Gibbs adsorption isotherm,[†]

$$\left(\frac{\partial \gamma}{\partial \mu_i} \right)_{T,p,\mu_j(j \neq i)} = -n_i \quad (11.1.3b)$$

but this contains two flaws. Firstly the right hand side is a meaningless layer quantity. Secondly the variation at constant temperature, pressure and chemical potential of all but one species cannot be made without violation of the Gibbs phase rule. We will see below how to write a correct adsorption isotherm. At the heart of this is that any variation must be made by reference to the Gibbs-Duhem equations (A2.11)—one for each phase,

$$\begin{aligned} -S^\alpha dT + V^\alpha dp - n_i^\alpha d\mu_i &= 0 \\ -S^\beta dT + V^\beta dp - n_i^\beta d\mu_i &= 0 \end{aligned} \quad (11.1.4)$$

Each Gibbs-Duhem equation states that for each phase in equilibrium independent variations of the intensive properties is not possible. You might think that if a system comprises C components, each having its chemical potential, and there are two additional intensive variables, T and p , then there will be $C + 2$ degrees of freedom (number of intensive variables that can be adjusted independently while remaining in equilibrium). However for each phase there is a Gibbs-Duhem equation which removes a degree of freedom so the Gibbs phase rule (discovered by Gibbs between 1875 and 1878) states that the number of degrees of freedom is actually

$$F = C + 2 - P$$

where P is the number of phases (number of constraints). For example a binary system consisting of two components and four phases (say, two solid alloy compositions and a liquid and gas phase all in equilibrium) can only exist at a single temperature and pressure and at fixed values of the two chemical potentials. Pure water has one component, so at the triple point where liquid, solid and vapour are in equilibrium there are $F = 1 + 2 - 3 = 0$ degrees of freedom and so the temperature and pressure at the triple point is fixed, as is well known. Along the coexistence line separating liquid and vapour between the triple point and the critical point there are two phases and so just *one* of either T or p can be adjusted and the other is then determined by the requirement to remain on the coexistence line. These are all examples of the phase rule at work.

The case of a single phase

Before embarking on the full machinery of Cahn's reformulation of Gibbs's interfacial thermodynamics, we follow Finnis and consider the special case of a single phase. This

[†] and indeed Schmickler and Santos do exactly that in their equation (8.5). What these authors actually write, in our notation, is $d\gamma = \sum_i n_i d\mu_i$ in which the right hand side depends on the size of the layer, but the left hand side doesn't. Butt, Graf and Kappl have a related issue in their (3.46), in which the excess depends on the placement of the dividing surface, but $d\gamma$ does not.

covers a number of simple cases: (*i*) a grain boundary, twin, stacking fault, antiphase boundary where there is the same crystal on either side of the interface, and components may segregate to the boundary in order to maintain the same chemical potential as in the homogeneous phases, resulting in an *excess* interfacial concentration; (*ii*) the surface separating phase α with vacuum; (*iii*) the electrolyte–metal interphase. The last example is perhaps surprising but as long as the components of interest (ions, water) do not dissolve in the metal then the quantities n_i^β are zero if we denote the metal phase by β so it effectively acts like vacuum. Therefore this simplification (and it's hard enough) is both revealing and relevant.

We have for the phase α the Gibbs adsorption equation, and the Gibbs-Duhem equation,

$$\begin{aligned} d\gamma &= -sdT + vdp - n_i d\mu_i \\ 0 &= -S^\alpha dT + V^\alpha dp - n_i^\alpha d\mu_i \end{aligned} \quad (11.1.5)$$

I cannot suppress the superscript α even though there is only one phase because I must distinguish between S , V and n_i which are the total entropy, volume and number of moles of component i of the system enclosed by the boundary in figure 11–1, and S^α , V^α and n_i^α which are the entropy, volume and number of moles of component i of the homogeneous phase, α , to the left of the dotted line in figure 11–1. Now, multiply the second equation (11.1.5) by n_1/n_1^α in which we choose for component 1 the solvent, or majority species. Then subtract the result from the first equation which evolves into,

$$d\gamma = -\left(s - \frac{n_1}{n_1^\alpha} S^\alpha\right) dT + \left(v - \frac{n_1}{n_1^\alpha} V^\alpha\right) dp - \left(n_i - \frac{n_1}{n_1^\alpha} n_i^\alpha\right) d\mu_i \quad (11.1.6)$$

Note that the implicit sum now effectively excludes $i = 1$ because that term is zero so the constraint has removed our freedom to vary the chemical potential of the (solvent) component 1. Now we are safe and the $C + 1$ variations of intensive quantities can be made independently; the chemical potential of the solvent will have to adjust itself in order to be consistent with the Gibbs-Duhem relation. The quantities in parentheses are known as the *interfacial excesses with respect to component 1*. In particular the molar excess of component i with respect to component 1 is

$$\Gamma_i = n_i - \frac{n_1}{n_1^\alpha} n_i^\alpha \quad (11.1.7a)$$

The solvent excess is by construction zero. Γ_i is exactly the same definition of excess as given by Schmickler and Santos in their equation (8.7). It also the same as the “relative adsorption” defined by Butt, Graf and Kappl in their equation (3.11) in the case that concentrations in the β -phase are zero. These latter authors derive this using the Gibbs dividing surface and so their chapter is instructive to read.

For a solution in contact with gas phase,[†] the Gibbs adsorption isotherm follows from (11.1.6), namely

$$\left(\frac{\partial \gamma}{\partial \mu_i}\right)_{T,p,\mu_j(j>1,j\neq i)} = -\Gamma_i \quad (i > 1) \quad (11.1.8)$$

[†] More strictly, vacuum. We have to assert that the surface is in equilibrium with the underlying phase but not with the gas phase—otherwise we may not assume a single, α , phase

This is a correct variation of γ which can be made without violating the phase rule, and the right hand side is independent of the size of the layer. An example of the Gibbs adsorption isotherm at work is the effect of soap on water. If a small amount of soap, component i , is dissolved in water then if it segregates to the surface its excess, Γ_i , will be positive. If I now add more soap then the chemical potential will increase in accord with the Appendix, equation (A.3.1). The minus sign in the adsorption isotherm then serves to decrease the surface tension, as expected from experience.

Let us journey to understand the excess quantities in greater inwardness (thanks to Volker Heine for the phrase). The first term in the parentheses in (11.1.6) is the amount of entropy per unit area in the interphase region, take away n_1/n_1^α times the total entropy contained in the homogeneous phase α to left of the interphase region. Now n_1/n_1^α is $1/A$ times the number of moles of component 1 (solvent) in the interphase region divided by the number of moles of solvent in the homogeneous region to the left. This leads us to the definition of the excess entropy as the difference between the amount of entropy of unit area of the layer, or interphase, and the amount of entropy of homogeneous phase having the same amount of solvent. Similarly, the excess, Γ_i , in (11.1.7a) is $1/A$ times the *number of moles of i in the layer*, take away the *number of moles of i in a piece of homogeneous phase, α , having the same number of moles of solvent as are in the layer*. Clearly if $i = 1$ this is zero by construction so there is no excess of solvent. We can express the excess in general using Cahn's notation, and our rule that layer quantities are denoted using lower case sans serif font,

$$[Z/X] = z - \frac{x}{X^\alpha} Z^\alpha \quad (11.1.9)$$

again: the amount of extensive quantity Z belonging to unit area of interphase compared to the amount of Z in homogeneous phase containing the same amount of extensive quantity X as the interphase. Hence this is the excess of Z with respect to X . It is helpful in view of what follows to point out that mathematically we may write,

$$[Z/X] = \frac{1}{X^\alpha} \begin{vmatrix} z & x \\ Z^\alpha & X^\alpha \end{vmatrix} \quad (11.1.10)$$

as you can see by expanding out the determinant. In view of this notation, I can write the excess in (11.1.7a) as

$$\Gamma_i = [n_i/n_1] \quad (11.1.7b)$$

meaning the excess of species i with respect to species 1 (the solvent usually). Alternatively it is the excess of species i having eliminated the extensive quantity conjugate to μ_1 . Strictly speaking we should indicate the constraint used in the notation: the right hand side is unambiguous, but the left hand side would be written more completely as, say, $\Gamma_i^{(1)}$ as done by Butt, Graf and Kappl in their equation (3.11). After all I could have decided to eliminate the pressure from equations (11.1.5) by multiplying the Gibbs-Duhem equation by v/V^α and subtracting from the Gibbs adsorption equation. The result, equivalent to (11.1.6) is

$$\begin{aligned} d\gamma &= - \left(s - \frac{v}{V^\alpha} S^\alpha \right) dT - \left(n_i - \frac{v}{V^\alpha} n_i^\alpha \right) d\mu_i \\ &= - [S/V] dT - [n_i/V] d\mu_i \end{aligned}$$

This introduces an excess

$$\Gamma_i^{(V)} = [n_i/V] = \left(n_i - \frac{v}{V^\alpha} n_i^\alpha \right)$$

which is the difference between the amount of component i in unit area of interphase and the amount of component i in an equal volume of homogeneous phase. This is independent of the volume of the interphase and so is a perfectly legitimate measure of the amount of segregation which does not single out a particular component as the solvent.

The case of two phases

As I have said, for the metal-electrolyte interphase it is enough to consider the case of a single phase. Therefore at a first reading this section may be skipped without loss in the following sections. However for completeness, let us now pursue the instance when the components i are distributed in equilibrium in both phases, α and β , to left and right of the interphase region in figure 11–1. Take a deep breath, and read on.

By comparison with (11.1.5) and taking in view that there are now two phases, there are two Gibbs-Duhem equations and hence two constraints on the variations that may be made in the intensive quantities. We have these three equations,

$$d\gamma = -sdT + vd\mu_i \quad (11.1.11a)$$

$$0 = -S^\alpha dT + V^\alpha dp - n_i^\alpha d\mu_i \quad (11.1.11b)$$

$$0 = -S^\beta dT + V^\beta dp - n_i^\beta d\mu_i \quad (11.1.11c)$$

We label the extensive quantities in the homogeneous phases with superscripts α and β ; we do not need to label T , p and the chemical potentials because they are the same for each component throughout the system in equilibrium. The two Gibbs-Duhem equations prescribe two constraints on the $C + 2$ intensive variables $\{T, p, \mu_i\}$; alternatively they may be solved to tell us how any two must evolve if we freely vary the others. Cahn analyses the three equations (11.1.11) using the properties of determinants. First, following Finnis, write these as three simultaneous equations,

$$\begin{aligned} d\gamma &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1C+2}x_{C+2} \\ 0 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2C+2}x_{C+2} \\ 0 &= a_{31}x_1 + a_{32}x_2 + \cdots + a_{3C+2}x_{C+2} \end{aligned} \quad (11.1.12)$$

where x stands for the differential of one of the $C + 2$ intensive variables and a stands for one of the $3C + 6$ extensive quantities $\{s, v, n_i, S^\alpha, V^\alpha, n_i^\alpha, S^\beta, V^\beta, n_i^\beta\}$. We decide to eliminate two of the intensive variables, $\{T, p, \mu_i\}$, represented by x_ℓ and x_k . First form a matrix,

$$\mathbf{c} = \begin{pmatrix} c_{11} & a_{1\ell} & a_{1k} \\ c_{21} & a_{2\ell} & a_{2k} \\ c_{31} & a_{3\ell} & a_{3k} \end{pmatrix}$$

whose entries c are formed from one of the other columns of equations (11.1.12). The determinant of \mathbf{c} is constructed from its cofactors,

$$\det \mathbf{c} = \begin{vmatrix} c_{11} & a_{1\ell} & a_{1k} \\ c_{21} & a_{2\ell} & a_{2k} \\ c_{31} & a_{3\ell} & a_{3k} \end{vmatrix} = c_{11}C_{11} + c_{21}C_{21} + c_{31}C_{31}$$

where the cofactors are

$$C_{11} = a_{2\ell}a_{3k} - a_{2k}a_{3\ell}; \quad C_{21} = -(a_{1\ell}a_{3k} - a_{1k}a_{3\ell}); \quad C_{31} = a_{1\ell}a_{2k} - a_{1k}a_{2\ell}$$

I now multiply each of equations (11.1.12) in turn by one of the cofactors,

$$\begin{aligned} C_{11}d\gamma &= (a_{2\ell}a_{3k} - a_{2k}a_{3\ell})(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1C+2}x_{C+2}) \\ 0 &= -(a_{1\ell}a_{3k} - a_{1k}a_{3\ell})(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2C+2}x_{C+2}) \\ 0 &= (a_{1\ell}a_{2k} - a_{1k}a_{2\ell})(a_{31}x_1 + a_{32}x_2 + \cdots + a_{3C+2}x_{C+2}) \end{aligned}$$

Add these three together and expand them out; and consider just the three terms in x_1 ,

$$\begin{aligned} C_{11}d\gamma &= (a_{2\ell}a_{3k} - a_{2k}a_{3\ell})a_{11}x_1 - (a_{1\ell}a_{3k} - a_{1k}a_{3\ell})a_{21}x_1 + (a_{1\ell}a_{2k} - a_{1k}a_{2\ell})a_{31}x_1 + \cdots \\ &= \begin{vmatrix} a_{11} & a_{1\ell} & a_{1k} \\ a_{21} & a_{2\ell} & a_{2k} \\ a_{31} & a_{3\ell} & a_{3k} \end{vmatrix} x_1 + \cdots \end{aligned}$$

and the second line follows from the same cofactor property of determinants as above. The same will hold if I single out the terms in x_2 , x_3 and so on, so it follows that

$$C_{11}d\gamma = \sum_{n=1}^{C+2} D_n x_n \quad (11.1.13)$$

where D_n is the determinant

$$D_n = \begin{vmatrix} a_{1n} & a_{1\ell} & a_{1k} \\ a_{2n} & a_{2\ell} & a_{2k} \\ a_{3n} & a_{3\ell} & a_{3k} \end{vmatrix} = a_{1n}C_{11} + a_{2n}C_{21} + a_{3n}C_{31}$$

This is the equivalent of (11.1.6) after we multiplied (11.1.5) by a factor and subtracted it from the adsorption equation. Now, when there are two phases, we have taken an appropriate linear combination of (11.1.11b) and (11.1.11c) and added it to (11.1.11a). In the case of one phase we eliminated one of the chemical potentials; now we have two intensive variables to eliminate, let's choose $x_\ell = d\mu_1$ and $x_k = d\mu_2$.[†] With a little thought, by comparison with (11.1.12) you can see that in this case

$$\begin{aligned} a_{1\ell} &= n_1 \\ a_{2\ell} &= n_1^\alpha \\ a_{3\ell} &= n_1^\beta \\ a_{2k} &= n_2 \\ a_{2k} &= n_2^\alpha \\ a_{3k} &= n_2^\beta \end{aligned}$$

[†] You don't have to choose 1 and 2. If you want you could use, say, m and n for 1 and 2 in what follows; but it's up to me how I number my components so I may as well place the ones I want to eliminate at the top of the list!

and substituting the actual names of variables into (11.1.13), we have

$$C_{11} = n_1^\alpha n_2^\beta - n_2^\alpha n_1^\beta = \begin{vmatrix} n_1^\alpha & n_2^\alpha \\ n_1^\beta & n_2^\beta \end{vmatrix}$$

and

$$d\gamma = - \frac{\begin{vmatrix} s & n_1 & n_2 \\ S^\alpha & n_1^\alpha & n_2^\alpha \\ S^\beta & n_1^\beta & n_2^\beta \end{vmatrix}}{\begin{vmatrix} n_1^\alpha & n_2^\alpha \\ n_1^\beta & n_2^\beta \end{vmatrix}} dT + \frac{\begin{vmatrix} v & n_1 & n_2 \\ V^\alpha & n_1^\alpha & n_2^\alpha \\ V^\beta & n_1^\beta & n_2^\beta \end{vmatrix}}{\begin{vmatrix} n_1^\alpha & n_2^\alpha \\ n_1^\beta & n_2^\beta \end{vmatrix}} dp - \frac{\begin{vmatrix} n_i & n_1 & n_2 \\ n_i^\alpha & n_1^\alpha & n_2^\alpha \\ n_i^\beta & n_1^\beta & n_2^\beta \end{vmatrix}}{\begin{vmatrix} n_1^\alpha & n_2^\alpha \\ n_1^\beta & n_2^\beta \end{vmatrix}} d\mu_i \quad (11.1.14)$$

This equation is rendered more compactly by Cahn's notation. But before showing that consider these three points.

1. The multiplier to each $d\mu_i$ is an excess concentration with respect to two components 1 and 2 (cf (11.1.9) and (11.1.10) expressed as 2×2 determinants divided by "1 \times 1 determinants").
2. The implicit sum over all components remains in the third term as it does in (11.1.6); but while in that case the first term, $i = 1$, vanished; now the *first two* terms, $i = 1$ and $i = 2$, vanish by a property of a determinant—namely that it is zero if any two rows or columns are identical.
3. A second property of determinants guarantees that the numerators in (11.1.14) are independent of the width of the layer. For the value of a determinant is unchanged if a multiple of one row (or column) is added to another row (or column). If I move the left hand dotted line in figure 11–1 to the left, then since this lies in homogeneous phase I am adding a multiple of the second row to the first row; or of the third row to the first if I move the right hand dotted line to the right.

In fact all the ratios of determinants in (11.1.14) are *excess* extensive quantities. Cahn extends the notation of (11.1.10) to allow the elimination of two variables,

$$[Z/XY] = \frac{\begin{vmatrix} z & x & y \\ Z^\alpha & X^\alpha & Y^\alpha \\ Z^\beta & X^\beta & Y^\beta \end{vmatrix}}{\begin{vmatrix} X^\alpha & Y^\alpha \\ X^\beta & Y^\beta \end{vmatrix}} \quad (11.1.15)$$

In Cahn's words, "The physical meaning of $[Z/XY]$ is the difference in the amount of Z between a unit area of the layer and two portions of the homogeneous phase having the same total amount of X and Y as the layer." He means, one portion each of the homogeneous phases α and β . Now we can write (11.1.14) more neatly,

$$d\gamma = - [S/n_1 n_2] dT + [V/n_1 n_2] dp - [n_i/n_1 n_2] d\mu_i$$

The sum over i is still implicit but by construction $[n_1/n_1 n_2]$ and $[n_2/n_1 n_2]$ are zero as observed above. For this case, our excess concentration is

$$I_i = [n_i/n_1 n_2] \quad (11.1.7c)$$

and the Gibbs adsorption isotherm becomes

$$\left(\frac{\partial \gamma}{\partial \mu_i} \right)_{T,p,\mu_j(j>2,j\neq i)} = -\Gamma_i \quad (i > 2) \quad (11.1.16)$$

This fixes both of the errors associated with equation (11.1.3b). Firstly in the left hand side the variation can be made without violating the phase rule because two of the intensive variables, μ_1 and μ_2 , will adjust themselves in order to maintain equilibrium and we do not insist that they are to remain constant. Secondly the right hand side is independent of the width of the layer as proved in point 3, above. The interfacial excess in (11.1.16) is not the same as the Γ_i in equation (11.1.7b)—as I have pointed out Cahn's notation is unambiguous, and really we should decorate our Γ 's with suitable superscripts. In fact the symbol Γ was used by Gibbs himself for the excess; his Γ corresponds to our n and Schmickler and Santos use Γ^* for this. Cahn points out that his extension has a number of benefits compared to Gibbs's treatment; one of which is that the “volume”, v , of Gibbs's interface is zero and indeed Gibbs regards the interface as a mathematical “dividing surface” (this can be seen in the development by Butt, Graf and Kappl). Another benefit is that for Gibbs, but not necessarily for Cahn, one of the X or Y in the excess (11.1.15) *must* be volume. Let's pursue that a bit. Suppose I would like to eliminate $d\mu_1$ and dp . Then my excess is

$$\Gamma_i = [n_i/n_1 V] = \frac{\begin{vmatrix} n_i & n_1 & v \\ n_i^\alpha & n_1^\alpha & V^\alpha \\ n_i^\beta & n_1^\beta & V^\beta \end{vmatrix}}{\begin{vmatrix} n_1^\alpha & V^\alpha \\ n_1^\beta & V^\beta \end{vmatrix}} \quad (11.1.7d)$$

Gibbs called this excess $\Gamma_{i(1)}$. If you set $n_i^\beta = 0$, $\forall i$, and then work out the determinants you'll find

$$\Gamma_i = n_i - \frac{n_1}{n_1} n_i^\alpha$$

which is the same as (11.1.7a). This is to be expected because $n_i^\beta = 0$ indicates that there are no components in the β -phase. This would be the case for the electrolyte–electrode interface if the components 1 and 2 are not to be found in the metal.

11.2 The electrocapillarity and Lippmann equations

The Lippmann equation states how the interfacial tension of an electrode depends on its Galvani potential. The interest arises from experiments made nearly 150 years ago on mercury electrodes. Almost all such experiments are made using mercury because this is the only pure liquid metal (and amalgams can be used to study effects of alloying) and hence the only metal whose surface tension can be measured using its capillarity. In an electrochemistry experiment the surface tension of mercury can be measured as a function of the composition of the electrolyte with which it is in contact and its Galvani potential compared to a standard potential. Important results which are derived below are the capillary equation

$$d\gamma = -\Gamma_i d\mu_i - \sigma d\phi$$

where σ is the surface charge density on the metal electrode, which relates the change in interfacial tension to changes in composition and inner potential; and from which follows the Lippmann equation

$$\left(\frac{\partial \gamma}{\partial \phi} \right)_{\mu} = -\sigma$$

and the result

$$\Gamma_i = \left(\frac{\partial \gamma}{\partial \mu_i} \right)_{\phi, \mu_j \neq i}$$

which permits measurement of the relative adsorption of species to the metal surface. An important measurable quantity is the *potential of zero charge* (PZC) which is the value of ϕ for which the left hand side of the Lippmann equation is zero. Details of how the measurements are made can be found in Schmickler and Santos, and Bockris. Unless the electrode is liquid Hg, it's not possible to measure the surface tension directly. However a relative surface tension can be measured by integration of the surface charge over the electric potential, starting at a reference potential more negative than the PZC,

$$\gamma(\phi) - \gamma(\phi_{\text{ref}}) = \int_{\phi_{\text{ref}}}^{\phi} \sigma(\phi') d\phi'$$

while the surface charge density, σ , is measured by integrating from the PZC (chrono-coulometric (!) method) or integration over the capacity—see ahead, equation (11.2.5).

There are at least two ways to derive the Lippmann equation and it will be worthwhile to include two here: the first as given by Bockris and the second adapted from Schmickler and Santos, and Butt, Graf and Kappl. A further key point is that the Lippmann equation permits the mapping of the electric properties of the interphase to a flat plate capacitor of capacity, C [$\mu\text{F}/\text{cm}^2$], which may or may not be a function of the Galvani potential. The measured dependence of C upon ϕ provides information on the extent of adsorption of ions from the electrolyte on the metal electrode surface.

Derivation number one

Recall the electrochemical cell in figure 2–1. Metal which is dipped into electrolyte may become charged. Imagine that figure 11–1 represents a section through the metal–electrolyte interphase. The system enclosed within the boundary in figure 11–1 is electrically neutral. I will re-derive equation (11.1.3a). The combined first and second laws read

$$dU = TdS - pdV + \mu_i dn_i + \gamma dA + \phi dq \quad (11.2.1)$$

which is the same as (11.1.1) but there is one additional work term which is the work done on the system by shovelling an amount of charge, dq , across the interphase through a potential difference $\phi = \phi_s - \phi_m$. As usual, we integrate (11.2.1),

$$U = TS - pV + \mu_i n_i + \gamma A + \phi q$$

Therefore

$$\gamma A = U - TS + pV - \mu_i n_i - \phi q$$

and taking the total differential and comparing to (11.2.1)

$$d(\gamma A) = \gamma dA + Ad\gamma = \gamma dA - SdT + Vdp - n_i d\mu_i - q d\phi$$

so that

$$Ad\gamma = -SdT + Vdp - n_i d\mu_i - q d\phi$$

We are allowed to construct excesses assuming a single phase as in (11.1.6) because none of the species i reside in the metal electrode. Therefore the Gibbs adsorption equation in the case of a charged electrode is

$$d\gamma = -[S/n_1]dT + [V/n_1]dp - [n_i/n_1]d\mu_i - q d\phi$$

and at constant T and p we have

$$d\gamma = -\Gamma_i d\mu_i - \sigma d\phi \quad (11.2.2)$$

where Γ_i is given by (11.1.7a) and $\sigma = q/A$ is the surface charge density. This is called the *electrocapillary equation*. There is no difficulty in asserting that σ is an unambiguous excess interfacial charge because the amount of charge transferred is exactly equal to the excess of electrons compared to the neutral metal and these all reside in a uniform sheet at the electrode surface. I repeat that ϕ is the Galvani or inner potential difference across the electrode–electrolyte interface and this is of course not measurable. However as in figure 2–4 we can imagine that the cell is completed by connecting to a perfectly reversible, non polarisable standard hydrogen electrode and we can assume that the left hand electrode is perfectly polarisable (that is, no current can leak into the electrolyte as the voltage, V , is observed to change). The measured voltage is

$$\begin{aligned} V &= (\phi_{Cu'} - \phi_{Cu}) \\ &= (\phi_{Cu'} - \phi_{Pt}) + (\phi_{Pt} - \phi_s) + (\phi_s - \phi_m) + (\phi_m - \phi_{Cu}) \end{aligned}$$

To be more general I replace Zn with an unspecified electrode metal with inner electric potential ϕ_m . The metal–metal inner potential differences are all differences in Fermi energy (see equation (3.2.1)) and therefore constant. This leaves

$$\begin{aligned} dV &= d(\phi_{Pt} - \phi_s) + d(\phi_s - \phi_m) \\ &= d(\phi_{Pt} - \phi_s) + d\phi \end{aligned}$$

to put into (11.2.2),

$$d\gamma = -\sigma dV - \Gamma_i d\mu_i + \sigma d(\phi_{Pt} - \phi_s)$$

For the last term on the right hand side we can use the Nernst equation (9.2.1),

$$d(\phi_{Pt} - \phi_s) = \frac{RT}{F} d \ln a_{H^+} = \frac{RT}{F} d\mu_{H^+}$$

where a_{H^+} and μ_{H^+} are the activity and chemical potential of the proton that carries the current leakage across the non polarisable electrode interphase. F is the Faraday constant. Finally, then,

$$d\gamma = -\sigma dV - \Gamma_i d\mu_i - \sigma \frac{RT}{F} d\mu_{H^+} \quad (11.2.3)$$

In an experiment, the concentration of electrolyte and activity of hydrogen ions in the SHE will be controlled and held constant and so (11.2.3) becomes a form of the Lippmann equation,

$$\left(\frac{\partial \gamma}{\partial V} \right)_\mu = -\sigma \quad (11.2.4)$$

If I differentiate a second time,

$$\left(\frac{\partial^2 \gamma}{\partial V^2} \right)_\mu = - \left(\frac{\partial \sigma}{\partial V} \right)_\mu = -C \quad (11.2.5)$$

this defines the “differential capacity” of the interphase between electrode and electrolyte. Until we study the molecular structure in detail we can only assert that the layer of electrons at the metal surface and the layer of equal and opposite charge somewhere in the solution are modelled as a flat plate capacitor with capacity C . Because this must be positive, (11.2.4) must have the shape of a parabola (if C is constant) having a *maximum* at $\sigma = 0$ —see figure 11–2, taken from the publication by David C. Grahame in *Chem. Rev.*, **41**, 441 (1947). The potential V at the maximum is called *the potential of zero charge*, PZC.

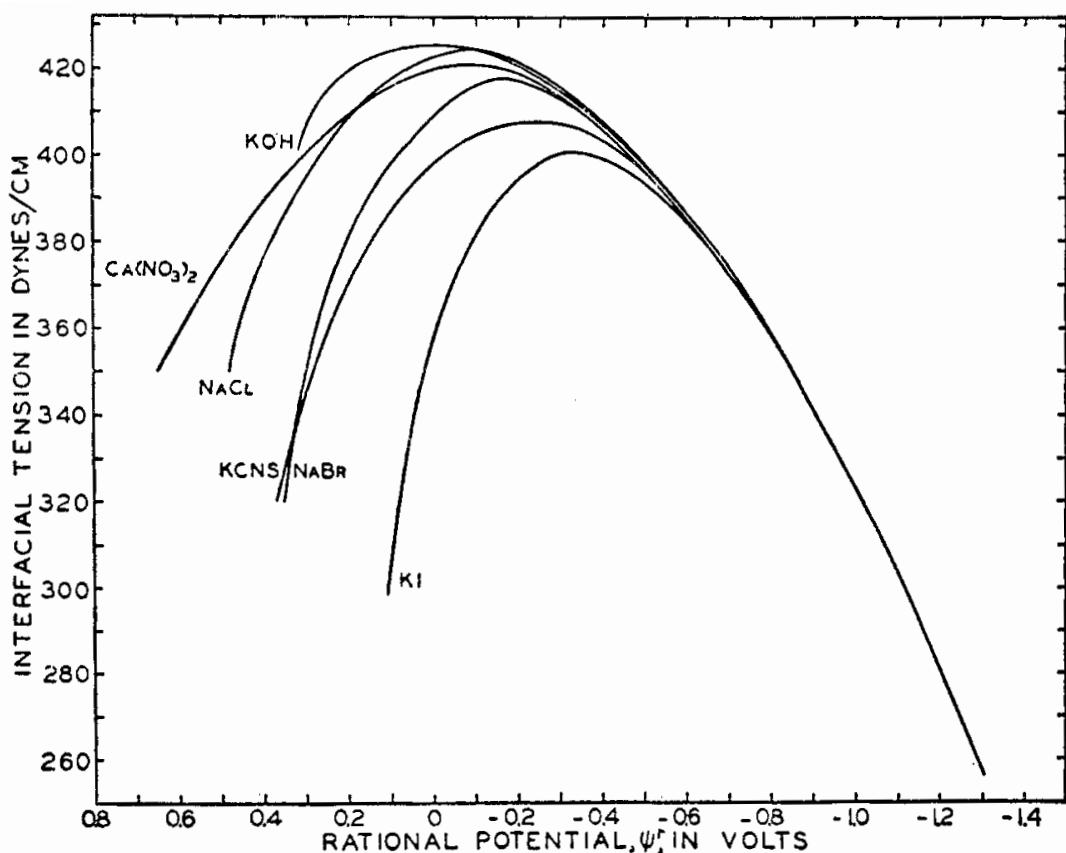


FIG. 1. Interfacial tension of mercury in contact with aqueous solutions of the salts named. $T = 18^\circ\text{C}$. Abscissas are measured relative to a “rational” scale in which the potential difference between the mercury and a capillary-inactive electrolyte is arbitrarily set equal to zero at the electrocapillary maximum.

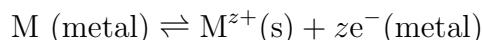
FIGURE 11–2

There follows a table of potentials of zero charge, in V/SHE, for metals in a 0.01 molal solution of either potassium chloride or sodium sulphate.

Al	Bi	Cd	Pb	Co	Cu	Au	Ag
KCl	KCl	KCl	KCl	Na ₂ SO ₄			
-0.52	-0.36	-0.92	-0.69	-0.32	+0.03	+0.23	-0.70

Derivation number two

Since we already have the Lippmann equation, this section can be skipped on a first reading. But it is instructive to follow this second derivation because it does not require the additional work term in (11.2.1). Instead the charge and potential are allowed to appear as a consequence of using electrochemical potential in the Gibbs adsorption isotherm. The picture that we need of the interphase requires us to imagine the metal phase to be inert and impenetrable to the molecules and ions in the electrolyte, so these all have excesses, Γ_i as given in (11.1.7b). But the metal is dissolving into its ions and so the equilibrium,



is established in the interphase. In terms of electrochemical potential the Gibbs adsorption isotherm is,

$$d\gamma = -\Gamma_i d\tilde{\mu}_i - \Gamma_e d\tilde{\mu}_e \quad (11.2.6)$$

The index i labels all of the species in solution, including M^{z+} , charged or not; $i = 1$ refers to water, and excesses, Γ_i , are relative to solvent. The “excess” of electrons per unit area is not controversial since this is clearly the number of moles of electrons over and above those that make the metal neutral, and are all confined to a sheet at the surface of the electrode. I have omitted the contribution from the “excess” of metal, $-\Gamma_M d\mu_M$, because the chemical potential of a pure metal is constant (standard state) $d\mu_M = 0$. We now break up (11.2.6) into its terms. The second term is, in view of (3.4.4),

$$-\Gamma_e d\tilde{\mu}_e = -\Gamma_e d\mu_e + F\Gamma_e d\phi_m = -\sigma d\phi_m$$

because μ_e is the Fermi level and is constant, independent of the inner potential (see figure 3–1), so $d\mu_e = 0$. $-F\Gamma_e = \sigma$ is the excess charge per unit area on the surface of the metal in contact with the solution. The first term on the right hand side of (11.2.6) is, in accord with (4.1.1),

$$\begin{aligned} -\Gamma_i d\tilde{\mu}_i &= -\Gamma_i d\mu_i - F\Gamma_i z_i d\phi_s \\ &= -\Gamma_i d\mu_i + \sigma d\phi_s \end{aligned}$$

where z_i are the charge numbers of species i . The second line follows because the total excess charge in the electrolyte is just sum of excesses of all the charged ionic species including M^{z+} , and this must equal σ to guarantee overall charge neutrality. In fact these are the ions that are accumulated in the double layer, attracted to the charged metal surface. Hence,

$$\sigma = -F\Gamma_i z_i$$

and we can retain the sum over all i because uncharged species have $z_i = 0$ and won’t contribute to the sum. (11.2.6) now becomes,

$$d\gamma = -\sigma d\phi - \Gamma_i d\mu_i$$

where $\phi = \phi_m - \phi_s$ is the Galvani potential difference between electrolyte and electrode. Once again this leads us to the Lippmann equation,

$$\left(\frac{\partial \gamma}{\partial \phi} \right)_\mu = -\sigma$$

this time without making contact with a reference electrode. Here, and in (11.2.4) and (11.2.5) we must be careful with the “constant μ ” subscript. I have been deliberately vague. Not all the chemical potentials can be held constant as we saw in section 11.1. So I mean the partial derivative to be taken with all chemical potentials being constant that can be held fixed. In the case that we regard the metal phase as being inert, this means the excesses are given by (11.1.7a) and the solvent chemical potential is eliminated and cannot be controlled.

11.3 Measurement

The principal difference in comparing derivations number one and two are in how the charges and potential turn up in the combined first and second law—firstly as an explicit work term, and secondly by writing the adsorption isotherm in terms of electrochemical rather than chemical potential. The first derivation, which I think was Lippmann’s original in outline, is based purely in macroscopic thermodynamics whereas the second has a closer focus on the ions in the electrolyte; in fact we included explicitly the presence of those ions which appear as a result of corrosion of the electrode. These could, of course, be absent; as they are in most electrocapillary experiments using mercury. An example of such an experiment is shown in figure 11–2.

It is typical that when the electrode potential is most negative and the electrode is negatively charged the curve is close to parabolic and is the same for all salts. This is because in this condition it is the *cation* that is attracted to the surface of the metal. Conversely when the metal is positively charged, one speaks of “specific adsorption” and the structure of the anions in the double layer causes a variation in the capacitance. The explanation usually given is that the solvation of a cation is much stronger than of an anion. This is because the cation is small and so is easily surrounded by water molecules. Therefore the cation is protected by its solvation shell from being physically adsorbed on the metal surface. The anion, on the other hand, is more easily freed from its solvation shell and so adsorbs rather strongly.[†] This is clearly observed in impedance spectroscopy and cyclic voltammetry. An example, taken from G. Beltramo and Elizabeth Santos, *J. Electroanalytical Chem.*, **556**, 127 (2003) is shown in figure 11–3. The abscissa shows the voltage relative to the standard calomel electrode (Appendix, section 3.6). The two upper figures show the capacity in the ordinate. The two lower figures are *cyclic voltammograms*, showing the current density as a function of voltage. The electrode is

[†] Schmickler and Santos make much mention of “inner” and “outer” sphere reactions. I think that inner sphere refers to the case in which the ion is not completely screened by its solvation shell—*i.e.*, it is adsorbed at the electrode; and outer sphere reactions involve a fully solvated ion, presumably some distance from the electrode, in the electrolyte.

a single crystal silver (111) surface; the left hand figures are measurements in chloride electrolyte and the right hand figures in bromide. The solid and broken lines correspond to concentrations of 0.1 milli-molar and 10 nM respectively. The dotted line is in the absence of chloride or bromide. The upper curves are examples of “impedance spectroscopy”: the voltage is increased at a rate of 10 mV s^{-1} and is impressed over an alternating voltage of 20 Hz. To me, the experiment is reminiscent of the LCR-circuit which is the electrical analogue of a typical resonant system and comprises a resistor, a capacitor and a choke in series. The differential capacity–potential curves in this experiment are modelled by a simple equivalent circuit of a resistor and capacitor *in series* (not in parallel as in figure 2–6) which represents the electrolyte by a resistor and the polarisable interphase by a capacitor. The paper by Beltramo and Santos is well worth reading. The capacity–potential curve will be discussed further in the next section. For brevity I will just state here that the interpretation of the broad and narrow peaks in the cyclic voltammograms are interpreted as the result of the growth of disordered and ordered adsorbed anions on the metal surface; and the ordered structure is confirmed by scanning tunnelling microscopy experiments.

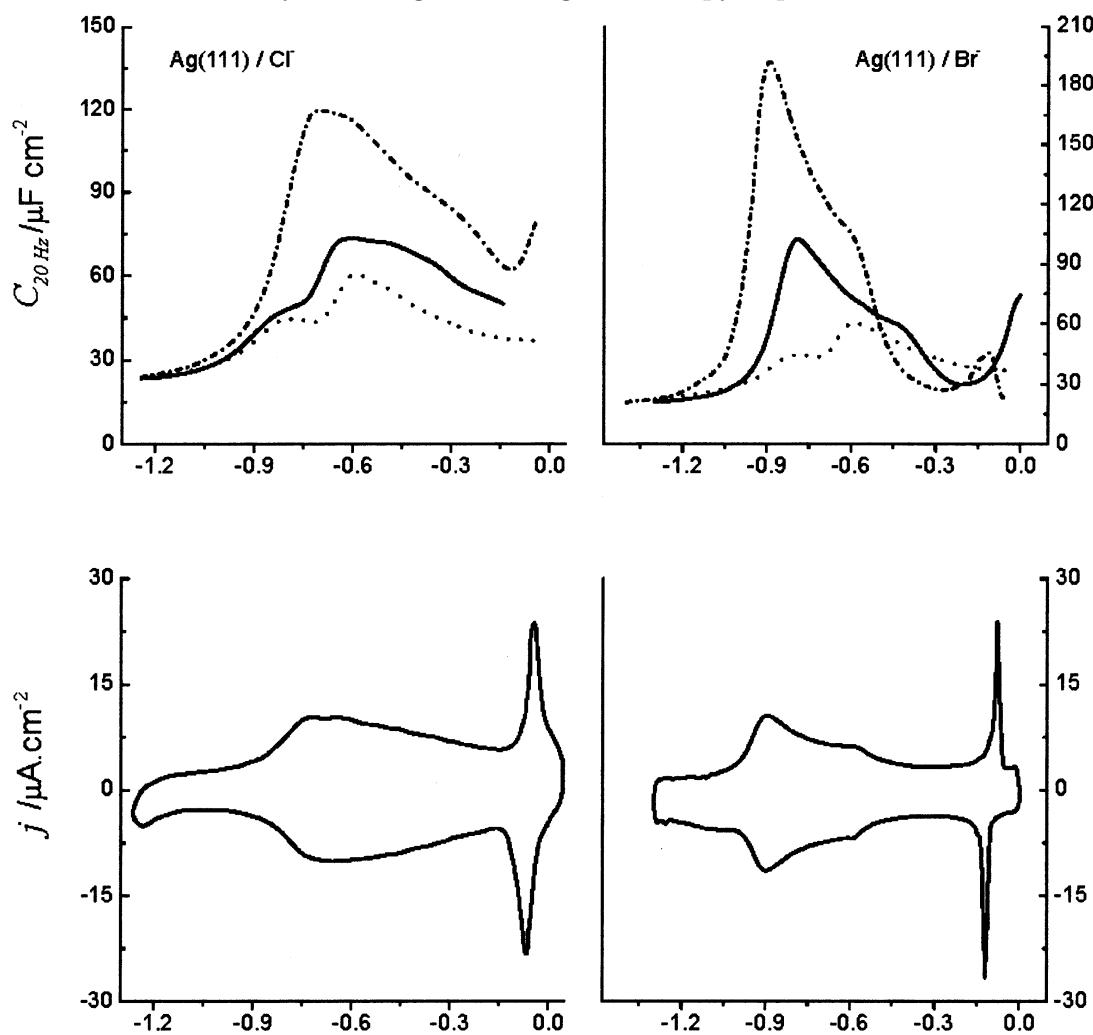


FIGURE 11-3

12. Atomistic models of the interphase and interphase capacitance

It is the experimental observation that the electric properties of the electrode-electrolyte interphase may be mapped to an equivalent circuit that suggests to us that the *ideally polarisable* interphase may be modelled as a flat plate capacitor. If we then ponder on what the atomic arrangement of ions and solvent may look like then we imagine, in the first instance, a *Helmholtz double layer* comprising an ordered layer of ions attached to the metal surface, just of sufficient concentration so as to neutralise the layer of electrons or holes residing at the surface of the electrode. This in fact, as mentioned above, can be observed by STM and possibly low energy electron diffraction; and deduced by the peaks in the cyclic voltammogram in figure 11–3.

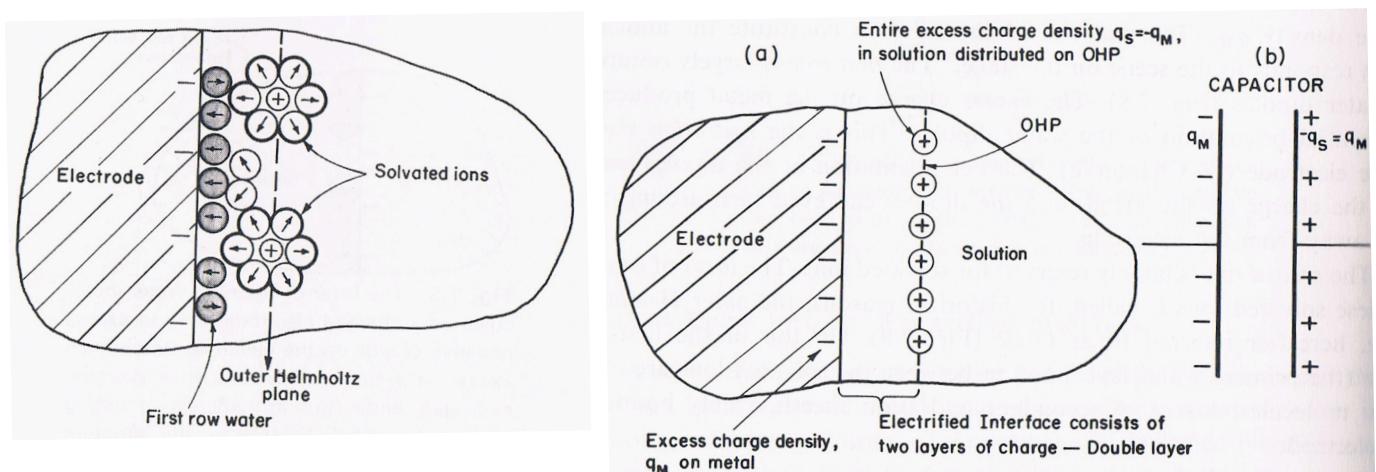


FIGURE 12-1

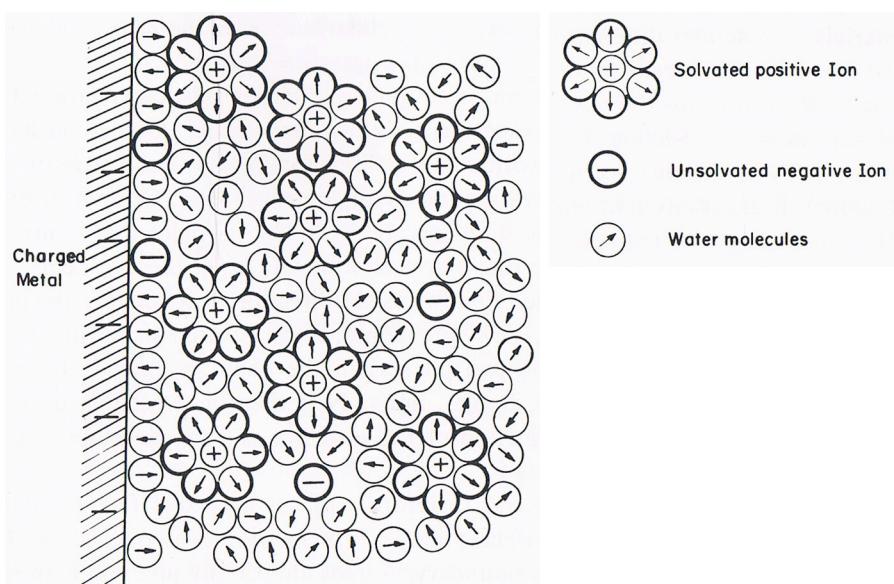


FIGURE 12-2

Figures 12–1 and 12–2 are taken from Bockris and show insightful cartoons into how the double layer and the interphase may look. Note how the cations are fully solvated (outer sphere) while the anions are adsorbed (inner sphere). However it is not sufficient to expect a static double (or triple layer). An accumulation of ions in solvent is bound

to be to some extent disordered at finite temperature. The model for this interphase was put forward by Gouy and by Chapman as early as 1910.

12.1 Gouy–Chapman theory (after Schmickler and Santos)

The model is of point ions solvated by a continuum dielectric of permittivity ϵ , distributed according to the principles of electrostatics and statistical mechanics. In the simplest case of a planar electrode and a one-to-one electrolyte (such as NaCl), we place the electrode surface at the coordinate $x = 0$ and consider a one-dimensional distribution of the electric potential, $\phi(x)$, which is given in terms of the charge density $\rho(x)$ by the Poisson equation,

$$\frac{d^2\phi}{dx^2} = -\frac{\rho(x)}{\epsilon}$$

The point ions may be either cations or anions and their densities will be denoted $n_+(x)$ and $n_-(x)$ respectively. Hence the total charge density is

$$\rho(x) = ze(n_+(x) - n_-(x))$$

where z is the charge number and e is the proton elementary charge. As mentioned above we assume $z = 1$ for simplicity but z can be easily retained in the more general case. If we set the zero of electric potential at $x = \infty$ then according to the Boltzmann statistics,

$$\begin{aligned} n_+ &= n_0 e^{-e\phi(x)/kT} \\ n_- &= n_0 e^{+e\phi(x)/kT} \end{aligned}$$

where n_0 is the bulk concentration of cations and anions, $n_0 = n_+(\infty) = n_-(\infty)$. The resulting Poisson–Boltzmann equation is

$$\frac{d^2\phi}{dx^2} = -\frac{en_0}{\epsilon} (e^{-e\phi(x)/kT} - e^{+e\phi(x)/kT})$$

A transparent solution exists if the Poisson–Boltzmann equation is linearised; that is, in the limit that $e\phi(x)/kT \ll 1$, namely,

$$\frac{d^2\phi}{dx^2} = \kappa^2 \phi(x)$$

where

$$\kappa = \lambda^{-1} = \sqrt{\frac{2e^2n_0}{\epsilon kT}}$$

and λ is the *Debye length*. Here is a table of Debye lengths for a one-to-one electrolyte at room temperature..

molar conc ⁿ	10^{-4}	10^{-3}	10^{-2}	0.1
$\lambda / \text{\AA}$	304	96	30.4	9.6

The solution to the linearised Poisson–Boltzmann equation, subject to the boundary condition, $\phi(x) \rightarrow 0$ as $x \rightarrow \infty$ is

$$\phi(x) = Ae^{-\kappa x}$$

and the constant A may be established by equating the total charge density from the anions and cations existing between $x = 0$ and $x = \infty$ with minus the charge density at the surface of the electrode in order to ensure overall charge neutrality,

$$\int_0^\infty \rho(x)dx = -\sigma$$

Combining this with the Poisson equation results in the following two formulas,

$$\begin{aligned}\phi(x) &= \frac{\sigma}{\epsilon\kappa} e^{-\kappa x} \\ \rho(x) &= -\sigma\kappa e^{-\kappa x}\end{aligned}$$

$\rho(x)$ describes a layer of *space charge* which exactly balances the charge on the electrode.

This leads us back to the notion that the interphase acts as a parallel plate capacitor. The capacitance is charge over voltage, or equivalently ϵ times area over distance between the plates. In this case we have a *capacity* defined as capacitance per unit area and this is clearly

$$C = \epsilon\kappa \quad [\mu\text{F} / \text{cm}^2]$$

a nice simple formula which identifies the “plate separation” of the equivalent circuit capacitor (section 11.3 and figure 12–1) as the Debye length. Since this is typically a few Angstroms the capacity is quite large.

While the linearised Poisson–Boltzmann gives useful insight into the equivalent capacitor, it is really necessary to use the solution from the non linearised equation. The maths is not very edifying and is given by Schmickler and Santos and by Bockris. The result for the double layer capacity is

$$C = \epsilon\kappa \cosh \frac{e\phi(0)}{2kT}$$

However the electric potential at the origin cannot be measured. On the other hand this potential differs by a constant from the inner, or Galvani potential of the electrode. When $\phi(0) = 0$ then there is no charge on the electrode and hence its potential is the potential of zero charge (PZC), ϕ_{pzc} . So the Gouy–Chapman capacity is

$$C = \epsilon\kappa \cosh \frac{e(\phi - \phi_{\text{pzc}})}{2kT}$$

The cosh function has the shape of a *catenary*—a chain hanging under its own weight. Hence the capacity in the Gouy–Chapman theory is a simple convex function with a minimum at the PZC. This is not really observed in experiments. The upper curves in

figure 11–3 do show minima at voltages which are probably close to the PZC, but there are additional maxima and minima which are not described by the Gouy–Chapman capacity. To overcome this deviation from Gouy–Chapman capacity at high ion concentrations, the interphase is modelled as two capacitors in series, having respectively the Gouy–Chapman capacity, C_{GC} , and the *Helmholtz capacity*, C_H . Hence the interphase has capacity, C , given by

$$\frac{1}{C} = \frac{1}{C_{GC}} + \frac{1}{C_H}$$

See Schmickler and Santos for more details.

13. Kinetics

13.1 Corrosion

Up to now in these notes we have confined ourselves to the considerations of electrochemistry *in equilibrium*. Probably the key finding is the *electrochemical series*, section 9.4, and the observation that *all metals* with the exception of gold are thermodynamically unstable with respect to their oxides (see section 9.1, page 33). From this there arise two questions as posed neatly by J. M. West. The first is, given this fact, why do not all metals spontaneously waste away? I already answered this question on page 2. If a piece of metal is immersed in water then some metal atoms will be converted to cations and migrate into the electrolyte, but *they will leave behind their electrons* and cause the metal to become charged; this accumulated charge will resist further dissolution. John West's next question is more challenging: in that case why do metals corrode at all? There are several answers to this. One is that in the case that a piece of metal is immersed in electrolyte and come to the equilibrium, then this state is achieved at a certain single electrode potential, ${}^m\Delta^s\phi = \Delta\phi_{eq}$; if the electric potential within the metal can be contrived to be raised then the equilibrium condition is shifted and more metal can dissolve (see footnote, p. 35). Alternatively the metal may be connected electrically to an electron drain such as in an electrochemical cell. In real life corrosion what mostly happens is that the macroscopic surface of the metal divides itself into patches: some anodes and some cathodes at which oxidation and reduction reactions take place simultaneously, the electrons and ions flowing within the metal and the electrolyte such as to complete the circuit and allow the anodic patches to waste away. This leads to the notion of a *mixed potential* which we will come to in section 16. In some particularly unfortunate cases a base metal may be electrically connected to more noble metal in order to create a natural electrochemical cell such as in the lemon lamp (page 1). This really does happen, usually as consequence of poor design, as in the union of an aluminium car body with a steel chassis leg for instance, or the use of fasteners and components of dissimilar metals (try to remove the bumper from a Jaguar XJ8 and you will find a demonstration of this probably accompanied by oaths and bruised knuckles as the fasteners will be seized to the brackets by corrosion).

Now follows the most crucial question of all from a practical point of view. How fast does a piece of metal corrode? In these notes we can only cover *uniform corrosion*. We are then interested in the uniform rate of metal loss, measured in, say mm per

year. By coincidence the rate in mm/yr is almost numerically the same as the corrosion current density in amps per metre squared. Again in real life, uniform corrosion is not of paramount interest since it can be designed against simply by specifying a life time based on, say, the wall thickness of a pipe. What is dangerous is *localised attack*. In the example in which the surface of the metal has negotiated its division into anodes and cathodes the conservation of charge is expressed as the equality between the cathode and anode currents. If the anode areas are much less than the cathode areas then the anode current density is much greater than the cathode current density, which leads to *pitting* holes appearing in the surface in which the electrolyte composition is vastly different compared to the bulk. When this is exacerbated by the presence of chloride ion the results will be catastrophic. Tony Entwistle used to add to the metallurgical dictum, “anodes are positive and they corrode,” (page 4) the statement, “if you investigate a corrosion failure, look for where the chlorine came from.”

However I do not wish to wander further into the engineering aspects; let's continue to look at the physics of uniform corrosion—specifically, can we formulate its *rate*? This means pedagogically that we move from the subject of chemical thermodynamics to chemical kinetics. Our first aim is to take the steps needed to derive the famous Butler–Volmer equation. In the spirit of these notes, I will assume a knowledge of some physics, in particular some statistical mechanics and solid state physics,[†] but no chemistry to speak of. The overarching framework is the “transition state theory” that is foremost associated with the name of Henry Eyring who flourished in the 1930s. In the transition state theory we talk of a rate coefficient, k_1 , (usually called a “rate constant”, but it's hardly constant, having an exponential dependence on temperature), in units of s⁻¹, which is described by the formula,

$$k_1 = \frac{kT}{h} e^{-\Delta G^\ddagger/RT} \quad (13.1.1)$$

which we first have to derive, and in particular enquire how the Planck and Boltzmann constants (h and k) arise, and what is meant by the free enthalpy, ΔG^\ddagger , of the “activated complex.” The subscript “1” to k_1 indicates that this is a particular case of the more general k_n rate coefficient, the n denoting the “order of the reaction”. The speed or rate of a chemical reaction is then k_n multiplied by (powers of) the concentrations or activities of the various reacting species. This means that the dimensions of k_n are not trivial; but in our case, k_1 , has units of s⁻¹. The order of a reaction is generally to be determined by experiment.

13.2 First order rate equation

Imagine I have one gram of metal which was at one time pure artificial radioactive cobalt-60. Its atoms are decaying to become nickel atoms by emitting gamma rays, an electron and an electron anti neutrino,



[†] But the reader unfamiliar with these topics may skip without loss of continuity from the end of section 13.3 to section 13.8.

The decay is a quantum mechanical, stochastic process and so the probability that any particular atom decays is governed by Poisson statistics. At some particular time, t , the fraction of metal that is cobalt-60 is, say, x (the remaining fraction being nickel is $1-x$). The positive *rate of decay*, the rate at which the amount of cobalt-60 is depleting, is $-\dot{x}$ and this depends only on the amount of cobalt-60 remaining in the sample. (This is the principle on which carbon dating is based.) The rate equation is thereby,

$$\frac{dx}{dt} = -k_1 x \quad (13.2.2)$$

The solution to this equation is,

$$x = x_0 e^{-k_1 t}$$

where x_0 is the fraction of cobalt-60 at time $t = 0$. The rate constant, k_1 , has units of s^{-1} . If at $t = 0$ we have pure cobalt-60 then after the half-life, $t_{\frac{1}{2}}$, has elapsed we have, $x = \frac{1}{2}$, and so the half life (5.3 years) is related to the rate constant through $t_{\frac{1}{2}} = -\ln \frac{1}{2}/k_1 = 0.69/k_1$. Hence $k_1 = 4.1 \times 10^{-9} \text{s}^{-1}$

In chemistry (13.2.2) is called a “first order rate equation” because the “rate of reaction” depends linearly on the concentration, x . (I am a bit lax in using concentration and atom fraction interchangeably.) Radioactive decay is about the only reaction I can think of in which the rate is easy to understand. In general chemical reactions involve any number of reactants and the rate is found experimentally usually to depend on integer powers of the concentrations of the reactants. For example in the reaction (9.1.1) the rate may depend on the n_A^{th} power of c_A times the n_B^{th} power of c_B . Rates of reaction are of course of central importance and come under the very difficult discipline of *kinetics* in physical chemistry. Luckily for us, in corrosion we only need to think about reactions like (13.2.1) of the type,



which we can take to have first order kinetics. On the other hand we have deep difficulties of a pictorial or conceptual type because the detachment of a metal atom from the surface of an electrode and its journey as a positive cation through the “double layer” and into the bulk of the electrolyte is hard to conceive of within the conventional picture of “molecular dynamics,” that is to say the random collision of molecules and the generation of “activated states” by drawing energy from the heat bath.

There are at least three ways to “derive” the equation (13.1.1) and I will try and outline these in the following sections. Let us start with a very appealing picture due to Bockris.

13.3 Rate constant according to Bockris

We plunge straight in and accept that the process whose rate we wish to calculate is the transfer of an ion from electrode to electrolyte. Bockris considers an idealisation of this as represented in figure 13–1. We imagine two “boxes” separated by a distance, d , and the “reaction” is the process in which an ion, initially in box 1 transfers across the gap into box 2. We will call d_1 the distance from the ion to box 1 and $d_2 = d - d_1$ its distance

from box 2. At what point do we say that “a jump has occurred” and the ion is on its way to box 2? If the process is completely symmetric then obviously that is the point at which the ion is half way across (and travelling at some speed, v , in the direction of box 2—more on the speed later). If the potential energy is not symmetric, say the energy of the ion in box 2 is less than that in box 1, then the “point of no return” is not midway between the two. This is captured in a number called the “symmetry factor” which we will find plays a major role in the Butler–Volmer equation. Suppose that we know the potential energy of the system as a function of the coordinates d_1 and d_2 . A contour map of this two dimensional function is illustrated in figure 13–2.

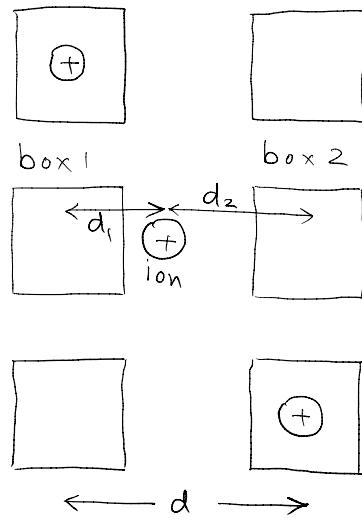


FIGURE 13-1

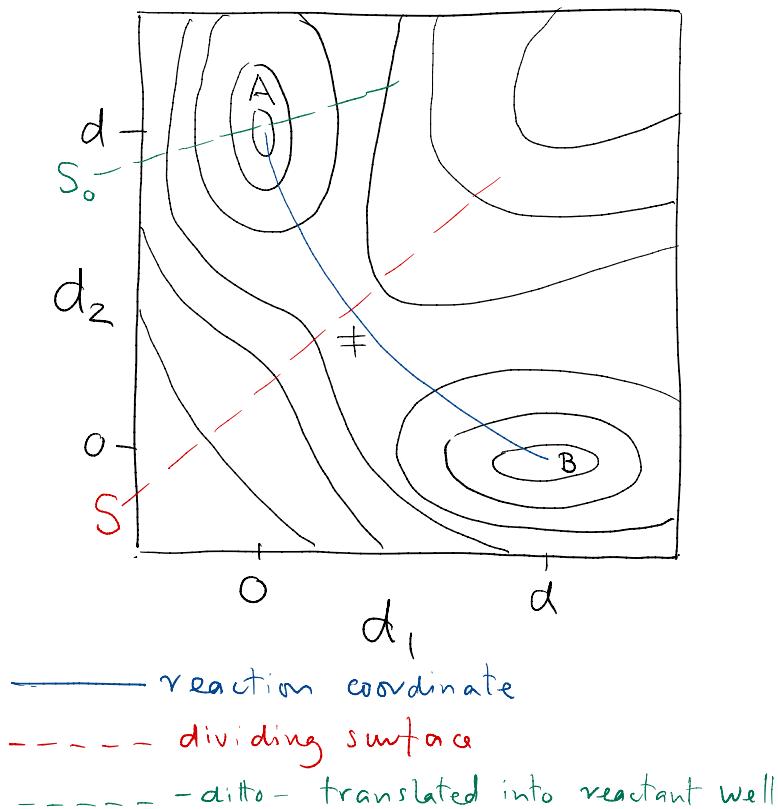


FIGURE 13-2

The path that the ion takes in the configuration space $\{d_1, d_2\}$ is called the “reaction coordinate”. We take it that the path chosen is the one which entails the least amount of potential energy increase—namely the path that takes the ion over the saddle point in the energy landscape. We can then plot the potential energy as a function of the reaction coordinate in a plot such as figure 13–3. In general there are not just two coordinates involved in such a process and so we are invited to try and imagine that a plot such as figure 13–2 is somehow representative of an energy landscape in a multi coordinate configuration space; if there are N atoms involved in some thermally activated process then we have to imagine a path in a $3N$ dimensional space. What characterises the saddle point is that there all the curvatures of the potential energy are positive *except for one*, namely the curvature along the direction of the reaction coordinate.[†] At that point we talk of the existence of an “activated complex” always indicated by the symbol \ddagger . The maximum potential energy, which is encountered at the saddle point in the $3N$ -dimensional space is called the critical activation enthalpy, ΔH^\ddagger ; if there is an entropy change involved in the change from the “reactant basin” to the saddle point then we talk of the free enthalpy of activation, $\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$. We will return to the complication of many coordinates later; for now let us pursue Bockris’s argument for the simple two-coordinate idealisation.

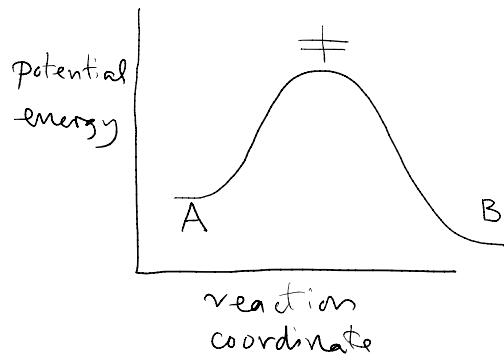
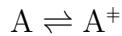


FIGURE 13–3

What characterises the state of the system at the saddle point? We think of this as a decisive situation: if the ion reaches the saddle point moving at a finite speed, v , then the reaction is bound to go to completion. The system will “fall down the energy hill”. What is the unique indicator of this “point of no return”? It is the loosening of the hold that box 1 has on the ion: the bond connecting the two breaks. One way to describe that is to say that vibration (or stiffness) of the bond vanishes and is replaced by a translational degree of freedom. Bockris argues that this condition is reflected in an equality of the vibrational energy $h\nu$ and the translational energy kT . Hence at the moment of the reaction happening there exists an activated state whose frequency of vibration in the direction of the reaction coordinate is $\nu = kT/h$ ($6 \times 10^{12} \text{ s}^{-1}$ at 300K) where h is the Planck constant. His argument now is that the rate of transfer of the ion from box 1 to box 2 is the concentration of activated complexes, c^\ddagger , times the number of times per second on average the complex makes “an attempt” to fall down the hill (from that language the name “attempt frequency,” ν , has been coined). This

[†] In the simple example here there are only two curvatures at the saddle point—one positive in sign and one negative; but we will generalise this picture in section 13.7.

is a central postulate of the transition state theory and sometimes it is formulated in terms of the speed, v , rather than the frequency, ν . So we have clearly moved from the notion of a single transfer event to the consideration of an ensemble of systems since we are enquiring into the concentration of activated complexes. Now we come to a second postulate of transition state theory and one which on the face of it is hard to take seriously. This is that we allow that in the ensemble there exists *an equilibrium between the reactants and the activated complexes*. For the case of reaction (13.2.3) this implies the equilibrium,



Following the arguments in section (9.1) there must be an equilibrium constant (let's use concentration for now),

$$K^\ddagger = \frac{c^\ddagger}{c} \quad (13.3.1)$$

and the molar standard free enthalpy of activation must be

$$\Delta G^\ddagger = -RT \ln K^\ddagger \quad (13.3.2)$$

This is quite an astonishing assertion in view of the thought that the activated complex is extremely short lived and has at least one imaginary frequency[†] of vibration—although in Bockris's argument we actually have $\nu = kT/h$ real and positive. Nonetheless we are assured by the textbooks that this is by no means a weak postulate of transition state theory.

Finally we can complete the derivation. We are looking for the rate of reaction in the form of (13.2.2), namely,

$$\dot{c} = -k_1 c$$

and if we take unit concentration of reactants, we find that the *rate of reaction* is, in moles m⁻³ s⁻¹,

$$r = k_1 c = k_1 \times \mathbf{one}$$

and hence

$$k_1 = r = \text{attempt frequency} \times \text{conc}^n = \nu c^\ddagger = \frac{kT}{h} e^{-\Delta G^\ddagger / RT} \quad (13.3.3)$$

as required in (13.1.1). In preparing these notes I may have become a bit obsessed with the various ways in which the formula for the rate coefficient may be derived so we will now go into a rather lengthy and possibly over detailed excursus. The reader who is content with Bockris's argument as it stands may profitably skip now to section 13.8.

13.4 Velocity at the saddle point

Sometimes the rate coefficient is couched in terms of the velocity, v , introduced in the last section, rather than the frequency, ν . Using figure 13–2 and some results from

[†] $\nu^2 < 0$ at the top of the potential energy curve in figure 13–3.

statistical mechanics we can develop a formula for this speed. Consider that the reaction path shown in figure 13–2 is parameterised by a coordinate x as $x(d_1, d_2)$. The ion in motion has a mass m and a speed \dot{x} and so its kinetic energy is $m\dot{x}^2/2$. If the speeds of an ensemble of these particles (ions in the example of section 13.3) have an equilibrium distribution then the average speed is given by a ratio of configuration integrals, namely,

$$\langle v \rangle = \frac{\int_0^\infty e^{-\frac{1}{2}\beta m\dot{x}^2} \dot{x} \, d\dot{x}}{\int_{-\infty}^\infty e^{-\frac{1}{2}\beta m\dot{x}^2} \, d\dot{x}} \quad (13.4.1)$$

in which we use the standard notation, $\beta = 1/kT$. Note that the lower limit in the numerator is zero and not $-\infty$ because we are only interested in states in which the activated complex is flying in the positive x direction and hence the reaction is occurring. We need these two standard integrals,

$$\int_{-\infty}^\infty e^{-ax^2} \, dx = \sqrt{\frac{\pi}{a}} \quad (13.4.2a)$$

$$\int_0^\infty x e^{-ax^2} \, dx = \frac{1}{2a} \quad (13.4.2b)$$

There results,

$$\langle v \rangle = \sqrt{\frac{1}{2\pi\beta m}} = \sqrt{\frac{kT}{2\pi m}} \quad (13.4.3)$$

Note that since the axes of figure 13–2 are distance the units of v are indeed m s^{-1} . However it is not necessary to require that v is actually the speed of some particle or ion. Indeed we will generalise this picture in section 13.7. If the reaction is dominated by the passage of a single particle in the presence of other particles which are essentially fixed in position, then we may take it that m really is the mass of the particle. An example would be the diffusion of atomic hydrogen through a lattice of metal atoms. Otherwise, and indeed usually, we must interpret m as an effective mass which characterises the reaction.

13.5 Equilibrium constant in terms of partition function

We consider the equilibrium



and imagine that the reactants and products are simple systems such as molecules having a set of fixed single particle energy levels as sketched in figure 13–4. Each then has a microcanonical partition function,

$$z'_A = \sum_{i=0}^{\infty} e^{-\epsilon_i(A)/kT}$$

and

$$z'_B = \sum_{i=0}^{\infty} e^{-\epsilon_i(B)/kT}$$

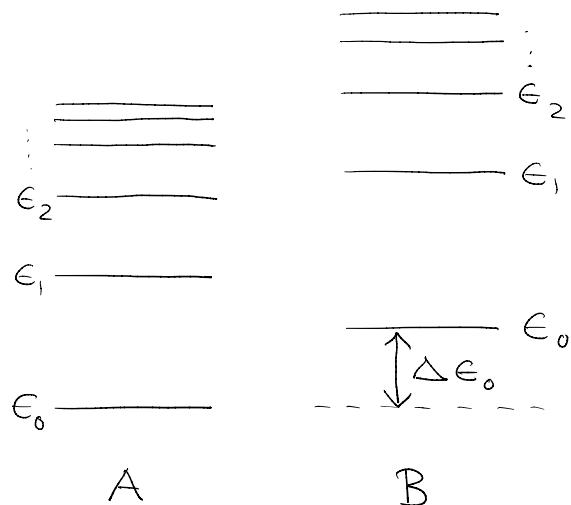


FIGURE 13-4

Each of these is relative to its own lowest eigenvalue ϵ_0 . We need to choose a common energy zero so we use $\epsilon_0(A) = 0$. When each partition function is expressed relative to this common zero of energy we get,

$$z_A = z'_A$$

and

$$\begin{aligned} z_B &= \sum_{i=0}^{\infty} e^{-(\epsilon_i(B)+\Delta\epsilon_0)/kT} \\ &= z'_B e^{-\Delta\epsilon_0/kT} \end{aligned}$$

The absence of a prime indicates partition functions referred to a common energy zero. z_B is the sum of all probabilities that the equilibrium (13.5.1) finds itself in one of the energy levels of B; and similarly for A. It follows that the concentrations of B and A are in the same ratio as their partition functions,

$$\frac{c_B}{c_A} = \frac{z_B}{z_A} = \frac{z'_B}{z'_A} e^{-\Delta\epsilon_0/kT} \quad (13.5.2)$$

Indeed this result is very general and not just confined to single particle partition functions. If the reactants or products are made of an assembly of N identical particles then their canonical partition function is,

$$Z = \frac{z^N}{N!}$$

and it becomes axiomatic that if both partition functions are referred to the same energy zero,

$$\frac{c_B}{c_A} = \frac{Z_B}{Z_A} \quad (13.5.3)$$

We should note for future reference that the single particle partition function can be approximately factored into its translational, rotational, vibrational and electronic contributions,

$$z' = z_{\text{tr}} z_{\text{rot}} z_{\text{vib}} z_{\text{el}}$$

(omitting primes on the right hand side) well known formulas being,

$$\begin{aligned} z_{\text{vib}} &= (1 - e^{-h\nu/kT})^{-1} \\ z_{\text{tr}} &= \frac{V}{\Lambda^3} \end{aligned} \quad (13.5.4)$$

where,

$$\Lambda = \frac{h}{\sqrt{2\pi mkT}}$$

is the thermal de Broglie wavelength and m is the mass of the particle.

The Helmholtz free energy is given by a central formula of statistical mechanics,

$$F = -kT \ln Z$$

Let us use a series of approximations to find a relation between the single particle partition functions and the equilibrium constant for the reaction (13.5.1). We proceed as follows. For one mole (L particles) of a substance,

$$F = -kT \ln Z \quad (13.5.5)$$

$$= -kT \ln \left(\frac{z^L}{L!} \right) \quad (13.5.6)$$

$$\begin{aligned} &= -LkT \ln z + kTL(\ln L - 1) \\ &= -RT \ln \frac{z}{L} \end{aligned}$$

in which L is the Avogadro constant: the number of particles per mole. I have used the slacker Sterling approximation, $\ln N! = N(\ln N - 1)$ and neglected one compared to $\ln L = 55$, and used the relation between the Boltzmann and gas constants, $R = Lk$. Finally observe that in condensed matter we may ignore the difference between free energy and free enthalpy, $G = F + pV$, since bulk properties of solids and liquids are virtually unchanged between ambient and near-zero pressure. We can thereby assert that the standard state free enthalpy per mole in condensed matter is,[†]

$$G^\circ = -RT \ln \frac{z^\circ}{L}$$

as long as we may factor the partition function into identical single particle partition functions as in (13.5.6). In view of this I can write the standard free enthalpy change of reaction (13.5.1) as,

$$\Delta G^\circ = -RT \ln \left(\frac{z_B^\circ}{z_A^\circ} e^{-\Delta\epsilon_0/kT} \right) = -RT \ln K$$

[†] This result is exact for an ideal gas as you can see by adding $pV = RT$ to (13.5.5). Then I don't need to neglect the one.

defining standard state partition functions referred to a common energy zero. This identifies an equilibrium constant for the reaction (13.5.1) as the ratio of the probability of finding the equilibrium system in one of the states of the reactants to the probability of finding the system in one of the states of the products. We see from (13.5.2) that

$$K = \frac{c_B}{c_A}$$

when reactant and product are in a standard state, which is indeed consistent with our definition of equilibrium constant in section 9.1 if we assume that activity coefficients are one, that is, that the system is in some sense “ideal.”

This observation throws up a serious disconnect in the relation between equilibrium constant and partition function. Suppose we return to the reaction of section 9.1,



Its equilibrium constant is,

$$K = \frac{a_C^{n_C} a_D^{n_D}}{a_A^{n_A} a_B^{n_B}} \quad (9.1.3)$$

(Clearly this can be extended to any number of reactants and products.) On the other hand the ratio of partition functions gives us concentrations since these relate directly to the probabilities of the equilibrium being found in one of the microcanonical states. Hence we must write,

$$\frac{c_C^{n_C} c_D^{n_D}}{c_A^{n_A} c_B^{n_B}} = \frac{(z_C/V)^{n_C} (z_D/V)^{n_D}}{(z_A/V)^{n_A} (z_B/V)^{n_B}}$$

This means that, in our simple “unimolecular” reaction (13.5.1) we should write the relation between equilibrium constant and partition function as,

$$K = \frac{\gamma_B c_B}{\gamma_A c_A} = \frac{\gamma_B z_B^\circ}{\gamma_A z_A^\circ}$$

In the next section when we apply this formula to the activated complex you will see how unsatisfactory this is. Not only do we need to imagine that we can apply equilibrium thermodynamics to a mechanically unstable object, we need to assert that we can somehow define an “ideal” state and furthermore to give a value to its activity coefficient. Some textbooks avoid this by defining equilibrium constant directly in terms of concentration, or in terms of partial pressure; but I’m not sure whether that doesn’t just sweep the problem under the carpet.

Ultimately the whole of rate theory rests on *experimental* measurements that show that the rate of a reaction is in proportion to some powers of the concentrations of the reactants. If the observations had shown a closer correlation with activities rather than concentrations then maybe the theory would have looked a little different.

13.6 Rate constant in terms of partition function

If we apply reasoning from the previous section to the equilibrium,



between an initial state and the activated complex as in section 13.3, then we may write (13.5.2) in the form

$$K^+ = \frac{z^+}{z} e^{-\Delta U/kT}$$

in which z stands for z_A and ΔU is the difference in energy between the lowest energy levels of the activated complex and the initial state: the equivalent of $\Delta\epsilon_0$ in figure 13–4. In this way ΔU is the zero temperature activation energy (in the sense that at $T = 0$ the species are all in their ground states). Now, by definition, the rate of change of the amount of reactant in the initial state of the system is, in comparison to (13.3.3),

$$\begin{aligned} -\dot{c} &= k_1 c = \text{conc}^n \text{ of activated complexes times} \\ &\quad \text{their frequency of crossing the saddle point} \\ &= c^\ddagger \nu^\ddagger \end{aligned}$$

Now from (13.5.2),

$$c^\ddagger = c \frac{z^+}{z} e^{-\Delta U/kT}$$

so it follows that,

$$k_1 = \nu^\ddagger \frac{z^+}{z} e^{-\Delta U/kT}$$

We factorise out from z^+ the contribution from the vibrational mode whose oscillations are in the direction of the reaction coordinate at the saddle point,

$$z^+ = z_{\text{vib}}^+ z^{*\ddagger}$$

(The star reminds us that this partition function has one fewer degree of freedom.) We recall the argument of Bockris in section 13.3 that this vibrational mode is a borderline translation so its frequency must be tending to zero. Therefore from (13.6.4) we have

$$z_{\text{vib}}^+ = \left(1 - e^{-h\nu^\ddagger/kT}\right)^{-1} = \frac{kT}{h\nu^\ddagger} \quad (h\nu^\ddagger \ll kT)$$

The result is,

$$k_1 = \frac{kT}{h} \frac{z^{*\ddagger}}{z} e^{-\Delta U/kT} \quad (13.6.2)$$

It is very typical, as we shall see, that the rate coefficient involves a ratio of partition functions in the numerator of which there is the partition function of the activated complex *with one degree of freedom removed*. This is handy because since we are at a saddle point there is one normal mode whose frequency is imaginary, so it's well that we are able to factor it out. Please note that the reduction by one degree of freedom as indicated by the asterisk, in this case does not entail a change in dimension since the

partition functions here are dimensionless. We will see in section 13.7 that in a more general case the reduction in degrees of freedom also results in a change in the dimension of a partition function as this has an impact on the interpretation of a formula equivalent to (13.6.2).

Finally, we can eliminate the partition functions and replace the Arrhenius activation energy with the activation free enthalpy in the following way. We write the van't Hoff isotherm for the reaction (13.6.1),

$$\Delta G^{\circ\ddagger} = -RT \ln K^\ddagger$$

and insert this into,

$$k_1 = \frac{kT}{h} K^\ddagger$$

and we get the expression for the rate coefficient in its familiar form (familiar that is to physical chemists),

$$\begin{aligned} k_1 &= \frac{kT}{h} e^{-\Delta G^{\circ\ddagger}/RT} \\ &= \frac{kT}{h} e^{-\Delta S^{\circ\ddagger}/R} e^{-\Delta H^{\circ\ddagger}/RT} \end{aligned} \tag{13.1.1}$$

$\Delta G^{\circ\ddagger}$, $\Delta H^{\circ\ddagger}$ and $\Delta S^{\circ\ddagger}$ are respectively the free enthalpy (Gibbs free energy), enthalpy and entropy of activation in a standard state. There is a glaring inconsistency in this last argument because it ignores the fact that we have already pulled out the imaginary vibration frequency. However Walter Moore insists that this "...has no appreciable effect on the practical applications of the equation." Given the messy nature of the problem, the difficulty in finding a pictorial representation of the process and obscurity in defining what is an "activated complex" it is probably not surprising that there are logical difficulties. However if you ask is there any rigorous way to write down the rate coefficient then read on to the next section.

13.7 Formulation from first principles

If you are in despair that a plausible explanation of the rate coefficient has so far eluded us, then join me as we turn to solid state physics for a more satisfactory development. This is associated with the names of C. A. Wert and the legendary Clarence Zener of Chicago and established in the seminal paper by George H. Vineyard (*J. Phys. Chem. Solids*, **3**, 121 (1957)). Return to figure 13–2 but now allow it to represent a case where $N/3$ atoms are interacting. The following development was originally made for the diffusion of an atom in a metal, so the point marked "A" may represent the state in which there is a vacancy next to an atom which is about to jump into the vacant site surmounting a barrier in the energy landscape. The point marked "B" may represent the crystal after the atom has jumped into the vacant site, but in general the points A and B need not be symmetrically disposed and there may be an energy change associated with the reaction A → B. Each atom has three position coordinates, denoted x_1, x_2, \dots, x_N so figure 13–2 must now be interpreted as a set of

contours in an N -dimensional, not 2-dimensional space. The red line is constructed to intersect the saddle point and to be perpendicular to all the contours. It is called a *dividing surface* and is a hyperplane of $N - 1$ dimensions. For the system to move from the basin A into the basin B through the dividing surface there must be some “flow of probability” which we try and calculate. Associated with each position coordinate, x_i , there is a velocity \dot{x}_i . In what follows we will simplify matters and insist that all atoms have the same mass, m . If this were not the case then we would rescale all the coordinates (as in the original work) to $y_i = \sqrt{m_i}x_i$.

Because we have migrated from physical chemistry to solid state physics (and who can judge which is closer to the detachment of an ion from a metal into an electrolyte?) we dispense with the rather artificial notion of “concentration”, and indeed we did this in the argument given by Bockris in section 13.3 by setting $c = 1$. Instead we ask a completely well posed question, what is the average lifetime, τ , of all configurations represented by that part of the configuration space to one side of the dividing surface—namely the side containing configuration A? Then the rate of transition from A to B is

$$k_1 = \frac{1}{\tau} = \frac{\mathcal{I}}{Z_A} \quad (13.7.1)$$

Here, Z_A is the number of points in the configuration space on the “A-side” of the dividing surface, and \mathcal{I} is number of such points crossing the dividing surface in the direction of B per unit of time. Z_A is a partition function that I write to within a multiplicative normalisation that will cancel in the end so I’ll leave it out,

$$Z_A = \int \dots \int_A e^{-\beta U} dx_1 dx_2 \dots dx_N$$

where $U = U(x_1, x_2 \dots x_N)$ is the potential energy surface whose contours in the N -dimensional space are sketched in figure 13–2. It is worthwhile to point out what are the *dimensions* of Z_A in the absence of a normalising constant. This is an integral over configuration space (restricted to that part which is on the reactant side of the dividing surface) and so has the dimensions of $[\text{length}]^N$. Now we need to calculate \mathcal{I} . Figure 13–5 shows a patch of the dividing surface (in $N - 1$ dimensions!), a point, P , and an element of surface at P whose normal is the vector $d\mathbf{S}$. If \mathbf{v} is the velocity of points in configuration space (see section 13.4) then we will have that the infinitesimal element of current crossing $d\mathbf{S}$ is

$$d\mathcal{I} = d\mathbf{S} \cdot \langle \mathbf{v} \rangle$$

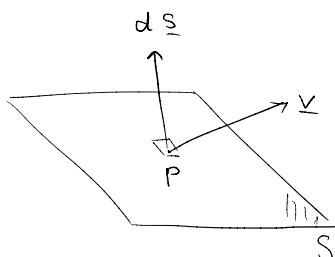


FIGURE 13–5

We need to know what is the statistical distribution of the velocity at P . The number of configuration space points at P having a speed v in the range $dv = d\dot{x}_1 d\dot{x}_2 \dots d\dot{x}_N$ is,

$$\rho(P, v) dv = \frac{e^{-\beta U(P)} e^{-\frac{1}{2}\beta mv^2} d\dot{x}_1 d\dot{x}_2 \dots d\dot{x}_N}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} d\dot{x}_1 d\dot{x}_2 \dots d\dot{x}_N e^{-\frac{1}{2}\beta m\dot{x}_1^2} e^{-\frac{1}{2}\beta m\dot{x}_2^2} \dots e^{-\frac{1}{2}\beta m\dot{x}_N^2}}$$

The denominator is a product of identical terms as in (13.4.1) so using (13.4.2a) we get

$$\text{denominator} = \left(\sqrt{\frac{2\pi}{\beta m}} \right)^N = \left(\frac{2\pi kT}{m} \right)^{\frac{1}{2}N}$$

We now need to find,

$$d\mathcal{I} = d\mathbf{S} \cdot \int \mathbf{v} \rho(P, v) dv$$

and the domain of integration is to be such that the flow is all in the direction A→B, that is $d\mathbf{S} \cdot \mathbf{v} > 0$. A neat way to achieve this is to imagine that we have rotated the coordinate system so that one of the \dot{x}_i , say \dot{x}_1 is directed parallel to $d\mathbf{S}$. This means that all other velocities are lying in the dividing plane so they factor out. We get,

$$\begin{aligned} d\mathcal{I} &= \left(\frac{2\pi kT}{m} \right)^{-\frac{1}{2}N} e^{-\beta U(P)} dS_1 \int_0^{\infty} \dot{x}_1 e^{-\frac{1}{2}\beta m\dot{x}_1^2} d\dot{x}_1 \\ &\quad \times \int_{-\infty}^{\infty} e^{-\frac{1}{2}\beta m\dot{x}_2^2} d\dot{x}_2 \int_{-\infty}^{\infty} e^{-\frac{1}{2}\beta m\dot{x}_3^2} d\dot{x}_3 \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\beta m\dot{x}_N^2} d\dot{x}_N \end{aligned}$$

As in (13.4.1) we only integrate over the velocities in the forward direction across the dividing surface and this first integral evaluates to kT/m (13.4.2a). The second line here is a product of $N - 1$ identical integrals exactly one fewer than in the denominator we evaluated just above. So these all but cancel the prefactor,

$$\left(\frac{2\pi kT}{m} \right)^{-\frac{1}{2}N} \left(\frac{2\pi kT}{m} \right)^{\frac{1}{2}(N-1)} = \sqrt{\frac{m}{2\pi kT}}$$

When this is multiplied by the factor kT/m from the integral over \dot{x}_1 we get a prefactor $\sqrt{kT/2\pi m}$. Now we obtain the total current by integrating $d\mathcal{I}$ over the dividing surface,

$$\mathcal{I} = \sqrt{\frac{kT}{2\pi m}} \int \dots \int_S e^{-U/kT} dS$$

Now we combine this with (13.7.1) and obtain our final result for the rate coefficient,

$$k_1 = \sqrt{\frac{kT}{2\pi m}} \frac{\int \dots \int_S e^{-U/kT} dS}{\int \dots \int_A e^{-U/kT} dV} = \sqrt{\frac{kT}{2\pi m}} \frac{Z_S^*}{Z_A} \quad (13.7.2)$$

The prefactor, $\sqrt{kT/2\pi m}$ is the velocity we calculated in (13.4.3). Again, this has the actual units of velocity, ms^{-1} , as long as the energy surface is constructed in direct space (not for example scaled by $\sqrt{m_i}$ as in the original work by Vineyard). However

one could imagine that in a chemical reaction the reaction coordinate does not have dimensions of length as measured in metre, in which case we may think of a generalised “velocity” along the reaction coordinate, possibly having unusual “length” dimensions. At any event the expression (13.7.2) has the expected units of s^{-1} . This is because the velocity prefactor has dimensions of [“length”][time] $^{-1}$ while the ratio of partition functions has dimensions of [“length”] $^{-1}$ since the numerator has one fewer degree of freedom and hence one few factor of dx than the denominator.

Equation (13.7.2) is reminiscent of our earlier formula (13.6.2), namely a ratio of two partition functions, this time multiplied by the velocity rather than by a frequency. It is very important to note that as in (13.6.2) the partition function in the numerator has one degree of freedom fewer than that in the denominator. I have indicated this again with an asterisk.

The formula (13.7.2) does not, on the other hand, have the familiar Arrhenius form involving the exponential of an activation energy, or free enthalpy, over kT . Nevertheless we can cast it into both such forms. Firstly, if we use the approximation of small vibrations then in the manner of the theory of lattice vibrations we expand the potential energy U both at the configuration point A at the minimum of energy in the “A-basin” and again at the point A^\ddagger where the reaction coordinate intersects the dividing surface,

$$\begin{aligned} U &\approx U(A) + \frac{1}{2} \sum_{i=1}^N m\omega_i^2 q_i^2 && \text{about A} \\ U &\approx U(A^\ddagger) + \frac{1}{2} \sum_{i=1}^{N-1} m\omega_i^{\ddagger 2} q_i^{\ddagger 2} && \text{about } A^\ddagger \end{aligned}$$

Here, the ω_i and ω_i^\ddagger are normal mode angular frequencies ($\omega = 2\pi\nu$) and the q_i and q_i^\ddagger are associated normal mode coordinates. In the second expansion the vibrations are constrained to remain in the dividing surface so there is one fewer mode, namely the mode that would correspond to vibration in the direction of the reaction coordinate and which would hence possess an imaginary frequency. In this harmonic approximation the partition functions can be evaluated.[†] They are,

$$\begin{aligned} Z_A^* &= \prod_{i=1}^{N-1} \sqrt{\frac{2\pi kT}{m\omega_i^{\ddagger 2}}} e^{-U(A^\ddagger)/kT} \\ Z_A &= \prod_{i=1}^N \sqrt{\frac{2\pi kT}{m\omega_i^2}} e^{-U(A)/kT} \end{aligned}$$

Putting this into (13.7.2) there is cancellation of some $2\pi s$ and ms and the result is,

$$k_1 = \nu' e^{-(U(A^\ddagger) - U(A))/kT} \quad (13.7.3)$$

where,

$$\nu' = \frac{\nu_1 \nu_2 \dots \nu_N}{\nu_1^\ddagger \nu_2^\ddagger \dots \nu_{N-1}^\ddagger} \quad (13.7.4)$$

[†] Put the expansion of U into $\int \dots \int e^{-U/kT} dq_1 dq_2 \dots$ and use (13.4.2a).

This is in the form of (13.6.2) in the sense that the Arrhenius term is the exponential of a *potential* energy, or zero temperature enthalpy.

To cast the rate coefficient in terms of an exponential of a free enthalpy of activation, so that we can finally make contact with the putative formula (13.1.1), we return to (13.7.2) and figure 13.2. In addition to the dividing surface, S , we imagine a hypersurface of the same reduced dimension, S_0 , as sketched in figure 13–2, which is the dividing surface *translated* so that it now intersects the minimum of the reactant basin at A. There are now two reduced dimension partition functions,

$$Z_S^* = \int \dots \int_S e^{-U/kT} dS$$

$$Z_0^* = \int \dots \int_{S_0} e^{-U/kT} dS$$

Now take equation (13.7.2) and multiply and divide by Z_0^* ,

$$k_1 = \sqrt{\frac{kT}{2\pi m}} \frac{Z_0^*}{Z_A} \frac{Z_S^*}{Z_0^*}$$

Now comes the *coup de grâce* (this is not french for lawnmower). Because the free energy is $F = -kT \ln Z$ we have,

$$\frac{Z_S^*}{Z_0^*} = e^{-\Delta F/kT}$$

and we interpret ΔF as the reversible work done in taking the configuration space points which are constrained to belong to S_0 into those points that are constrained to belong to S . That is to say, we carry the system from the reactant basin to the saddle point. In this sense $\Delta F = \Delta U - T\Delta S$ is the free energy of activation, and ΔU and ΔS are potential energy and entropy of activation, exactly as in (13.1.1) allowing that we need not distinguish enthalpy and internal energy in condensed matter (p = ambient and $p \approx 0$ lead to little difference in bulk properties). Finally then we have for the rate coefficient,

$$k_1 = \sqrt{\frac{kT}{2\pi m}} \frac{Z_0^*}{Z_A} e^{-\Delta F/kT} = \nu e^{-\Delta F/kT} \quad (13.7.5)$$

Now the ratio is of partition functions both belonging to the reactant basin: the denominator being the number of all configuration space points to the reactant side of the dividing surface, and the numerator being the same but constrained to that part of the space that excludes those configurations which are in some sense “parallel” to the dividing surface as indicated in figure 13–2. It is most significant that in this formulation of the rate coefficient no reference is made to the properties of the activated complex in contrast to (13.7.2) which requires us to know the partition function at the dividing surface. Instead, in (13.7.5) the nature of the activated complex is relegated to an activation free energy and only partition functions evaluated at the reactant state are called for. (This is particularly evident in quantum transition state theory in which the activated state becomes almost irrelevant since an atom may anyway *tunnel* through the barrier.) I should point out that the choice of S_0 is not unique—but then neither is the

choice of dividing surface if the energy landscape is not symmetrical. In the small vibration approximation it is possible as in (13.7.3) to identify the prefactor, ν , in (13.7.5) with a ratio of frequencies similar to (13.7.4). But this time there is no reference to the vibrations of the activated complex since the frequencies belonging to A and to S_0 are those of the reactant basin. The frequency prefactor, ν is related to the frequency ν' in (13.7.3) via,

$$\nu = \nu' e^{-\Delta S/k}$$

Then, in the small vibration approximation, if you use the result in statistical mechanics that $S = k \ln Z + U/T + \text{const.}$, you find in contrast to (13.7.4),

$$\nu = \frac{\nu_1 \nu_2 \dots \nu_N}{\nu_1^0 \nu_2^0 \dots \nu_{N-1}^0} \quad (13.7.6)$$

I can only see one way in which this ratio can be cast into a simple kT/h as required by (13.1.1). This is if we assume all these frequencies, which belong to the reactant basin A, are single Einstein frequencies all but one of which cancel in the ratio since there is one more frequency factor in the numerator that will not cancel. If this mode is fully excited then it could be expressed as kT/h , but the contrast with the earlier argument leading to (13.6.2) is that in that case we identify kT/h with a mode of the activated complex, not of the reactants. For this reason the formulation leading to (13.7.5) is appealing because it does not deal in unmeasurable, unthermodynamic quantites such as frequencies of, or equilibria with, fictitious “activated complexes.”

We can now leave rate theory and return to electrochemistry, but the reader may be interested to know that whereas in the 1950s (13.7.5) was mostly a theoretical construction, nowadays with computational methods for calculation of partition functions, prediction of rate coefficients is feasible even using reliable electronic structure based total energy and force methods. Indeed the theory can be extended by replacing the classical partition functions in (13.7.5) with quantum partition functions which can be calculated using Monte Carlo methods. This then allows a calculation of transition rates that include quantum tunnelling and zero point energy. An example of this in the context of hydrogen diffusion in iron can be found in A. T. Paxton and I. H. Katzarov, *Acta Materialia*, **103**, 71 (2016).

As a final comment, we have noted that one of the postulates of the classical transition state theory is that if the system finds itself at the saddle point with a finite velocity in the direction away from the reactant basin, then the reaction *will* occur. There is a caveat to this which arises from quantum mechanics, namely that there is a finite probability that even under this condition the atoms will draw back from the brink and return to the reactant state. This possibility is expressed in terms of a *transmission coefficient* which is a number less than one which multiplies the rate coefficient. Nowadays transmission coefficients can also be calculated using molecular dynamics techniques.

14. Single electrode in equilibrium and at an overpotential

14.1 Detailed balancing

We return to what is really a thought experiment of placing a coupon of pure metal into a solution of its ions. This establishes an equilibrium in the reaction,



which is a special case of (9.5.1) on page 39. Depending on the initial concentration of M^{n+} ions in the solution this reaction will proceed to the left or to the right until metal ions in both phases come to equilibrium, that is the electrochemical potentials of M^{n+} ions are equal in the electrolyte and in the metal.[†] As I have pointed out at the beginning of section 13, even if the free enthalpy of oxidation dictates that this reaction should continue until all the metal has dissolved, the charging of the metal by the electrons liberated by oxidation of the metal prevents the backward reaction from proceeding; this is because of the electric self energy associated with a charged object and indeed there is an additional self energy due to charging of the electrolyte. In addition the transfer of M^{n+} into the solution becomes increasingly more costly as the metal's negative charge grows. In fact it continues until the metal becomes charged to the extent of establishing a certain equilibrium electric potential difference across the interphase,

$$\Delta\phi_{\text{eq}} = \phi_m - \phi_s \quad (14.1.2)$$

In principle this single electrode potential cannot be measured. On the other hand if the interphase were ideally polarisable and we connect it to an ideally reversible electrode such as a standard hydrogen electrode then we could measure it. Indeed if the electrolyte solution is at standard activity of M^{n+} ions then the equilibrium potential may be read off from the electrochemical series table on page 38. If the concentration of M^{n+} ions is different from unit activity then $\Delta\phi_{\text{eq}}$ can be found by application of the Nernst equation (9.5.2).

As a consequence of the equilibrium potential difference, $\Delta\phi_{\text{eq}}$, there will exist an electric field within the interphase, and the M^{n+} ion will have to do work against this electric field (the positive metal is repelling the ion and the negative electrolyte is attracting it back since it leaves behind a positive anion).

“Equilibrium” of this reaction does not mean that nothing is happening. What it does mean is that the forward and backward reactions are happening at the same rate (like the Angels ascending and descending Jacob’s ladder). This is called the *principle of detailed balancing* and is a direct consequence of the quantum and statistical mechanics (see a textbook for an explanation).

We shall examine the forward and backward rates of the reaction (14.1.1) separately, since unlike many activated processes both in chemistry and physics, there is a very

[†] What we mean by the “electrochemical potential of ions in the metal” is understood within the decomposition of the chemical potential of atoms in the metal as $\mu_M = \tilde{\mu}_{M^{n+}} + \tilde{\mu}_e$; admittedly a strange way to think of metallic bonding.

large asymmetry when we compare the detachment of a metal atom from a surface to the migration of an ion from solution onto the surface of the metal. The former may well be better described by solid state kinetics and the latter by solution kinetics. You may have arrived here from section (13.3) in which case we are agreed that the rate of some process is the product of a “concentration” or activity and a “rate coefficient” which may be written reasonably generally as,

$$k_1 = \nu e^{-\Delta G^\ddagger/RT}$$

in which in the “frequency prefactor” is usually given in textbooks as,

$$\nu = \frac{kT}{h}$$

or you may prefer a more rigorous form such as in equation (13.7.6).

14.2 Electronation

The forward reaction of (14.1.1) is electronation (see section 1). The M^{n+} ion moves from the outer Helmholtz plane (OHP) through the “double layer” and attaches to the metal surface, which then acquires a charge of $+ne$. Accompanying this process there is necessarily an electric current and if one mole of M^{n+} ions make the journey across a unit area of metal surface per second then there arises an electronation current density, i_e , which we can find if we know the rate of the electronation reaction. If the concentration of M^{n+} ions is $\bar{c}_{M^{n+}}$ moles per unit area (not volume) then the rate of the reaction in moles per square metre per second is,

$$r_e^0 = \nu_e \bar{c}_{M^{n+}} e^{-\Delta G_e^\ddagger/RT}$$

where ΔG_e^\ddagger is the free enthalpy of activation of electronation, and the superscript, zero, denotes zero electric field. Bockris calls ΔG_e^\ddagger the “chemical activation energy”.

Indeed this neglects the effect of the electric field. If the ion is moving against the field generated by the electric potential difference, $\Delta\phi_{eq}$, then this effectively *raises* the activation barrier. If the total potential difference between the OHP and the metal surface is $\Delta\phi_{eq}$ then the amount of work done against the electric field per mole of ions is *not* $nF\Delta\phi_{eq}$. This is because the ion only needs to do work against the electric field until it reaches the point somewhere between the OHP and the metal surface at which the free enthalpy of activation is at its maximum. After that the ion may traverse the remainder of the double layer without doing further electric work. This leads us to a central parameter in electrochemistry, the so called *symmetry factor*, β . This is the *fraction of the total distance from the OHP travelled over which electrical work must be done by the ion*. Hence the additional energy that needs to be supplied to the ion to complete the electronation reaction is $\beta nF\Delta\phi_{eq}$ and this essentially *lifts* the height of the free enthalpy barrier. The rate becomes,

$$\begin{aligned} r_e &= \nu_e \bar{c}_{M^{n+}} e^{-(\Delta G_e^\ddagger + \beta nF\Delta\phi_{eq})/RT} \\ &= r_e^0 e^{-\beta nF\Delta\phi_{eq}/RT} \end{aligned} \tag{14.2.1}$$

It follows that the equilibrium electronation current density in A m⁻² is,

$$i_e^{\text{eq}} = r_e nF \quad (14.2.2)$$

where F is the Faraday constant. We have made two principal approximations here. One is that the electric field is constant across the double layer so that the potential drop is linear. The second is that the only source of electric field is the surface charge on the metal and the neutralising surface charge density distributed over the outer Helmholtz plane of the electrolyte. Actually, even if the metal were neutral, there will be alignment of water molecules and ions in the solution reacting to image forces from the metal which may result in a dipole electric field in the double layer. Actually we will see in a moment that all the numerous approximations we have made all get swept up into a quantity which we call equilibrium exchange current density, which can anyway be measured and we won't have to worry any more where it comes from!

14.3 De-electronation

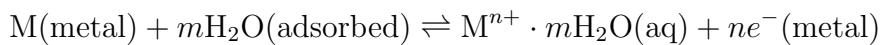
For the moment we are considering the equilibrium state, so that simultaneously there is a current in the opposite direction corresponding to de-electronation—the backward reaction of (14.1.1). We should write something like,

$$r_d^0 = \nu_d \text{ (areal concentration)} e^{-\Delta G_d^\ddagger / RT}$$

where ΔG_d^\ddagger is the free enthalpy of activation of de-electronation. But what is the “concentration” of the metal atoms in the electrode? If we instead use activity, then we might expect to insert the activity of a pure metal, namely one. But if a metal atom is to detach from the surface it should have a higher energy than a bulk atom: firstly there is the surface free energy that it possesses compared to the bulk, secondly, surfaces are not vicinal or flat. The most likely atom to detach is an atom at a kink site of an edge between two terraces. Let us denote its activity \bar{a}_{kink} in mole m⁻². So we should write,[†]

$$r_d^0 = \nu_d \bar{a}_{\text{kink}} e^{-\Delta G_d^\ddagger / RT}$$

[†] West prefers to write the backward reaction of (14.1.1) in the more illustrative form,



to indicate that a surface metal atom reacts with m water molecules adsorbed on the surface and migrates through the double layer as a hydrated ion. Water has the wonderful ability to point its charged ends towards ions effectively screening their electric field and indeed stabilising charged or zwitterionic species that would otherwise be unstable in the vacuum (A. T. Paxton and J. B. Harper, *Mol. Phys.*, **102**, 953 (2004), *loc. cit.* p. 1981). This must have a large role to play in the structure of the activated complex and in deciding what is the free enthalpy of activation of the reaction (14.1.1). A further complication that I shall ignore is that if we take this more realistic picture of the de-electronation, or metal oxidation, reaction then it is no longer described by first order kinetics. In that case the rate coefficient has different dimensions to k_1 . In addition it is not clear what charge to associate with a screened cation so the electric work is also problematic. Such complications are best avoided, especially considering the comment at the end of the previous section.

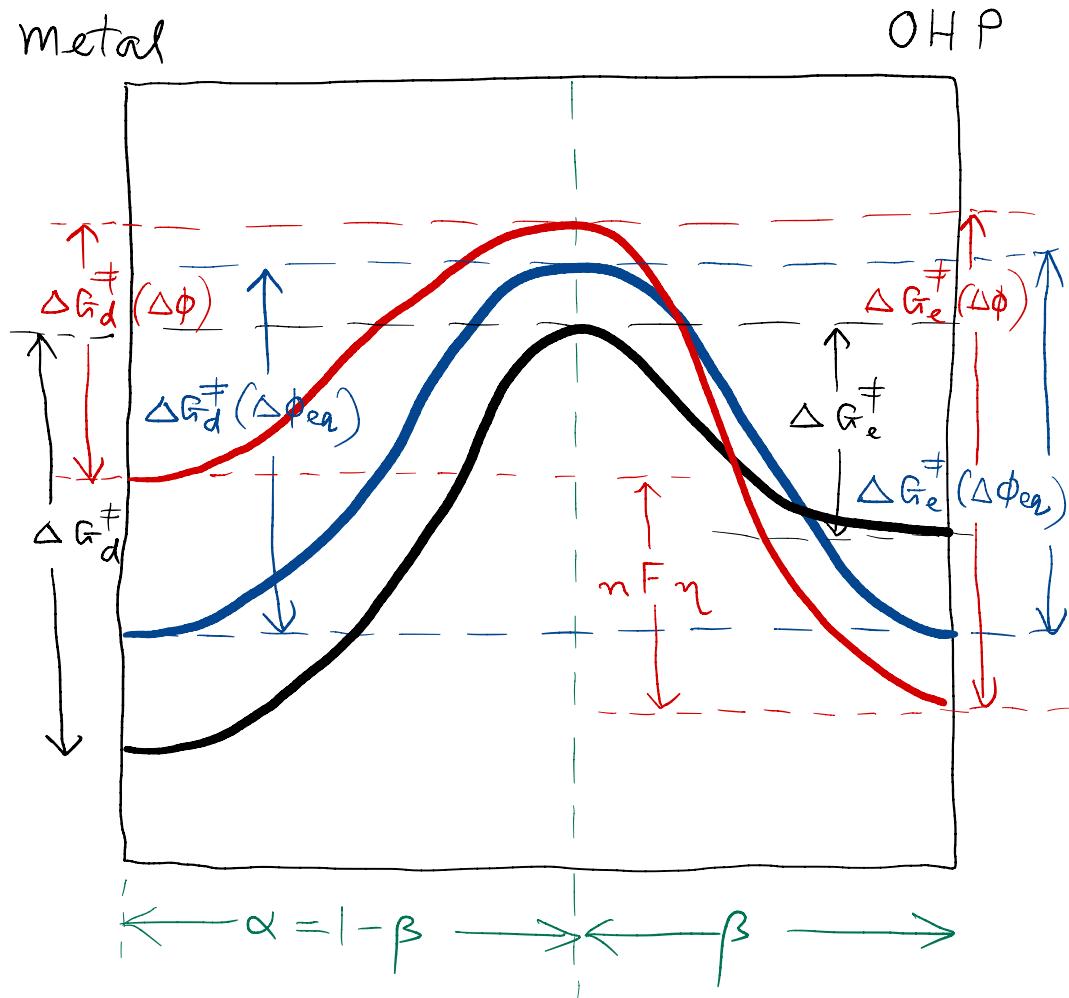


FIGURE 14–1: A cartoon of the free enthalpy profile over the reaction coordinate in which an M^{n+} ion is carried across the double layer from right to left, from the OHP to the metal, where it picks up an electron and is reduced to an M atom; conversely an M atom leaves an electron behind at the electrode and is carried to the OHP. The black line represents the zero field case, $\Delta\phi = 0$, in which the electric potential is the same in the metal and at the OHP. The blue line corresponds to the equilibrium of (14.1.1), $\Delta\phi = \Delta\phi_{eq}$. If, *in addition* a potential difference $\eta = \Delta\phi - \Delta\phi_{eq}$ is applied across the double layer (say, by attaching the electrode to a battery) then the free enthalpy profile is as shown by the red line. In either case, the potential difference alters the free enthalpy difference between the minima at the left and right by an amount $nF\Delta\phi$. In effect the curves are “tilted” by an amount $nF\Delta\phi$; this alters the height of the maximum—not by the full amount of tilt, $nF\Delta\phi$, but by roughly half that if the maximum is roughly half way along the reaction coordinate. In equilibrium we have the free energy of product and reactant equal and hence, $\Delta G_d^\ddagger(\Delta\phi_{eq}) = \Delta G_e^\ddagger(\Delta\phi_{eq})$ as shown. Symmetry factors, α and β are shown for a total reaction coordinate of one.

Now if we also account for the electric field in the double layer, we must have

$$\begin{aligned} r_d &= r_d^0 e^{+\alpha nF\Delta\phi_{eq}/RT} \\ &= r_d^0 e^{(1-\beta)nF\Delta\phi_{eq}/RT} \end{aligned} \quad (14.3.1)$$

where α is the fraction of the total distance travelled by the metal ion from the metal surface until it reaches the point in the double layer where there is a maximum in the free enthalpy. Rather obviously $\alpha = 1 - \beta$. Much of the corrosion literature calls α the *symmetry factor* (or even “transmission coefficient”—which is a phrase I have already used at the end of section 13.7, or “transfer coefficient”). The equilibrium de-electronation current density is,

$$i_d^{\text{eq}} = r_d nF \quad (14.3.2)$$

The signs in (14.2.1) and (14.3.1) make sense. If the metal is noble, for example copper, and hence positively charged with respect to the electrolyte then $\Delta\phi_{eq} > 0$ and so the barrier is raised for the electronation (metal deposition—reduction) and lowered for de-electronation (metal oxidation). This situation is illustrated by the black and blue markings in figure 14–1 (ignore the red for now). Conversely if the metal is base, for example magnesium, it is negatively charged with respect to the electrolyte then $\Delta\phi_{eq} < 0$ and so the barrier is lowered for the electronation and raised for de-electronation. This sounds counterintuitive but see the footnote on page 35, and note that this is a discussion of the *equilibrium* rates of electronation and de-electronation—there is neither overall reduction nor oxidation occurring. In the next section we see how a positive “overpotential” applied to the metal will upset the equilibrium and drive the oxidation of metal atoms to their ions.

14.4 Further interpretations of the symmetry factor—overpotential

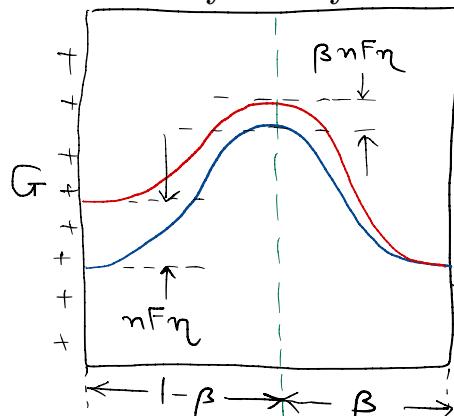


FIGURE 14–2

A simple, and often described, interpretation of the symmetry factor, α , arises from a simplified version of figure 14–1, sketched in figure 14–2. Imagine that the blue and red lines merge at the right. This implies that the electric potential shift due to the electric field is entirely accomplished by the change of potential on the electrode. In fact, with a view to what follows in section 15, we have allowed, by the use of red markings in figure 14–1, an interphase potential difference of an amount $\Delta\phi$ volts: not

necessarily the equilibrium potential, $\Delta\phi_{\text{eq}}$, since in a general case the electrode may be connected to power supply. Imagine then that the electrolyte is very concentrated so that any change in potential in the solution is screened out by counter ions. Consider the electronation reaction in terms of figure 14–2. The M^{n+} ion moving from right to left encounters a barrier in equilibrium shown by the maximum in the blue curve. By how much is that barrier raised if the free enthalpy of the ion on arrival at the metal is raised by the application of an additional potential difference,

$$\eta = \Delta\phi - \Delta\phi_{\text{eq}} \quad (14.4.1)$$

You can see that the nearer the barrier is to the OHP side of the double layer the less it is raised; by linear interpolation the increase in free enthalpy of activation is $\beta nF\eta$ as shown. Compare this with the argument that led to equation (14.2.1). Conversely, if the M^{n+} ion is moving from left to right, having left its electron(s) behind in the metal, then the additional electric field supplied by the “overpotential”, η , only assists it as far as the maximum, after which it is downhill anyway; hence the rate is enhanced as in (14.3.1) by only that fraction $\alpha = 1 - \beta$ of the total distance travelled times the increase in free enthalpy, $nF\eta$, of the ion at the electrode surface.

A third interpretation is mathematical. By reference to figure 14–1 we may regard $\Delta G_e^\ddagger(\Delta\phi)$ and $\Delta G_d^\ddagger(\Delta\phi)$ as functions of the electric field, or equivalently the potential difference $\Delta\phi = \phi_m - \phi_s$ across the double layer (not necessarily the equilibrium potential difference, $\Delta\phi_{\text{eq}}$). Then the “chemical” activation free enthalpies are $\Delta G_d^\ddagger(0)$ and $\Delta G_e^\ddagger(0)$. We now expand the activation free enthalpies at an arbitrary $\Delta\phi$ about the equilibrium. To first order,

$$\Delta G_d^\ddagger(\Delta\phi) = \Delta G_d^\ddagger(\Delta\phi_{\text{eq}}) - \alpha nF\eta \quad (14.4.2a)$$

$$\Delta G_e^\ddagger(\Delta\phi) = \Delta G_e^\ddagger(\Delta\phi_{\text{eq}}) + \beta nF\eta \quad (14.4.2b)$$

in which $\eta = \Delta\phi - \Delta\phi_{\text{eq}}$ is called the *overpotential*. For now, α and β are just positive numbers. The signs reflect that an increase in electric potential at the electrode *lowers* the barrier for de-electronation, while it *raises* the barrier for electronation. By the Taylor expansion the symmetry factor is interpreted as the first order change in activation free enthalpy due to a change in electrode potential,

$$\alpha = - \left. \frac{\partial \Delta G_d^\ddagger}{\partial \Delta\phi} \right|_{\Delta\phi=\Delta\phi_{\text{eq}}} \quad (14.4.3a)$$

$$\beta = \left. \frac{\partial \Delta G_e^\ddagger}{\partial \Delta\phi} \right|_{\Delta\phi=\Delta\phi_{\text{eq}}} \quad (14.4.3b)$$

If we were interested in the equilibrium as in sections 14.2 and 14.3, then we'd expand instead about the zero electric field free enthalpies and in (14.4.2) we would replace the left hand sides by $\Delta G_d^\ddagger(\Delta\phi_{\text{eq}})$ and $\Delta G_e^\ddagger(\Delta\phi_{\text{eq}})$, and on the right hand sides we'd have $\Delta G_d^\ddagger(0)$ and $\Delta G_e^\ddagger(0)$ and replace η with $\Delta\phi_{\text{eq}}$, that is, we now consider the black and blue not the blue and red lines. Then you'd see that we obtain the rates (14.2.1) and (14.3.1) confirming that the first derivatives are indeed the symmetry factors as we have defined them, aside from the subtlety of exactly where we evaluate the derivatives

in (14.4.3) which doesn't matter to lowest order. Finally I can point out that, in reference to figure 14–1, we have

$$\Delta G_d^\ddagger(\Delta\phi_{\text{eq}}) = \Delta G_e^\ddagger(\Delta\phi_{\text{eq}})$$

is the equilibrium free enthalpy of activation of the redox reaction (14.1.1). We also see from figure 14–1 that,

$$\Delta G_d^\ddagger(\Delta\phi) - \Delta G_e^\ddagger(\Delta\phi) = nF\eta$$

By substitution into (14.4.2) there results

$$\alpha + \beta = 1$$

which is the required property of the symmetry factors.

In this section I have covered two new matters, both of which will take centre stage in what follows, and are particularly prominent in the theory of corrosion. I have extended the equilibrium considerations of sections 14.2 and 14.3, to a more general situation in which the coupon of metal is not merely inserted in the solution of its ions, it is also connected to an electrical source or drain. This introduces the *overpotential*, η , in volts. Secondly you have three ways of thinking about the symmetry factors.

- (i) α and β represent a fraction of the total amount of electrical work, $nF\Delta\phi$, that has to be done on the ion in de-electronation and electronation, respectively; reflecting the idea that once the ion reaches that point in its travel at which it encounters the activation free enthalpy maximum, from that point on the journey is “downhill” anyway so no further electrical work needs to be done.
- (ii) There is a simple tilting picture as sketched in figure 14–2. If a free enthalpy profile, such as, say, the blue curve in figure 14–1, is held fixed at the solution OHP and the left hand end is pivoted about that point then if the left hand point is shifted by $nF\eta$ then the maximum in the curve is shifted by a fraction β of that if it is situated a fraction β of the total width of the the double layer from the outer Helmholtz plane. Similarly if the curve is pivoted from the left and the solution potential is shifted by $nF\eta$ then the maximum changes by only the fraction α of that.
- (iii) The first order expansion of the activation free enthalpy (14.4.1) clearly defines the symmetry factors α and β as in (14.4.3), namely they are minus the equilibrium first order change in the “chemical” activation free enthalpy for de-electronation, ΔG_d^\ddagger , with potential, and the equilibrium first order change in the “chemical” activation free enthalpy for electronation, ΔG_e^\ddagger , with potential, respectively. This definition with respect to Taylor expansion about equilibrium ensures the condition $\alpha+\beta=1$ which otherwise we regard as a geometric construction as in figure 14–1.

14.5 Equilibrium exchange current density

I will write out equations (14.2.2) and (14.3.2) in full, following the derivations in the previous sections,

$$i_d^{\text{eq}} = nF\nu_d \bar{a}_{\text{kink}} e^{-\Delta G_d^\ddagger/RT + \alpha nF\Delta\phi_{\text{eq}}/RT} \quad (14.5.1a)$$

$$i_e^{\text{eq}} = nF\nu_e \bar{c}_{M^{n+}} e^{-\Delta G_e^\ddagger/RT - (1-\alpha)nF\Delta\phi_{\text{eq}}/RT} \quad (14.5.1b)$$

These are the *equilibrium* electronation and de-electronation current densities, and on account of detailed balancing they must be numerically identical.[†] This leads us to a central quantity in electrochemistry, namely the *equilibrium exchange current density*,

$$i_0 = i_e^{\text{eq}} = i_d^{\text{eq}}$$

In view of all the uncertainties (and the twenty-odd pages of notes) leading to equations (14.5.1) you may be either happy or depressed to learn that we never need to write them down again! It is enough that we have “understood” processes at the atomic level that are occurring in equilibrium at the surface of a metal immersed in a solution of its cations. An important consequence of the appearance of the activity of a metal atom at a kink site in (14.5.1b) is the role that surface *microstructure* plays. A perfect, ideal low index surface will possess a very different equilibrium current density compared to a real metal surface. This is particularly important in corrosion in assessing the roles of emerging dislocations and grain boundaries and the different crystal faces that a typical polycrystalline metal surface presents to the outside.[‡] All the same, from now on we accept the reality of an exchange current density and I will show in what follows that this is a *measurable* property of a single electrode and moreover that even the crazy *symmetry factor* can be shown to be a measurable reality. Exchange current density may vary from about 10,000 Am⁻² to as little as 10⁻⁵ Am⁻². It is surprising that there is a nine orders of magnitude variation across different metals.

14.6 Interpretation of the equilibrium potential

Before leaving equations (14.5.1) let us set the two right hand sides equal to each other, as indeed they are in equilibrium. Then we get,

$$nF\Delta\phi_{\text{eq}} = \Delta G - G^0 \quad (14.6.1)$$

in which we may regard as “constant”,

$$G^0 = -RT \ln \frac{\nu_e \bar{c}_{M^{n+}}}{\nu_d \bar{a}_{\text{kink}}}$$

and, in reference to figure 14–1,

$$\begin{aligned} \Delta G &= \Delta G_d^\ddagger - \Delta G_e^\ddagger \\ &= G_M - G_{M^{n+}} \end{aligned}$$

[†] I do not use plus or minus signs to indicate the direction of current flow—that is obvious from the context. Often, i_d is called the “oxidation current density”, i_{ox} , and i_e the “reduction current density”, i_{red} ; and if signs *are* used then it is conventional to take i_{ox} and positive and i_{red} as negative. Remember that if you reduce a species then you donate electrons to it, and that oxidation is the taking away of electrons.

[‡] In their chapter 15, Schmickler and Santos have some much deeper insights into the atomic scale processes of metal deposition and wastage and show some detailed models of how these impact on the kinetics.

is a sort of free enthalpy change for the reaction (14.1.1). Here are some points to note.

- (i) It is the equality of (14.4.1a) and (14.4.1b) that establishes the value of $\Delta\phi_{\text{eq}}$, since all other quantities are fixed by the nature of the experiment. This tells us something fundamental about the equilibrium electric potential that is established across the double layer when a coupon of pure metal is placed into a solution of its ions.
- (ii) The free enthalpy difference, ΔG , is not the thermodynamic free enthalpy change of the reaction (14.1.1) because it involves the free enthalpy change of the specific reaction in which a hydrated metal cation is transferred through an electric field to become adsorbed at a surface kink site of the metal.
- (iii) From that perspective we may attempt the sort of thought experiment beloved of Cheng and Sprik, and Schmickler and Santos (see section 7, page 25). We take a metal atom to the Bockris point outside a metal vacuum interface, this costs us free enthalpy per atom (or metal cohesive energy) $-\Delta G_{\text{coh}}^M$; we strip off n electrons, which costs us the first n ionisation energies; we return these electrons to the metal, gaining us n times the metal's work function; we transfer the M^{n+} cation through the vacuum from the Bockris point of the metal to the Bockris point of the solution, the work done is $-ne(\psi_m - \psi_s)$; finally we plunge the cation into the solution requiring an amount of work $-\Delta_{\text{sol}}^r G(M^{n+})$ which is minus the real free enthalpy of solvation of the M^{n+} ion (see page 25). The result (per cation) is,

$$\Delta G \approx \Delta G_{\text{coh}}^M - (I_1 + \dots + I_n) + nW_M + ne(\psi_m - \psi_s) + \Delta_{\text{sol}}^r G(M^{n+}) \quad (14.6.2)$$

This is an approximate formula because it does not account for the difference in free enthalpy of bulk atoms and cations and those at the metal surface and outer Helmholtz plane, respectively; and it does not account for the electric charges of the two phases.

- (iv) All the same this formula throws some light on the factors advanced by Fawcett (see page 39) that determine the position of a particular metal in the electrochemical series. If we associate $\Delta\phi_{\text{eq}}$ with standard half cell electromotance, \mathcal{E}_M° , and if we then compare (14.6.1) and (14.6.2) we see that what controls a metal's position in the electrochemical series are indeed the free enthalpy of solvation of the M^{n+} cation, the free enthalpy (cohesive energy) of the metal, and the ionisation energies of the metal atom as stated by Fawcett. Furthermore we find that further factors are involved, namely the work function of the metal and difference in Volta potentials between the metal and the electrolyte.

14.7 Nernst equation from the point of view of kinetics

Please return to section 9.5, page 39. Reconsider the half cell reaction,



Let us suppose we can write its equilibrium exchange current density in terms of electronation and de-electronation rate coefficients in first order kinetics; the rate of the forward and back reactions being given by the rate coefficient times the *activity* of the

species. The electronation and de-electronation (reduction and oxidation) rate coefficients are, by reference to (14.2.1) and (14.3.1),

$$k_e = \nu_e e^{-\Delta G_e^\ddagger / RT}$$

$$k_d = \nu_d e^{-\Delta G_d^\ddagger / RT}$$

and the equilibrium exchange current density is,

$$i_0 = i_e^{\text{eq}} = i_d^{\text{eq}} = nFk_e a_{\text{Ox}} e^{-(1-\alpha)nF\Delta\phi_{\text{eq}}/RT} = nFk_d a_{\text{Red}} e^{\alpha nF\Delta\phi_{\text{eq}}/RT} \quad (14.7.1)$$

Rather as in section 14.6 we find that the symmetry factor cancels when we resolve this equality into,

$$\Delta\phi_{\text{eq}} = \frac{RT}{nF} \ln \frac{k_d}{k_e} + \frac{RT}{nF} \ln \frac{a_{\text{Ox}}}{a_{\text{Red}}}$$

In the standard state in which the activities are one, the second term on the right hand side vanishes and we are led to define a standard equilibrium potential,

$$\Delta\phi_{\text{eq}}^\circ = \frac{RT}{nF} \ln \frac{k_d}{k_e}$$

so bearing in mind that the equilibrium potential difference (14.1.2) is defined for a particular concentration of M^{n+} ions in solution, we can express the equilibrium electrode potential difference as,

$$\Delta\phi_{\text{eq}} = \Delta\phi_{\text{eq}}^\circ + \frac{RT}{nF} \ln \frac{a_{\text{Ox}}}{a_{\text{Red}}} \quad (14.7.2)$$

This is the half cell or *single electrode* form of the Nernst equation. You can recover equation (9.5.2) which is the usual form of the Nernst equation expressed with respect to the standard hydrogen electrode (SHE) simply by imagining that our half cell is connected to a SHE as in figure 9–1. In that case add to both sides of (14.7.2),

$$(\phi_{\text{Cu}'} - \phi_m) + (\phi_s - \phi_{\text{Pt}}) + (\phi_{\text{Pt}} - \phi_{\text{Cu}})$$

if we are using copper leads, and you end up with,

$$\mathcal{E} = \mathcal{E}^\circ + \frac{RT}{nF} \ln \frac{a_{\text{Ox}}}{a_{\text{Red}}} \quad (9.5.2)$$

which is the Nernst equation expressing the equilibrium single electrode potential on the SHE scale as a function of the activities.

What we've achieved in this section is to derive the Nernst equation from the kinetics of the redox reaction, rather than the thermodynamics, by setting the equilibrium condition that equates the forward and backward current densities. I repeat that the symmetry factor has cancelled, which indicates that there is *no equilibrium property* or measurement that can be used to extract symmetry factors—they are a strictly *kinetic* quantity (which we will see can be measured).

14.8 Measurement of the exchange current density

In figure 2–4, page 5, I sketched the “next simplest electrochemical cell”. In figure 14–3 I sketch the “next-next simplest electrochemical cell”. On the left is the familiar standard hydrogen electrode (see section 9.4). This is connected through a high impedance voltmeter (to prevent any current flowing) to a copper test electrode which is immersed in hydrochloric acid (HCl). The finger-like probe is a “Luggin capillary”; its purpose is to minimise any electrical resistance in the electrolyte so that the SHE measures exactly the potential of the electrolyte very close to the interphase, relative to the SHE. With the switch open this is equivalent to the electrochemical cell in figure 9–1, page 34; and you do not for a moment imagine that this gives us a direct measure of $\Delta\phi_{eq}$ because of the usual bloody mindedness of electrochemical experiments causing additional potential differences to intervene. With no current flowing, the potential difference measured at the voltmeter is,

$$\begin{aligned}\mathcal{E}_{eq} &= (\phi_{Cu'} - \phi_{HCl}) + (\phi_{HCl} - \phi_{Pt}) + (\phi_{Pt} - \phi_{Cu}) \\ &= \Delta\phi_{eq} + \Delta\phi_{SHE} + (\phi_{Pt} - \phi_{Cu})\end{aligned}$$

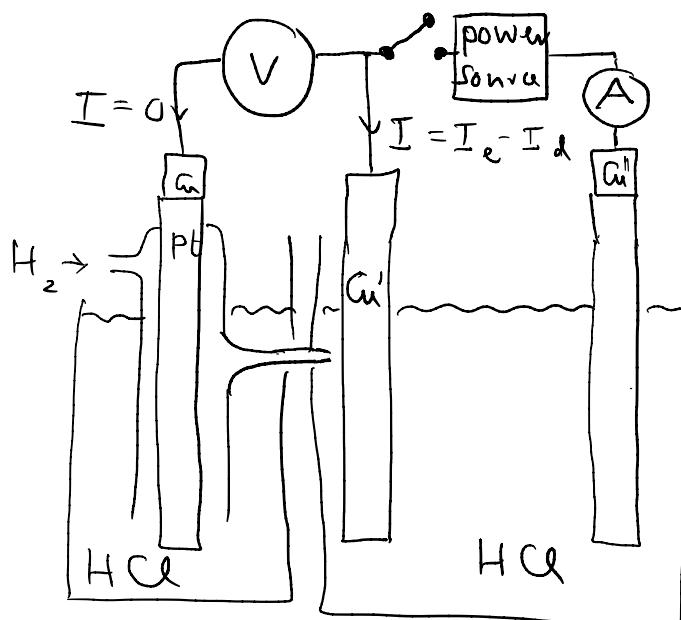


FIGURE 14–3

When the switch is closed, the cell has *three electrodes* in order to be able to control the current $I = I_e - I_d$ by varying the voltage of the power supply independently of the voltage on the SHE. The right hand electrode is a so called auxiliary or *counter electrode*. It is made of an inert conductor such as graphite so that the only electrochemical reaction involves the electrolyte: for example the oxidation of chloride ions to chlorine gas. If the power source is turned off, then the voltmeter records the equilibrium potential difference, \mathcal{E}_{eq} . As the voltage in the power source is ramped up, the current, I , is measured at the ammeter and the voltage recorded at the voltmeter is,

$$\mathcal{E} = \Delta\phi + \Delta\phi_{SHE} + (\phi_{Pt} - \phi_{Cu}) - IR$$

where $\Delta\phi$ is the non equilibrium potential difference across the double layer of the test electrode, and R is the resistance of the electrolyte between the tip of the Luggin capillary and the test, or *working* electrode. The point is that we can eliminate the “ IR -drop” by means of careful assembly and glass blowing. We may also take it that as the voltage in the power source and the resulting current are ramped up, the standard potential of the SHE and the inner potential difference at the copper–platinum interface in the SHE remain constant.

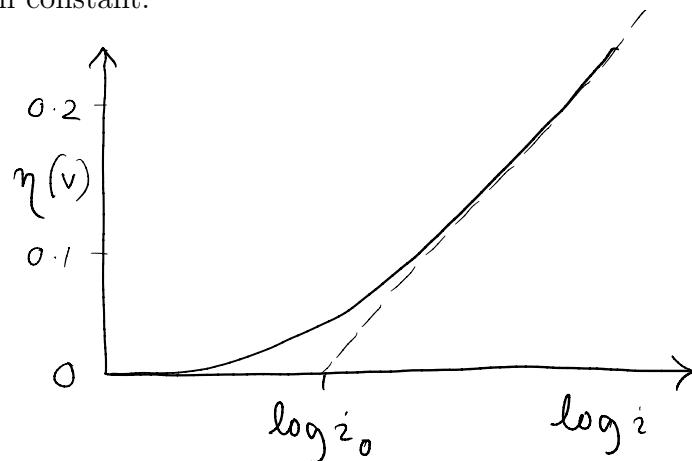


FIGURE 14-4

If the logarithm of the current density, i , (I divided by the surface area of the electrode) is plotted against the voltage, $\mathcal{E} - \mathcal{E}_{\text{eq}} = \Delta\phi - \Delta\phi_{\text{eq}} = \eta$, a plot as figure 14-4 is typically obtained. An extrapolation to zero voltage permits a measurement of the equilibrium exchange current density, i_0 . This is an elementary example of *voltammetry* which we encountered in section 11.3 (page 59).

So the point is that i_0 is *measurable*. On the other hand it is not possible to manipulate equations (14.4.1) as we could to find a simple form for $\Delta\phi_{\text{eq}}$. Indeed, unlike $\Delta\phi_{\text{eq}}$, i_0 is seen to depend on the height of the activation barrier and on at least one of the vibrational and concentration terms. It must be because of the exponential dependence on the activation free enthalpy that there is a nine order of magnitude variation in i_0 across a range of metals. Mind you, that still translates into a factor of about twenty in comparing the smallest and largest activation barrier, which is rather puzzling. The frequency and concentration prefactors cannot vary that much between metals. This implies that the reality is more complicated than the picture I have drawn. Below is a table of some exchange current densities in water at 25°C. (Please just look at the first four lines and four columns of data for now.) It is also puzzling that there is a variation of ten or a hundred even for the same metal and electrolyte. This may reflect the importance of the state of the surface as I have mentioned already. There is a rough correlation between exchange current density and equilibrium potential (page 38). The more noble metals have the largest exchange current density.

redox system	metal	i_0 (Am ⁻²)	\mathcal{E}_M° (V/SHE)	Tafel slope, b (V/decade)
$\text{Ag}^+ \mid \text{Ag}$	Ag	10^4	+0.80	0.03–0.12
$\text{Cu}^{2+} \mid \text{Cu}$	Cu	1–10	+0.34	≈ 0.06
$\text{Zn}^{2+} \mid \text{Zn}$	Zn	10^{-3} – 10^{-1}	-0.76	0.03–0.06
$\text{Fe}^{2+} \mid \text{Fe}$	Fe	10^{-5} – 10^{-4}	-0.44	0.05–0.08
$\text{H}^+ \mid \text{H}_2$	Pt	10^2	0	≈ 0.12
$\text{H}^+ \mid \text{H}_2$	Fe, Cu	10^{-3} – 10^{-2}	0	≈ 0.12
$\text{H}^+ \mid \text{H}_2$	Zn	10^{-7}	0	≈ 0.12
$\text{H}^+ \mid \text{H}_2$	Pb	10^{-9}	0	≈ 0.12
$\text{O}_2 \mid \text{OH}^-$	Pt	10^{-6}	+0.4	0.10–0.15
$\text{O}_2 \mid \text{OH}^-$	Fe	$\approx 10^{-10}$	+0.4	>0.12

14.9 Interpretation of the potential of zero charge

I have said that an isolated piece of metal placed in water, or in a solution of its ions, will spontaneously acquire a charge. As a consequence, the solution acquires an equal and opposite charge and a potential difference, $\Delta\phi_{\text{eq}}$, is developed across the interphase between metal and solution (electrolyte). This is of course accompanied by an electric field across the interphase, which we have taken as constant. I have then asserted that the conditions of charge and the potential difference can be modified by connecting the metal to ground, or to a source or drain of charge such as a battery. By connecting the metal to a negative or positive terminal of a battery we effectively supply an *overpotential*, η , (itself a potential difference),

$$\eta = \Delta\phi - \Delta\phi_{\text{eq}} \quad (14.1.1)$$

and $\Delta\phi$ is the resulting potential difference across the interphase. Therefore I can find that overpotential which results in zero potential difference across the interphase; and in that condition the metal is neutral. Since the metal now carries zero charge, I have arrived at that overpotential which corresponds to the potential of zero charge that was encountered first in these notes in section 11.2, page 55. In that case we have,

$$\eta_{\text{pzc}} = -\Delta\phi_{\text{eq}}$$

since $\Delta\phi$ is zero. Remember that in section 11.2, page 56, the PZC, ϕ_{pzc} , is $\phi_s - \phi_m$ whereas η is $\phi_m - \phi_s$ so that we expect to find,

$$\phi_{\text{pzc}} = \Delta\phi_{\text{eq}}$$

You have seen in section 14.7 that the equilibrium potential difference, $\Delta\phi_{\text{eq}}$, is related to the standard electrode potential if both are measured with respect to the standard hydrogen electrode and if the metal ions in the electrolyte are at unit activity. This implies a relation between the PZC and the equilibrium standard electrode potential.

So I can combine data from the electrochemical series on page 38 with some potentials of zero charge on page 58, which I took from Bockris. You do not expect equality between these two because although both are referred to the SHE, the electrolytes in the measurements of the PZC are not at unit activity of the metal ions. The comparison is presented graphically in figure 14–5. The correlation is not wholly convincing, but the trend is there.

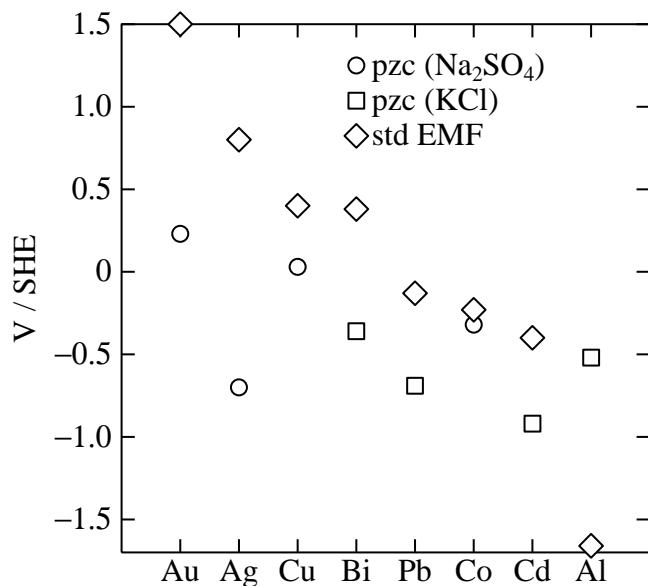


FIGURE 14–5

14.10 Final remarks on the baseness and nobility of metals

I have now attempted to attach some physical substance to the electrochemical series—perhaps the most important feature of electrochemistry.

- (i) I have remarked on page 39 upon the essential fact that *some metals are base* and there is nothing the materials designer can do about this depressing thought. In section 17 we will come to the subject of passivation and the existence of “stainless” metals.
- (ii) Fawcett has given a list of thermodynamic properties that influence the position of a metal in the electrochemical series (see page 39), which I repeat and extend on page 90.
- (iii) I have now indicated a possible correlation between potential of zero charge, equilibrium potential difference and position in the electromotive series. Application of the Nernst equation should bring these into closer coincidence if the conditions of the experiment are known.
- (iv) I have also pointed out on page 92 that there is a rough correlation between position in the electromotive series and exchange current density, although the latter is also affected by the catalytic properties of the electrode metal.

15. Overpotential and the Butler–Volmer equation

We are ready to derive the Butler–Volmer equation. By now you are aware that an isolated metal in a solution of its ions inevitably sets up an equilibrium potential difference between metal and solution which we have called $\Delta\phi_{\text{eq}}$ (see equation 15.1.2). We have also seen that it represents an *unmeasurable* surrogate for the standard electrode potential, E_M° , on the SHE scale (see the electromotive force series on page 38) once adjusted to standard conditions using the Nernst equation. In addition I have indicated that the equilibrium will be upset if an additional overpotential, η , is applied to the electrode (or the electrolyte). To my mind there is an analogy between free charge and bound charge in classical electrostatics, and overpotential and equilibrium potential in electrochemistry. In the study of ponderable electric media we learn that an applied electric field will cause matter to polarise and that so called bound charge will appear within the dielectric and on its surfaces as a result of atomic scale charges becoming shifted from equilibrium as a result of the applied field. The applied field of course is created by us by the manipulation of electronic circuitry, for example by moving charge onto the plates of a capacitor. In that sense we call the charge that we manipulate and control the *free charge*, and the *bound charge* or *inaccessible* charge is not measurable and is out of our control being the inevitable result of matter having an electric susceptibility. In the same way, we control the *overpotential* in an electrochemical experiment, but we do not control the unmeasurable potential drops across the interphase which result from atomic scale movements of charge, which again are dictated by the properties of the electrode and the electrolyte which are outside our control. Indeed just as in electrostatics, it is not clear to me that we may always unambiguously identify which part of $\Delta\phi$ is equilibrium potential and which part is overpotential.

However let us take the point of view that the electrode potential, $\Delta\phi$, can be divided into the equilibrium part and the overpotential, as introduced in (14.1.1),

$$\Delta\phi = \Delta\phi_{\text{eq}} + \eta$$

The electrode doesn't care about this division—to it $\Delta\phi$ is just a potential drop across the interphase; but to us it's important because η is the part that we can control: the rest comes along for the ride.

If the potential drop is $\Delta\phi_{\text{eq}}$ then the interphase is in equilibrium; there is no net current flowing and $i_d^{\text{eq}} = i_e^{\text{eq}} = i_0$. A current *will* flow if an overpotential is applied, and the question is what is then the current density, $i = i_d - i_e$. It is the convention that positive charge flowing from electrode to electrolyte is designated a positive current, see the footnote on page 88. We can calculate $i = i_d - i_e$ using an obvious extension of equation (14.7.1) on page 90 to non equilibrium by replacing $\Delta\phi_{\text{eq}}$ with the non equilibrium potential difference across the interphase, $\Delta\phi$,

$$\begin{aligned} i &= nFk_{\text{d}}a_{\text{Red}} e^{\alpha nF\Delta\phi/RT} - nFk_{\text{e}}a_{\text{Ox}} e^{-(1-\alpha)nF\Delta\phi/RT} \\ &= nFk_{\text{d}}a_{\text{Red}} e^{\alpha nF\Delta\phi_{\text{eq}}/RT} e^{\alpha nF\eta/RT} - nFk_{\text{e}}a_{\text{Ox}} e^{-(1-\alpha)nF\Delta\phi_{\text{eq}}/RT} e^{-(1-\alpha)nF\eta/RT} \\ &= i_d^{\text{eq}} e^{\alpha nF\eta/RT} - i_e^{\text{eq}} e^{-(1-\alpha)nF\eta/RT} \\ &= i_0 (e^{\alpha nF\eta/RT} - e^{-(1-\alpha)nF\eta/RT}) \end{aligned}$$

The wonderful thing is that it doesn't really matter how I formulate the equilibrium electronation and de-electronation current densities—for example I could have used (14.5.1)—or how I argue what exactly is the right way to express the “frequency” and “concentration” terms, because by pulling out the equilibrium potential from $\Delta\phi$ and using $i_d^{\text{eq}} = i_e^{\text{eq}} = i_0$ in equilibrium then all the mess cancels and I can fulfill my promise at the end of section 15.2 that we will never need them again.

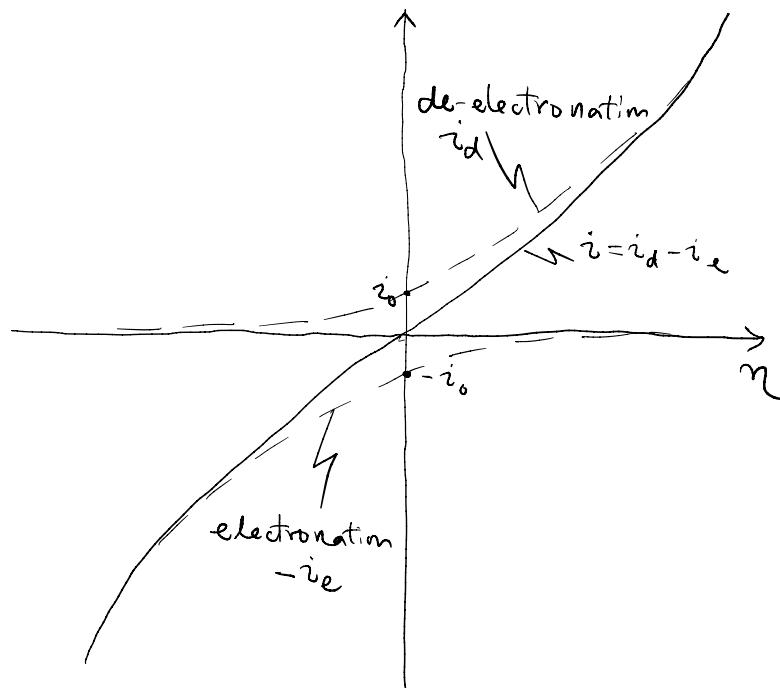


FIGURE 15–1

Just because it's so important I'll repeat the Butler–Volmer equation,

$$i = i_0 (e^{\alpha nF\eta/RT} - e^{-(1-\alpha)nF\eta/RT}) \quad (15.10.1)$$

in fact, this is what is plotted in figure 14–4, and in more detail in figure 15–1. This and the Nernst equation may be the most important statements in electrochemistry, one expressing the kinetics and the other the equilibrium of the interphase. The Butler–Volmer equation is of particular importance in corrosion[†] and in the next few paragraphs I will make some analysis of the limits of small and large overpotential. Points to note in figure 15–1 are the individual de-electronation and electronation current densities which add to give the total current density; the intersections of i_d and i_e with the ordinate at $\eta = 0$ to indicate their equality to the equilibrium exchange current density, i_0 ; and the observation that at large magnitudes of the overpotential the total current density merges with i_d if $\eta > 0$ and i_e if $\eta < 0$, which we will return to at equations (15.10.2) below.

[†] It is fascinating that in the new science and engineering of fuel cells the design is often driven by competition between kinetic and thermodynamic effects which tend to pull in opposite directions. These notes would double in length if I went into fuel cells and batteries, but I hope the reader now has enough background to engage in these fields of vital importance to the climate emergency.

But I should end this section on a strict caveat. We have swept away the details, approximations and imprecise arguments into the exchange current density, i_0 , and obtained a clear physical law that can be readily interpreted and which has great power in the study of corrosion, electroplating, galvanising and so on. But as Schmickler and Santos are careful to point out, the validity of the Butler–Volmer equation must be established on a case-by-case basis and in many situations it is far from being justified. In some cases, a detailed theory can be cast into Butler–Volmer form; however this may result in apparent, and indeed measured symmetry factors that are not between zero and one, for example. So caution is needed.

The limit of small overpotential

If I expand the exponentials in (15.10.1) to first order, the terms in the symmetry factor cancel and I get simply,

$$i = i_0 \frac{nF\eta}{RT} ; \quad nF\eta \ll RT$$

You can see by a quick calculation that at 300K and $n = 1$ this will be good to about a 1% error if the overpotential is less than 0.01 V, which really is quite small. The symmetry factor cannot be measured at such low overpotential.

In this linear regime the current–voltage relation is *ohmic*. We could write,

$$\left(\frac{\partial \eta}{\partial i} \right)_{\text{activity},T} = \frac{RT}{i_0 n F} = \rho d$$

in which ρ is the differential resistivity of the interphase of width d . This throws some light on polarisable and non polarisable (reversible) electrodes which I first introduced in section 2, page 6. This is exactly the resistivity that dictates the resistance in the model of figure 2–6 of the interphase as a resistor and capacitor in series. An ideally reversible electrode such as the SHE is intended to have $\rho \rightarrow 0$ so that it is perfectly “leaky” of current. So it cannot sustain a change in potential difference ensuring that any change in voltage in an experimental cell such as in figure 2–4 is entirely due to the polarisation of the test electrode and not the reference electrode. For this measurement to be valid it is also necessary that the test electrode is ideally polarisable, meaning that $\rho \rightarrow \infty$. In this way in the model of figure 2–6, the polarisable interphase is modelled by a capacitor only and the reversible interphase by a resistor only. Recall that in section 11.3 and 12.1 we get a brief reference to the modelling of an interphase as an equivalent circuit of resistors, capacitors and indeed inductors.

The *reversibility* of the SHE is clearly reflected in the very high equilibrium exchange current density for proton reduction on platinum which is evident in the table on page 93. Conversely as is evident from the above equation a signature of a *polarisable* electrode is one having a very small i_0 .

The limit of large overpotential

It is instructive to make a sketch of the current against overpotential. Note that if $\alpha = 0.5$ then (15.10.1) is,

$$i = i_0 \sinh \frac{1}{2} \frac{nF\eta}{RT}$$

So mathematically, α dictates the asymmetry of the $i - \eta$ relation under a change in polarity: a reversal in the sign of the voltage. Conversely, if α is not equal to one half then as shown in figure 15–2 the $i - \eta$ relation is not symmetric: the interphase is non ohmic—it is *rectifying* (like a diode). This admits a direct measurement of the symmetry factor, simply by reversing the polarity in an experiment such as shown in figure 15–3 and discussed above in section 14.8. This means that at high overpotential the equivalent circuit modelling of the interphase must include some non linear elements.

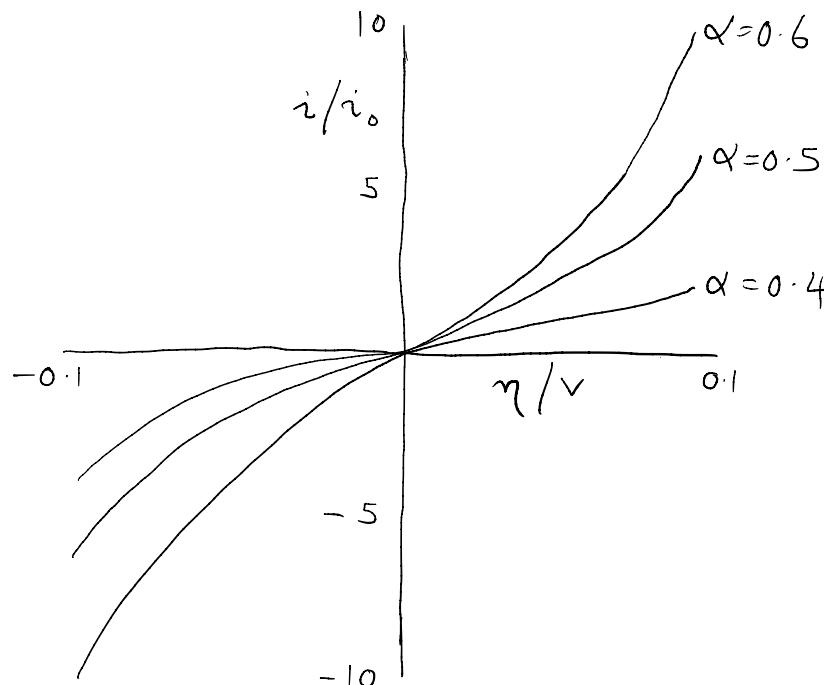


FIGURE 15–2

In the limit of large $|\eta|$ either the first or the second term in the Butler–Volmer equation is neglected (see figure 15–1). In that case the electronation and de-electronation current densities display the asymmetry since,

$$i_d = e^{\alpha nF\eta/RT} ; \quad nF\eta \gg RT \quad (15.10.2a)$$

$$i_e = e^{-(1-\alpha)nF\eta/RT} ; \quad -nF\eta \gg RT \quad (15.10.2b)$$

are only equal if $\alpha = \frac{1}{2}$.

The approximations (15.10.2) are generally correct for $n = 1$ with less than 0.1% error for overpotentials that are greater than 1.2 V in magnitude. Equations (15.10.2) are

frequently expressed in the form of the empirical Tafel's law,

$$|\eta| = a + b \log \frac{|i|}{i_1} \quad (15.10.3)$$

using logarithms to the base of 10 and in which i_1 is a constant of unit current density in order to render the argument to the logarithm dimensionless. a and b are the "Tafel coefficients" and b is frequently expressed in units of volts per tenfold change in i , or "volts per decade". Tafel's law predates by some decades the foregoing theory, which may have been developed, as Professor Bilby used to say, "with an eye to the final result." The Tafel b -coefficient is often called the Tafel slope and often the sign is taken as understood. Magnitudes of some Tafel slopes are given in the table on page 93. We come to the second two sets of data in that table in a moment.

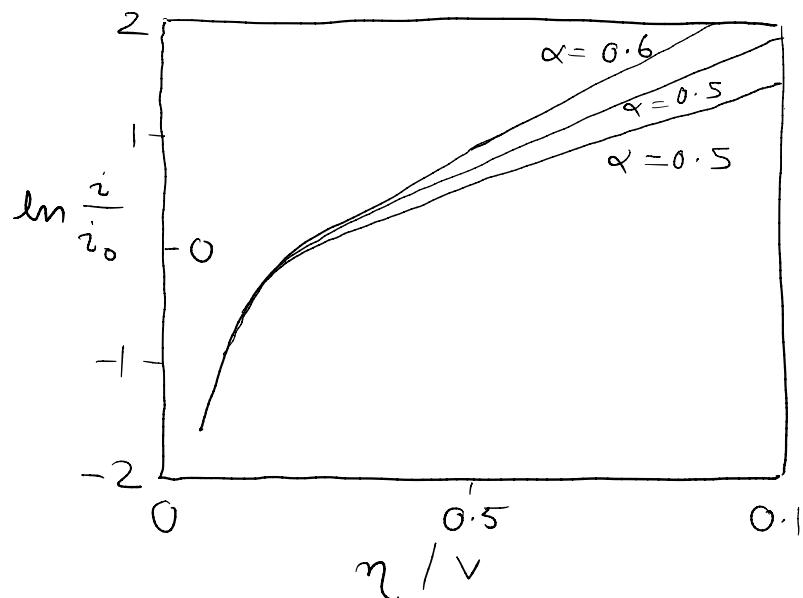


FIGURE 15-3

I have copied figures 15-2 and 15-3 from Schmickler and Santos since they show quantitative graphs of $i - \eta$ and $\ln i - \eta$ relations. Note that for overpotentials greater than about 0.1 V the graph of $\ln i$ versus η is essentially *linear*. This is taken to be valid almost universally in the treatment of corrosion and leads to the widespread use of the Tafel plot and Evans diagrams, which we come to next.

16. The Evans diagram and the corrosion potential

16.1 Corrosion

I wish now to turn to corrosion probably for the remainder of these notes. We need to revisit John West's question of why metals corrode at all in the case that the metal is not directly connected to a source or drain of electrons. As I mentioned before what will happen is that patches of the metal surface will become anodes and patches will become cathodes and current will flow between the two to maintain charge neutrality while

the anode wastes away and the cathode supports some reduction reaction—often the reduction of dissolved oxygen, or in acid solutions the reduction of protons to hydrogen gas. This situation is illustrated in figure 16–1.

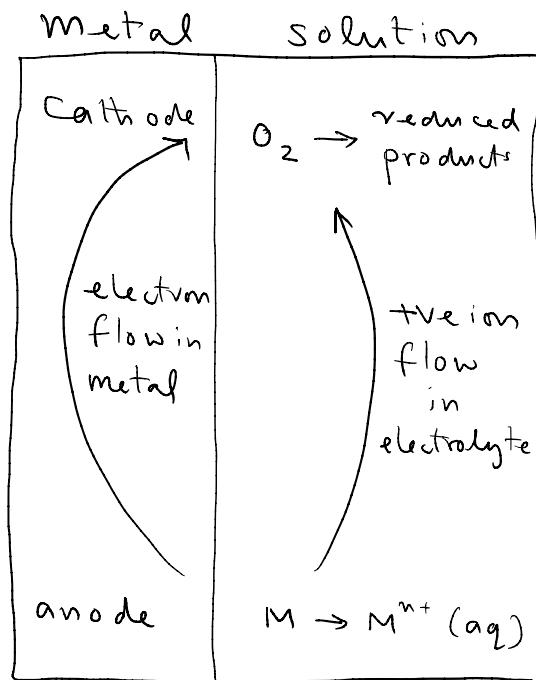


FIGURE 16–1

You could carry out the following experiment if you have access to two chemicals. Take a big lump of iron (say from a scrapyard) and clean and degrease a flat surface and rough it up a bit with emery cloth. Then make up a solution of water with a little salt (optional) and some potassium hexacyanoferrate and some phenolphthalein indicator. Shake it up a bit to saturate it with dissolved oxygen. You place a big drop of this onto the iron surface, maybe about half an inch or so across. The iron will begin to corrode having chosen a number of patches to become anodes; you will see these patches turning blue as the potassium hexacyanoferrate indicates the appearance of Fe^{++} ions in the water. To conserve charge cathodic patches also appear and the cathodic reaction is the reduction of oxygen (see (16.2.1) below). These patches will show up pink which is an indicator of alkali—the presence of OH^- ions. After a while the dissolved oxygen will run out and the appearance will change: the blue patches will turn into a ring of blue around the edge of the drop where further oxygen can be supplied from the air in contact with both the water edge and the metal, and the centre of the drop will become a red cathode. Eventually a ring of rust will appear that separates the outer blue ring and the central pink patch and the experiment will be over.

16.2 Redox equilibrium

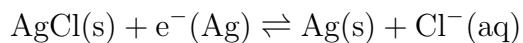
The Nernst equation as we have encountered it up to now applies to electrochemical reactions that actively involve the electrode as it is its atoms that are involved in the so

called *redox reaction* most generally expressed as before, pp. 39 and 89,



However it is more general than this in the sense that the species being oxidised or reduced need not belong to the metal electrode. For example Fe^{++} and Fe^{+++} ions may be in equilibrium in a solution but if an inert electrode such as platinum is inserted into the solution then the ratio of the two ions in equilibrium can be affected by the electric potential imposed by a battery on the platinum. This is because in the absence of an electrode the relative concentrations of Fe^{++} and Fe^{+++} must be fixed by charge neutrality, but the intervention of a metal that might supply or drain electrons can upset this balance. From that point of view a bulk redox reaction becomes a surface redox reaction, and this is a most important point to bear in mind in what follows.

Let me give an example of each of these. Firstly return to the electrochemical cell in figure 9–1 on page 34. There exists an equilibrium,



in which “s” stands for “in solid phase”, “Ag” stands for “in the silver electrode” and “aq” is “in aqueous solution”; and in this case the species being reduced (in the forward reaction) and oxidised (in the backward reaction) is an atom of the electrode metal itself. The standard electromotance of this equilibrium can be worked out as,

$$\mathcal{E}^\circ = \frac{1}{F} (\mu_{\text{AgCl}}^\bullet - \mu_{\text{Cl}^-}^\circ) = \frac{1}{96.48} (-109.8 - 131.1) = -2.5 \text{ V/SHE}$$

where I have used the standard chemical potentials in kJ/mole at 25°C relative to the hydrogen ion from thermodynamic tables. Application of the Nernst equation allows me to calculate the equilibrium potential at different temperature and concentration of chloride ions in the solution,

$$\mathcal{E} = \mathcal{E}^\circ + \frac{RT}{F} \ln \frac{a_{\text{AgCl}}}{h_{\text{Cl}^-}}$$

Now consider this redox equilibrium,



Here, “g” means “gas”, “l” means “liquid” and “Pt” in an inert platinum electrode. So here the electrode plays no part in the reaction *except to act as a source or sink of electrons* and also possibly to catalyse the reaction (increase its rate without itself being altered). Again, I can work out the equilibrium potential of this extremely important reduction of oxygen,

$$\begin{aligned} \mathcal{E}^\circ &= \frac{1}{4F} (2\mu_{\text{H}_2\text{O}}^\bullet - 4\mu_{\text{OH}^-}^\circ) \\ &= \frac{1}{385.9} (-474.4 + 628.8) \\ &= +0.40 \text{ V/SHE} \end{aligned} \quad (16.2.2)$$

Application of the Nernst equation gives us,

$$\begin{aligned}\mathcal{E} &= \mathcal{E}^\circ + \frac{RT}{4F} \ln \frac{p_{O_2}}{h_{OH^-}^4} \\ &= 0.40 + 0.015 \log p_{O_2} - 0.06 \log h_{OH^-}\end{aligned}$$

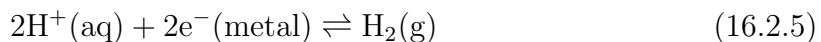
which includes a change from natural to base-10 logarithms. The partial pressure of oxygen must be in bar. Now you need to know what is meant by pH. This is a measure of *acidity* which may be characterised by the presence of H^+ ions (protons) in solution. The measure used is minus the logarithm to the base ten of the concentration (Henrian activity) of hydrogen ions.[†] In terms of the pH then the equilibrium electromotance of the oxygen reduction at a metal surface at 25°C is,

$$\mathcal{E} = 1.23 + 0.015 \log p_{O_2} - 0.06 \text{ pH} \quad (16.2.3)$$

in volts on the standard hydrogen electrode scale. In neutral water with dissolved oxygen in equilibrium with air whose partial pressure is 0.2 bar (20% of one atmosphere), the equilibrium potential for oxygen reduction on metal is,

$$\mathcal{E} = 1.23 + 0.015 \log 0.2 - 0.06 \times 7 = +0.8V/\text{SHE} \quad (16.2.4)$$

In acid solutions the cathodic reaction is likely to be reduction of hydrogen ions rather than dissolved oxygen. In that case the reaction is,[‡]



[†] Consider the equilibrium



The proton does not really exist in water, by H^+ we really mean the hydronium cation as on the right. This equilibrium expresses the fact that water is always slightly dissociated into ions (that's why it conducts electricity even when pure). Pure water at 25°C is *neutral* and the concentration of each ion is equal and such that its activity is 10^{-7} (from observation). We write the equilibrium constant for the dissociation as,

$$K_W = h_{OH^-} h_{H_3O^+}$$

assuming the activity of water is one since there are so few ions. We *define*,

$$pH = -\log h_{H_3O^+} \quad \text{and} \quad pOH = -\log h_{OH^-}$$

Chemists also define the so called *pK* of water which is $pK = -\log K_W$. Because neutral water has $pH = pOH = 7$ the pH is always between zero and 14 and an equivalent measure is the *alkinity*, $pOH = pK - pH$.

[‡] Probably I should write this as,



this is the “hydrogen evolution reaction” (HER) that occurs at the standard hydrogen electrode (SHE) and since the electromotance series is constructed with reference to the SHE, the equilibrium standard electromotance of reaction (16.2.5) is, by construction, zero volts. If the activity of hydrogen ions is not one, then the electromotance is,

$$\begin{aligned}\mathcal{E} &= \frac{RT}{2F} \ln \frac{h_{\text{H}^+}^2}{p_{\text{H}_2}} \\ &= -2.303 \frac{RT}{F} (\text{pH} + \log p_{\text{H}_2})\end{aligned}\quad (16.2.6)$$

again, the partial pressure of hydrogen gas is in bar; and here and in (16.2.4) we assume the gas is ideal. The number $\ln x / \log x = 2.303$ turns up a lot in electrochemistry calculations. Note, also, that $RT/F = kT/e = 0.026$ V at 300K. For reference in figure 16–4 and section 17, below, note that slope of \mathcal{E} versus pH in both HER (16.2.6) and oxygen reduction (16.2.3) is $-2.303RT/F = -0.06$ V.

The table on page 93 shows equilibrium exchange current densities and Tafel slope for the two important reduction reactions occurring next to an electrode, namely reactions (16.2.1) and (16.2.5). I need to emphasise that these are redox processes central to electrochemistry and corrosion but they differ from those we have studied up to now because they do not involve ions of the electrode metal. All the same the inert electrode serves a number of functions. It may *catalyse* the reaction; it may offer a source or drain of electrons (generally a source as these are electronation reactions in the forward direction); it is clear from the table that whereas the standard equilibrium single potential is given by thermodynamics, namely zero for HER and +0.4 V for reduction of dissolved oxygen, the equilibrium exchange current density varies over eleven orders of magnitude for the HER. It seems that the excellent property of platinum as a catalyst is reflected in its large i_0 for these reactions.

16.3 Corrosion potential

Let us recap briefly. A piece of metal placed in a solution of its ions will acquire an electric potential difference across the interphase between its surface and the bulk solution. We usually refer to this interphase loosely as the “double layer” (see section 3, page 6). The numerical value of this potential difference can be calculated by reference to the electromotive force series (page 38) and by use of the Nernst equation (if the concentration of ions is not unimolar at 25°C). We have seen in section 15 that this equilibrium potential difference may be modified by an *overpotential*, η , by connecting the metal electrically to an electron drain or source. We wrote, on page 93,

$$\Delta\phi = \Delta\phi_{\text{eq}} + \eta \quad (16.3.1)$$

and $\Delta\phi_{\text{eq}}$ is given by equation (14.7.2) on page 90. This potential difference is not measurable but we complete a circuit via high impedance voltmeter to a SHE by means of which we effectively subtract from both sides of (16.3.1) the potential difference of the SHE and we get,

$$\mathcal{E} = \mathcal{E}_{\text{eq}} + \eta \quad (16.3.2)$$

A simple question to ask is, will a metal dissolve in an acid solution at unit molality? The answer is yes if the metal's equilibrium electromotance is less (more negative) than that of the hydrogen evolution reaction which at unit activity is zero. So it's easy to find which metals will dissolve in acid at normal strength: it's those metals in the electrochemical series whose electromotance is negative—that is all metals more base than copper. The next question is how can this happen if the metal is electrically isolated? The answer, as I've revealed already, is that the metal divides itself into patches: at the anodes the metal dissolves and at the cathodes hydrogen is evolved (see figure 16–1). But, wait! you say: the metal and the electrolyte are electrically conducting, how can there be different electric potentials at different parts of the metal? Or different parts of the electrolyte for that matter. Well, there can't. In addition to negotiations arriving at which patches will be anodes and which cathodes, there must also be agreement on a common $\Delta\phi$ across the interphase everywhere. So the metal chooses for itself what will be the potential everywhere. In effect this is chosen so that when (16.3.2) is applied at the anode the single potential is still greater than the equilibrium potential of the $M \rightleftharpoons M^{n+}$ equilibrium, while at the cathode it is such that the potential is smaller (or more negative) than the equilibrium potential for the oxidation reaction—either HER or oxidation of dissolved oxygen. This negotiated potential difference is a so called *mixed potential* and in corrosion science it is called the *corrosion potential*, \mathcal{E}_{cor} . Our task is to find what it is and I will show a neat graphical method.

Meanwhile let's do another example. Consider immersing a coupon of pure silver into neutral aerated water. Will it corrode? At first sight the question is moot since the equilibrium potential of oxygen reduction is +0.8 volts (16.2.4) and so is the the value of \mathcal{E}_M° for Ag in the electrochemical series (page 38). However we have not accounted for the concentration of silver ions in the water. Let's suppose there is a trace amount so that their activity is 10^{-6} . Then using the Nernst equation,

$$\mathcal{E}_{\text{Ag}} = \mathcal{E}_{\text{Ag}}^\circ + \frac{RT}{F} \ln h_{\text{Ag}^+} = +0.44 \text{ V/SHE} \quad (16.3.3)$$

so the silver is more base than the oxygen reduction on the metal surface (it doesn't matter that now the oxygen reduction is on silver and not platinum as in my previous calculation starting from (16.2.1) because in both cases the metal is inert since at the cathode patches the silver ions are not playing a role). I now calculate the “driving electromotance” of the combined reactions,

$$\mathcal{E} = 0.8 - 0.44 = 0.36 \text{ V}$$

Combine this with the important equation from section 9.3,

$$\Delta G = -nF\mathcal{E} \quad (9.3.6)$$

and we see that the free enthalpy change associated with corrosion of the silver coupon is,

$$-0.36F = -34.7 \text{ kJ/mole}$$

so indeed the free enthalpy decreases as the metal corrodes.

I hope this gives a clear picture of how metals corrode normally, without being connected to an external power supply as in all our previous examples. The key step is to ask whether an anodic de-electronation can occur simultaneously with a cathodic electronation (usually HER or dissolved oxygen reduction) and then work out the mixed potential and confirm that the overall combined processes at given concentrations (and temperature) will lead to an overall decrease in free enthalpy.

16.4 Evans diagram

Right. The next question is, how fast does the metal corrode? Finally I can put together the equilibrium thermodynamics of section 14.7 and before, and the kinetics of section 15 and I can show you a graphical construction that leads to a prediction of the corrosion potential and the rate of uniform corrosion.

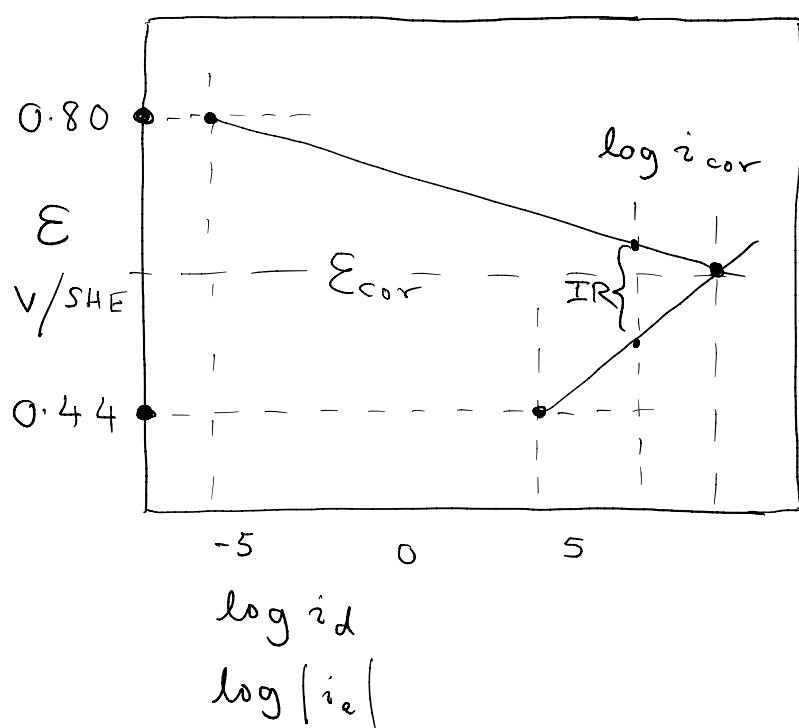


FIGURE 16-2

Return to the example in which I place a coupon of silver into neutral, aerated water. A corrosion cell as sketched in figure 16-1 is set up. I have all the data I need in (16.2.4), (16.3.3) and the table on page 93. I construct a plot of electromotance, \mathcal{E} , versus the logarithm of the current density as in figure 16-2. I assume that the anode and cathode areas are equal—otherwise I must plot current, not current density. I mark off on the ordinate the equilibrium single electrode potentials of the metal de-electronation reaction (16.3.3), namely 0.44 V/SHE and the oxygen reduction, 0.8 V/SHE (16.2.4). I mark off on the abscissa the logarithms of the corresponding equilibrium exchange current densities, i_0 , taken from the table on page 93. I then draw straight lines, assuming the Tafel law (15.8.3) and using the Tafel slopes taken from the table on page 93. The Tafel line for the de-electronation, metal wastage, slopes upwards since the

overpotential increases with increasing current density (15.8.2a) and conversely for the electronation (oxygen reduction). These are called the anodic and cathodic polarisation curves, respectively. Now you see how the metal negotiates its corrosion potential, \mathcal{E}_{cor} , and corrosion current density, i_{cor} . The electronation and de-electronation current densities must be equal to conserve charge (I am assuming equal areas) and so this must be found by a vertical line drawn on the diagram. If the electrolyte is perfectly conducting (as of course is the metal) then the corrosion current density must be at the intersection of the two Tafel lines as shown. In case the electrolyte is not perfectly conducting there may be an "IR" drop in the electrolyte between anode and cathode, in which case, as shown, i_{cor} must be common to both anode and cathode, but now the interphase potentials are not common to anode and cathode. But in the case of zero resistance then the intersection of the two Tafel lines indicates not only i_{cor} , but also the common corrosion potential, \mathcal{E}_{cor} , as I described it in section 16.3.

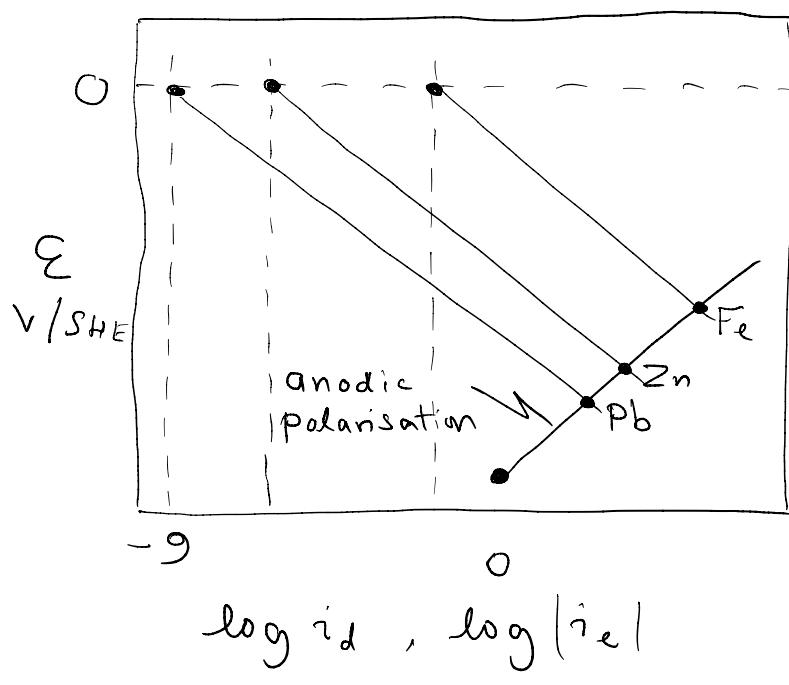


FIGURE 16-3

Figure 16-2 is called an "Evans diagram" after the pioneering corrosion scientist, Ulick R. Evans (1889-1980). There are a million uses of the Evans diagram to illustrate corrosion processes: crevice attack, differential aeration and countless others. Here is a crude example which is nonetheless insightful. In figure 16-3 I show an anodic polarisation curve for a generic metal assuming it's the same for, say, lead, zinc and iron (although you see from the table on page 93 that they actually are very different). I then show cathodic polarisation curves for the hydrogen evolution reaction on these three metals taking account that they all have the same Tafel slope, but very different exchange current densities, taken from the table on page 93. This is also very rough and ready since these are i_0 values in water and the HER is only the relevant cathode reaction in acids—exchange current densities in normal strength sulphuric acid, say, are typically about ten times what they are in pure water. All the same the Evans diagram illustrates a key fact that the rate of corrosion of these three metals will vary by several orders of magnitude in the same strength acid. This is the basis of *galvanisation* in

which a coating of zinc is used to protect steel. Not only is the corrosion rate greatly reduced but so is the rate of hydrogen evolution, and this provides a means of protection of the steel from hydrogen embrittlement.

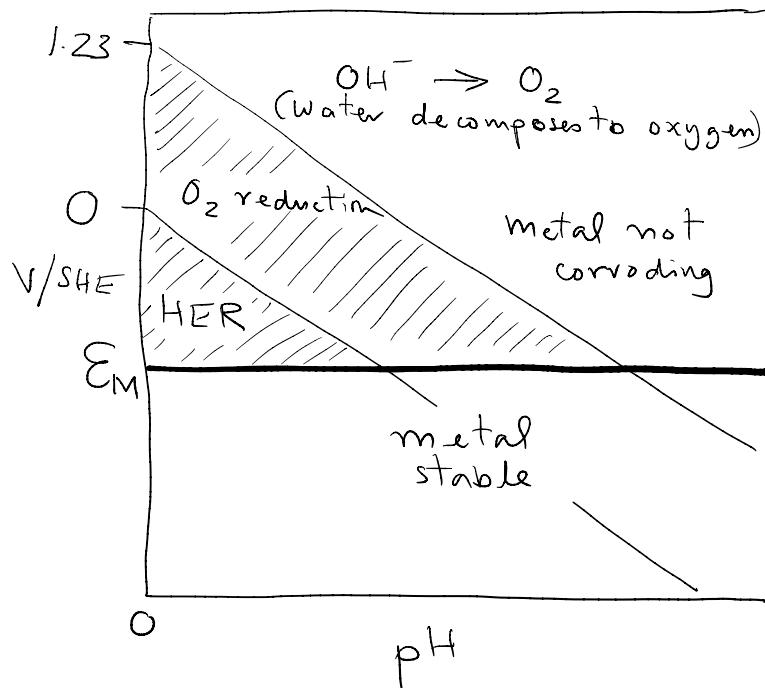


FIGURE 16-4

In figure 16-2 I treat a case in which the cathode reaction as in figure 16-1 is oxygen reduction, while in figure 16-3 I expect it to be hydrogen evolution because conditions in the electrolyte are acidic. Can I decide in general what will be the cathode process in a given corrosion situation? In figure 16-4 I plot electromotance against pH, not current density; so the single electrode potential on the SHE scale, \mathcal{E}_M , appears as a thick horizontal line as it does not depend on pH. At any electrode potential below this line, the metal is immune from corrosion. Now I plot the equilibrium potentials of the oxygen reduction and HER which do depend on pH, as expressed in equations (16.2.3) and (16.2.6). At electrode potentials *above* the oxygen reduction equilibrium in figure 16.4-4, the reaction (16.2.1) is driven to the left and indeed water decomposes to oxygen at such a large positive potential. Moreover the metal may not corrode because its electrode potential is more positive than either possible electronation reactions so there can be no mixed potential established. Below this line but above $\mathcal{E} = \mathcal{E}_M$ the metal can corrode and oxygen can be reduced at the cathode. At potentials below the HER sloping line, water is again unstable—this time with respect to hydrogen gas (you can see how this determines the conditions for the electrolysis of water to either oxygen or hydrogen gas, essential to the “hydrogen economy”). In the smaller shaded triangle the electrode potential is sufficiently base that the hydrogen evolution reaction may displace oxygen reduction at the cathode; this will happen in acidic conditions (low pH) and in the absence of dissolved oxygen. Therefore you will often read a statement such as to the effect that a metal at an electrode potential lower than the potential for oxygen reduction or hydrogen evolution will become oxidised.

17. Surface film and Pourbaix diagram

I have called the experiment of placing a coupon of pure metal into water, or a solution of its ions, a *thought experiment* partly because there is no pure metal, but much more significantly because all metals, except gold, are likely to have surface films of oxide or hydroxide and so the pure metal electrolyte interface is only achieved under certain conditions as we will now discover. You are particularly aware that aluminium, titanium and chromium, among many other metals, are protected from corrosion in damp air not because they are noble (in fact they are base—see page 38) but because they are protected by a film of oxide, possibly only a few nanometres thick, which coheres tenaciously to the surface. Other base metals, such as iron, are not protected in this way because their oxides, possibly due to a mismatch in atomic volume (the so called Pilling–Bedworth ratio) flake off and expose new metal to be oxidised. Alloying of steel with chromium led to the fantastically important discovery of stainless steel by Harry Brearley. If you can discover stainless magnesium you will make your fortune.

17.1 Passivation

The benefits of the formation of an oxide or hydroxide film can be sketched out in an Evans diagram. The details are quite complicated but I can give a qualitative description.

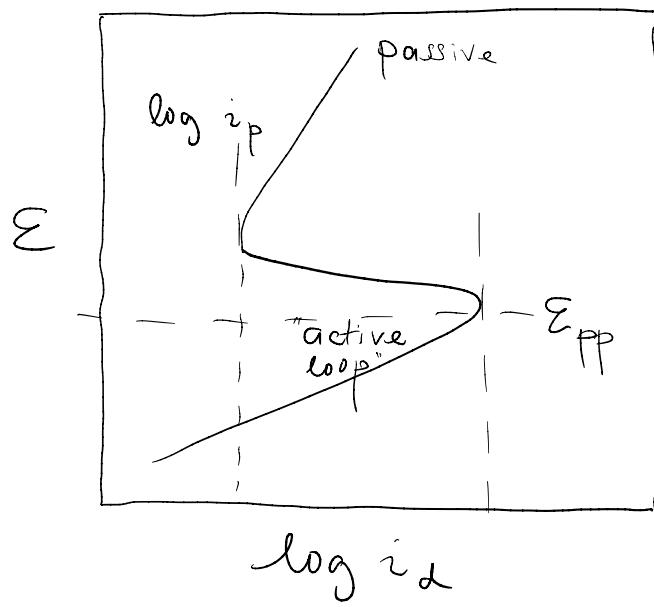


FIGURE 17-1

In reference to figure 17-1, begin at the lower left of the anodic polarisation curve of a metal such as stainless steel or aluminium. As the potential is raised so the current density increases with a normal Tafel slope just as in figure 16-2. As the metal oxidises, M^{n+} ions may detach and migrate through the double layer as I have already described, or under the right conditions the M^{n+} ions may combine with oxygen or hydroxide ions from the solution to form a solid oxide or hydroxide film. This may become a tenacious

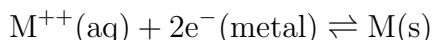
ceramic coating a few ones or tens of nanometres thick. Clearly this will greatly affect the kinetics and we now must regard the film and double layer acting in parallel, each with its de-electronation rate coefficient and symmetry factor. In fact transport across the film of electrons and cations is a matter of electrical conduction or quantum tunnelling of electrons and vacancy mediated solid state diffusion of, indeed, anions as well as cations. Without doing the maths, the result is that once the electrode potential is sufficiently raised to what is called the peak passivation potential, \mathcal{E}_{pp} , at some maximum current density as indicated in figure 17–1, the kinetics become a great deal more sluggish and the current density drops to a value i_p , called the passivation current density, and this may represent several orders of magnitude reduction in corrosion rate. Thereafter, as may be demonstrated by calculation (see J. M. West) the Tafel slope is increased by as much as a factor two so that even if the electrode potential is further raised, the increase in current density is much less rapid. For example, in an 18–8 stainless steel in de-aerated sulphuric acid at 25°C, the peak passivation potential is quite base: –0.2 V/SHE; and the passivation current density is as small as 0.006 A/m², which is not much if you remember that equates to a similar number of mm per year metal wastage.

17.2 Pourbaix diagram

Figure 16–4 is a simple kind of “Pourbaix diagram.” This is an equilibrium diagram, or phase diagram if you like, indicating the state of equilibrium of an electrode–electrolyte system in the space of electromotance (single potential) and acidity, pH. It allows you to predict whether a metal will corrode and what will be the corrosion products for a given state represented by an electrode potential (on the SHE scale) and the pH of the solution—all other things being equal, of course, including concentrations of ions in solution and temperature. In what follows I will show how the simplest Pourbaix diagram is constructed from thermodynamic data, and then take a short tour of the Pourbaix diagram for iron.

Separation into fields in the Pourbaix diagram

Assume for simplicity a divalent M^{n+} ion, so $n = 2$, which allows the formation of a simple stoichiometric oxide, MO. Then there are two possible electronation reactions. The first is,



for which the equilibrium potential is,

$$\mathcal{E}_{M^{++}} = \mathcal{E}_{M^{++}}^\circ + \frac{RT}{2F} \ln h_{M^{++}} \quad (17.2.1)$$

and you could look up $\mathcal{E}_{M^{++}}^\circ$ in the electrochemical series on page 38. The second possible reaction involves the formation of the oxide film,



for which the equilibrium potential is,

$$\begin{aligned} \mathcal{E}_{MO} &= \mathcal{E}_{MO}^\circ + \frac{RT}{F} \ln h_{H^+} \\ &= \mathcal{E}_{MO}^\circ - 0.06pH \quad \text{V/SHE at } 25^\circ\text{C} \end{aligned} \quad (17.2.2)$$

I can plot on a graph of \mathcal{E} against pH, the lines (17.2.1) and (17.2.2). This is sketched in figure 17–2. Thermodynamics dictates that for conditions of pH and electrode potential to the right of the sloping line, the metal is unstable with respect to its solid oxide. This means that as long as the oxide is able to form a tenacious and continuous surface layer then we will be in the passive region of figure 17–1 and the metal will not corrode. If the electrode potential is smaller (more negative) than \mathcal{E}_{ion} ($\mathcal{E}_{M^{++}}$) the metal is at a lower potential than its equilibrium $\Delta\phi_{\text{eq}}$ and will not corrode as long as we are to the left of the sloping line so that it is also stable with respect to its oxide. This region of behaviour is sometimes called “immunity”. The third domain of behaviour is the closed triangle in which a metal is not passive nor immune and will corrode.

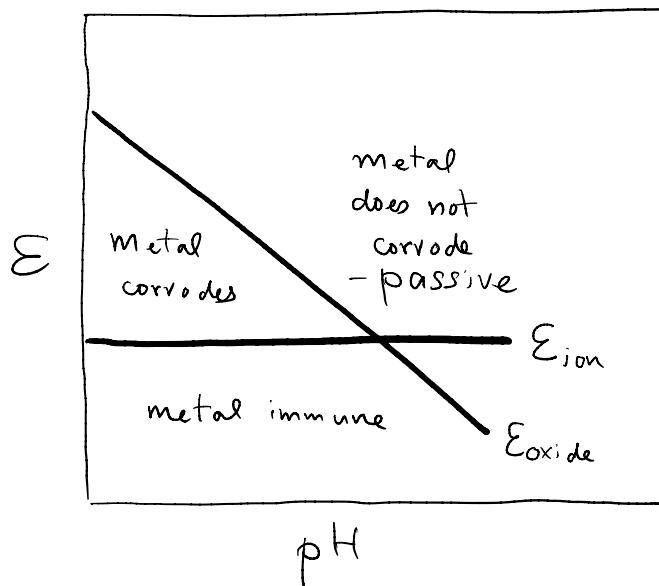


FIGURE 17–2

You can imagine superimposing the oxide sloping line onto the Pourbaix diagram in figure 16–4, which will add additional data since that previous diagram was drawn without taking into consideration the possibility of passivity. You then have a full picture of the prevailing conditions that will determine corrosion, passivity and immunity; and if corrosion is indicated, what is the likely accompanying cathode reaction (see figure 16–1).

Pourbaix diagram for iron

I have trawled the internet for a Pourbaix diagram for iron and I found figures 17–3 and 17–4 at chem.libretexts.org. This will provide us with sufficient pedagogical material for a brief tutorial. Reading a Pourbaix diagram is not unlike reading a binary temperature composition phase diagram in which for a given T and c the thermodynamically stable phase or phase mixture is indicated by the field in which the T and c combination falls. The phase boundaries in a Pourbaix diagram may be vertical, horizontal or straight with a constant slope.

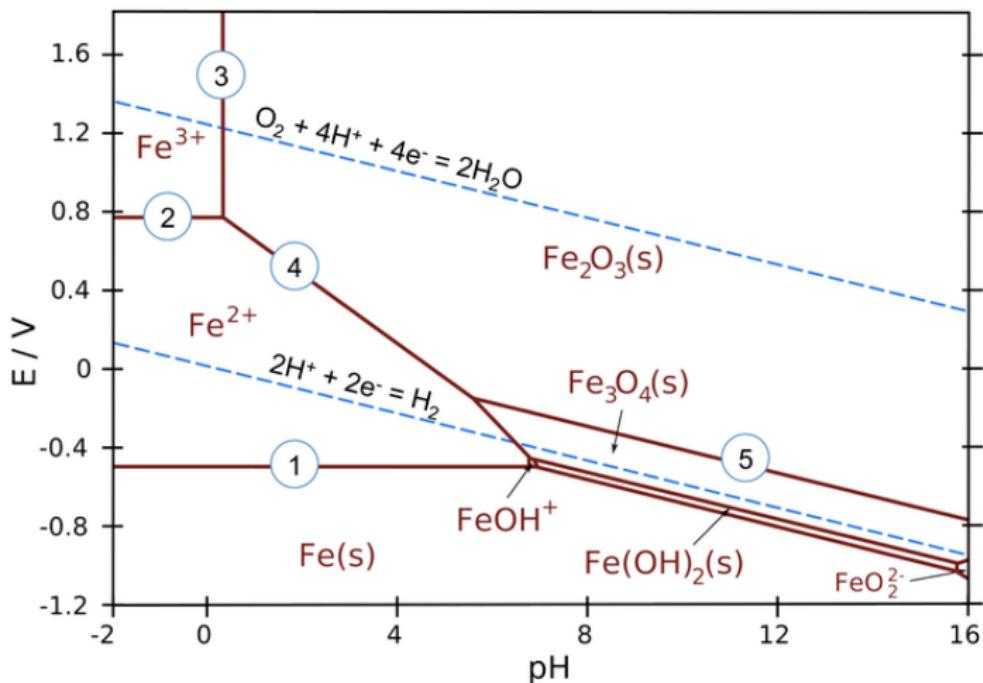
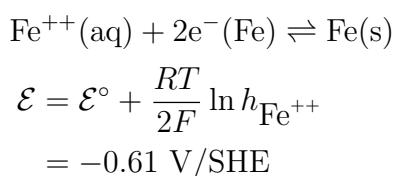


FIGURE 17–3

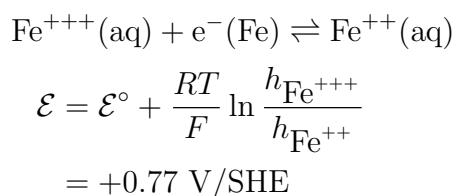
The Pourbaix diagram for iron is sufficiently complicated to include all three types of phase boundary. It is complicated because iron can exist in at least two forms of oxide, Fe_2O_3 and Fe_3O_4 . The Pourbaix diagram I have found usefully labels five of these lines and the equilibria whose right and left hand sides are separated by such phase boundaries are described now. In these calculations I will assume that Fe^{++} and Fe^{+++} ions in solution are at henrian activity (molal concentration) of 10^{-6} . Also recall that the activities of pure substances (including water, as it's very nearly pure) are one.

1. Reduction (electronation) of Fe^{++} to metal iron.

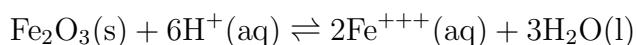


This reaction's electromotance is independent of pH so appears as a horizontal line.

2. Reduction (electronation) of Fe^{+++} to Fe^{++} (also independent of pH, so horizontal).

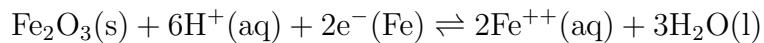


3. Equilibrium of Fe_2O_3 and Fe^{+++}



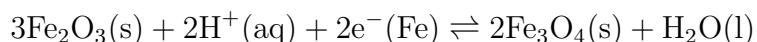
In this reaction there is no charge transfer to or from the electrode, no reduction or oxidation, hence the boundary is independent of \mathcal{E} and appears at a fixed pH.

4. Reduction (electronation) of Fe_2O_3 (slope -0.177 V/pH)



$$\begin{aligned}\mathcal{E} &= \mathcal{E}^\circ + \frac{RT}{2F} \ln \frac{h_{\text{H}^+}^6}{h_{\text{Fe}^{++}}^2} \\ &= +0.73 - 0.117\text{pH} - 0.06 \log 10^{-6} \\ &= +1.08 - 0.177\text{pH V/SHE}\end{aligned}$$

5. Reduction (electronation) of Fe_2O_3 to Fe_3O_4 (slope -0.06 V/pH)



$$\begin{aligned}\mathcal{E} &= \mathcal{E}^\circ + \frac{RT}{2F} \ln h_{\text{H}^+}^2 \\ &= +0.22 - 0.06\text{pH V/SHE}\end{aligned}$$

Also shown on figure 17–3 are the oxygen reduction and hydrogen evolution lines which I showed in figure 16–4. So the Pourbaix diagram for iron is complete and you can read off the domains of behaviour. The passive field is now strangely shaped because of the different stoichiometries of iron oxides and the iron hydroxides that I have not discussed. Above the oxygen reduction line, there may be neither corrosion nor passivation since the water is unstable with respect to gaseous oxygen. The corroding field is bounded by the lines 1, 2 and 4, and the field at the bottom is the region of immunity. This is summed up in the second illustration I got from chem.libretexts.org, figure 17–4.

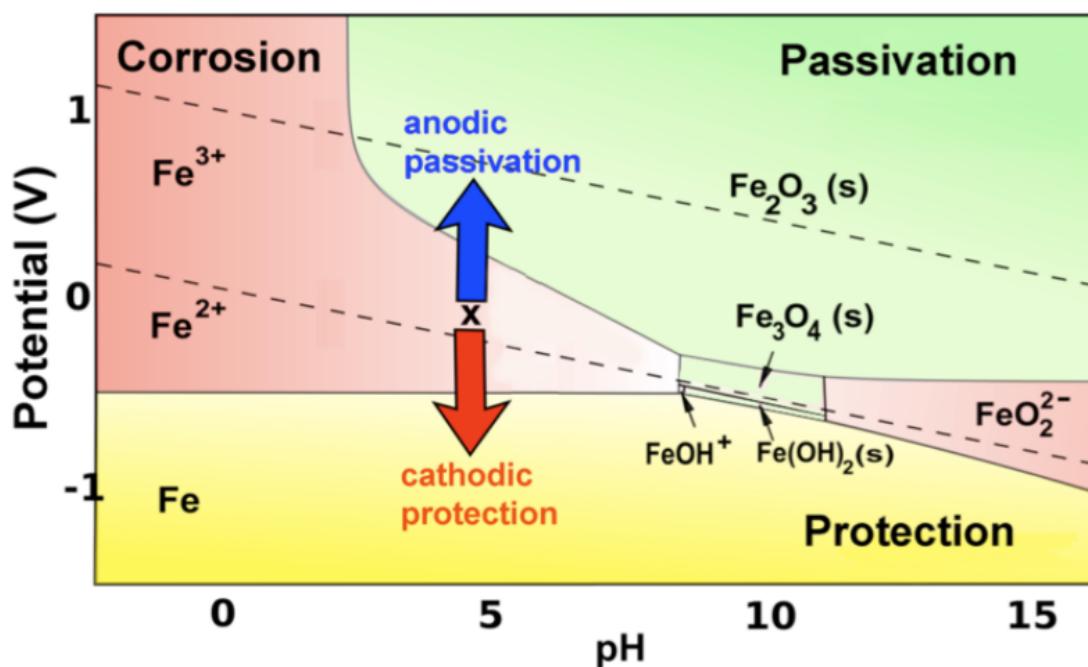


FIGURE 17–4

18. The evils of chloride

The engineer can deal with uniform corrosion and the benefits of passivation are obvious. The reason as I stated in the beginning of section 13 that we cannot leave it at that is that almost all corrosion failures are a consequence of localised attack at crevices, dissimilar metal joints, regions of reduced aeration and so on. But *physics* cannot take us there and so we reach the limits of our scope except for some brief remarks which follow. While what we have learned has great value in many branches of modern science in particular batteries and fuel cells for e-mobility, I have avoided those topics and concentrated in these last sections on corrosion. My hope is that these notes have equipped the reader with the tools to proceed into these areas; and of course the climate emergency calls upon us rapidly to develop better batteries and fuel cells.

When I was an undergraduate I was taught by John West, and I also took a job in the long vacations at the Sheffield University Metals Advisory Centre (SUMAC) in which Graeme Davies was Chairman and Tony Entwistle was Managing Director, not to mention the colourful John Devine. Our business was mostly failure investigation with a bit of alloy design. I remember two corrosion problems and in one instance I suggested to Tony that we ask John West for an opinion. Tony didn't think that was called for and that's when he issued his wonderful remark: you only need to know two things in a corrosion investigation; anodes are positive and they corrode; and find where the chloride came from. He was right and I found it! It is of course common in seawater and greatly complicates marine corrosion.

In hydrogen embrittlement, Dave Rugg of Rolls-Royce famously said, hydrogen is intelligent and it's evil. Almost the same can be said of the chloride ion. It plays at least two roles in the devastation of engineering structures. One is that it attacks passive films. So Pourbaix diagrams must be re-drawn in sea water, with the passive regions shrinking dramatically.

Second is its role in pitting corrosion. Uniform corrosion can be handled, but if corrosion becomes localised, say due to patches of metal in regions of electrolyte depleted of dissolved oxygen, then if these regions have a small area the anode current density greatly exceeds the cathode current density and the metal is effectively drilled by anodic dissolution. Once a pit forms the process is self sustaining and irrecoverable. The electrolyte in the pit constitutes a "micro climate". Detailed thermodynamic calculations reveal the effect of chloride ions in the pit is to lower the pH, to raise the partial pressure of hydrogen and to lower the equilibrium potential. Drilling will continue until a plate or pipe is perforated and it fails.

Further Reading

- John M. West, *Basic Corrosion and Oxidation*, (Ellis Horwood, 1980)
- J. O'M. Bockris and A. K. N. Reddy, *Modern Electrochemistry*, (Plenum, New York, 1970)
- W. Ronald Fawcett, *Liquids, Solutions, and Interfaces*, (OUP, Oxford, 2004)
- Wolfgang Schmickler and Elizabeth Santos, *Interfacial Electrochemistry*, 2nd Edition, (Springer, Berlin, 2010)
- Walter J. Moore, *Physical Chemistry*, 5th Edition, (Longman, London, 1972)
- Charles Kittel, *Introduction to Solid State Physics*, 8th Edition, (John Wiley, New Jersey, 2004)
- Neil W. Ashcroft and N. David Mermin, *Solid State Physics*, (Holt, Reinhart and Winston, Philadelphia, 1976)
- David J. Griffiths, *Introduction to Electrodynamics*, 3rd Edition, (CUP, Cambridge, 2017)
- N. Cusack, *The Electrical and Magnetic Properties of Solids*, (Longmans, London, 1958)
- Norton D. Lang and Walter Kohn, "Theory of Metal Surfaces," *Physical Review B*, **1** 4555 (1970); *loc. cit.*, **3**, 1215 (1971)
- Jun Cheng and Michiel Sprik, "Alignment of electronic energy levels at electrochemical interfaces," *Phys. Chem. Chem. Phys.*, **14**, 11245 (2012)
- M. Weinert, "Resolution of an apparent paradox in the physics of Fermi levels," *Solid State Commun.*, **54**, 513 (1985)
- Josiah Willard Gibbs, *Collected Works*, (New Haven Conn., Yale University Press, 1957), vol. 1, pp. 219–331
- John W. Cahn, *Thermodynamics of solid and fluid surfaces*, in *Interfacial Segregation*, Editors, W. C. Johnson and J. M. Blakely (Metals Park Ohio, American Society for Metals, 1979) p. 3
- Mike Finnis, *Thermodynamics of Interfaces*, (unpublished lecture notes, Imperial College London)
- Hans-Jürgen Butt, Karlheinz Graf and Michael Kappl, *Physics and Chemistry of Interfaces*, second edition, (Weinheim, WILEY-VCH Verlag GmbH, 2006), chapter 3

Appendix—outline of thermodynamics

1. Closed systems

1.1 State functions

All thermodynamic thinking begins with a definition of the portion of the universe under study, i.e., the system.

John O'M. Bockris

I will not start right at the beginning but I will try and introduce what we will need in thermodynamics using just what you have already learned about the first and second laws.

The first law states that for a “closed” system changes in the heat content and the amount of work done *on* the system amount to a change in *internal energy*[†] written as

$$dU = \delta q + \delta w \quad (\text{A1.1})$$

If this is combined with the second law then this becomes

$$dU = T dS - p dV \quad (\text{A1.2})$$

This is true as long as the only work done involves a change in volume, V , against an external pressure, p . The absolute temperature is T and S is the entropy. A *closed* system is one that does not exchange matter with its surroundings (although it may transfer heat). We will come on to *open* systems in a while.

Before that, we note that (A1.2) is a complete statement of the combined first and second laws for a closed system if only pV -work is done.[‡] However it may not be the

[†] From a mechanical point of view we often state that the total energy of a body, say a sphere of Cu–Ni alloy, is equal to its potential energy, E_{pot} , if it's in a gravitational field, say, plus its kinetic energy, E_{kin} , if it's moving relative to some inertial frame. This neglects the *internal* state of the body, that is to say, what its atoms are doing and whether any diffusion, phase transformations or chemical reactions are taking place, whether it is expanding or getting hotter or colder. So in thermodynamics as opposed to mechanics we express the total energy as $E_{\text{tot}} = E_{\text{pot}} + E_{\text{kin}} + U$ and since thermodynamics concerns only *changes* in total energy we consider systems for which the potential and kinetic energies are unchanging, for example that the body is not in motion, and focus purely on the internal energy.

[‡] We have to extend this in cases where there is surface tension, non uniform stress, changes in electric or magnetic state and so on. For example work done on a body by a non hydrostatic stress is $\sigma_{ij}d\varepsilon_{ij}$, or work done on a body by increasing the area, A , of an internal interface of energy γ per unit area (say between the ferrite and austenite phases in steel) is γdA .

most convenient. For example, we may like to simplify things by doing experiments so that one of the two terms on the right hand side of (A1.1) is zero. So if we want to know just the result of the system doing some work, then working with the internal energy as a function of entropy and volume, $U(S, V)$, requires us to do experiments adiabatically, that is under the condition $\delta q = 0$, and this is not very easy.

So without making any additional postulates we devise a number of *auxiliary functions*. These are,

$$\begin{aligned} \text{enthalpy, } & H = U + pV \\ \text{Helmholtz function, } & F = U - TS \\ \text{Gibbs function, } & G = H - TS \end{aligned}$$

G is also known as the *free enthalpy* or *Gibbs free energy*; F is also known as the *free energy* or *Helmholtz free energy*. We are most interested in the Gibbs function so I'll show you how we get it from equation (A1.2). We do it in two steps. First we have that U is a function of S and V as indicated by (A1.2). Can we find an auxiliary function that depends on T and V instead? What is the relation between T and S ? From (A1.2) we have

$$T = \left(\frac{\partial U}{\partial S} \right)_V$$

We say that S and T are “conjugate variables”, and we invent the new function

$$F = U - TS \quad (\text{A1.3})$$

that is, the first function, U , take away the product of the two conjugate variables (the one we are trying to get rid of and replace with the other). Now by taking the total derivative (that means, asking what does a change in F result in?)

$$\begin{aligned} dF &= dU - d(TS) = dU - TdS - SdT \\ &= -SdT - pdV \end{aligned} \quad (\text{A1.4})$$

using (A1.2). Now we have a function F which depends on T and V rather than U which depends on S and V so that we can more readily interpret experiments done at constant volume or constant temperature. Suppose we are working at constant pressure rather than constant volume, then we need yet another auxiliary function, which we obtain from F by replacing the variable V with its conjugate $-p$. Indeed we know from (A1.4) that

$$p = - \left(\frac{\partial F}{\partial V} \right)_T$$

so $-p$ and V are conjugate variables, meaning that I can find a function that depends on p and T by writing down

$$\begin{aligned} G &= F + pV \\ &= H - TS \end{aligned} \quad (\text{A1.5})$$

(the original function take away the product of the two conjugate variables[†]). Then similar to before,

$$\begin{aligned}
 dG &= dF + pdV + Vdp \\
 &= dU - TdS - SdT + pdV + Vdp \\
 &= TdS - pdV - TdS - SdT + pdV + Vdp \\
 &= -SdT + Vdp
 \end{aligned} \tag{A1.6}$$

and is thereby shown to be the required auxiliary function of T and p .

1.2 Conditions for equilibrium in closed systems

According to the second law, in an *isolated system* (one that can exchange neither matter nor energy with the environment) if any change takes place it must be such that the entropy increases or remains constant,[‡]

$$dS \geq 0$$

Strictly in thermodynamics when we write

$$d(\text{something})$$

we really mean

$$\frac{d(\text{something})}{dt}$$

but we accept that the “arrow of time” always runs in the positive direction even if we don’t quite understand why and we leave out the dt denominator. If the isolated system is made up of one or more parts then we have

$$\sum_m dS_m \geq 0$$

each part being labelled by a subscript m . So some part may suffer a decrease in entropy as long as the total entropy does not decrease. So we can think of a closed system as an isolated system made up of two parts—the part we are interested in (say, a block of alloy) and a large reservoir of heat at a fixed temperature T . In this way we can keep our body at constant temperature if it is kept in contact with the reservoir. So if the entropy of our body is S and that of the reservoir is S_r then, for a natural process,

$$dS + dS_r \geq 0 \tag{A1.7}$$

[†] In mathematics, this is called a *Legendre transformation*. Note that for a pair of conjugate variables, one is intensive and the other extensive.

[‡] We believe that the Universe is an isolated system, hence the famous statement of Rudolf Clausius, *Die Energie der Welt ist konstant. Die Entropie der Welt strebt einem Maximum zu*: The energy of the world is constant. The entropy of the world is striving to a maximum. This will of course lead to the “heat death” of the universe, but don’t worry the sun will have grown into a red giant and swallowed the earth by then.

Now for an infinitesimal change in heat content of the body, δq , at a temperature T , the change in entropy is $\delta q/T$;[†] this means that the change in entropy of the reservoir where the heat came from is $dS_r = -\delta q/T$ (because that amount of heat has been taken out of the reservoir at constant temperature T) and so putting this into (A1.7) the second law has it that

$$dS - \frac{\delta q}{T} \geq 0$$

Now using (A1.2) we have

$$dS - \frac{dU - \delta w}{T} \geq 0$$

I multiply through by T , rearrange and note that since this change is at constant temperature $d(U - TS) = dU - TdS$ and I get

$$d(U - TS) \leq \delta w \quad (\text{A1.8})$$

By comparison with (A1.3) we have $dF \leq \delta w$ so that if the body does work ($\delta w < 0$) then the Helmholtz function may increase, but otherwise and especially if no work is done either by or on the body, then the Helmholtz function can only decrease during a natural change, or remain constant. Thus when all changes have happened and the body is in equilibrium then the Helmholtz free energy is at a minimum. This is a condition for equilibrium.

I can add $d(pV)$ to both sides of equation (A1.8) and I get

$$d(U + pV - TS) = dG \leq \delta w + d(pV)$$

If as well as working at constant temperature, my body is maintained under a constant external pressure, p , then $d(pV) = pdV$ and so if I define $\delta w' = \delta w + pdV$ as the work done not including work done, pdV , by the body against the external pressure[‡] then the condition for equilibrium is that the Gibbs free energy of a body at constant temperature and pressure is a minimum if no work other than that either *against* or *by* the external pressure is done either on or by the body,

$$dG \leq \delta w' \quad \text{in a closed system}$$

If no such work is done ($\delta w' = 0$) then $dG \leq 0$ in which case the Gibbs free energy can only decrease or remain constant. This means that when all changes have happened

[†] You can think of this as the *definition* of entropy if you like: if an infinitesimal quantity of heat, δq , is added to a body reversibly and at constant temperature T , then the body's entropy increases by the amount $\delta q/T$.

[‡] In these notes δw is the work done on the body, not the work done by the body. The latter convention is commonly used by engineers who are interested in, say, an internal combustion engine for which pdV is the work done by the explosion of a fuel moving a piston against the external pressure p . So for us, if the body does work against the external pressure then the work done *on* the body is $-pdV$. So $\delta w' = \delta w - (-pdV) = \delta w + pdV$.

that are going to happen the system is in equilibrium and G is a minimum, $dG = 0$. This is the condition for equilibrium of a closed system at constant temperature and pressure.

In metallurgy when we deal with solids the amount of work done say by the thermal expansion of a body against atmospheric pressure is so tiny as to be negligible, and we can regard the Helmholtz and Gibbs functions to be interchangeable and both minimal at equilibrium; but since we usually work at constant pressure and not constant volume the focus is always on the Gibbs free energy.

2. Open systems, chemical potential

An “open” system, such as the lump of metal that we have been thinking about, can exchange both energy and matter with its surroundings. Under these circumstances we must modify our statement of the combined first and second laws (A1.2). We have to ask about the chemical composition of our body and to identify how many different “components” it is made up from. It is sufficient for our purposes to identify each chemical element as one of the components, so that for example a piece of Cu–Ni alloy has two components. In addition to components the body may be divided into phases. For example a piece of Fe–C alloy at equilibrium within the $\alpha + \gamma$ field of the Fe–C phase diagram may have a certain volume fraction of ferrite, having a very small concentration of carbon and a certain volume fraction of austenite with a much larger concentration of dissolved carbon. These two phases will be intimately in contact sharing one or more interfaces which divide the body into its phases. For now we consider a homogeneous, single phase body having N components and we will use a subscript i to label these in our mathematical formulas. The internal energy is now no longer a function of only the two variables S and V it is also a function of the number of moles,[†] n_i , of each component that currently make up the body,

$$U = U(S, V, n_1, n_2 \cdots n_N)$$

The total differential of the internal energy is, now,

$$dU = \left(\frac{\partial U}{\partial S} \right)_{V, n_i} dS + \left(\frac{\partial U}{\partial V} \right)_{S, n_i} dV + \sum_{i=1}^N \left(\frac{\partial U}{\partial n_i} \right)_{S, V, n_j \neq i} dn_i \quad (\text{A2.1})$$

In the case of constant composition, this is obviously still valid, and so from (A1.2)

$$\left(\frac{\partial U}{\partial S} \right)_{V, n_i} = T \quad (\text{A2.2})$$

[†] If you’re a physicist you tend to think in terms of the number of particles or number of atoms of each component, but since we deal with macroscopically sized bodies, in metallurgy and chemistry we use the number of moles to keep the numbers of reasonable size. A mole is nothing other than an Avogadro number of objects and the Avogadro number is $N_A = 6.023 \times 10^{23}$. Physicists use the Boltzmann constant, k_B , while we use the gas constant $R = N_A k_B$.

and

$$\left(\frac{\partial U}{\partial V} \right)_{S,n_i} = -p \quad (\text{A2.3})$$

We now define the *chemical potential* of component i as the term under the summation sign in (A2.1)

$$\mu_i = \left(\frac{\partial U}{\partial n_i} \right)_{S,V,n_j \neq i} \quad (\text{A2.4})$$

and here the partial derivative is taken of U with respect to the number of moles of component i while keeping all other variables constant: namely the volume and entropy and the numbers of moles of all the other components. You might ask how can I do this in practice? The answer is to imagine that to the body in question you take an infinitesimal number of moles of component i (from a reservoir in which its chemical potential has some standard value—more on that later) and add it to the body. While the body's volume necessarily changes you readjust that by application of an infinitesimal increase in pressure. Since the added quantity of matter may bring with it some heat, to ensure that the entropy doesn't change in this process it is necessary then to remove that heat, say by placing it in contact with a heat bath at the appropriate temperature.

Maybe you'd like to think of the chemical potential as the reversible work done in bringing an infinitesimal amount of matter from a reservoir to the body in question. Sometimes I like the analogy with electrostatics; the *electric potential* is the work done in taking a positive test charge from infinity (the reservoir) to a place in space where there is an electric field—that is, the influence of other charges. Loosely speaking if the electric potential is large and negative then positive charges are attracted to that place and the work done in bringing the test charge from infinity is negative; conversely, if the electric potential $\phi(\mathbf{r})$ is positive at position \mathbf{r} , then I need to do work to bring in extra positive charges. In this vein if the chemical potential of, say, carbon in Fe–C is large and positive then I need to expend energy to increase its concentration. If the concentration, or rather the chemical potential, of C in Fe varies from place to place then if the carbon is mobile it will diffuse from regions of high chemical potential to regions of low chemical potential. This is rather obvious if we equate chemical potential with concentration. Later we'll see how these two are actually related. Whereas in standard diffusion theory you are told that the carbon will travel down a concentration gradient, to be properly precise you should say that it diffuses down the gradient in *chemical potential*.

In view of equations (A2.1)–(A2.4) we have for the total differential of the internal energy of a body,

$$dU = TdS - pdV + \sum_{i=1}^N \mu_i dn_i \quad (\text{A2.5})$$

which is the modification of the combined first and second laws (A1.2) for the case of an open system.

Now let me begin this argument again, but this time starting with the modification of the Gibbs function (A1.5) for an open system. The Gibbs free energy is a function of

temperature, pressure and the numbers of moles of each of the components,

$$G = G(T, p, n_1, n_2 \cdots n_N)$$

and so its total differential is

$$dG = \left(\frac{\partial G}{\partial T} \right)_{p,n_i} dT + \left(\frac{\partial G}{\partial p} \right)_{T,n_i} dp + \sum_{i=1}^N \left(\frac{\partial G}{\partial n_i} \right)_{T,p,n_j \neq i} dn_i \quad (\text{A2.6})$$

We now define the *chemical potential* of component i as

$$\mu_i = \left(\frac{\partial G}{\partial n_i} \right)_{T,p,n_j \neq i} \quad (\text{A2.7})$$

and (A2.6) being still valid in the special case of all the dn_i being zero (that is, fixed composition) equation (A1.6) gives

$$\left(\frac{\partial G}{\partial T} \right)_{p,n_i} = -S \quad ; \quad \left(\frac{\partial G}{\partial p} \right)_{T,n_i} = V \quad (\text{A2.8})$$

which when put into (A2.6) results in

$$dG = -SdT + Vdp + \sum_{i=1}^N \mu_i dn_i \quad (\text{A2.9})$$

with μ_i defined by equation (A2.7). This is the modification of (A1.6) for an open system. Have I now made two different definitions of the chemical potential? That would be a real mess. Well, luckily, no! If I add to both sides of (A2.5) the quantity $d(pV - TS)$ then this equation is transformed into equation (A2.9) by virtue of the fact that $G = U + pV - TS$.

We will use equation (A2.7) as our expression for the chemical potential (numerically it is identical to (A2.4), but we are concerned in metallurgy with the Gibbs function whose independent variables are p and T which are easily controlled—it's not easy to measure let alone control the entropy!). In the case of there being a single component then the chemical potential is the “partial free energy per mole”. Usually for all extensive state functions we define a partial amount, being the amount per mole. We use lower case for these. In this way, we have for a single component homogeneous body,

partial molar volume,	$v = V/n$
partial molar enthalpy,	$h = H/n$
partial molar free energy,	$f = F/n$
partial molar free enthalpy,	$g = G/n$

Equation (A2.7) is the more general expression of the partial free enthalpy of component i when it finds itself in a body having mole numbers $n_1, n_2, \dots n_N$.

Finally in this section let's find a way to express the total Gibbs free energy of a body in terms of the mole numbers and chemical potentials of the components, equation (A2.10) below. Now, the Gibbs free energy is an *extensive* property so if I know $G(T, p, n_1, n_2 \cdots n_N)$ for a body and then I assemble some number λ of these bodies together then p and T do not change as these are intensive properties and the mole numbers are all multiplied by λ and the total Gibbs free energy is also multiplied by λ (because it is an extensive property),

$$G(T, p, \lambda n_1, \lambda n_2, \cdots \lambda n_N) = \lambda G(T, p, n_1, n_2 \cdots n_N)$$

(actually λ does not need to be a whole number—if I extend the body by an amount λ then the free energy is increased in the same ratio). We now invoke Euler's theorem for homogeneous functions of first order. Suppose a function f of N variables satisfies

$$f(\lambda x_1, \lambda x_2, \cdots \lambda x_N) = \lambda f(x_1, x_2, \cdots x_N)$$

We first exchange left and right hand sides and write $u_i = \lambda x_i$

$$\lambda f(x_1, x_2, \cdots x_N) = f(u_1, u_2, \cdots u_N)$$

then differentiate each side with respect to λ , using the rule for differentiating a function of a function for the right hand side,

$$\begin{aligned} f(x_1, x_2, \cdots x_N) &= \sum_{i=1}^N \frac{\partial f}{\partial u_i} \frac{du_i}{d\lambda} \\ &= \sum_{i=1}^N \frac{\partial f}{\partial u_i} x_i \\ &= \sum_{i=1}^N \frac{1}{\lambda} \frac{\partial f}{\partial x_i} x_i \end{aligned}$$

This is true for any λ but if I choose $\lambda = 1$ then I get

$$f(x_1, x_2, \cdots x_N) = \sum_{i=1}^N \frac{\partial f}{\partial x_i} x_i$$

This is Euler's theorem which I now apply to the Gibbs function,

$$G(T, p, n_1, n_2 \cdots n_N) = \sum_{i=1}^N \left(\frac{\partial G}{\partial n_i} \right)_{T, p, n_j \neq i} n_i$$

and by comparison with (A2.7) I get

$$G = \sum_{i=1}^N \mu_i n_i \tag{A2.10}$$

This states that the total free energy of a phase having N components is equal to the sum of the number of moles of each component times its chemical potential.

From this I can obtain the famous *Gibbs-Duhem equation*. If I take the total differential of (A2.10),

$$dG = \sum_{i=1}^N \mu_i dn_i + \sum_{i=1}^N n_i d\mu_i$$

and if I compare this with (A2.9) there results

$$-SdT + Vdp - \sum_{i=1}^N n_i d\mu_i = 0 \quad (\text{A2.11a})$$

This is the Gibbs-Duhem equation which you can use to prove the *Gibbs phase rule* which is central to the interpretation of phase diagrams (but for brevity I'll leave that out of these notes). At constant T and p , we then find

$$\sum_{i=1}^N n_i d\mu_i = 0 \quad (\text{A2.11b})$$

2.1 Conditions for equilibrium in open systems

Possibly the most significant statement in metallurgical thermodynamics that you need to remember is this, *in equilibrium the chemical potential of each component is the same in all phases*. That is to say, if the body comprises more than one phase (and we will label the phases with subscripts α , β , \dots) and the components are distributed among the phases so that the number of moles of component i in phase α is $n_{i\alpha}$, then as long as the phases are in contact and each component can diffuse throughout the body, then at equilibrium the chemical potential μ_i for component i is the same in each phase. This makes sense because it implies that if the chemical potential of component i is the same everywhere then there's no gradient to drive diffusion and so all atoms stay where they are—equilibrium is reached. For example a piece of Fe–C alloy at equilibrium within the $\alpha + \gamma$ field of the Fe–C phase diagram has a particular volume fraction of ferrite, having a very small concentration of carbon and a certain volume fraction of austenite with a large concentration of dissolved carbon. The concentrations of C in the two phases is very different because of the big difference in solubility of C in α -Fe and γ -Fe, but in equilibrium $\mu_{C\alpha} = \mu_{C\gamma}$ —the chemical potential of carbon is the same everywhere.

For the curious, I will try and prove this for you. Suppose in a multiphase, multicomponent body I take an infinitesimal amount of component i , $dn_{i\alpha}$, from phase α and transfer it reversibly into phase β .[†] The amount of increase of component i in the β

[†] For example I could have some Fe–C at the temperature at which austenite and ferrite are in equilibrium (at the “no name” line in the Fe–C phase diagram). Then keeping

phase is, rather obviously, $dn_{i\beta} = dn_{i\alpha}$ while the change of mole number of component i in phase α is negative *viz.*, $-dn_{i\alpha}$. If I do this at constant temperature and pressure, then according to equation (A2.9) the change in Gibbs free energy is

$$dG = -\mu_{i\alpha}dn_{i\alpha} + \mu_{i\beta}dn_{i\beta} = (\mu_{i\beta} - \mu_{i\alpha})dn_{i\alpha}$$

since (A2.9) must hold separately in each phase as each phase may be regarded as an open system and a part of the whole body. If the system was originally in equilibrium and the transfer is done reversibly then

$$dG = 0 = (\mu_{i\beta} - \mu_{i\alpha})dn_{i\alpha}$$

and so, since $dn_{i\alpha} \neq 0$, it must be true that

$$\mu_{i\alpha} = \mu_{i\beta}$$

if the two phases are in equilibrium. We can extend this argument to any number of phases.

2.2 Chemical potential in a closed system

Paradoxically we need to ask, what is the meaning of chemical potential in a closed system? This is because even if matter is not exchanged with the environment or a reservoir the composition of a closed system may nonetheless change reversibly. A reversible chemical reaction may happen, for example as a consequence of a change of temperature or pressure. In fact Denbigh quotes Gibbs's definition of chemical potential: "If to any homogeneous mass we suppose an infinitesimal quantity of any substance to be added, the mass remaining homogeneous and its entropy and volume remaining unchanged, the increase of the energy of the mass divided by the quantity of substance added is the *potential* for the substance in the mass considered." This statement is exactly equivalent to the definition (A2.4). From a practical point of view Gibbs's thought experiment is entirely possible if we imagine that although the additional matter will serve to increase the entropy of the mass, its total entropy can be brought back to what it was by the extraction of the appropriate amount of heat through contact with a heat bath. Similarly the volume can be kept constant during the addition by the adjustment of the external pressure. A similar argument applies to (A2.7); in fact the

T , p and the carbon concentrations fixed, I could take $dn_{Fe\alpha}$ moles of body centred cubic Fe (ferrite) and transfer them across the α / γ interface and rearrange them into the face centred cubic austenite. Alternatively I could reversibly move the α / γ interface in such a way that $dn_{Fe\alpha}$ moles of ferrite are rearranged into austenite. Actually there is a complication in this thought experiment because once I have transformed some ferrite into austenite I will have to also transfer an amount of carbon so that the austenite retains the carbon concentration appropriate to the tie line in the phase diagram; I gloss over this somewhat here, but as long as the amount I am transferring is *infinitesimal* in the sense of the differential calculus then I can neglect the need to re-equilibrate the carbon in this process.

statement amounts almost trivially to the assertion that the free enthalpy is an extensive quantity—if I keep the *concentrations* of all components constant by simply extending the size of a closed system then I will increase its free enthalpy. For example if I add reagents in fixed proportion to the contents of a galvanic cell, I increase the amount of electrical work that it is able to do (see section 9.3). In fact a rather artificial device that we will need in section 11.1 is to imagine a closed system as being simply a piece, of arbitrary size, taken by drawing a boundary within a larger whole. If I choose to extend that boundary, and hence increase the size of my system I am changing the mole numbers of components, but *not* by exchange of matter with external reservoirs.

Finally let us consider again the formula

$$dU = TdS - pdV + \sum_{i=1}^N \mu_i dn_i \quad (\text{A2.5})$$

but now as applied to a *closed system*. The final terms then correspond to changes in the mole numbers, or concentrations, of the components due a chemical reaction (or even, presumably, a *nuclear* reaction). Of course for (A2.5) to apply to the closed system, before and after any such change the system must be initially in equilibrium, and remain in equilibrium (or quasi- or static-equilibrium) throughout the process of change. In that case, since by definition the first term TdS is the change in heat content, the final term must be regarded as a “chemical” work term. This is exactly the work done by the passage of charge in a galvanic cell. Indeed, even at constant volume, it is clear that a system can perform work, lift a weight, in the absence of any change of volume as would be required in the operation, say, of a piston engine. The interpretation of the last term in (A2.5) as work is only unambiguous in the case of a closed system (hence the significance of this sub-section); if the system is open and the change in mole numbers is as a consequence of exchange of matter with the environment or reservoir, then as matter is introduced or removed it brings its own entropy and hence heat, so that TdS is no longer the only heat term in (A2.5). The additional heat introduced by mass transport is sometimes called the “heat of transport” in diffusion problems.

3. Activity

Next, we need to know how does the chemical potential relate to the concentration. Obviously we design and process metal alloys using concentrations of the components but as we have seen what controls the thermodynamics and the kinetics at the fundamental level is the chemical potential. The best way to start this is simply to write down the answer. The chemical potential of component i in a particular phase (I’ll leave out the subscript α unless it’s needed) is

$$\mu_i = \mu_i^\circ + RT \ln a_i \quad (\text{A3.1})$$

μ_i° is the chemical potential in some “standard state” (mathematically it’s an integration constant) and depends only on temperature and pressure; and a is the “activity”. Where does this equation come from? At the most shallow level it’s simply a *definition* of

activity, and still we need to find the connection between activity, a_i and concentration, x_i expressed as the mole fraction, n_i/n . This is most generally expressed like this,

$$a_i = \gamma_i x_i$$

which may be a very complicated function since the so called “activity coefficient” γ_i is a function of temperature, pressure and the concentrations of all the components, not just component i ,

$$\gamma_i = \gamma_i(T, p, x_1, x_2, \dots, x_N)$$

Thermodynamics can tell us nothing about how the activity depends on the concentrations. Ultimately it is up to experiment or detailed atomic-scale theory to determine the relationship. For example, it follows from the Gibbs-Helmholtz equation that[†]

$$R \left(\frac{\partial \ln \gamma_i}{\partial (1/T)} \right)_{p, n_j \neq i} = h_i - h_i^\circ \quad (\text{A3.2})$$

where h_i° is the partial molar enthalpy in the same standard state as is applied to μ_i° . For an ideal solution, $\gamma = \text{constant}$ and so the left hand side is zero, implying that the enthalpy of the component in solution is the same as in the pure substance at the same T and p , if the pure substance standard state is used. By integration of the Gibbs-Helmholtz equation, activity coefficients can be inferred from measured heats of solution.

The results of such experiments are shown in figure A1. You can see that the activity does not at all show a straightforward relationship to the concentration, except to say that as the concentration increases so does the activity. However you see two very clear limits: for large concentrations we see that

$$a_i = x_i \quad (x_i \rightarrow 1)$$

meaning that in the concentrated limit the activity of the *solvent* is equal to its concentration, the activity coefficient being one; and

$$a_i = \gamma_i^\infty x_i \quad (x_i \rightarrow 0)$$

[†] Combine (A1.5) with (A2.8),

$$G = H + T \left(\frac{\partial G}{\partial T} \right)_{p, n_i}$$

which can be recast as follows,

$$\left(\frac{\partial (G/T)}{\partial T} \right)_{p, n_i} = -\frac{H}{T^2}$$

which is the Gibbs-Helmholtz equation. Just divide G and H by the number of moles and insert (A3.1) with $a_i = \gamma_i x_i$ to get (A3.2).

with γ_i^∞ constant, that is, in the infinitely dilute limit the activity of the *solute* is proportional to its concentration. In this limit the activity coefficient is a proportionality constant. These two limits are statements of, respectively, Raoult's law and Henry's law (first stated to refer to the partial pressures of gasses).

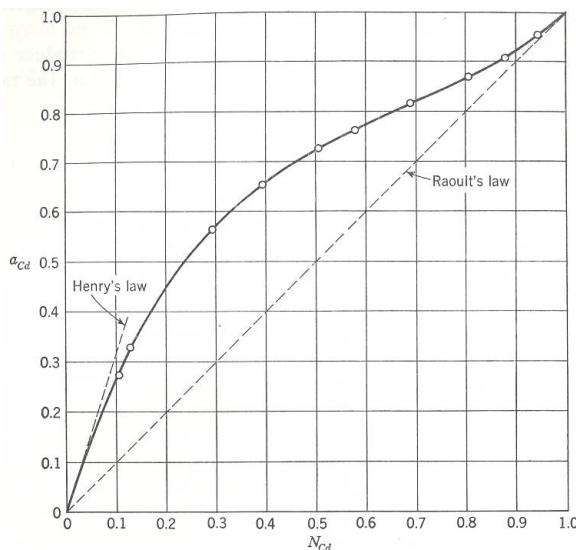


FIG. 10-5. Activity of cadmium in the Cd-Pb system at 500°C.

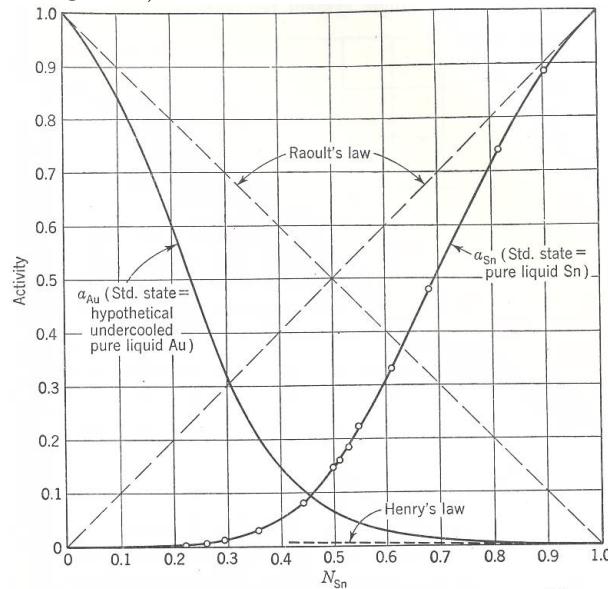


FIG. 10-10. Activity of tin and gold in the Sn-Au system at 600°C.

Figure A1. Activity for some liquid binary alloys as functions of concentration. The left hand figure shows a *positive* deviation from Raoult's law and the right hand figure shows *negative* deviation. When the concentration of a component tends to one we then call this component the *solvent* for obvious reasons.

In solid solutions with positive deviation, $a_i > x_i$, the solute behaves as if it were in a greater concentration than it actually is, and *vice versa*. Equation (A3.2) indicates that if $a_i > x_i$ (activity coefficient greater than one) then the mixing of the solute in the solvent is endothermic, meaning that the solute doesn't want to form a solid solution and the mixture has a tendency to phase separation or possibly spinodal decomposition. Conversely if $\gamma_i < 1$ then the mixing is exothermic, very roughly meaning that the component *i* likes to form bonds with the atoms of the solvent.

Incidentally a consequence of the Gibbs-Duhem equation (A2.11) is that

$$\sum_{i=1}^N x_i d \ln a_i = 0$$

This can often be used to find the activity of one component if the activity of the others is known.

3.1 Ideal mixture

The “ideal mixture” or ideal solution, or ideal solid solution, like the ideal gas, is an idealisation but which fits the facts in many cases. They are characterised by a special form of the entropy of mixing, namely this: if the two components are initially separated into two containers and subsequently allowed to mix at constant temperature then the change in entropy is

$$\begin{aligned}\Delta S &= R \sum_{i=1}^N n_i \ln \frac{n}{n_i} \\ &= R \left(n \ln n - \sum_{i=1}^N n_i \ln n_i \right)\end{aligned}\tag{A3.3}$$

Here, n is the total number of moles and n_i is the number of moles of component i . For an ideal gas, we have from the ideal gas law (Dalton’s Law) that $n/n_i = p/p_i$ where p is the total pressure and p_i is the partial pressure of component i . So you can define the ideal solution as one whose entropy of mixing is taken by analogy with the ideal gas. You can see (A3.3) derived in a textbook on thermodynamics (look up “Gibbs’s paradox”).[†]

So when I take n_i moles of each of N components in the pure state at a given temperature and pressure, and I mix them to produce an *ideal* homogeneous single phase mixture of atoms, then the total Gibbs free energy of the body that I have created is the sum of the numbers of moles times the molar free energies, $\mu_i^{\alpha,\bullet}$, of the pure substances in phase α plus the free enthalpy of mixing which is $-T\Delta S$ since the solution is ideal. Hence,

$$G(T, p) = \sum_{i=1}^N n_i \left(\mu_i^{\alpha,\bullet}(T, p) - RT \ln \frac{n}{n_i} \right)$$

[†] In the case of a binary mixture you can derive (A3.3) from a statistical point of view. If the N atoms are arranged on the lattices of two perfect crystals then the number of ways of arranging the atoms into two separated bodies is exactly **one**. When the atoms are randomly mixed on a common lattice in the alloy body then the number of ways of arranging the N_a A-atoms and the N_b B-atoms is

$$W = \frac{(N_a + N_b)!}{N_a! N_b!} = \frac{N!}{N_a! N_b!}$$

So from the Boltzmann formula for the entropy, $S = k_B \ln W$ the entropy of mixing is

$$\begin{aligned}\Delta S &= k_B \ln \frac{W}{\text{one}} = k_B (\ln N! - \ln N_a! - \ln N_b!) \\ &= -k_B \left(N_a \ln \frac{N_a}{N} - N_b \ln \frac{N_b}{N} \right)\end{aligned}$$

using Sterling’s approximation $\ln x! \approx x \ln x - x$ for large x , and this is consistent with (A3.3).

By comparison with equation (A2.10) I see that the chemical potential of component i in the ideal mixture is

$$\begin{aligned}\mu_i &= \mu_i^{\alpha,\bullet}(T, p) + RT \ln \frac{n_i}{n} \\ &= \mu^\bullet + RT \ln \frac{n_i}{n}\end{aligned}\quad (\text{A3.4})$$

so that μ^\bullet is the chemical potential of component i in its reference state: that is, in the pure substance in phase α (presumably but not necessarily the crystal structure adopted by the pure substance) at the temperature and pressure specified. So you can at least see where the structure of (A3.1) comes from: x_i is n_i/n and we replace x_i with a_i to recognise that not all mixtures are ideal in real life. In the case of the ideal gas, we have

$$dG = Vdp - SdT$$

and so at constant temperature

$$dG = Vdp = nRT d\ln p \quad ; \quad \text{using the ideal gas law } pV = nRT$$

which on integration results in

$$\left(\frac{\partial G}{\partial n} \right)_{T,p} = \mu = \mu^\ominus + RT \ln \frac{p}{p^\ominus} \quad (\text{A3.5})$$

in which if the pressure is measured in bar ($p^\ominus = 10^5 \text{ N/m}^2$) then μ^\ominus is the chemical potential (free enthalpy per mole) in the standard state of one bar pressure and the temperature in question, and the pressure p is measured in units of one bar.

3.2 Non ideal mixture

Up to now I have used the superscript \circ applied to μ to denote some unspecified “standard state”. In the case of the ideal gas once we have integrated dG then the integration constant will depend on the boundary condition. If we choose to measure the pressure in bar, then (A3.5) indicates that μ^\ominus is the chemical potential of the ideal gas at one bar pressure and the temperature in question. We could have chosen different units of pressure, atmospheres or Pascal, in which case the integration constant would be different. We are transferring weight from the constant term to the log term without altering the chemical potential in the right hand side of (A3.5). This is called “choosing a standard state.” Then, rather than the generic superscript \circ we use the superscript \ominus to indicate that we are using the “one-bar standard state.” The only restriction on the choice of standard state is that the standard chemical potential (the integration constant) should depend *only* on temperature and pressure.

In (A3.4) we used the superscript \bullet to denote the “pure substance standard state” so we rewrite (A3.1),

$$\mu_i = \mu_i^\bullet + RT \ln \gamma_i^\bullet x_i = \mu_i^\bullet + RT \ln a_i^\bullet$$

In the limit that the mixture tends to pure component i , γ_i^\bullet tends to one,

$$\mu_i = \mu_i^\bullet + RT \ln \gamma_i^\bullet x_i \quad (\text{A3.6a})$$

$$\rightarrow \underbrace{\mu_i = \mu_i^\bullet + RT \ln x_i}_{\text{Raoult's law}} \quad (x_i \rightarrow 1) \quad (\text{A3.6b})$$

At $x_i = 1$, the last term is zero ($\log 1 = 0$) and so $\mu_i = \mu_i^\bullet$ which identifies μ_i^\bullet as the chemical potential of pure substance i at the temperature and pressure of interest. From now on, we will refer to a_i^\bullet as the *Raoultian activity* of component i , understanding that the Raoultian activity is referred to the pure solvent standard state. In the *infinitely dilute limit*, $x_i \rightarrow 0$, the activity coefficient becomes a constant, independent of concentration.[†]

$$\mu_i = \mu_i^\bullet + RT \ln \gamma_i^\bullet x_i \quad (\text{A3.7a})$$

$$\rightarrow \underbrace{\mu_i = \mu_i^\bullet + RT \ln \gamma_i^\infty x_i}_{\text{Henry's law}} \quad (x_i \rightarrow 0) \quad (\text{A3.7b})$$

We would like, if possible, to be able to use concentration as a surrogate for the activity as then we don't have to worry about activity coefficients. If the concentration of a particular species is close to one, for example if it is the *solvent*, then we can invoke Raoult's law and use (A3.6b) as our chemical potential and μ^\bullet as our standard chemical potential. If on the other hand the species has a very small concentration it may be that we can invoke Henry's law and assume a constant activity coefficient, γ^∞ . We could then use (A3.7b) for the chemical potential. It is useful therefore to create a new, infinitely dilute, standard state in order to describe the activity of dilute solutes in a way that allows us to use concentration in place of activity in the dilute limit. To this end we transfer the constant γ^∞ from the log term to the standard chemical potential so that Henry's law (A3.7b) becomes

$$\mu_i = \mu_i^\infty + RT \ln x_i \quad (\text{A3.7c})$$

$$= \mu_i^\infty + RT \ln h_i^\infty \quad (\text{A3.7d})$$

and

$$\mu_i^\infty = \mu_i^\bullet + RT \ln \gamma_i^\infty$$

is the chemical potential in the *Henrian* standard state and, since γ_i^∞ does not depend on composition, depends only on T and p . Of course the species may not be sufficiently dilute to obey Henry's law, but still it's worthwhile to refer its chemical potential to the Henrian standard state. Just as in acknowledging that if the solvent is not sufficiently concentrated its chemical potential may be given by (A3.6a) in which the activity coefficient, γ_i^\bullet depends on T , p and the concentrations of all species; so we write for the chemical potential of the *solute*

$$\mu_i = \mu_i^\infty + RT \ln \gamma^r x_i \quad (\text{A3.8a})$$

$$= \mu_i^\infty + RT \ln h_i^r \quad (\text{A3.8b})$$

[†] The mathematically astute will notice that as x_i goes to zero so μ_i goes to minus infinity.

in which h^r is the *rational Henrian activity* and γ^r is the activity coefficient which tends to one in the dilute limit just as γ^\bullet tends to one in the concentrated limit of the Raoultian activity.[‡]

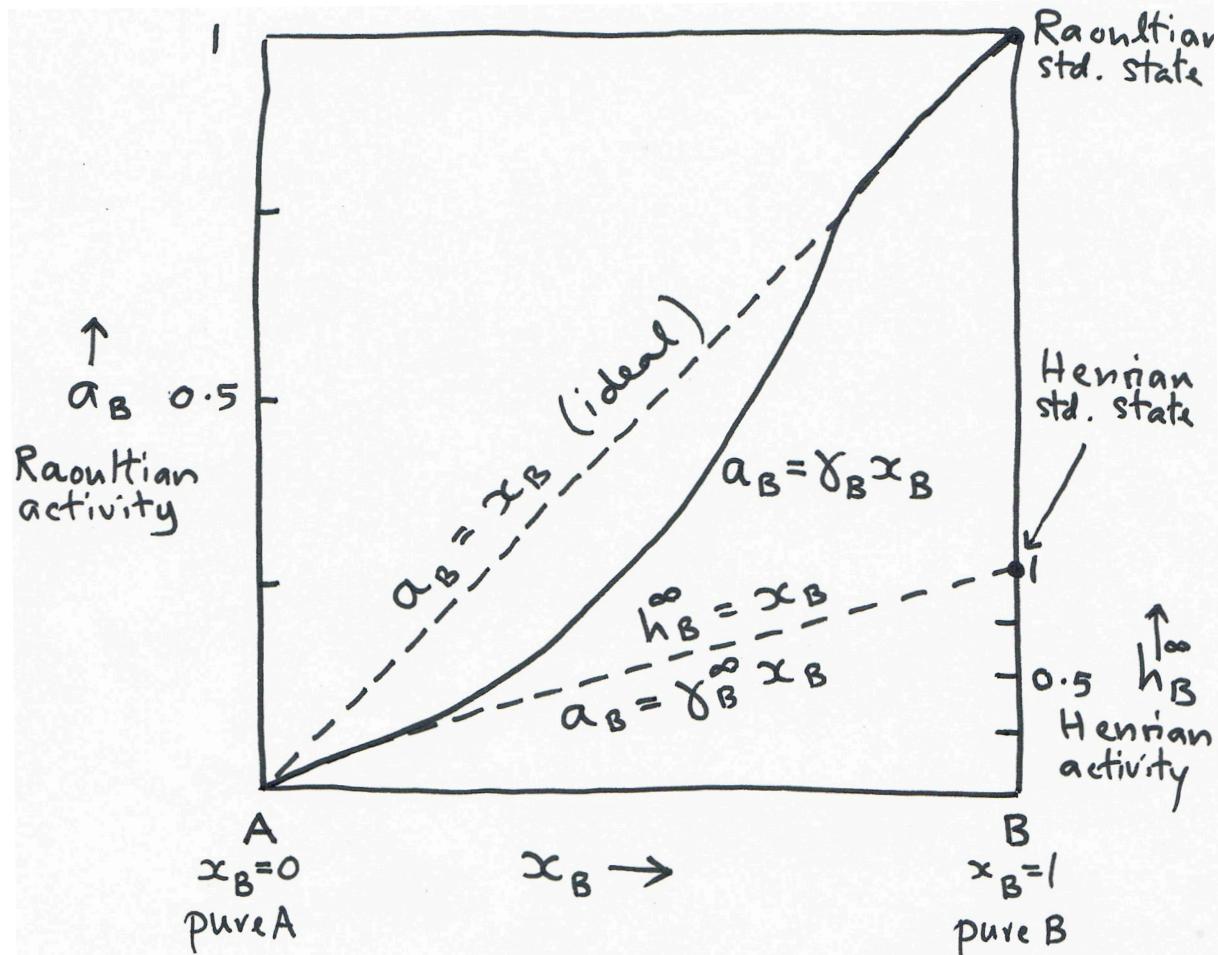


Figure A2. (After John West) Cartoon of the activity as a function of concentration (solid line) for a binary AB solution. An ideal solution has $a = x$. The Raoultian activity, a , is equal to the concentration, x in the limit as $x \rightarrow 1$. The Henrian activity h^∞ is equal to x as $x \rightarrow 0$. Because the width of the abscissa is one, the slope of the Henrian activity as a function of x is γ^∞ and the straight line intersects the right hand ordinate at the Henrian standard state.

These concepts are illustrated in figure A2. The Raoultian standard state (A3.7a)

[‡] The ‘r’ superscript stands for *rational* (see below). I have been a bit unconventional and decorated my activities and activity coefficients with superscript symbols. This is because in electrochemistry we want to use the *unimolar* standard state to be described below and I therefore reserve un-superscripted symbols for that scale so that they can be raised to mathematical powers without notational clutter.

may be used over the whole range of concentration since the Raoultian activity, $a_i^* = \gamma_i(T, p, x_1, x_2, \dots, x_N)x_i$ describes the whole solid curve and contains all of the non ideality of the mixture. On the other hand the Henrian standard state (A3.7c) is very convenient because we can use concentration directly as a surrogate for activity, but it is *only correct as long as the component is sufficiently dilute* for Henry's law to be valid; that is, the solid curve in figure A2 must coincide with the Henry's law straight line. Otherwise, we must multiply the atom fraction, x by a Henrian activity coefficient γ^r , defined in (A3.8a), which like the Raoultian activity coefficient depends on T and p and on the concentration of all the components. Also the Henrian standard chemical potential, μ^∞ , is impossible to visualise: it is the chemical potential for a solute at a mole fraction of one—at which point it has become pure solvent! On the other hand the Raoultian standard chemical potential is clearly the molar free enthalpy of the pure substance, which may be conveniently be set to zero at standard temperature and pressure.

3.3 Unimolar standard state

The mole fraction, x_i , is the same as the atom fraction and is the number of moles (or atoms) of component i divided by the total number of moles (or atoms) in a substance. This is obviously a good measure of concentration in physical or metallurgical thermodynamics, but in electrochemistry and solution chemistry it is usual to measure concentrations in moles per litre or moles per kilogram of solvent. By definition one mole of a substance has a mass equal to its relative molar mass (molecular weight), M_W , in grams. Hence one kg of solvent contains $10^3/M_W$ moles. The mole fraction is

$$x_i = \frac{n_i}{n} \quad (\text{A3.9})$$

and the concentration in moles per kg is

$$c_i = \frac{n_i}{M} \quad (\text{A3.10})$$

where M is the mass of the solvent in kg. If there are n moles of solvent the mass of the solvent in kg is

$$M = 10^{-3} \times n M_W$$

and hence

$$c_i = \frac{10^3}{M_W} \frac{n_i}{n} = x_i \frac{10^3}{M_W}$$

This is not dimensionless and so we finally define a quantity called the “molality”

$$m_i = \frac{c_i}{c^\circ} \approx x_i \frac{10^3}{c^\circ M_W} \quad (\text{A3.11})$$

in which c° is the standard concentration of one mole per kilogram.

There is a small error which I have indicated in comparing x_i and c_i . In (A3.9) the denominator is the *total number of moles* $\sum_j n_j$ including the component i , whereas

in (A3.10) the denominator is just the mass of the solvent. So (A3.10) is only correct in the dilute limit. All the same we can still move weight from the log term to the standard chemical potential term, and introduce the *unimolar standard state*,[†]

$$\mu_i = \mu_i^\circ + RT \ln h_i \quad (\text{A3.12a})$$

$$= \mu_i^\circ + RT \ln \gamma_i m_i \quad (\text{A3.12b})$$

In this way, μ_i° is the standard unimolar chemical potential, h_i is the unimolar Henrian activity and γ_i is the unimolar Henrian activity coefficient. While μ^∞ is truly a mythical state of a solute in unit mole fraction, μ° is the chemical potential of the solute in the concentration of one mole per kg of solvent at standard T and p —this is far from dilute.

For rather obvious reasons, the activity coefficient, γ^r , on the mole fraction scale is sometimes called *rational*; and the activity coefficient, γ , on the molality scale is sometimes called *practical*. The relations between the two scales are easily derived, the standard chemical potentials are,

$$\mu_i^\infty - \mu_i^\circ = RT \ln \frac{10^3}{M_W} \frac{n}{n_0}$$

where n_0 is the number of moles of solvent. In the dilute limit,

$$\mu_i^\infty - \mu_i^\circ = RT \ln \frac{10^3}{M_W}$$

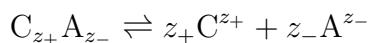
The relation between practical and rational activity coefficients is

$$\frac{\gamma_i}{\gamma_i^r} = \frac{10^3}{M_W} \frac{x_i}{m_i}$$

In view of (A3.11), the two are equal in the dilute limit.

3.4 Chemical potential and activity of ionic species

Some substances, acids, alkalis and salts for example, *dissociate* into their ions when dissolved in water. In the case of a dissolved salt, such as sodium chloride, this will be completely ionised in water to its constituent ions, Na^+ and Cl^- . We need to be careful of what can and cannot be known of the individual ions—their chemical potentials and activities are not *measurable*: only certain combinations are. We consider a salt or acid which dissociates into anions and cations,



in which the cation, C, has a charge number z_+ , and the anion, A, has a charge number z_- . We imagine that m moles of the substance are dissolved in n_0 moles of solvent. If

[†] I have been putting superscripts on h and γ up to now, in order to reserve the naked versions for the unimolar scale. This is because it is this scale that is mostly used in the main text; and in the following, where these are sometimes raised to some power so that a superscript would be clumsy. This is a bit unconventional.

not all the substance is dissociated then let m_u be the number of moles of undissociated substance. The numbers of moles of the ions in solutions will then be,

$$m_+ = z_+ (m - m_u) \quad (\text{A3.13a})$$

$$m_- = z_- (m - m_u) \quad (\text{A3.13b})$$

The total differential (A2.9) in this case will be

$$dG = -SdT + Vdp + \mu_u dm_u + \mu_+ dm_+ + \mu_- dm_- + \mu_0 dn_0$$

According to (A2.7) the chemical potential of the ions will be,

$$\begin{aligned} \mu_+ &= \left(\frac{\partial G}{\partial m_+} \right)_{T,p,m_+,n_0} \\ \mu_- &= \left(\frac{\partial G}{\partial m_-} \right)_{T,p,m_-,n_0} \end{aligned}$$

the subscript 0 referring to the solvent (water). This change in mole numbers cannot be realised in practice because even an infinitesimal increase in the number of positive ions while keeping the number of negative ions fixed will give rise to a large electric field that will swamp the change in energy. We now need to look into what can be known. In terms of (A3.13) at constant T and p ,

$$\begin{aligned} dG &= \mu_u dm_u + z_+ \mu_+ (dm - dm_u) + z_- \mu_- (dm - dm_u) + \mu_0 dn_0 \\ &= (\mu_u - z_+ \mu_+ - z_- \mu_-) dm_u + (z_+ \mu_+ + z_- \mu_-) dm + \mu_0 dn_0 \end{aligned} \quad (\text{A3.14})$$

Now all the differentials refer to quantities that *can* be varied independently without charging up the solution. *Firstly* consider the case where n_0 and m are fixed. Then at equilibrium

$$\left(\frac{\partial G}{\partial m_u} \right)_{T,p,m,n_0} = 0$$

from which according to (A3.14)

$$\mu_u = z_+ \mu_+ + z_- \mu_-$$

which expresses the condition after m moles of the substance have been added to n_0 moles of the solvent and the dissociation has proceeded to equilibrium. *Secondly* now relax the constraint on m and n_0 but allow that as these are varied this is done slowly enough that equilibrium is maintained in the process so the degree of dissociation is always as in the previous condition. Then the first term in (A3.14) vanishes by virtue of the last equation and

$$dG = (z_+ \mu_+ + z_- \mu_-) dm + \mu_0 dn_0$$

Now it is possible to define the chemical potential of the substance through the variation of m at fixed n_0 while maintaining the equilibrium dissociation of the substance,

$$\mu = \left(\frac{\partial G}{\partial m} \right)_{T,p,n_0}$$

By comparison of the last two equations it is evident that

$$\mu = z_+ \mu_+ + z_- \mu_- \quad (\text{A3.15})$$

The conclusion is that individual chemical potentials of ions cannot be defined but a combination weighted by charge numbers amounts to the chemical potential of the dissolved substance.

As to the activity coefficients of individual ions, these could be defined via the following equations,

$$\mu_+ = \mu^\circ_+ + RT \ln \gamma_+ m_+ \quad (\text{A3.16a})$$

$$\mu_- = \mu^\circ_- + RT \ln \gamma_- m_- \quad (\text{A3.16b})$$

Putting (A3.16) into (A3.15) we get,

$$\mu = z_+ \mu^\circ_+ + z_- \mu^\circ_- + RT \ln \gamma_+^{z+} \gamma_-^{z-} m_+^{z+} m_-^{z-}$$

Since the chemical potentials on the left hand sides of (A3.16) are not measurable neither are the quantities on the right hand sides. However since μ is measurable then so is the linear combination

$$\mu^\circ = z_+ \mu^\circ_+ + z_- \mu^\circ_- \quad (\text{A3.17})$$

and so is the product $\gamma_+^{z+} \gamma_-^{z-}$. We let $z = z_+ + z_-$ and we define the *mean ion activity coefficient*, γ_\pm , through the formula

$$\mu = z_+ \mu^\circ_+ + z_- \mu^\circ_- + RT \ln \gamma_\pm^z m_+^{z+} m_-^{z-} \quad (\text{A3.18})$$

In this way we have identified

$$\gamma_\pm^z = \gamma_+^{z+} \gamma_-^{z-} \quad (\text{A3.19})$$

as a proper measurable combination of ion activity coefficients. Even though the individual ion molalities, m_+ and m_- , are known from the concentration of our electrolyte, we can go further and define a *geometric mean ion molality*

$$m_\pm = (m_+^{z+} m_-^{z-})^{1/z} \quad (\text{A3.20})$$

and finally (A3.18) takes on the simple form

$$\mu = z_+ \mu^\circ_+ + z_- \mu^\circ_- + zRT \ln \gamma_\pm m_\pm \quad (\text{A3.21a})$$

$$= \mu^\circ + zRT \ln \gamma_\pm m_\pm \quad (\text{A3.21b})$$

3.5 Electrolyte chemical potential and activity

Electrochemistry is fraught with difficulty because the student would like to use equilibrium thermodynamics and is hampered by the problem that there are no well defined thermodynamic properties of charged species, or substances which carry a net charge, so that the charge neutrality demanded by thermodynamic theory cannot be guaranteed.

What is most needed is a measure of the activity of an ion. We achieve a “working value” of this in (A3.26), below. Suppose that a substance is totally dissociated as is for example NaCl or hydrochloric acid, HCl, or an alkali such as NaOH, then (A3.13) simplifies to $m_+ = z_+ m$ and $m_- = z_- m$ and these are known from the amount of salt or acid we’ve thrown into the water. Suppose we are dealing with an electrolyte such as sodium chloride, then as in (A3.11) we can work out the molality of the electrolyte, m_{el} . If the solution is not ideal, then the activity in the *practical Henrian scale* is

$$h_{\text{el}} = \gamma_{\text{el}} m_{\text{el}}$$

and γ_{el} is the activity coefficient of the electrolyte in that scale. It is surely uncontroversial to write for the chemical potential of the electrolyte in solution using the unimolar practical standard,

$$\mu_{\text{el}} = \mu_{\text{el}}^\circ + RT \ln h_{\text{el}} \quad (\text{A3.22a})$$

$$= \mu_{\text{el}}^\circ + RT \ln \gamma_{\text{el}} m_{\text{el}} \quad (\text{A3.22b})$$

We accept from the previous section which is thermodynamically rigorous that *at least in principle* there exist single ion chemical potentials and activities. And, as in (A3.16), that we can also write for the individual ions,

$$\mu_+ = \mu_+^\circ + RT \ln h_+$$

$$\mu_- = \mu_-^\circ + RT \ln h_-$$

which defines “thermodynamic” ion activities,

$$h_+ = \gamma_+ z_+ m_{\text{el}} = \gamma_+ m_+ \quad (\text{A3.23a})$$

$$h_- = \gamma_- z_- m_{\text{el}} = \gamma_- m_- \quad (\text{A3.23b})$$

So if we now write (A3.17) for the electrolyte, we’ll have,

$$\mu_{\text{el}} = z_+ \mu_+ + z_- \mu_-$$

$$\mu_{\text{el}}^\circ = z_+ \mu_+^\circ + z_- \mu_-^\circ$$

then in terms of activities, after subtracting,

$$\begin{aligned} \mu_{\text{el}} - \mu_{\text{el}}^\circ &= z_+ (\mu_+ - \mu_+^\circ) + z_- (\mu_- - \mu_-^\circ) \\ &= z_+ RT \ln h_+ + z_- RT \ln h_- \\ &= RT \ln h_{\text{el}} \end{aligned}$$

whence

$$h_{\text{el}} = h_+^{z_+} h_-^{z_-} \quad (\text{A3.24a})$$

$$\equiv h_\pm^z \quad (\text{A3.24b})$$

where $z = z_+ + z_-$. This means that we can express the activity of the electrolyte in terms of a *geometric mean ion activity*, h_\pm . As we saw in the previous section we can

define a mean ion activity coefficient (A3.19), $\gamma_{\pm}^z = \gamma_+^{z+} \gamma_-^{z-}$, and using (A3.23) it follows by simple algebra that

$$h_{\pm} = \gamma_{\pm} z_+^{z+} z_-^{z-} m_{\text{el}} \quad (\text{A3.25})$$

We also understand that (A3.22b) and (A3.12b) are expressions for the same chemical potential. It then follows that as long as the electrolyte is fully dissociated,

$$\gamma_{\text{el}} = \gamma_{\pm}^z z_+^{z+} z_-^{z-} m_{\text{el}}^{z-1}$$

In this way, the relation between the activity coefficient of the electrolyte and the mean ion activity coefficient varies with the concentration (molality). For a one-to-one electrolyte, such as NaCl or HCl, at a molality of one (the standard state)

$$\gamma_{\pm} = \sqrt{\gamma_{\text{el}}}$$

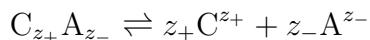
So far so good, and thermodynamically rigorous. However the activity of individual ions cannot be determined experimentally. It is sometimes necessary to assume some values of these, in which case *leaving thermodynamics aside* momentarily a modification of (A3.23) is possible in order to define “working values of the ion activities,”

$$h_+ = \gamma_{\pm} z_+ m_{\text{el}} = \gamma_{\pm} m_+ \quad (\text{A3.26a})$$

$$h_- = \gamma_{\pm} z_- m_{\text{el}} = \gamma_{\pm} m_- \quad (\text{A3.26b})$$

3.6 Measurement of mean ion activity coefficient

We consider again the dissociation of an electrolyte into its cations and anions in solution in water,

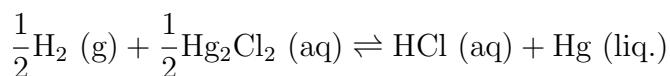


The mean ion activity coefficient can be measured in an electrochemical cell, as long as one electrode is reversible with respect to the anion and one with respect to the cation.[†]

The cell diagram



represents a hydrogen electrode connected to a “calomel electrode”. (Hg_2Cl_2 is the substance calomel which was used widely as a medicine for a huge number of diseases, until it was realised in the late nineteenth century that mercury was poisoning the patients.) The cell reaction is



[†] See section 2.1 in the main text. *Reversible* means that the electrode reaction is unique and rapid. For example in the lemon lamp (section 1 in the main text) the Zn electrode is not reversible because the electrode reaction may either be $\text{Zn} \rightarrow \text{Zn}^{++} + \text{e}^-$ or $2\text{H}^+ + \text{e}^- \rightarrow \text{H}_2$, which can short circuit the cell so that the “resistance” of the interphase is not infinite (see figure 2–3 in the main text).

According to equation (9.3.6) in the main text,

$$\mu_{\text{HCl}} + \mu_{\text{Hg}} - \frac{1}{2}\mu_{\text{H}_2} - \frac{1}{2}\mu_{\text{Hg}_2\text{Cl}_2} = -F\mathcal{E}$$

Using (A3.21) we have,

$$\mu_{\text{HCl}} = \mu_{\text{H}^+} + \mu_{\text{Cl}^-} + 2RT \ln \gamma_{\pm} m_{\pm}$$

and according to (A3.20), in this instance,

$$m_{\pm} = m$$

the molality of the electrolyte, HCl. Assuming that the hydrogen is an ideal gas,

$$\frac{1}{2}\mu_{\text{H}_2} = \frac{1}{2}\mu_{\text{H}_2}^{\ominus} + \frac{1}{2}RT \ln p_{\text{H}_2}$$

where p_{H_2} is the pressure in bar. We now have

$$RT \ln \frac{\gamma_{\pm}^2 m^2}{\sqrt{p_{\text{H}_2}}} - F\mathcal{E}^{\circ} = -F\mathcal{E}$$

where the *standard e.m.f.* of the cell is defined by

$$-F\mathcal{E}^{\circ} = \mu_{\text{H}^+}^{\ominus} + \mu_{\text{Cl}^-}^{\ominus} + \mu_{\text{Hg}}^{\bullet} - \frac{1}{2}\mu_{\text{H}_2}^{\ominus} - \frac{1}{2}\mu_{\text{Hg}_2\text{Cl}_2}^{\bullet}$$

which is the standard free enthalpy change of the cell reaction. If we maintain the pressure of hydrogen at one bar then,

$$2RT \ln m + F\mathcal{E} = F\mathcal{E}^{\circ} - 2RT \ln \gamma_{\pm}$$

and the two terms in the left hand side are *measurable*. The constant $F\mathcal{E}^{\circ}$ can be determined by varying the molality (concentration) of the acid and recording the open circuit voltage (e.m.f.) of the cell. This can be extrapolated to zero molality at which $\gamma_{\pm} = 1$ (although there are subtleties needed to avoid the infinity in the logarithm—see Denbigh) and it is found that \mathcal{E}° at 25°C is 0.26796V. Once that is known then at any concentration, the equation can be used to find γ_{\pm} of the electrolyte.

Here are measurements for HCl which show how deviation from Henry's law is observed as the concentration increases.

m_{HCl}	0.001	0.005	0.01	0.05	0.1	0.5	1	2
x_{HCl}	3.6×10^{-5}	1.8×10^{-4}	3.6×10^{-4}	0.0018	0.0036	0.018	0.036	0.072
γ_{\pm}	0.966	0.928	0.905	0.830	0.796	0.757	0.809	1.009

In this table, (A3.11) is used to compare the practical molality with the rational mole fraction concentrations. Note that the activity coefficient is less than one in the dilute limit (of course tending to one in the infinitely dilute limit), that it becomes initially smaller as the concentration increases, but that at large concentrations it becomes greater than one. Note also that one molal HCl has a mean molal activity of

about 0.8. This is why the concentration of HCl in the standard hydrogen electrode is used at 1.2 molal to achieve “unit activity of the hydrogen ion,” see footnote, page 37.

Further Reading

- K. Denbigh, *The principles of chemical equilibrium*, Fourth edition (CUP, 1981)
- L. S. Darken and R. W. Gurry, *Physical chemistry of metals*, (McGraw-Hill, 1953)
- Arnold Sommerfeld, *Thermodynamics and statistical mechanics*, Volume 5 of *Lectures on Theoretical Physics*, Translated by J. Kestin, (Academic Press, New York, 1956)