2 Electrochemistry Basis

The performance of a fuel cell is mainly determined by the kinetics of the reactions occurring at the electrode, and by the transport processes through the solution and in the electrode. It is to be emphasized that the ideal or maximum efficiency of an electrochemical energy converter depends upon electrochemical thermodynamics, whereas the real efficiency depends on electrode kinetics. The fact that the real efficiencies are less than the ideal efficiencies instils interest in understanding the fundamental theory governing this direct electrochemical conversion of chemical energy to electricity. Further, in fuel cell applications, the power per unit weight as well as the efficiency are important factors to be considered. Since power is the rate of producing energy, the kinetics of the electricity producing interfacial charge transfer reactions, that is, electrode kinetics, becomes important.

2.1 Thermodynamic Aspects of Electrochemical Energy Conversion

Three possible routes for the conversion of the chemical energy of fuels into electrical energy are given in Fig. 2.1

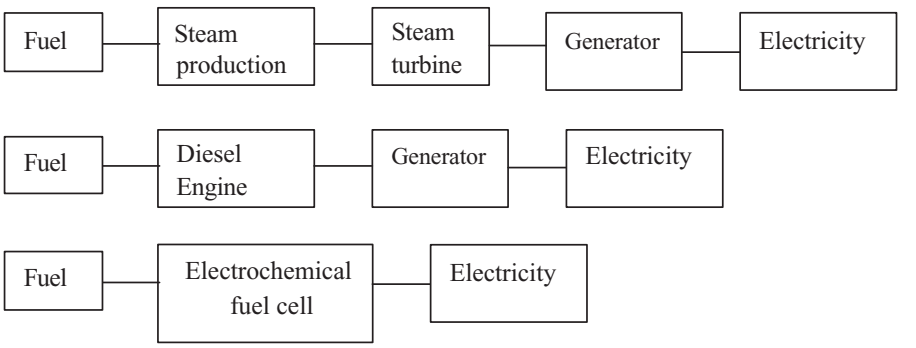

Figure 2.1 Possible routes in the conversion of chemical energy of fuels into electrical energy

The first two processes involve the indirect conversion of chemical energy into electricity by way of high temperature heat energy and mechanical energy. It is known that considerable loss in energy occurs in the conversion of high-temperature heat energy into mechanical work.

According to the Second law of thermodynamics, the maximum work, W_{max}, obtainable from an ideally reversible heat engine, working by the natural flow of heat from a higher temperature T_1 to a lower temperature T_2, depends on the temperatures between which heat is transferred.

$$W_{\max} = \frac{T_1 - T_2}{T_1} \cdot q_1 = \eta_c q_1 \tag{2.1}$$

where q_1 is the heat adsorbed at the higher temperature and η_c is the so-called Carnot factor: $\eta_c = (T_1 - T_2)/T_1$. If the heat supplied to the engine at the higher temperature comes from a chemical reaction (for example, from the combustion of a fuel, q_1 can be identified with the heat liberated by the reaction proceeding at constant pressure, $-\Delta H$, where ΔH is the enthalpy of reaction). It is clear that in principle, only for $T_2 \to 0\,\mathrm{K}$ can an efficiency of 100% be approached. In practice, the optimum efficiency of power stations are: 45% for steam turbines; 30% for diesel and 20% for petrol driven generators.

The situation is quiet different for the direct conversion of chemical energy into electrical energy by means of electrochemical fuel cells (the third scheme in Fig. 2.1). In principle, an efficiency of 100% or more (in relation to $-\Delta H$) is obtainable.

2.1.1 Free energy change of a chemical reaction

Since it is more practical to carry out reactions at a constant temperature and pressure rather than at constant temperature and volume, the change in Gibbs free energy is more often considered than the change in Helmholtz free energy.

The Gibbs free energy change, ΔG, of the reaction

$$aA + bB \to cC + dD \tag{2.2}$$

is given by the equation

$$\Delta G = c\mu_C + d\mu_D - d\mu_A - d\mu_B \tag{2.3}$$

where μ is the chemical potential ($\mu_i = (\partial G/\partial n_i)_{T,P,nj};\ j \neq i$) of the indicated species.

The free energy change of a chemical reaction is a measure of the maximum net work obtained from the reaction. It is equal to the enthalpy change of the reaction only if the entropy change, ΔS, is zero, as may be seen from the following equation,

$$\Delta G = \Delta H - T\Delta S \tag{2.4}$$

In this regard, it is interesting to note that if in a chemical reaction the number of moles of gaseous products and reactants are equal, the entropy change of such a reaction can be zero. This is because the main contribution to the entropy change in a reaction is from the change in translational entropies of reactants and products, and it can be approximately zero for a reaction involving no change in the number of molecules in the gas phase during the reaction. An example of a chemical reaction of this type is oxidation of carbon to carbon dioxide.

$$\mathrm{C} + \mathrm{O_2} \to \mathrm{CO_2}, \quad \Delta G = -394.6\,\mathrm{kJ/mol} \tag{2.5}$$

If the number of moles of gaseous products exceeds that of the gaseous reactants, the entropy change of the reaction is positive, as a result of the increase in translational modes. Oxidation of carbon to carbon monoxide is a good example of this type:

$$2C + O_2 \rightarrow 2CO, \quad \Delta G = -137.3\,\text{kJ/mol} \tag{2.6}$$

For such a reaction, it follows from (2.4) that the free energy change is more negative than the enthalpy change of the reaction.

In general, reactions in which the number of moles of gaseous reactants exceeds that of the products are common. In such cases, the entropy change is negative and the free energy change is less negative than the enthalpy change. The gas phase reaction between hydrogen and oxygen to produce water is of this type:

$$2H_2 + O_2 \rightarrow 2H_2O, \quad \Delta G = -237.3\,\text{kJ/mol} \tag{2.7}$$

The free energy and enthalpy changes for the various types of reactions carried out in fuel cells are presented in Table 2.1

2.1.2 Standard free energy change of a chemical reaction

The chemical potential(μ), of any substance may be expressed by an equation of the form

$$\mu = \mu^{\circ} + RT \ln a \tag{2.8}$$

where a is the activity of the substance and μ has the value μ° when a is unity. The standard free energy change ΔG^0 of the reaction (2.2) is then given by (2.3) with the chemical potentials of all species replaced by their standard chemical potentials:

$$\Delta G^0 = c\mu_{c^0} + d\mu_{D^0} - a\mu_{A^0} - b\mu_{B^0} \tag{2.9}$$

Substitution of (2.8) for each of the reactants and products, and (2.9) into (2.3) gives

$$\Delta G = \Delta G^0 + RT \ln \frac{a_C^c a_D^d}{a_A^a a_B^b} \tag{2.10}$$

For a process at constant temperature and pressure at equilibrium, the free energy change is zero. With the free energy change in (2.10) as zero, it follows that

$$\Delta G^0 = -RT \ln \left(\frac{a_C^c, a_D^d}{a_A^a a_B^b} \right)_e = -RT \ln K \tag{2.11}$$

The suffix e in the activity terms indicates the values of the activities at equilibrium, and K is the equilibrium constant for the reaction.

The knowledge of ΔG° is essential to calculate the value of ΔG for any composition of a reaction mixture. The value of ΔG indicates whether a reaction will take place or not: if ΔG is positive, a reaction cannot occur for the assumed composition of reactants and products; if ΔG is negative, a reaction can occur.

Table 2.1 Thermodynamic data for some fuel cell reactions at temperature 298 K

Reaction	ΔG°, kcal mole^{-1}	ΔH°, kcal mole^{-1}	n	Δn	E_r°, volts	$\partial E_r^\circ/\partial T$, mV/°C	$\partial E_r/\partial \log P$, mV	ε_i
1. $H_2 + \frac{1}{2}O_2 \rightarrow H_2O$	−56.69	−68.32	2	−1.5	1.229	−0.84	45	0.830
2. $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$	−195.50	−212.80	8	−1	1.060	−0.31	15	0.919
3. $C_2H_6 + \frac{3}{2}O_2 \rightarrow 2CO_2 + 3H_2O$	−350.73	−372.82	14	−2.5	1.087	−0.23	10.7	0.941
4. $C_3H_8 + 5O_2 \rightarrow 3CO_2 + 4H_2O$	−503.93	−530.61	20	−3	1.093	−0.19	9	0.950
5. $C_4H_{10} + \frac{13}{2}O_2 \rightarrow 4CO_2 + 5H_2O$	−656.74	−687.99	26	−3.5	1.095	−0.17	8	0.955
6. $C_3H_{18} + 8O_2 \rightarrow 5CO_2 + 6H_2O$	−809.48	−845.17	32	−4	1.097	−0.16	7.5	0.958
7. $C_8H_{18}(g) + \frac{25}{2}O_2 \rightarrow 8CO_2 + 9H_2O$	−1268.43	−1317.46	50	−5.5	1.100	−0.14	6.6	0.963
$C_8H_{18}(1) + \frac{25}{2}O_2 \rightarrow 8CO_2 + 9H_20$	−1265.87	−1307.54	50	−4.5	1.098	−0.12	5.4	0.968
8. $C_{10}H_{22}(g) + \frac{31}{2}O_2 \rightarrow 10CO_2 + 11H_2O$	−1574.42	−1632.35	62	−6.5	1.101	−0.13	6.3	0.965
9. $CH_3OH(g) + \frac{3}{2}O_2 \rightarrow CO_2 + 2H_2O$	−168.05	−182.61	6	−1.5	1.215	−0.35	15	0.920
$CH_3OH(1) + \frac{3}{2}O_2 \rightarrow CO_2 + 2H_2O$	−167.91	−173.67	6	−0.5	1.214	−0.13	5	0.967
10. $NH_3 + \frac{3}{2_2} \rightarrow \frac{1}{2}N_2 + \frac{3}{2}H_2O$	−85.04	−91.44	3	−1.25	1.225	−0.31	25	0.930
11. $N_2H_4 + O_2 \rightarrow N_2 + 2H_2O$	−143.83	−148.69	4	−1	1.560	−0.18	15	0.967
12. $C + \frac{1}{2}O_2 \rightarrow CO$	−32.81	−26.42	2	+0.5	0.711	+0.46	−15	1.24
13. $C + O_2 \rightarrow CO_2$	−94.26	−94.05	4	0	1.022	0	0	1.002
14. $CO + \frac{1}{2}O_2 \rightarrow CO_2$	−61.45	−67.63	2	−0.5	1.333	−0.44	15	0.909
15. $Li + \frac{1}{2}Cl_2 \rightarrow LiCl(g)$	−58	−53.00	1	0.5	2.515	−0.72	−30	1.094

Table 2.1 (*continued*)

Reactions	ΔG°, kcal mole^{-1}	ΔH°, kcal mole^{-1}	n	Δn	E_r°, volts	$\partial E_r^\circ/\partial T$, mV/°C	$\partial E_r/\partial \log P$, mV	ε_i
Temperature, 423 K								
1. $H_2 + \frac{1}{2}O_2 \rightarrow H_2O$	−52.94	−58.142	2	−0.5	1.14799	−0.25	21	0.911
2. $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$	−191.29	−191.42	8	0	1.03702	0	0	0.999
3. $C_2H_6 + \frac{3}{2}O_2 \rightarrow 2CO_2 + 3H_2O$	−346.99	−340.66	14	0.5	1.07491	+0.04	−3	1.019
4. $C_3H_8 + 5O_2 \rightarrow 3CO_2 + 4H_2O$	−499.54	−487.82	20	1	1.08324	+0.05	−4.2	1.024
5. $C_4H_{10} + \frac{13}{2}O_2 \rightarrow 4CO_2 + 5H_2O$	−651.94	−634.29	26	1.5	1.08747	+0.06	−4.9	1.028
6. $C_8H_{12} + 8O_2 \rightarrow 5CO_2 + 6H_2O$	−804.98	−781.19	32	1	1.09099	+0.07	−5.4	1.030
7. $C_6H_{18} + \frac{25}{2}O_2 \rightarrow 8CO_2 + 9H_2O$	−1263.72	−1221.70	50	3.5	1.09614	+0.08	−5.9	1.034
8. $C_{10}H_{32}(g) + \frac{31}{2}O_2 \rightarrow 10CO_2 + 11H_2O$	−1569.62	−1515.37	62	4.5	1.09796	+0.08	−6.2	1.036
9. $NH_3 + \frac{3}{2}O_2 \rightarrow \frac{1}{2}N_2 + \frac{3}{2}H_2O$	−47.28	−77.28	3	0.25	0.6835	−0.96	−7.1	0.612
10. $C + \frac{1}{2}O_2 \rightarrow CO$	−36.09	−26.31	2	0.5	0.782	0.47	−21	1.372
11. $C + O_2 \rightarrow CO_2$	−94.09	−94.08	4	0	1.02309	0	0	1.003
12. $CO + \frac{1}{2}O_2 \rightarrow CO_2$	−58.26	−67.77	2	−0.5	1.26335	−0.46	21	0.860
13. $Li + \frac{1}{2}Cl_2 \rightarrow LiCi(g)$	−81.38	−52.01	1	0.5	3.52942	+2.84	−42	1.565

Thermodynamic data presented are for the oxidation of one mole of fuel.

2.1.3 Relation between free energy change in a cell reaction and cell potential

The enthalpy change of the reaction may be written as

$$\Delta H = \Delta E + P\Delta V = Q - W + P\Delta V \tag{2.12}$$

where, ΔE, Q and W are the change in internal energy, heat absorbed and work done respectively, by the system. If equation (2.7) was carried out in a heat engine, then the only work done by the system would be the work of expansion, and equation (2.12) reduces to

$$\Delta H = Q \tag{2.13}$$

The enthalpy change of the reaction is therefore equal to the heat evolved from the reaction, which is then the heat absorbed by the system.

If the same reaction is carried out electrochemically, W in (2.12) is not only the work of expansion of the gases produced, but is also the electrical work involved in transporting the charges around the circuit from the anode to the cathode—the potentials of which are $V_{\text{rev,a}}$ and $V_{\text{rev,c}}$, respectively. The maximum electrical work that can be done by the overall reaction carried out in a cell involving the transfer of n electrons is given as:

$$W_{el} = ne(V_{\text{rev,c}} - V_{\text{rev,a}}) \tag{2.14}$$

for a hypothetical case, in which the internal resistance of the cell and the overpotential losses are negligible. To convert to molar quantities, it is essential to multiply W_{el} by N, the Avogadro number. Since the product of electronic charge and Avogradro's number is the Faraday, it follows that

$$W_{el} = nF(V_{\text{rev,c}} - V_{\text{rev,a}}) \tag{2.15}$$

The only forms of work involved in the operation of the electrochemical cell are electrical work and work of expansion. Therefore, one can write

$$W = W_{el} + P\Delta V \tag{2.16}$$

If the process is carried out reversibly,

$$Q = T\Delta S \tag{2.17}$$

Using equations (2.12), (2.15) and (2.17), it follows that

$$\Delta H = T\Delta S - nF(V_{\text{rev,c}} - V_{\text{rev,a}}) \tag{2.18}$$

Comparison of equation (2.18) with (2.4) which holds for an isothermal process, results in:

$$\Delta G = -nF(V_{\text{rev,c}} - V_{\text{rev,a}}) \tag{2.19}$$

Noting that

$$(V_{\text{rev,c}} - V_{\text{rev,a}}) = E \tag{2.20}$$

equation (2.19) becomes

$$\Delta G = -nFE \tag{2.21}$$

where E is the electromotive force of the cell.

If the reactants and products are all in their standard state, it follows that

$$\Delta G^\circ = -nFE^\circ \tag{2.22}$$

where E° is the standard electromotive force, commonly referred to as the standard reversible potential of the cell.

2.1.4 Effect of temperature on free energy change

At any temperature, the free energy change of a chemical reaction is given by (2.4). Using the relation

$$\left(\frac{\partial \Delta G}{\partial T}\right)_p = -\Delta S \tag{2.23}$$

equation (2.4) becomes

$$\Delta G = \Delta H + T\left(\frac{\partial \Delta G}{\partial T}\right)_P \tag{2.24}$$

Another useful form of equation (2.24) can be obtained by utilizing equation (2.23) and remembering that ΔS is almost independent of temperature. Dividing equation (2.24) by T and subsequent differentiation with respect to T yields:

$$\left(\frac{\partial \Delta G/T}{\partial T}\right)_P = -\frac{\Delta H}{T^2} \tag{2.25}$$

Equations (2.24) and (2.25) are referred to as the Gibbs–Helmholtz equations.

The standard free energy change of a reaction may easily be expressed as a function of temperature by the substitution of $\Delta G = \Delta G^\circ$ and $\Delta H = \Delta H^\circ$ in equations (2.24) and (2.25) respectively. Since $\Delta G^\circ = -RT$. In K, the variation of the equilibrium constant K with temperature is also obtained.

2.1.5 Effect of pressure on free energy change

From equation (2.3), it follows that

$$\left(\frac{\partial \Delta G}{\partial P}\right)_T = c\left(\frac{\partial \mu_C}{\partial P}\right)_T + d\left(\frac{\partial \mu_D}{\partial P}\right)_T - a\left(\frac{\partial \mu_A}{\partial P}\right)_T - b\left(\frac{\partial \mu_B}{\partial P}\right)_T \tag{2.26}$$

Also,

$$\left(\frac{\partial \mu_i}{\partial P}\right)_T = \left(\frac{\partial}{\partial P}\right)_T \left(\frac{\partial G}{\partial n_i}\right) = \left(\frac{\partial^2 G}{\partial P\ \partial n_i}\right)_{T,n1,n2\ldots} = \left(\frac{\partial^2 G}{\partial n_i\ \partial P}\right)_{T,n1,n2\ldots} \tag{2.27}$$

and

$$\left(\frac{\partial G}{\partial P}\right)_T = V \tag{2.28}$$

Thus, using equation (2.28) in equation (2.27)

$$\left(\frac{\partial \mu_i}{\partial P}\right)_T = \left(\frac{\partial V}{\partial n_i}\right)_{n1,n2,\ldots} = \overline{V}_i \tag{2.29}$$

where V_i, is the partial molar volume of the i^{th} constituent of the mixture. Thus, introducing equation (2.29) into equation (2.26),

$$\left(\frac{\partial \Delta G}{\partial P}\right)_T = -c\overline{V}_C + d\overline{V}_D - a\overline{V}_A - b\overline{VB} \tag{2.30}$$

If the reactants are gases and the products are liquids (V_C and $V_D < V_A$ and V_B, as for the H_2–O_2 fuel cell), the net change in partial molar volumes will be considerable. If the products behave ideally, it follows that

$$V = \sum n_i \frac{RT}{P} \tag{2.31}$$

Thus, the partial molar volume of any component, i, of an ideal gas mixture is given by

$$\left(\frac{\partial V}{\partial n_i}\right)_{n1,n2\ldots} = \frac{RT}{P} = \overline{V}_i \tag{2.32}$$

where i denotes the particular constituent under consideration

Substitution of an equation of the form equation (2.32) for each of the species into equation (2.30) gives

$$\left(\frac{\partial \Delta G}{\partial P}\right)_T = \Delta n \frac{RT}{P} \tag{2.33}$$

where:

$$\Delta n = c + d - a - b = n_C + n_D - n_A - n_B \tag{2.34}$$

Integration of equation (2.33) yields

$$\Delta G_{P_2} = \Delta G_{P_i} + \Delta n\ RT\ \text{In}\ \frac{P_2}{P_1} \tag{2.35}$$

Thus, if ΔG_{P_i} is known at a pressure P_1, ΔG_{P_2} at pressure P_2 may be calculated.

If the reactants and products are all liquids or solids, the influence of pressure on the free energy change is small. The effect of pressure on the standard free energy change of a reaction is expressed by an equation similar to (2.30) but with the substitutions $\Delta G = \Delta G^{\circ}$ and $\overline{V}_i = \overline{V}_i^0$ where $\overline{V}_i^0$ is the partial molar volume of the i^{th} component when all the reactants and products are in their standard states.

2.1.6 Temperature and pressure coefficients of thermodynamic reversible cell potentials

The variation of E with temperature at constant pressure can be obtained by using

$$\left(\frac{\partial \Delta G}{\partial T}\right)_p = -\Delta S$$

and using equations (2.21) and (2.22) one obtains

$$E = -\frac{\Delta H}{nF} + T\left(\frac{\partial E}{\partial T}\right)_p \tag{2.36}$$

From equations (2.11), (2.23) and (2.36), it can be seen that the second term in equation (2.36) is related to the entropy change of the cell reaction by the equation

$$nF\left(\frac{\partial E}{\partial T}\right)_p = \Delta S \tag{2.37}$$

Equations (2.36) and (2.27) show that cell potentials at any temperature can be obtained theoretically, if entropy changes of the corresponding cell reactions are known.

Using equations (2.35) and (2.21), the thermodynamic reversible cell potential may be expressed as a function of pressure:

$$E_p = E_{P_0} - \Delta n \frac{RT}{nF} \mathrm{In} \frac{P}{P_0} \tag{2.38}$$

Where E_p and E_{P_0} are the cell potentials at (total) pressure of P and P_0, respectively, and Δn is the change in the number of gaseous molecules during the reaction (assuming that the gaseous reactants and products obey the ideal gas laws). The general expression for the thermodynamic reversible cell potential as a function of pressure, as obtained from equation (2.38) and $((\partial G/\partial P)_T = V)$, is

$$E_p = E_{p_0} - \frac{1}{nF}\int_{p_0}^{p} \Delta V \, dP \tag{2.39}$$

where ΔV is the volume change of the reaction. From this equation, it follows that the effect of pressure on the thermodynamic reversible cell potential is small for reactions involving only liquids and solids. However, in cases where gaseous reactants and products are involved, the volume change is significant (for example, in the hydrogen–oxygen fuel cell), and therefore pressure effects must be taken into account. For instance, increasing the total pressure from 1 to 10 atm in the H_2–O_2 fuel cell changes the cell potential by 45 mV.

2.2 Theoretical Efficiency for Conversion of Liberated Heat in a Chemical Reaction into Mechanical Energy

2.2.1 Physical picture of heat engine and derivation of the theoretical efficiency

A common method of producing electrical energy is by converting the heat evolved during a chemical reaction into mechanical energy, which is then transformed into electrical energy. It is therefore essential to consider the theoretical limitations associated with this method of energy conversion before making similar considerations for the electrochemical method.

One can visualize this conversion process as follows. In a combustion reaction (for example, the combustion of a hydrocarbon to carbon dioxide), a certain amount of heat is evolved since the heat content of the products is less than that of the reactants. This evolved heat (the enthalpy change of the reaction) causes a rise in temperature of the products and of any unconverted reactants. Since most of these are gases, their expansion results in an increase of translational energy. In a heat engine, this expansion of gases produces mechanical work by making pistons move, whereby the mechanical energy is communicated to the wheels.

The efficiency of this conversion is less than unity, assuming an approach to an ideal system of low friction. The essential act in the conversion of the heat energy to work is the collision of gas molecules, which possesses a high translational energy due to the enthalpy of the reaction. However, when two bodies collide, they do not transfer their translational energy completely to each other. The energy that is not transferred to the piston in a heat engine appears as heat energy of the rejected gases. Hypothetically, the efficiency of a heat engine could be unity, only if the kinetic energy of the rebounding gas molecules is zero. If this were to be the case, the gases would leave the system at absolute zero.

A simple derivation for the theoretical efficiency of a heat engine is as follows. The translational energy per mole of gas entering the piston is $3RT_1/2$, where T_1 is the source of temperature; the energy per mole of gas leaving the piston after the expansion is $3RT_2/2$, where T_2 is the sink temperature. Thus, the amount of energy converted to useful work, W_u, during the expansion of the piston per mole of gas is given by

$$W_u = \frac{3R\,(T_1 - T_2)}{2} \tag{2.40}$$

The energy input per mole of gas (W_i) is given by

$$W_i = \frac{3RT_1}{2} \tag{2.41}$$

Therefore, the theoretical efficiency $\in$ of a heat engine—the Carnot efficiency—is given by

$$\in = \frac{W_u}{W_i} = \frac{T_1 - T_2}{T_1} \tag{2.42}$$

Thus, the efficiency can be unity only if T_2 is the absolute zero.

2.2.2 Thermodynamic derivation of the theoretical efficiency of a heat engine

The thermodynamic deduction of the theoretical, that is, the maximum efficiency of a heat engine, is well-known.[1–6] The derivation is based on the assumption that the gas inside the piston first expands isothermally from a volume V_1 to V_2 at a temperature T_1 and then adiabatically from V_2 to V_3 during which process the gas cools from temperature T_1 to T_2. The gas returns to its original state, again in two stages, that is, first isothermal contraction at T_2, followed by adiabatic contraction. The theoretical efficiency for the entire process was derived first by Carnot, as shown by equation (2.42). In general, it is difficult to have high source temperatures because of material problems and it is not practicable to have T_2 much less than T_1 because of conduction of heat between the high- and low-temperature zones of the machine. Thus, under these practical considerations, the maximum Carnot efficiency is generally found to be 40%–50%. The observed efficiencies of heat engines are about half these values.

2.2.3 Carnot limitation for some direct energy conversion devices

The conventional form of a heat engine that produces electrical energy, works indirectly in the sense that in the first step, mechanical energy is produced, which is then used to drive a generator and hence produce electrical energy.

There are several types of energy converters that produce electrical energy directly from heat energy—thermoelectric, thermoionic, and magnetohydrodynamic energy conversion devices. These systems are also subject to the Carnot limitation for the same reason as are conventional heat engines, where the gas molecules enter and leave the system with average kinetic energies per mole of $3RT_1/2$ and $3RT_2/2$, respectively. Thus, all the heat energy cannot be converted to electrical energy since in practice T_2 cannot be zero. The only direct energy conversion methods that are free of the Carnot limitation are those using photovolatic, electrochemical and gravitational devices.

2.3 Efficiency of Electrochemical Energy Conversion

2.3.1 Intrinsic maximum efficiency

For an electrochemical energy converter working ideally, it has been shown that the free energy change of the reaction may be converted totally into electrical energy. Thus, an electrochemical energy converter has an intrinsic maximum efficiency given by

$$\varepsilon_i = \Delta G/\Delta H = 1 - (T\Delta S/\Delta H) \tag{2.43}$$

It is perhaps not appropriate to regard this equation as indicative of an intrinsic maximum efficiency of less than 100 percent, since there is a possibility that in some reactions ΔG can exceed ΔH. Overpotential losses reduce the practical efficiencies of fuel cells to values less than the intrinsic maximum efficiencies. However, when sufficient advances in electrocatalysis are made (thereby

reducing over-potential) it might be possible to attain practical efficiencies of over 100 percent and close to the maximum intrinsic efficiencies, particularly in situations where the power density does not have to be high, so that the current density, and hence overpotential, can be low.

Effect of temperature on intrinsic maximum efficiency

If the entropy change of a reaction is negative, it follows from equation (2.43) that (recalling the negative sign of ΔH) the intrinsic maximum efficiency would *decrease* with increase of temperature. Thus, a hydrogen–oxygen fuel cell has intrinsic maximum efficiency of 0.83 at 298 K and 0.78 at 373 K. There is an important difference from this result to that for a thermal engine. As the temperature of the hot stage of the thermal energy converter increases, the difference between the temperatures of the heat source and the heat sink also increases and the corresponding Carnot efficiency consequently *increases*. Therefore, one can consider that the efficiency of a heat engine remains almost constant with change in the temperature while that of a fuel cell either decreases or increases with the increase of temperature.

At high temperatures, the need for expensive electrocatalysts is partly diminished, because the temperature itself accelerates the reaction rate and hence makes the overpotential necessary for a given current density (or power) less than that at lower temperatures.

Let us consider the oxidation of carbon to carbon monoxide. The entropy change for this reaction is positive and the value of ε_i, according to equation (2.43), exceeds unity and increases with temperature. For the complete oxidation of carbon to carbon dioxide, ΔS is approximately equal to zero and ε_i is nearly unity with hardly any temperature variation. The entropy changes in the overall reaction for several hydrocarbon–oxygen fuel cells operating at temperatures of over 373 K are also greater than zero.

2.3.2 Voltage efficiency

The cell voltage E under load is less than the thermodynamic reversible potential E_r. Factors affecting this deviation of E from E_r are dealt with in the following section. The voltage efficiency is defined as:

$$\varepsilon_e = E/E_r \tag{2.44}$$

Voltage efficiencies observed in some fuel cells (for example, hydrogen–oxygen) are as high as 0.9 at low current densities and decrease only slowly with increasing current drawn from the cell, until a limiting value is attained.[7]

2.3.3 Faradaic efficiency

The Faradaic efficiency is defined as

$$\varepsilon_f = I/I_m \tag{2.45}$$

where I is the observed current from the fuel cell and I_m is the maximum theoretically expected current on the basis of the amount of reactants consumed, assuming that the overall reaction in the fuel cell proceeds to completion. Faradaic efficiency is analogous to current efficiency in conventional electrochemical cells. In most fuel cells, ε_f can be unity. ε_f may be less than 1 because of: (1) parallel electrochemical reactions yielding fewer electrons per mole of reactant consumed,[8] (2) chemical reaction of reactants, catalyzed by electrodes[9] and (3) direct chemical reaction of the two electrode reactants.[10]

2.3.4 Overall efficiency

The overall efficiency ε in electrochemical energy conversion is the product of the efficiencies considered

$$\varepsilon = \varepsilon_i \, \varepsilon_e \, \varepsilon_f \tag{2.46}$$

For an electrochemical reaction under the chosen conditions of temperature, pressure, and concentration of reactants and products, ε_i is a definite quantity, and the maximum possible value of ε is ε_i. The main target in electrochemical energy conversion is to make both, ε_e and ε_f tend to unity. It is not a difficult task to make ε_f tend to unity in many of the fuel cell reactions.

Fuel cell systems are often used as co-generation systems, and total efficiency often includes the thermal and electrical efficiency. It is necessary to make an energy analysis that also includes application of the second law of thermodynamics.

2.4 Factors Affecting the Efficiency of Electrochemical Energy Conversion

Useful work (electrical energy) is derived from a fuel cell only when a reasonably large current is drawn and as a consequence, the cell potential will be decreased from its equilibrium potential due to irreversible losses. The losses in galvanic elements under the conditions of delivering current have been termed ‘polarization’ and this expression implies a multitude of phenomena, which are related to both the kinetics and the electrode reactions or any one of them. Polarization is influenced by the ionic mobility of the reaction partners (thermodynamic equilibrium) and are even responsive to the factors of the cell design, for example, *the dependence of cell resistance on the geometrical shape of the electrodes*.

‘Polarization’ is also termed as ‘overvoltage’. In a practical view point, the ‘overvoltage’ is the voltage difference that is measured between the ‘open circuit voltage’ and the ‘terminal voltage’ under the conditions of current flowing in either directions. Terminal voltage is also termed as ‘closed circuit voltage’. During the discharge process, the terminal voltage is lower and while charging, the terminal voltage is higher than the open circuit voltage. The ‘overvoltage’ can be considered to be a measurable value for the losses appearing as a result of the current flow in either direction. Thus, overpotential implies ‘the potential of the electrode when a net current flows through the electrode ($i \neq 0$) diminished by the equilibrium potential (when $i = 0$)’.

Fundamental types of overpotential

When electrons or metallic ions pass through the interface of a simple electrode, a thermodynamically defined reversible electrode potential E_{rev} is set up. The equilibrium involved is a dynamic one, and the rates at which charge carriers pass through the interface in either direction are equal. In other words, the exchange current density i_0 is same in both directions. If the anodic and cathodic partial current densities are respectively denoted by i_+ and i_-, then at equilibrium,

$$i_+ = \left|i_-\right| = i_0 \tag{2.47}$$

Overpotential may be observed when an additional external current is superimposed on the system by means of a suitable counter electrode, as occurs, for example, during electrolysis or during withdrawal of current from a galvanic cell. The electrode will then assume a potential E_i, which will differ from the equilibrium value to an extent governed by the current density. The differences between E_i and the equilibrium potential E_{rev} is known as the electrical overpotential, which is generally denoted by η, so that

$$\eta = E_i - E_{\text{rev}} \tag{2.48}$$

When no thermodynamic equilibrium potential is established for zero external current, the difference of the potential at $i = 0$ and $i \neq 0$ is termed polarization:

$$\eta = E_i - E_{(0)} \tag{2.49}$$

In some cases $E_{(0)} = E_{\text{rev}}$, that is, polarization and overpotential are equal to each other. $RE_{(0)} \neq E_{\text{rev}}$ is realized whenever the electrode potential is a so-called mixed potential.

Overpotential and polarization are determined with the aid of a suitable reference electrode (for example, a calomel electrode), connected to the electrode under test to form a galvanic cell. The voltage of such a cell is then measured in the presence and in the absence of a polarizing current. Two values are obtained, whose difference is the overpotential or the polarization, respectively.

The overpotential, which is generally positive when the experimental electrode works as anode ($i > 0$) and negative when $i < 0$, will in general assume a stationary value after an initial period of adjustment. A plot of these stationary values of η as a function of i is termed the 'overpotential–current curve'.

The initial slope of the overpotential–current curve,

$$R_p = \left(\frac{d\eta}{di}\right)_{i\to 0} \tag{2.50}$$

is known as polarization resistance. Since the charge e is equal to it, the polarization capacitance per unit area is defined by

$$\frac{de}{d\eta} = \frac{i\,dt}{d\eta} \equiv C_t \tag{2.51}$$

The initial capacitance when time $t = 0$ is identical to the capacitance of the electrochemical double layer. The e–η curve must therefore possess a finite initial slope, corresponding to the

finite capacitance of the double layer. The differential polarizability at time t is defined by the reciprocal of C_t, that is, by

$$\frac{1}{C_t} \equiv P_t = \frac{d\eta}{i\,dt} \tag{2.52}$$

The overpotential can be ascribed to three principal causes.

1. During the passage of a current, the activities (or concentrations) of the reactants near the electrode surface undergo a change, resulting in a corresponding alteration in the equilibrium electrode potential.
2. An additional overpotential may be required to enable the charge carriers to cross the interface (transition reaction).
3. The incorporation of atoms into the lattice of the electrode (or the reverse process) may be hindred. The resulting overpotential is termed as crystallization overpotential η_y.

The observed total overpotential, η, may thus be resolved into three additive components: the concentration overpotential η_c, the transition overpotential η, and the crystallization overpotential η_y. The concentration overpotential can be further subdivided into diffusion overpotential η_d and reaction overpotential η_r.

The *diffusion overpotential*, η_d, is caused by depletion or enrichment of the reactants at the surface of the electrode. When $AgNO_3$ solution is electrolyzed using silver electrodes, the silver is dissolved from the anode and gets deposited at the cathode, and the total concentration of Ag ions remains unchanged throughout. An increase in the concentration of silver ions will however be observed at the anode, while a corresponding decrease occurs at the cathode. Even when the electrolyte is vigorously stirred, concentration gradients of this type within a layer adhering to the electrodes may only be overcome by diffusion.

To understand the cause of the *reaction overpotential*, η_r, it should be realized that the potential determining transition reaction may be preceded or followed by chemical reactions whose rates are lower than its own. The equilibrium concentrations required to produce E_{rev} will not be established in time. The slow step may therefore be regarded as inhibiting (for establishment of E_{rev}).

Crystallization overpotential, η_y, is treated as a special type of overpotential, and is observed when the inhibiting step consists of the addition of metallic ions to a crystal lattice, or their separation from it.

The *resistance overpotential*, η_Ω, corresponds to an ohmic potential drop at the electrode and can be due to the presence of surface films.

The *total overpotential* is given by the algebraic sum of the various components, since several of them usually operate simultaneously.

$$\eta = \eta_t + \eta_y + \eta_d + \eta_r + \eta_\Omega \tag{2.53}$$

The resistance polarization, η_Ω, and the transition overpotential (η_t) are characteristic of irreversible reactions and are therefore collectively termed ‘irreversible overpotential’. On the other hand, since deviations from the equilibrium potential due to changes in the concentrations of the reactants are largely reversible, the diffusion overpotential (η_d) and the reaction overpotential (η_r) are collectively known as ‘reversible polarization’.

2.4.1 Open circuit voltage

It is the voltage that can actually be measured at the terminals of an idling cell and is a commonly used value for a cell at a given time under no-load conditions. This value is generally different from the often calculated—on the basis of thermodynamic data—'theoretical open circuit voltage', which is equivalent to the 'electromotive force' (EMF) of a cell and determined under conventionally defined standard conditions.

2.4.2 Activation overvoltage

Activation overvoltage is voltage loss caused by the charge transport/transfer of any material or process that has a limited speed. That is, activation over potential results from the slowness of one or more of the intermediate steps in either one or both of the electrode reactions. The open circuit voltage, E_a, is generally less than the thermodynamically reversible potential, E_r, for the specified conditions of temperature, pressure, activities of reactants and products, because of interference caused by adventitious reactions of impurities. When the net current drawn from the cell is small, cell potentials may tend to be controlled by these impurity reactions. However, due to their low concentrations, impurity effects are small at higher currents generated from the cell. Further, it is found that at low current densities there is a very sharp decrease of E with i. This type of behaviour is characteristic of highly irreversible processes and is caused by activation over potential. These limitations can generally be overcome by employing an 'activation catalyst' and by increasing the temperature.

2.4.3 Concentration overvoltage

It is the voltage difference caused by diffusion processes due to pressure gradients, and changes in the utilization rates of gases or liquids. The delay caused in reaching steady-state conditions, or the absence of equilibrium conditions as a result of the current flow using up or producing materials, are sources of concentration differences. Other factors could be the porosity of materials that influence the gas or liquid flow, or the permeability of membrane changing the ionic flow.

2.4.4 Reaction overvoltage

A voltage difference arises when a preceding or simultaneous (related) chemical reaction produces another compound which changes the operating conditions. For instance, water produced in a hydrogen–oxygen cell dilutes the electrolyte, which then causes an electrolyte concentration change at the electrode interface.

2.4.5 Transfer overvoltage

At sufficiently high rates, most heterogeneous reactions pass over into a region, where they are controlled by the rate of transport of reactants to, or products away from the electrode.

There are two causes of mass-transport control in electrochemical energy converters, which often use porous electrodes. The potential of one of the electrodes (rarely both together) reaches a value at which it demands a greater rate of supply of the reactant than the rate at which the diffusional and convectional processes can supply. Under these conditions, the current can no longer increase with change of potential. The other cause of mass-transfer polarization is characteristic only of porous gas diffusion electrodes. It is caused by a hold-up in the supply of gas to the electrode–electrolyte interface. This occurs when the rate of diffusion of the gas through the electrolyte-free part of the pores becomes equal to or less than the subsequent steps, with respect to, dissolution of gaseous reactant in the electrolyte, and electrochemical reaction of the dissolved reactant at the electrode–electrolyte interface within the pore.

2.4.6 Resistance overvoltage

It is not related to any chemical process at the electrodes and is simply the voltage drop across the resistive components of the cell. It is given as

$$\eta_r = ir_i \tag{2.54}$$

where r_i is the internal resistance per square centimeter of cross section of the cell. Resistivities of aqueous, molten and solid electrolytes are of the order of 1.0, 0.1, and 10 ohm cm, respectively. When porous electrodes are used, as in the case of most electrochemical energy converters, not only the layer of electrolyte between the anode and the cathode but also the electrolyte within the pores, contributes to the electrolyte resistance.

The ohmic overpotential is the simplest cause of loss of potential in an electrochemical energy converter. It may be thought that reduction in the thickness of the electrolyte layer between anode and cathode would eliminate ohmic overpotential. However, thin electrolyte layers may cause intermixing of anodic and cathodic reactants, thereby reducing Faradaic efficiencies. Even with these electrodes, ohmic overpotential may arise due to sections of porous electrodes. In the case of high-temperature fuel cells, ohmic overpotential ($i^2 r_i$ Watt cm^{-2}) may be a necessary part of the heat balance.

An electric double layer exists whenever two phases are in contact, except in the case of gas–gas, nonpolar–nonpolar liquid, or gas–nonpolar liquid interface. In general, all deductions on the rates of charge transfer reactions assume that the surfaces where reactions occur are smooth and planar in a macrosense, that is, electrodes are in the form of foils or wires. In fuel cells, however, due to the need for improved mass-transport conditions and probably greater effective area per apparent geometric area of the surface, electrodes are always porous.

2.5 Electrode Kinetics of Electrochemical Energy Conversion

2.5.1 General properties of the electrode double layer

At an interface, there are regions where the electrical field strength is not equal to zero. The electric field arises from the presence of excess of electrically charged particles such as ions, electrons

and oriented dipoles. The region in which this excess of charge is present is called the *electrical double layer*.

The presence of electric charges in the interface affects the interfacial tension. If one of the phases is a metal or semiconductor and the other is an electrolyte solution, then the changes of interfacial tension due to the presence of charges in the interface is termed *electrocapillarity*.

Let us consider the electrode–electrolyte interface. In this case, the electrical double layer is called the electrode double layer. The excess charge in the metal $Q^{(m)}$ has to be balanced in the solution by a charge of identical magnitude but with opposite sign ($q^{(1)}$). This charge is extracted from the solution by electrostatic and other forces (van der Waals and chemical). Then, the following relationship will be applicable.

$$Q^{(m)} + q^{(1)} = 0 \tag{2.55}$$

The excess charge on the electrolyte side of the interface is dispersed perpendicular to the interface towards the bulk of the solution while that on the metallic side has, even in molecular dimensions, the properties of a plate condenser. However, in a semiconductor electrode, the excess charge on the electrode side of the interface is also dispersed.

The space–charge density, $\ell(x)$, due to the diffuse dispersion of charge is related to $q^{(1)}$ by the equation

$$q^{(1)} = \int_0^\infty \ell(x)dx \tag{2.56}$$

where x is the distance from the interface. The limits of this integral are determined by considering the electrolyte as a semi-infinite medium, with the space–charge on the solution side of the double layer spreading to distances, x, of the order of hundreds of Angstroms in dilute solutions. The space–charge is usually expressed in $\mu C\ \text{cm}^{-2}$.

In the simple case of electrostatic attraction, the ions in the electrolyte can approach the interface to within distances given by their inner solvation sheaths so that only a monolayer of solvent molecules is situated between the atoms of the electrode and the 'bare' ions. The plane through the centers of ions at their minimum distance from the phase boundary is called the outer Helmholtz plane while the region between the outer Helmholtz plane and the surface of the electrode is called the Helmhotlz layer.

The electrostatic forces are unable to hold all the ions that form the excess charge $Q^{(1)}$ at the minimum distance from the phase boundary, since thermal motion continually disperses them away from the electrode. This is how the diffuse part of the double layer is formed. It is defined as that region of the double layer lying between the outer Helmholtz plane and the bulk of the solution. When only electrostatic forces are exerted on the ions the whole charge $Q^{(1)}$ is situated at this part of the electrode double layer (Fig. 2.2). The ions can also be influenced by forces other than the coulombic forces, such as van der Waals and chemical forces which cause the so-called *specific adsorption* of ions.

At a given charge on the electrode, the excess charge in the solution would increase, due to specific adsorption, to a value greater than that in the metal; however, this is balanced by a change in the amount of charge in the diffuse double layer. Specific adsorption depends on the properties

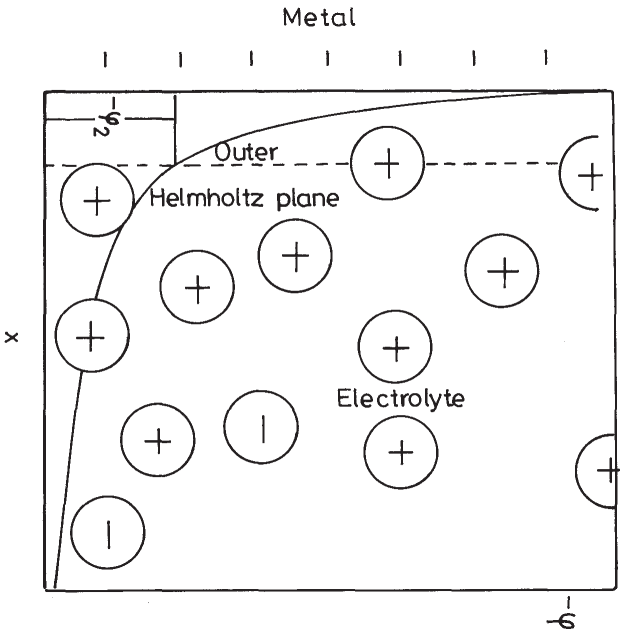

Figure 2.2 Double layer structure at metal–electrolyte interface and dependence of innerpotential, ϕ, on the distance x in the electrical double layer

of the ions as well as the electrode and it is influenced also by the electrode potential. The plane through the centers of adsorbed ions is called the inner Helmholtz plane.

Some ionic species (for example, ClO_4^-, NO_3^-) destroy the tetrahedral structure of water. The free energy of the solution decreases if these ions accumulate at the interface. This kind of adsorption occurs at the electrode–solution interface. In an analogous manner to ionic adsorption, the uncharged components of the solution will accumulate at the interface if they are less polar than the solvent (or if they are attracted to the metal by van der Waals or chemical forces). In case of electrode–solution interface, the adsorption of these substances is also affected by the electric field in the double layer acting on their dipoles. The substances that accumulate at the interface due to forces other than electrostatic ones are called surface active substances or surfactants.

As a basis for the derivation of the fundamental relationships between quantities characterizing the electrical double layer, it is necessary to introduce the concept of an *ideal polarized electrode*. The *reversible electrodes* are ones at a given temperature and pressure in equilibrium with the components of the solution, and their potentials are unambiguously determined by the activities of the components of the electrode and the solution. If a charge passes in an *ideal reversible electrode*, the subsequent processes instantly restore the original equilibrium; in this sense it is an *ideal non-polarizable electrode*. The *ideal polarized electrode* on the other hand, is capable of acquiring an electrical potential difference with respect to a reference electrode by the application of an external voltage source, and of keeping this potential difference even if the voltage source is disconnected. This potential difference will again be called the *electrode potential*. Like an EMF

of a reversible chemical cell, it is an equilibrium quantity. In contrast to the case of a galvanic cell, the potential of an ideal polarizable electrode may be varied in an arbitrary way by changing its charge without disturbing the electrode from the equilibrium state. This electrode thus has one degree of freedom more than the electrode whose potential is determined by the composition of the solution. The ideal polarized electrode is equivalent to a perfect condenser without leakage.

Obviously, real electrodes, whose equilibrium potentials are determined by the activities of the ions in solution, have the properties of a condenser (with leakage), since on their phase boundary with the electrolyte, an electrode double layer is also formed. However, this property can only be detected during the passage of current through the interface.

Another definition of the ideal polarized electrode originates from the properties of a model of this electrode. On an ideal polarized electrode, either no exchange of charged particles between the electrode and the solution takes place, or, if it is thermodynamically feasible, exchange occurs only slowly, since the value of the activation energy for this process is high.

According to Grahame, a suitable example of an ideal polarized electrode is a mercury electrode in 1 M potassium chloride solution. At the electrode potential – 0.556 V (vs. NCE) there are several reactions possible. With regard to the reactions $2Hg \quad Hg_2^{2+} + 2e^-$, $K^+ + e^- \quad K(Hg)$ and $2Cl \quad Cl_2 + 2e^-$ the equilibrium concentration of mercurous ions in solution is 10^{-36} mol l^{-1} and of potassium in the amalgam 10^{-45} mol l^{-1}; the partial pressure of chlorine is 10^{-56} atm. Obviously, such minute quantities will not influence the electrode potential. Another possible reaction is $2H_2O + 2e^- \rightarrow H_2 + 2OH^-$. The corresponding equilibrium pressure of hydrogen is rather high, 1.6×10^{-5} atm, but in terms of the high overvoltage of the hydrogen evolution reaction at the mercury electrode, this reaction does not take place at all. Thus, this electrode fulfils in an excellent way the condition that the transfer of charged particles between the electrode and the electrolyte cannot occur. This is a reason for employing mercury electrode for most of the investigations in the electrode double layer studies. But in a real situation it is usually difficult to satisfy the condition that the charge on the electrode remains constant after disconnecting the external voltage. A negatively charged electrode, for example, is discharged due to the reduction of adventitious impurities such as metal ions or oxygen.

2.5.2 The electrolyte double layer: surface tension, charge density and capacity

The interfacial region in solution is the region where the value of the electrostatic potential ϕ, differs from that in bulk solution. The basic concept is of an ordering of positive or negative charges at the electrode surface and ordering of the opposite charge and in equal quantity in solution to neutralize the electrode charge. For electrical double layer, see the next section also. The function of the electrode is to supply electrons to, or remove electrons from the interface; the charge at the interface depends on applied potential.

The proportionality constant between the applied potential and the charge due to the ordering of species in the solution interfacial region is the double layer capacity. The double layer capacity at different applied potentials can be studied by using various techniques. An often used method is the *impedance technique*, which is applicable to any type of electrode, solid or liquid. Another

method that uses *electrocapillary measurements* was developed for the mercury electrode. It is only applicable to liquid electrodes, and is based on measurement of surface tension.

The principle of electrocapillary measurements was described more than a century ago by Lippmann.[11] It is a null-point technique that counterbalances the force of gravity and surface tension, and highly accurate results can be obtained. It consists of a capillary column containing mercury upto a height 'h', regulated so that on altering the applied potential, the mercury–solution interface stays in the same position. Under these conditions, surface tension counterbalances the force of gravity according to the equation:

$$2\pi r_c \gamma \cos\theta = \pi r_c^2 \rho_{Hg} hg \tag{2.57}$$

where r_c is the radius of the capillary, θ is the contact angle, γ is the surface tension, and ρ_{Hg} the density of mercury. The contact angle is measured with a microscope. A plot of γ vs. E is called an electrocapillary curve and has the form shown in Fig. 2.3

A variation of this method consists in using the dropping mercury electrode.[12] The mass flux, m_1, is given as,

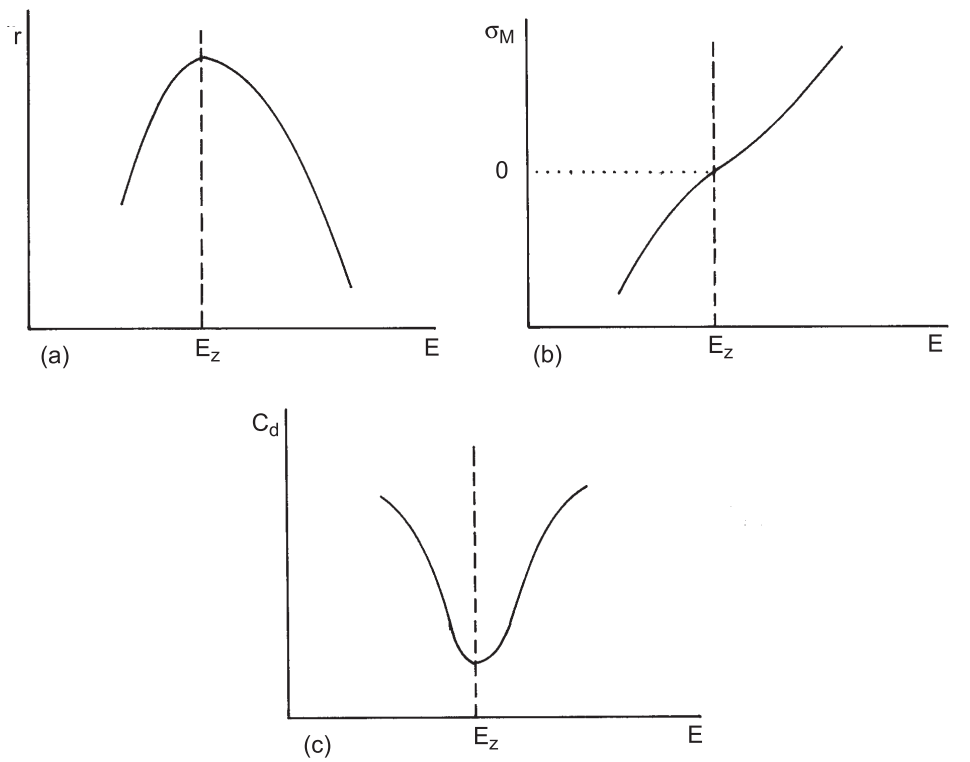

Figure 2.3 (a) Typical electrocapillary curve (plot of surface tension γ as a function of potential E) (b) Charge density on the electrode σ_M vs potential (obtained as the derivative of plot a) (c) Differential capacity, C_d vs potential (obtained by differentiating curve in b)

$$m_1 = \frac{\pi r_c^2 \rho_{Hg} h}{\tau} \tag{2.58a}$$

where τ is the drop lifetime. Substituting in equation (57) one can get

$$2\pi r_c \gamma = m_1 g t \tag{2.58b}$$

Thus, a plot of τ vs E gives a curve of the same form as the electrocapillary curve.

Conversion of values of γ into capacities is done by double differentiation in relation to the electrostatic potential difference $\Delta\phi$ between its value in the metal, ϕ_M, and that in the solution, ϕ_S. The first derivative gives the charge on the interface, and is known as Lippmann equation.

$$\frac{\partial \gamma}{\partial \Delta\phi} = -\sigma_{_M} = \sigma_{_s} \tag{2.59}$$

where σ_M is the charge on the metal and σ_s the charge of the solution such that $\sigma_M + \sigma_s = 0$. If we define an arbitrary reference potential, and the potential in relation to that reference is E_Δ then $\partial(\Delta\phi) \sim \delta(E_\Delta)$ and

$$\frac{\partial \gamma}{\partial E_\Delta} = -\sigma_{_M} \tag{2.60}$$

Therefore, the Lippmann equation is the derivative of the electrocapillary equation. The charge σ_M is zero when the slope of the electrcapillary curve is zero. The potential at which this occurs is called the point of zero charge, $E_{z.}$. Therefore, the maximum of the electrocapillary curve corresponds to the point of zero charge E_z.

A second differentiation of the electrocapillary curve gives the value of the interfacial capacity. This can be defined in two ways:

1. *The differential capacity* C_d. This is the derivative of the curve of σ_M vs E (Fig 2.3*c*). The minimum value corresponds to the point of zero charge E_z

$$C_d = \frac{\partial \sigma_M}{\partial E_\Delta} \tag{2.61}$$

2. *The integral capacity*, C_i. Consider σ_M values for two reasonably different potentials. The calculated capacity is then the average value in that zone. C_d varies with E, and the integral capacity is zero at $E = E_z$, the point of zero charge

$$C_i = \frac{\sigma_M}{E - E_z} \tag{2.62}$$

$$= \frac{\int_{E_z}^{E} C_d dE}{\int_{E_z}^{E} dE} \tag{2.63}$$

The values of C_d can be measured using impedance technique. It consists of the application of a small sinusoidal perturbation superimposed on a fixed applied potential, the component of the resulting current that is out of the capacity of the interface. Thus, value of σ_M and γ are obtained by integration.

2.5.3 Double layer models

Helmholtz model (compact layer model) (1879)

The first double layer model, due to Helmholtz,[13] considered the ordering of positive and negative charges in a rigid fashion on both sides of the interface (double layer or compact layer), the interactions not extending any further into the solution side. This situation is similar to that of a parallel plate capacitor (Fig. 2.4). x_H corresponds to the closest approach distance of the charges, point charges, that is, sum of ionic radius. The capacity would then be

$$C_H = \frac{\varepsilon_r \varepsilon_0}{x_H} \tag{2.64}$$

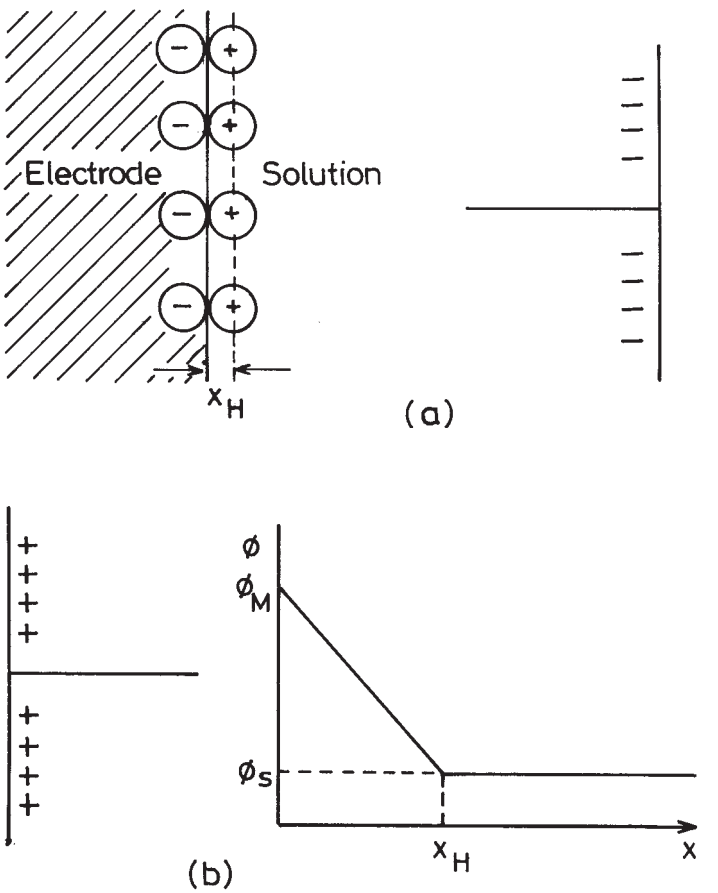

Figure 2.4 (a) Pictorial representation of Helmholz model of double layer with arrangement of point charges to give rise to a situation similar to parallel plate capacitor
(b) The variation of electrostatic potential ϕ as a function of distance x from the electrode

where ε_r is the relative permitivity (which is assumed not to vary with distance) and ε_0 the permitivity of vacuum. A typical value of ε_r is 6–7, leading to $C_H = 10\mu\,\mathrm{Fcm}^{-2}$. The decay of the electrostatic potential from ϕ_M to ϕ_s is linear and C_H remains constant in the applied potential.

This model has two serious short comings:

1. The interactions with ions in the subsequent layers to the first layer is neglected.
2. The dependence of electrolytic concentration on the accumulation of charges in the double layer has not been taken into account.

Gouy–Chapman model (diffuse layer model) (1910-1913)

Gouy[14] and Chapman[14] independently developed a double layer model. In this model they considered that both the applied potential and electrolyte concentration influenced the value of the double layer capacity. Thus, the double layer would not be compact as in Helmholtz's description, but of variable thickness. Since the ions are free to move (diffuse double layer) [Fig. 2.5(a)], there would be an equilibrium of the ions due to thermal and electrical fields in the double layer.

The distribution of species with distance from the electrode obeys Boltzmann's law

$$n_1 = n_i^0 \exp\left[\frac{-z_i e\phi_\Delta}{k_B T}\right] \tag{2.65}$$

where $\phi_\Delta = \phi - \phi_s$ and n_i^0 is the numerical bulk concentration of ions i in the bulk solution. Dividing the solution into slices of thickness dx, at distance x form the electrode, the charge density is

$$\rho\,(x) = \sum_i n_i z_i e \tag{2.66}$$

$$= \sum_i n_i^0 z_i e\ \exp\left[\frac{-z_i e\phi_\Delta}{k_B T}\right] \tag{2.67}$$

for all ions.

The Poisson equation relates the potential with the charge distribution

$$\frac{\partial^2 \phi_\Delta\,(x)}{\partial x^2} = -\frac{\rho\,(x)}{\varepsilon_r \varepsilon_0} \tag{2.68}$$

Combining equations (2.67) and (2.40), one obtains the Poisson-Boltzmann equation

$$\frac{\partial^2 \phi_\Delta\,(x)}{\partial x^2} = -\frac{e}{\varepsilon_r \varepsilon_0} \sum_i n_i^0 z_i \exp\left(\frac{z_i e\phi_\Delta}{k_B T}\right) \tag{2.69}$$

Using equation (2.69) as well as involving the property of derivatives, modulus x_{DL} is characteristic of the diffuse layer thickness and is expressed as

$$x_{DL} = \left(\frac{\varepsilon_r \varepsilon_0 k_B T}{2 n_i^0 z^2 e^2}\right)^{1/2} \tag{2.70}$$

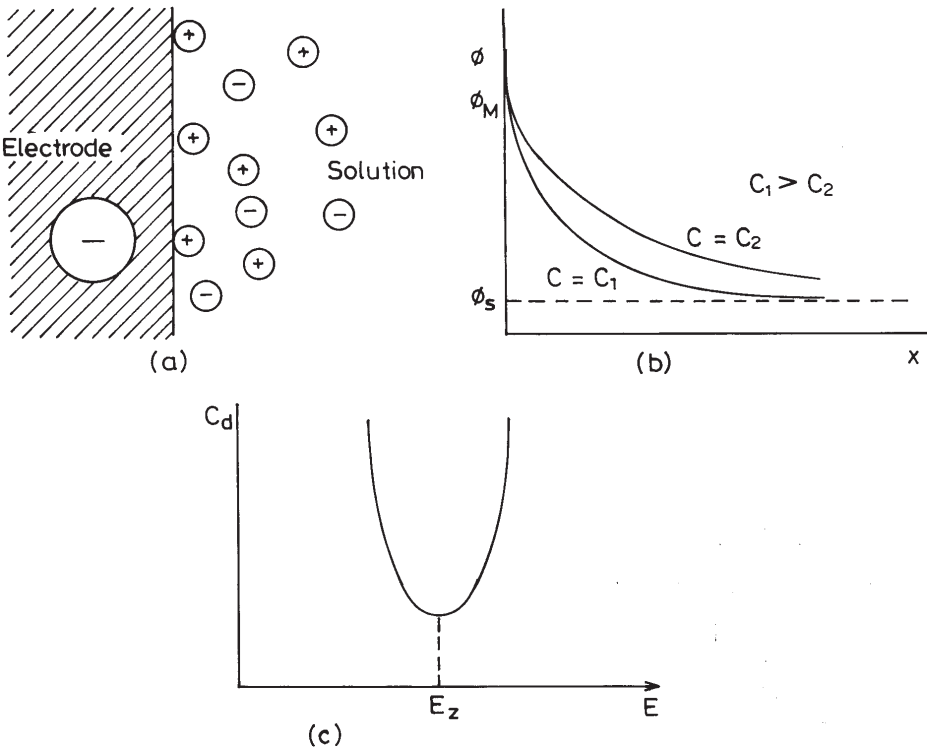

Figure 2.5 (a) Pictorial representation of ions in the diffuse double layer according to Gouy–Chapman model
(b) The variation of electrostatic potential ϕ as a function of distance x from the electrode. The variation of potential as a function of concentration of the ion is also shown in this figure
(c) Variation of C_d with potential, showing the minimum at the point of zero charge E_z

For water ($\varepsilon_r = 78$) at 298 K, $x_{DL} = 3.04 \times 10^{-8} z^{-1/2}$ cm. It is to be noted that the drop in potential with distance is faster for higher concentrations and that $x_{DL} \propto T^{1/2}$ reflects the thermal energy of the ions. If $C_\infty = 1.0$ M, and $z = 1$, then, from equation (2.70), $x_{DL} = 0.3$ mm. The decay of potential is shown in Fig. 2.5(b)

The charge density of the diffuse layer is given by

$$\sigma_M = \varepsilon_r \varepsilon_0 \left(\frac{\partial \phi_\Delta}{\partial x} \right)_{x=0} \tag{2.71}$$

$$= \left(8kT \varepsilon_r \varepsilon_0 n_i^0 \right)^{1/2} \sinh \left(\frac{ze\phi_{\Delta,0}}{2k_B T} \right) \tag{2.72a}$$

The capacity of diffuse double layer is given as:

$$C_{d,GC} = \frac{\partial \sigma_M}{\partial \phi_{\Delta,0}} = \left(\frac{2x^2 e^2 \varepsilon_r \varepsilon_0 n_i^0}{k_B T} \right)^{1/2} \cosh \left(\frac{ze\phi_{\Delta,0}}{2k_B T} \right) \tag{2.73}$$

The capacity varies systematically with potential. The point of zero charge E_z corresponds to the minimum in the curve.

For dilute aqueous solutions at 298 K,

$$C_{d,GC} = 228zc_{\infty}^{1/2} \cosh\left(19.5z\phi_{\Delta,0}\right) \mu F \text{ cm}^{-2} \tag{2.74}$$

This model is better than a parallel plate capacitor but only close to E_z: in reality, far from E_z, C_d is, to a first approximation, independent of potential.

Stern model (compact–diffuse layer model) (1924)

Stern[16] considered that the double layer was formed by a compact layer of ions next to the electrode followed by a diffuse layer extending into the bulk solution. The pictorial representation is shown in Fig. 2.6. According to this theory, the total charge on the solution side is divided between the compact and diffuse layers, and is equivalent to two capacitors in series, with capacities, C_H

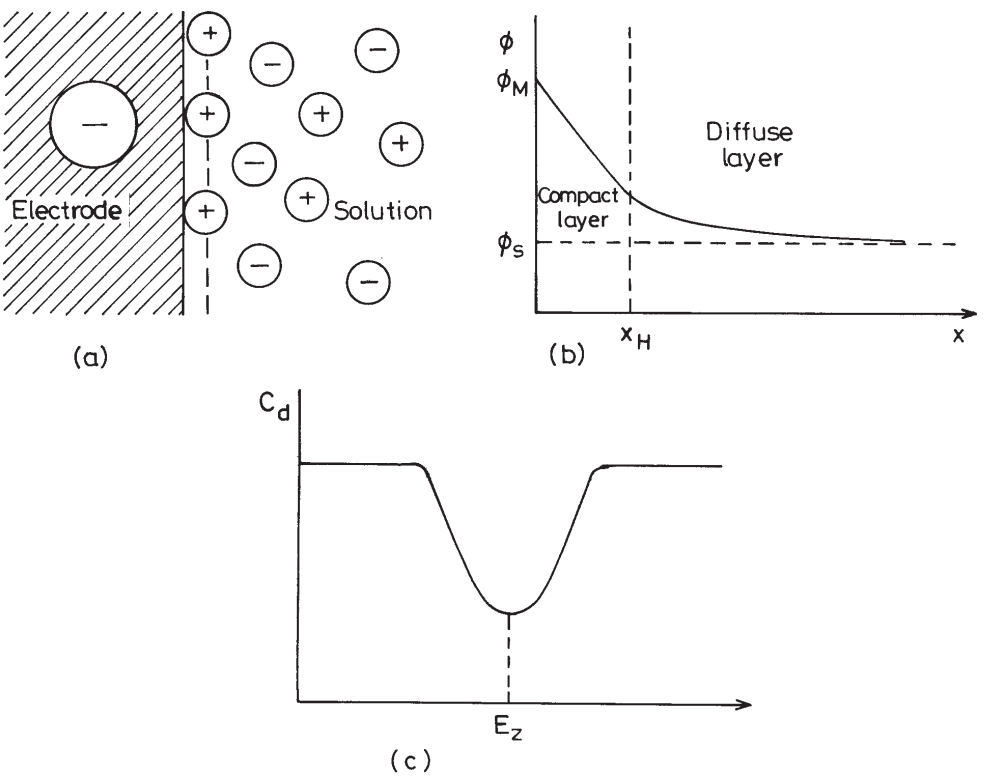

Figure 2.6 Pictorial representation of Stern model of double layer (a) compact layer (similar to Helmholz double layer model) close to the electrode and a diffuse layer extending in the solution side (Guoy–Chapmann model); (b) and (c) the variation of potential ϕ and C_d with distance and potential respectively

representing the compact layer and C_{GC} representing the diffuse layer. The smaller of the two capacities determines the observed behaviour:

$$\frac{1}{C_d} = \frac{1}{C_H} + \frac{1}{C_{GC}} \tag{2.75}$$

The total potential difference ϕ between the metal and the bulk of the solution drops at first in a linear fashion. ϕ_M in the metal drops till it meets $x_{_H}$ and thereafter decays exponentially to ϕ_S in the bulk of the solution.

There are two extreme cases for the variation of capacitance with potential:

1. close to E_z, $C_H \gg C_{GC}$ and so $C_d \sim C_{GC}$
2. far from E_z, $C_H \ll C_{GC}$ and $C_d \sim C_H$

For concentrated electrolyte solutions, the potential drop is rapid, and hence the importance of the diffused double layer is reduced. At distance x_H there is a transition from the compact to the diffuse layer. The separation plane between the two zones is called the outer Helmholtz plane (OHP).

Grahame model (triple layer model) (1947)

In spite of the fact that Stern had already distinguished between ions adsorbed on the electrode surface and those in the diffuse layer, it was Grahame[17] who developed a model that is constituted by three regions. The difference between this and the Stern model is the existence of specific adsorption: a specifically adsorbed ion loses its solvation, approaching closer to the electrode surface—with strong bonding. The inner Helmholtz plane (IHP) passes through the centers of these ions. The outer Helmholtz plane (OHP) passes through the centers of the solvated and non-specifically adsorbed ions. The diffuse region is outside the OHP.

In both the Stern and Grahame models, the potential varies linearly with distance upto the OHP and then exponentially in the diffuse layer.

Bockris, Devanathan and Muller model (1963)

More recent models of the double layer have taken into account the physical nature of the interfacial region. In dipolar solvents, such as water, it is clear that an interaction between the electrode and the dipoles must exist. This is reinforced by the fact that solvent concentration is always higher than solute concentration. For example, pure water has a concentration of 55.5 mol dm^{-3}.

The Bockris, Devanathan and Muller[18] model recognizes this situation and shows the predominance of solvent molecules near the interface. A pictorial representation is given in Fig. 2.7. The solvent dipoles are oriented according to the charge on the electrode where they form a layer together with the specifically adsorbed ions.

Regarding the electrode as a giant ion, the solvent molecules form its first solvation layer; the IHP is the plane that passes through the center of these dipoles and specifically adsorbed ions. In a similar fashion, OHP refers to adsorption of solvated ions that could be identified with a second solvation layer. Outside this comes the diffuse layer. The profile of electrostatic potential variation with distance is shown in Fig. 2.7(b) and is the same in qualitative terms as in Grahame's model.

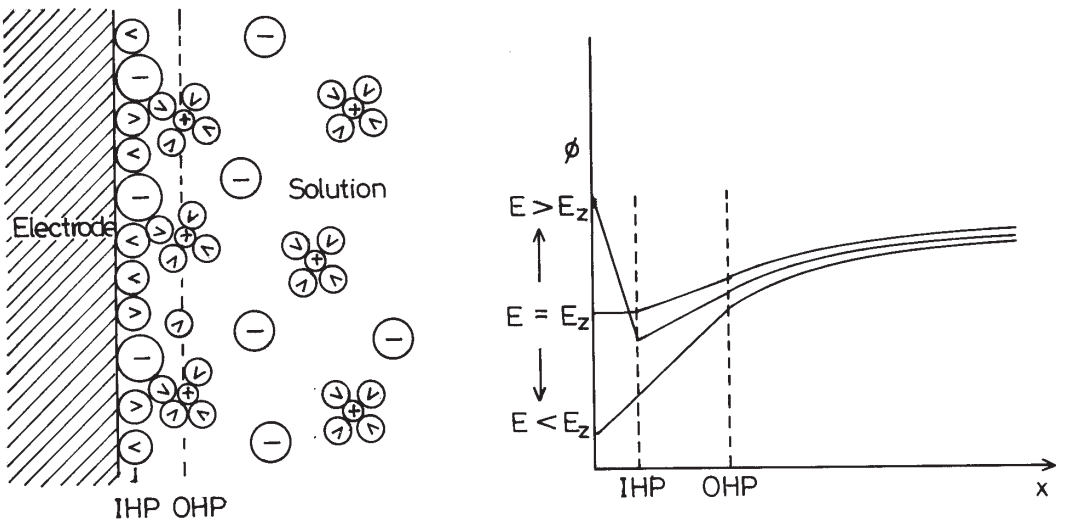

Figure 2.7 The double layer model of Bockris et al: (a) arrangement of ions and solvent molecules (^) represents a water molecule, and (b) the variation of electrostatic potential ϕ as a function of distance x from the electrode

These authors also defined a shear plane, not necessarily coincident with the outer Helmholtz plane, which is important in electrokinetic effects. The shear plane corresponds to the planes where the ions are no longer influenced by the charge on the electrode.

'Chemical' models

The concept of double layer structure is still evolving. The models developed so far emphasize electrostatic considerations. 'Chemical' models that have been developed take into consideration the distribution of the atoms in the electrode (especially solid electrodes) which is related to their work function. The variation of potential corresponding to the point of zero charge with the work function of the metal shows that sp metals follow a different linear relationship compared to transition metals.

The first model of this kind was proposed by Damaskin and Frunkin,[19] and based on these principles there has been a gradual evolution in the models, reviewed by Trasatti[20] and recently by Parsons.[21] The break in the structure of the solid causes a potential difference that begins within the solid itself.

The interfacial region of a metal upto the IHP has been considered as an electronic molecular capacitor. This model has successfully explained many experimental results.[22]

Summarizing remarks on models for the double layer

The first model for the structure of the double layer is analogous to a parallel plate condenser with a plane of charges on the metal and a second plane of opposite charges in the solution. According to this model, the capacity of the double layer should be independent of the potential across the metal–solution interface, which contradicts the experimentally observed behaviour.

The next model considered that the charges in the solution side are not located in one plane but diffuse into the bulk of the solution. This model yields a parabolic dependence of capacity on charge. Though this model is satisfactory for very dilute solutions (concentration $< 10^{-3}$ mol l^{-1}), the predicted values of the capacity are far too high in concentrated solutions.

A combination of the compact and diffuse layer models proved to be satisfactory. However, (1) in very dilute solutions, it became in practice identical with Helmholtz model (2) in the region of constant capacity, the dependence of the capacity on the radius of the ion present in the solution is not seen, though the model shows this, and (3) it cannot also explain the increase in the capacity which occurs in the anodic region.

The three layer model—the metal surface, an inner Helmholtz layer that is the locus of centers of specifically adsorbed ions, and an outer Helmholtz plane that is the locus of centers of the first layer of hydrated ions—has several satisfactory features. However, it does not provide a satisfactory explanation for the constant capacity observed in the negative branch or the hump on the anodic side.

In all these models, the role of water was ignored, even though it is the predominant species in solution. The current accepted model is that the electrode is covered with a layer of completely oriented water molecules. Specific adsorption of ions occurs in certain regions of potential by a replacement of some of the water molecules by the partially desolvated ions. The second layer of water molecules is not oriented, because these water molecules are under the influence of both the electric field and thermal fluctuations. They are like the secondary hydration sheath around an ion. Some of these water molecules can belong to the hydration sheaths of a layer of ions that are present in the outer Helmholtz plane. The dielectric constant of the first layer of oriented water molecules can be taken to be 6 to 7 and that of the second—partly oriented water layer—can be about 30 to 40. Considering the double layer as two capacitors in series, one with a low value of $\varepsilon = 6$ and the other a high value of $\varepsilon \sim 40$, the region of constant capacity with potential may be understood.

2.5.4 The solid metallic electrode

Mercury is not a typical electrode material: it is liquid, and there is constant alleviation of atoms on the surface in contact with the solution. A solid electrode has a well-defined structure (polycrystalline or monocrystalline). In solid metallic electrodes, the conduction is mainly due to the movement of free electrons at the Fermi surface, with energy E_F and following a Fermi-Dirac statistical distribution.

For a metal, the occupation of the electronic levels close to E_F is given[24] by the expression:

$$f = 1/[1 + \exp(E - E_F)/k_B T] \tag{2.76}$$

where f is the probability of occupation of a level of energy E, and k_B is the Boltzmann constant. The value of f is 0.50 when $E = E_F$ according to equation (2.76). Fermi energy therefore corresponds to the electrochemical potential of the electrons in the metal electrode. When $E = E_F + k_B T$, $f = 0.27$ and when $E = E_F - k_B T$, $f = 0.73$. The variation of f with energy E is shown in Fig. 2.8. At $T = 0$, the cut off of occupancy drops sharply at E_F, while for higher temperatures ($T > 0$ K) the occupancy is smeared out.

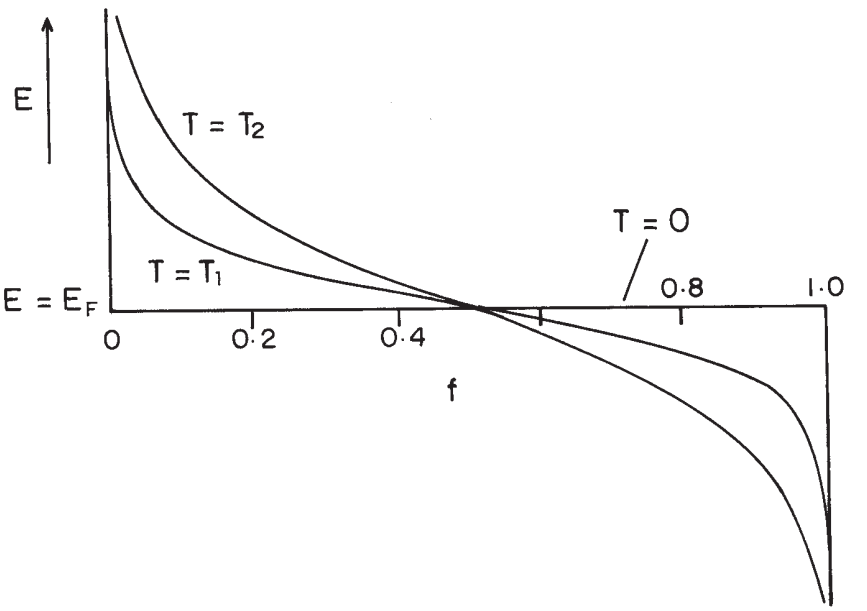

Figure 2.8 The density of occupied states in a metal in the region of the Fermi level, E_F at $T = 0$, T_1, and T_2, where $T_2 > T_1$

By convention, for a metal, only electrons with energies within $k_B T$ of E_F can be transferred. The interfacial structures of a solid electrode are of various types:

1. The inter-atomic distances vary with the exposed crystallographic planes.
2. In a polycrystalline material one can have various kinds of defects.
3. Adsoprtion of species can be specific or non-specific.

The structure of a double layer associated with the interface of gold and platinum monocrystals with solution has been investigated.[25] A clear difference between crystallographic faces is noted, as manifested in the values of differential capacity and in the evidence of adsorption in voltammograms. Cyclic voltammograms suggest that there is reorganization on the metal surface to give the equivalent of a surface layer of low Miller index. The identity of this face depends on the applied potential.[26] These studies are still in their infancy and concrete conclusions concerning a possible restructuring of the surface cannot yet be stated.

The effects of the crystallographic face and the difference between different metals are evidence for the incorrectness of the classical representation of the interface with a universal type of potential decay within the solution. In fact, a discontinuity at the interface is improbable and experimental evidence shows that the variation of electrostatic potential with distance may deviate from the value in the bulk electrode—even at a certain distance from the interface inside the electrode—as is usually considered in the Jellium model. This corresponds to the 'chemical' models and reflects the fact that the electrons from the solid penetrate a tiny distance into the solution (due to wave properties of the electron). In this treatment the Galvanic (or inner electric) potential, ϕ (associated with E_F), and the Volta (or outer electric) potential, ψ, that is the potential

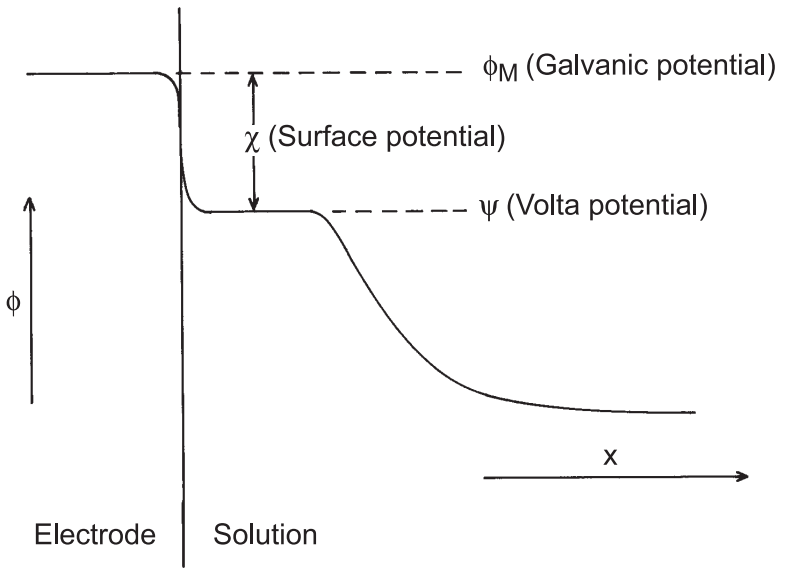

Figure 2.9 Potential variation as a function of distance from a metal electrode that is separated from electrolyte solution without charge modification (not to scale). ψ is the potential caused by the charge distributions when the electrode and solution are in contact

outside the electrode's electronic distribution (approximately at the IHP, 10^{-5} cm from the surface), are distinguished from each other. The difference between these potentials is the surface potential (Fig. 2.9).

2.5.5 The semiconductor electrode: the space–charge region

The essence of electrical conductivity is that charges must be able to move under an applied electrical field. In solids, conduction requires the movement of electrons. But for an electron to move there must be a partially vacant energy band. The electronic properties of solids are usually described in terms of the band model, in which the behaviour of an electron moving in the field of the atomic nuclei and all the other electrons are also treated. In isolated atoms, which are characterized by filled and vacant orbitals that are assembled into a lattice containing $\sim 5 \times 10^{22}$ atoms cm^{-3}, new molecular orbitals form. These orbitals are so closely spaced that they essentially form conduction bands: the filled bonding orbitals form the valence band and the vacant antibonding orbitals form the conduction band (Fig. 2.10). These bands are separated by a forbidden region or band gap of energy E_g, which has the unit of electron volts. When $E_g \ll kT$ or when the conduction and valence bands overlap, the material is a good conductor of electricity.

In a semiconductor there is a separation between the occupied valence band and the unoccupied conduction band. By convention, if the separation is greater than 3 eV the solid is called an insulator (for example, diamond 5.4 eV) and if it is less, it is a semiconductor. In a semiconductor

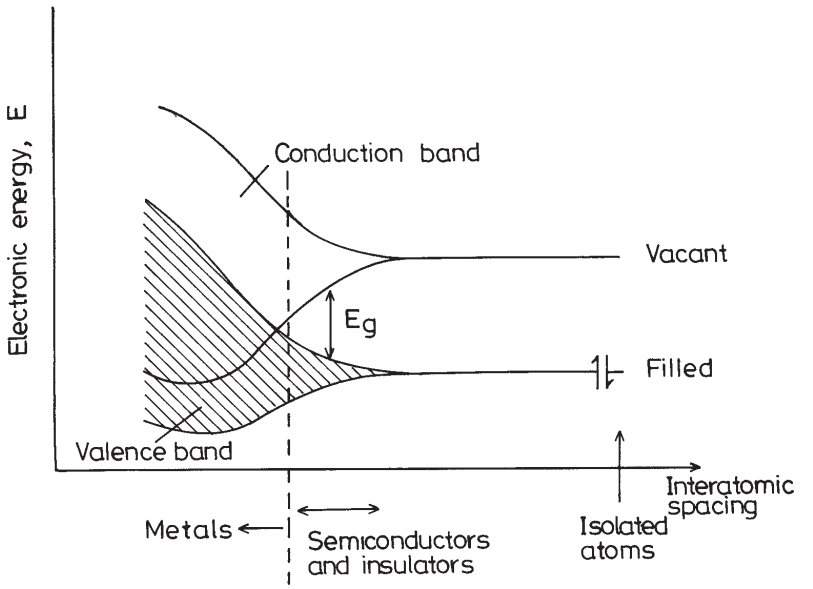

Figure 2.10 Formation of bands in solids by assembly of isolated atoms (characterized by orbitals far right of X-axis) into a lattice

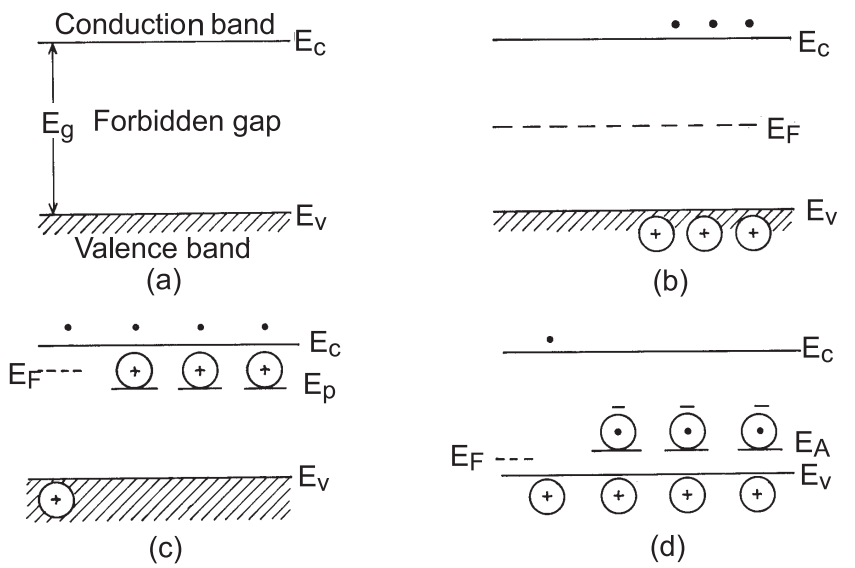

Figure 2.11 Schematic diagrams of energy levels of an intrinsic semiconductor lattice
(a) At absolute zero (or $E_g \gg kT$), assuming a perfect lattice; no holes or electrons exist
(b) At a temperature where some lattice bonds are broken, yielding electrons in the conduction band and holes in the valence band. Extrinsic semiconductor lattices
(c) *n*-type semiconductor
(d) *p*-type semiconductor

electrode[24,27,28] the accessible electronic levels are restricted, and this has important consequences. Promotion of an electron from valence to conduction band leaves a hole (lack of electron), a positive charge. It is therefore essential to consider not only electron movement but also hole movement. Conduction occurs by movement of electrons in the conduction band or of holes (lack of electrons) in the valance band.

In an intrinsic semiconductor, electron promotion to the conduction band occurs through thermal or photon excitation. An important concept in discussion of solid state material is the Fermi level. This is defined as the energy level at which the probability of occupation by an electron is $\frac{1}{2}$; for example, for an intrinsic semiconductor the Fermi level lies at the mid-point of the band gap, as shown in Fig. 2.11(b). Doping changes the distribution of electrons within the solid, and hence changes the Fermi level. For an n-type semiconductor, the Fermi level lies just below the conduction band Fig. 2.11(c), whereas for a p-type semiconductor, it lies just above the valence band Fig. 2.11(d).

In terms of the bandgap energy, E_g, substituting $E = E_c$ in (2.76), and considering the case where $E_g \gg kT$, the number of excited electrons, n, is given as:

$$n \propto exp(-E_g/2k_BT) \tag{2.77}$$

We now need to consider what happens at the (idealized) interface between a semiconductor electrode and an electrolyte solution. Figures 2.12 and 2.13 show schematically the charge distribution at the interface. The anions that form the electrolyte region of the double layer can

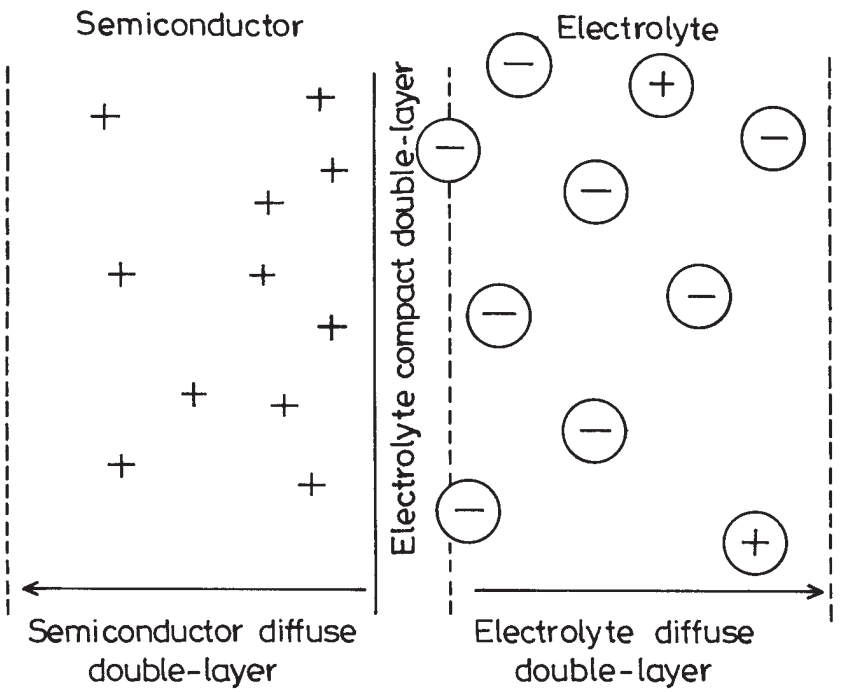

Figure 2.12 Double layer structure at a semi–conductor–electrolyte interface

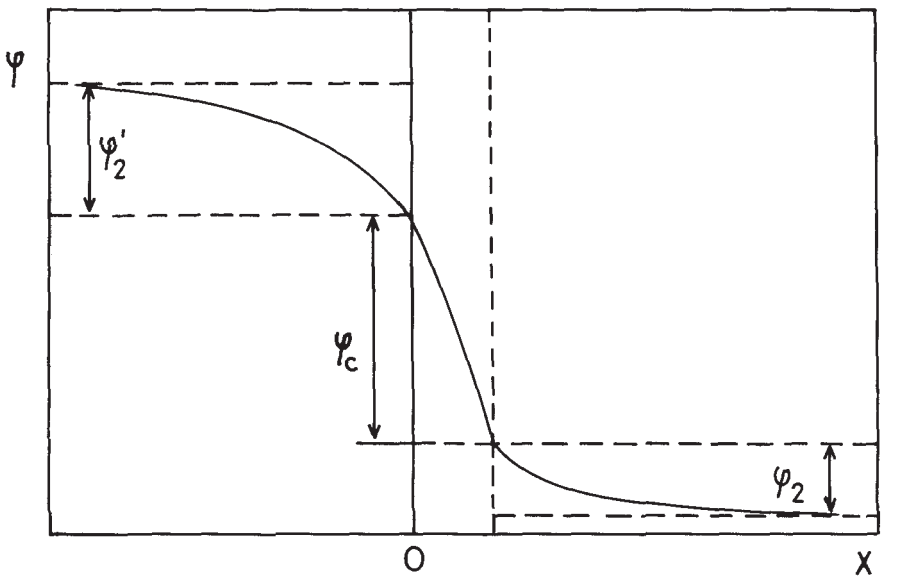

Figure 2.13 Dependence of the inner potential ϕ on the distance from the surface of the semiconductor x

approach the interface to the distance of the outer Helmoltz plane. Between this plane and the surface of the semiconductor there is a potential difference ϕ_c, so that the overall Galvani potential difference between the bulk of the semiconductor and the bulk of the electrolyte is given by the electrical potential differences ϕ_2 and $\phi_2^{'}$ in both diffuse double layers and ϕ_c. At the surface of a semiconductor, surface energy levels may also form. An additional electrical potential difference due to the charge corresponding to these surface levels may contribute to the potential difference in the compact part of the double layer.

The drop in the solution side is modified by a change in the composition of the solution, pH, and specifically adsorbed species, and may not vary by the introduction of a redox couple which does not interact with the semiconductor. An equilibrium situation is attained with a potential drop, U_{SE}, in the space–charge region below the semiconductor surface, leading to a certain bending of the energy bands (Fig. 2.14),[29] that is, in order for the two phases to be in equilibrium, their electrochemical potential must be the same. The electrochemical potential of the solution is determined by the redox potential of the electrolyte solution, and the redox potential of the semiconductor is determined by the Fermi level. If the redox potential of the solution and the Fermi level do not lie at the same energy, a movement of charge between the semiconductor and the solution is required in order to equilibrate the two phases. The excess charge that is now located on the semiconductor does not lie at the surface as it would for a metallic electrode, but extends into the electrode for a significant distance (100–10,000 Å). This region is referred to as *space–charge region*, and has an associated electric field. Hence, there are two double layers to consider: the interfacial (electrode–electrolyte) double layer, and the space–charge double layer.

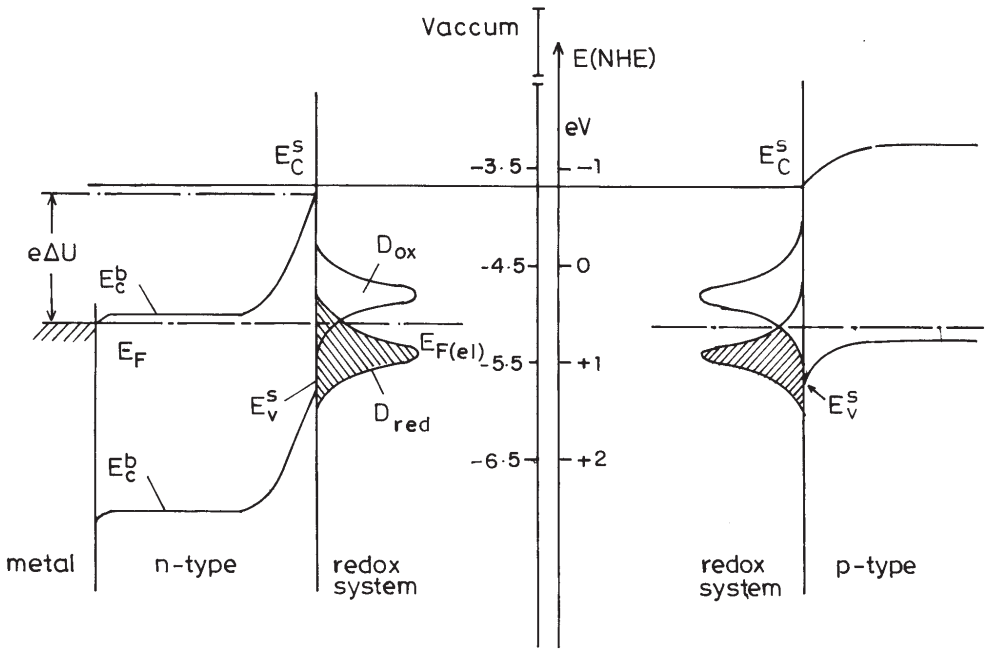

Figure 2.14 Energy states of semiconductor electrodes and redox systems [Reproduced from Ref. 29]

The redox system is characterized by its standard potential referred to the standard hydrogen electrode (NHE = 0). However, a vacuum level is chosen as a point of reference in these studies, as is common in the study of solids. The E_{redox} is a Fermi level E_F and the point of reference is shifted up by 4.5 eV.

$$E_{\text{NHE}} = -4.5\,eV$$
$$E_{F(el)} = -(4.5\,eV + E_{\text{redox}})$$

At equilibrium,

$$E_F = E_{F(el)}$$

For an *n*-type semiconductor electrode at open circuit, the Fermi level is typically higher than the redox potential of the electrolyte, and hence electrons will be transferred from the electrode into the solution. Therefore, there is a positive charge associated with the space–charge region, and this is reflected in an upward bending of the band edges. Since the majority charge carrier of the semiconductor has been removed from this region, this region is also referred to as a *depletion layer*. For a *p*-type semiconductor, the Fermi layer is generally lower than the redox potential and hence electrons must transfer from the solution to the electrode to attain equilibrium. This generates a negative charge in the space–charge region, which causes a downward bending in the band edges. Since the holes in the space–charge region are removed by this process, this region is again a depletion layer.

As for metallic electrodes, changing the potential applied to the electrode shifts the Fermi level. The band edges in the interior of the semiconductor (that is, away from the depletion region) also vary with the applied potential in the same way as the Fermi level. However, the energies of the band edges at the interface are not affected by changes in the applied potential. Therefore, the change in the energies of the band edges are prevented from going from the interior of the semiconductor to the interface, and hence the magnitude and direction of band bending, varies with the applied potential. There are three different situations to be considered.

1. At a certain potential, the Fermi energy lies at the same energy as the solution redox potential. There is no net transfer of charge, and hence there is no band bending. This potential is therefore referred to as the *flatband potential*, E_{fb}.
2. Depletion region arises at potentials positive of the flat band potential for an n-type semiconductor and at potentials negative of the flatband potential for a p-type semiconductor.
3. At potentials negative of the flatband potential for an n-type semiconductor, there is now an excess of the majority charge carrier (electrons) in this space–charge region, which is referred to as an *accumulation region*. An accumulation region arises in a p-type semiconductor at potentials more positive than the flatband potential.

The charge transfer abilities of a semiconductor electrode depend on whether there is an accumulation layer or a depletion layer. If there is an accumulation layer, the behaviour of a semiconductor electrode is similar to that of a metallic electrode, since there is an excess of the majority of charge carriers available for charge transfer, and electron transfer reactions occur slowly, if at all.

However, if the electrode is exposed to radiation of sufficient energy, electrons can now be promoted to the conduction band. If this process occurs in the interior of the semiconductor, recombination of the promoted electron and the resulting hole typically occurs, together with the production of heat. However, if it occurs in the space–charge region, the electric field in this region will cause the separation of the charge. For example, for an n-type semiconductor at positive potentials, the band edges curve upwards and hence the hole moves towards the interface, and the electron moves to the interior of the semiconductor. The hole is a high energy species that can extract an electron from a solution species: the n-type semiconductor electrode acts as a photoanode. The ideal behaviour of a n-type semiconductor electrode in the dark and under illumination is shown in Fig. 2.15. At the flat band potential, there is no current, either in the dark or upon irradiation, since there is no electric field to separate any generated charge carriers. At potentials negative of the flatband potential, an accumulation layer exists, and the electrode can act as a cathode, both, in the dark and upon irradiation (the electrode is referred to as a dark cathode under these conditions). At potentials positive of the flatband potential, a depletion layer exists, so there can be no oxidative current in the dark. However, upon irradiation, a photocurrent can be observed at potentials negative of the redox potential of the analyte (which lies at E^0), since some of the energy required for the oxidation is provided by the radiation (via the high energy hole). Using similar reasoning, it can be taken that p-type semiconductor electrodes are dark anodes and photo cathodes.

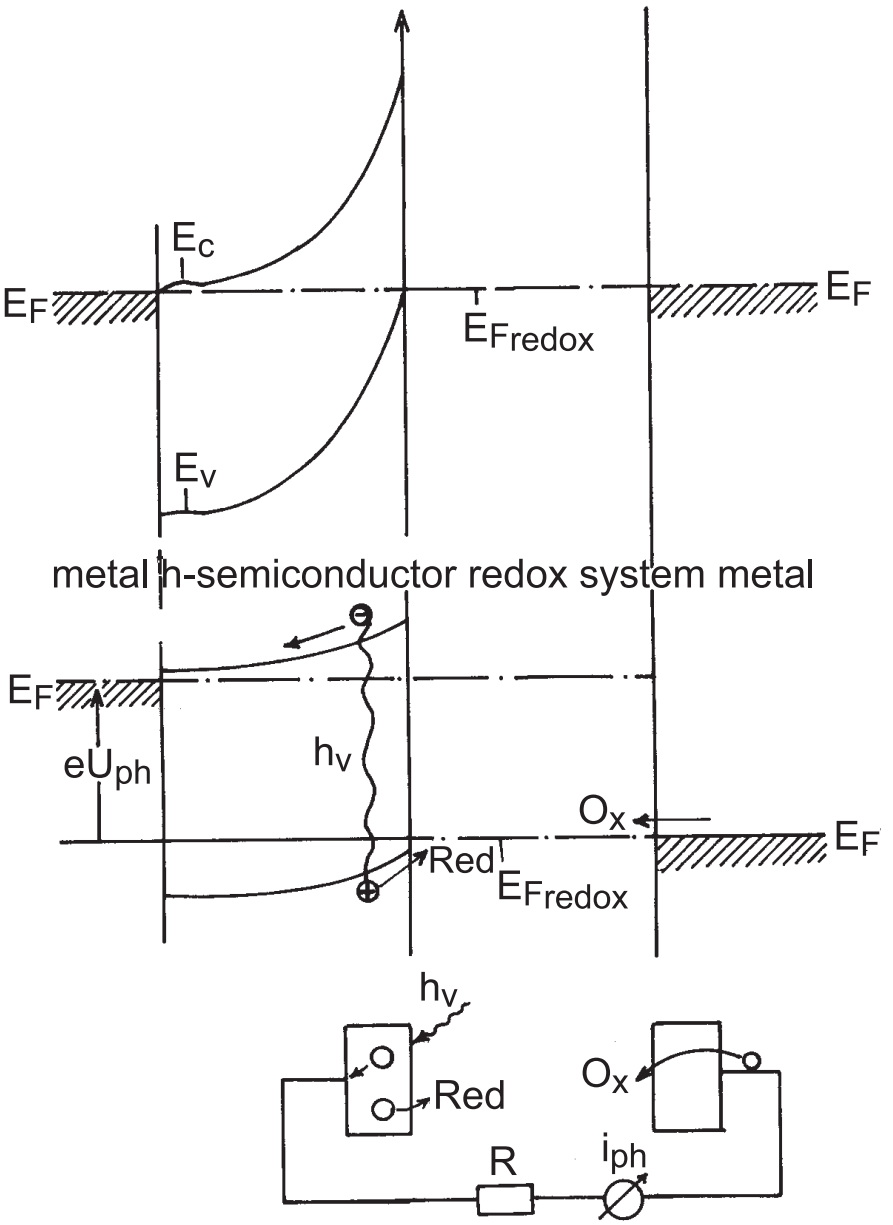

Figure 2.15 Energy scheme of a regenerative solar cell, with *ideal behavior* of an n-type semiconductor in the dark and under irradiation [Reproduced from Ref. 29]

In a semiconductor electrode, almost all the potential variation in the interfacial region occurs in the space–charge region. This is because the values for the space–charge capacity, C_{sc}, are from 0.001–1 $\mu F\,\mathrm{cm}^{-2}$, whilst those for C_d are from 10–100 $\mu F\,\mathrm{cm}^{-2}$, so that C_{sc} dominates. The

theory of the space–charge region was developed by Schottky,[30] Mott,[31] Davydov,[32] and more completely, by Brittain and Garrett.[33]

The important situations one should stress are:

- For a certain value of applied potential, there is equality between the number of electrons removed from and supplied to the electrode. In this situation there will be no space–charge region and the potential is called the flat-band potential E_{fb};
- Electrons accumulate in the space–charge layer by injection, giving rise to an accumulation layer;
- Electrons are removed from the space–charge layer, creating a depletion layer;
- The force necessary to extract electrons from the electrode is so great that they are extracted not only from the conduction band but also from the valence band (equivalent to hole injection). An inversion layer is formed, because the n-type semiconductor is converted into a p-type semiconductor at the surface. Adsorbates can facilitate this process.

To have the passage of current it is necessary that E_F is within the conduction or within the valence band in the space–charge region, that is, accumulation layer in an n-type semiconductor (oxidation).

There is an analogy with the Schottky diode. When $E_{\mathrm{c,sur}} > E_{fb}$ (negative voltage bias), and possibly for E, slightly negative in relation to E_{fb}, there will be a large current flux, assuming that there are electronic levels in solution to accept the electrons from the electrode. When $E_{\mathrm{c,sur}} < E_{fb}$ the current will be almost zero.

Another important aspect refers to adsorbates. These have their own associated energy levels, known as surface states, and can aid electron transfer if there is superposition of the conduction band and that corresponding to the surface state.

For a p-type semiconductor the arguments are analogous; in this case the majority carrier is the hole.

Due to the great extension of the space–charge region, almost all the potential drop occurs across it. There are a number of experiments used to measure the various parameters discussed above. The flatband potential can be determined by measuring the photopotential as a function of radiation intensity, the onset of the photocurrent or the capacitance of the space–charge region. The simplest method is to measure the open-circuit potential (photopotential) of the electrochemical cell under irradiation of varying intensity. For a system under equilibrium, the photopotential is the change in the Fermi level due to the promotion of electrons to the conduction band, and it reaches a maximum at the flatband potential. Therefore, a plot of photopotential versus light intensity will attain a limiting plateau at the flatband potential. For the second method, although the onset of the photocurrent might be simplistically considered to be the flatband potential, it is actually the potential at which the dark current and photocurrents are equal. Therefore, such measurements should be used with caution.

The third method involves measuring the apparent capacitance as a function of potential under depletion condition and is based on the Mott–Schottky relationship. So, one can measure its capacity, C_{sc}, and calculate E_{fb} from the Mott–Schottky relation

$$1/C_{sc}^2 = \frac{2}{e\varepsilon_r\varepsilon_0 N_D}\left[\left(E - E_{fb}\right) - \frac{k_B T}{e}\right] \tag{2.78}$$

where,

C_{sc} = capacitance of the space–charge region

ε_r = dielectric constant of the semiconductor

ε_0 = permittivity of free space

N_D = donor density (electron donor concentration for an n-type semiconductor or hole acceptor concentration for a p-type semiconductor)

E = applied potential

E_{fb} = flatband potential.

A plot of $1/C_{sc}^2$ vs E is a straight line if all other voltage drops are unaffected by the applied potential. E_{fb} can be calculated from the intercept and N_D from the slope. The presence of absorbates modifies C_{sc} and is manifested in the non-linearity of the plots. Once E_{fb} is known, one can calculate E_v and E_c from the relations

$$E_v = E_{fb} + k_B T \ln \frac{N_D}{N_v} \quad (p\text{-type semiconductor})$$

$$E_c = E_{fb} + k_B T \ln \frac{N_D}{N_c} \quad (n\text{-type semiconductor})$$

where N_v and N_c are the effective density of states in the valence and conduction bands, respectively, for p- and n-type semiconductors. Knowing the values of the bandgap energy, $E_{g,}$ one can then calculate E_c for p-type and E_v for n-type semiconductors, respectively.

The energy level of the lower edge of the conduction band of a semiconductor can be considered to be a measure of the strength of the photoexcited electrons, whereas that of the upper edge of the valence band is a measure of the oxidation strength of the hole. Based on the potential of these semiconductors for oxidation and reduction reactions, the system can be classified into four groups from the point of view of the water splitting reaction.[34]

1. *OR type*: OR indicates a strong ability for both oxidation and reduction. The oxidation and reduction power is strong enough (in principle) for the evolution of both, hydrogen and oxygen. Examples are TiO_2, $SrTiO_3$ and CdS.
2. *R type*: Only the reduction power is strong enough to reduce water. The oxidation power is too weak to reduce water. Examples are CdTe, CdSe and Si.
3. *O type*: The oxidation power is strong enough to oxidize water, since the valance band is located at a more positive value compared to the O_2/H_2O level. The reduction power is not strong enough to reduce water. Examples are WO_3, Fe_2O_3 and Bi_2O_3.
4. *X type*: The conduction and valence bands are located between the H^+/H_2 and O_2/H_2O levels. Therefore, both the oxidation and reduction powers are weak and neither oxygen nor hydrogen can be evolved.

The energy levels of various semiconductors are shown in Fig. 2.16 together with the redox potentials of various species.[35] Depending on the relative positions of the valence band and the conduction band and the redox potentials of the species undergoing reaction and those of the semiconductor for undergoing photo-corossion, either the redox chemical reaction will be promoted or the dissolution of the semiconductor (photo-corossion) will predominate. The energy level diagram in Fig. 2.17 shows that in the case of ZnO, the dissolution potential of ZnO is more negative compared to water oxidation and will be the preferred route.[35] Thus, the relative redox potentials of the various species present in the reaction medium decides which reaction will be favoured under photo-irradiation conditions.

2.5.6 Specific adsorption

Adsorption can be broadly classified into two types, namely, *nonspecific adsorption*, in which long range electrostatic forces perturb the distribution of ions near the electrode surface, and *specific adsorption*, in which strong interaction between the adsorbate and the electrode material causes the formation of a layer (partial or complete) on the electrode surface. Nonspecific adsorption of an electroactive species can affect the electrochemical response since it affects the concentration of the species as well as the potential distribution near the electrode. Specific adsorption can have several effects. For instance, if an electroactive species is adsorbed, the theoretical treatment of a given electrochemical method must be modified to account for the presence of the reactive species at the surface of the electrode with respect to an amount higher than the bulk concentration when time $t = 0$. Specific adsorption can also affect the energetics of the reaction, for example, adsorbed O may be more difficult to reduce than dissolved O specis. Specific adsorption of an electro-inactive species can also alter the electrochemical response by forming a blocking layer on the electrode surface. Sometimes, adsorption may also increase the reactivity of the species by causing dissociation of a nonreactive material into reactive fragments.

As explained in Grahame's model for the double layer, specific adsorption is the adsorption of ions at the electrode surface after losing their solvation partially or completely. These ions can have the same charge or the opposite charge to that of the electrode. Bonds formed with the electrode in this way are stronger than the bonds formed with solvated ions.

The existence of specific adsorption was suggested as an explanation for the fact that electrocapillary curves at mercury electrodes are different for different electrolytes at the same concentration. For sodium and potassium halides in water, the differences arise at potentials positive of E_z, which suggests an interaction with the anions. Larger this effect, the smaller is the ionic radius of the anion. So the idea of specific adsorption with partial or total loss of hydration was proposed.

The degree of specific adsorption should vary with electrolyte concentration, just as there should be a change in the point of zero charge due to specific adsorption of charges. This is the Esin–Markov effect, expressed by the Esin–Markov coefficient, β, where

$$\beta = \frac{1}{RT}\left(\frac{\partial\,(\delta E_z)}{\partial \ln a}\right)_{\sigma_M} = \left(\frac{\partial\,(\delta E_z)}{\partial \mu}\right)_{\sigma_M} \tag{2.79}$$

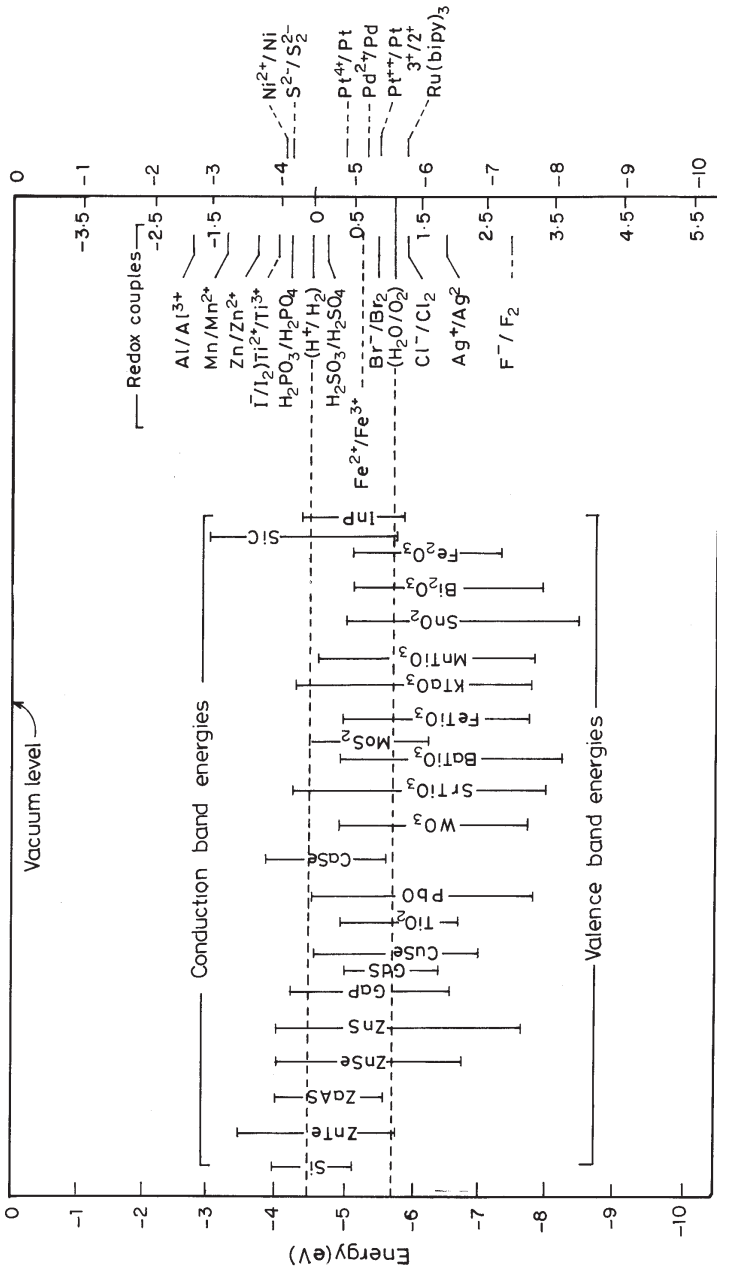

Figure 2.16 Relative energy levels of some common semiconductor electrode materials and redox systems [Reproduced from Ref. 35]

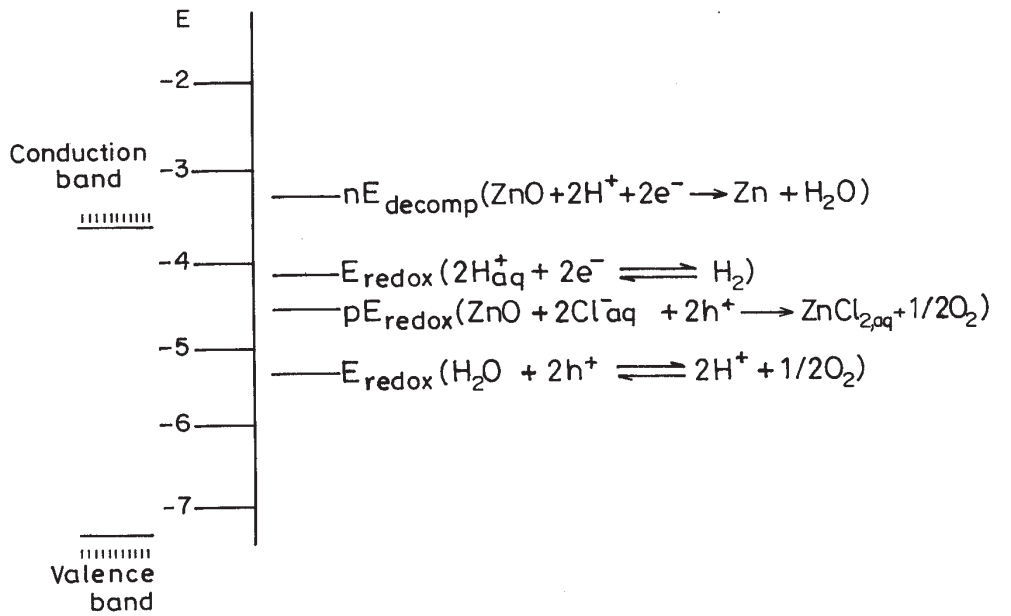

Figure 2.17 Relative positions of the redox potentials of the substrate and semiconductor dissolution for ZnO at $pH = 0$ [Reproduced from Ref. 35]

This derivative is equal to zero in the absence of specific adsorption. For anion adsorption, and constant charge density, the point of zero charge moves in the negative direction in order to counterbalance adsorption. For cations, E_z moves in the positive direction, assuming constant charge density. In aqueous solution, specific adsorption only occurs close to E_z. Far from E_z, solvent molecules are attracted so strongly that it is difficult to push them away.

Experimentally, it is observed that specific adsorption occurs more with anions than with cations. This is in agreement with chemical models of the interfacial region. Since, according to the free electron model, a metallic lattice can be considered as a cation lattice in a sea of electrons in free movement, it is logical to expect a greater attraction for anions in solution.

An isotherm gives the relation between bulk and surface concentration. The relationship between the amount of substance i adsorbed on the electrode per unit area, Γ_i, the activity in bulk solution a_i^b, and the electrical state of the system, E or q^M, at a given temperature, is given by the *adsorption isotherm*. This relationship is usually based on model assumptions. It is obtained from the condition of equality of electrochemical potential for bulk and adsorbed species i at equilibrium

$$\mu_i^A = \mu_i^b \tag{2.80}$$

where the superscripts A and b refer to adsorbed i and bulk i, respectively. Thus

$$\mu_i^{o,A} + RT \ln a_i^A = \mu_i^{o,b} + RT \ln a_i^b \tag{2.81}$$

where μ_i^o terms are the standard electrochemical potentials. The standard free energy of adsorption, ΔG^0, which is a function of the electrode potential, is defined as

$$\Delta G^0 = \mu_i^{o,A} - \mu_i^{o,b} \tag{2.82}$$

Thus

$$a_i^A = a_i^b e^{-\Delta G_i^\circ / RT} = \beta_i a_i^b \tag{2.83}$$

where,

$$\beta_i = \exp(-\Delta G_i^\circ / RT) \tag{2.84}$$

Equation (2.83) is a general form of an adsorption isotherm, with a_i^A a function of a_i^b and β_i. Different specific isotherms result from different assumptions or models for the relationship between a_i^A and Γ_i. Some frequently used ones are given below.

In the simplest model it is assumed that every adsorption site is equivalent (that is, no heterogeneity of the surface), only monolayer adsorption occurs and there is no interaction between molecules at adjacent sites. The saturation coverage of the electrode by adsorbate (to form a monolayer) is of amount Γ_s. Thus,

$$\frac{\Gamma_i}{\Gamma_s - \Gamma_i} = \beta_i a_i^b \tag{2.85}$$

Isotherms are sometimes written in terms of the fractional coverage of the surface, $\theta = \Gamma_i / \Gamma_s$.

Langmuir isotherm: The Langmuir isotherm is in the form

$$\frac{\theta}{1 - \theta} = \beta_i a_i^b \tag{2.86}$$

Langmuir isotherm is useful for a first approximation or for simple systems.

Interactions between adsorbed species complicate the problem by making the energy of adsorption a function of surface coverage. Isotherms that include this possibility are the logarithmic Temkin isotherm and the Frumkin isotherm.

Temkin isotherm: This considers that the adsorption energy is a function of the degree of coverage. According to it,

$$\Gamma_i = \frac{RT}{2g} \ln\left(\beta_i a_{i,\infty}\right) \tag{2.87}$$

where 'Γ_i' is the excess of species i, and 'g' is a parameter that treats the interaction energy between the adsorbed species, varying with coverage. The parameter g has dimensions of J(mol)$^{-1}$ per mol cm^{-2}.

Frumkin isotherm: This considers interactions in a different way

$$\beta_i a_{i,\infty} = \frac{\Gamma_i}{\Gamma_s - \Gamma_i} \exp \frac{2g\Gamma_i}{RT} \tag{2.88}$$

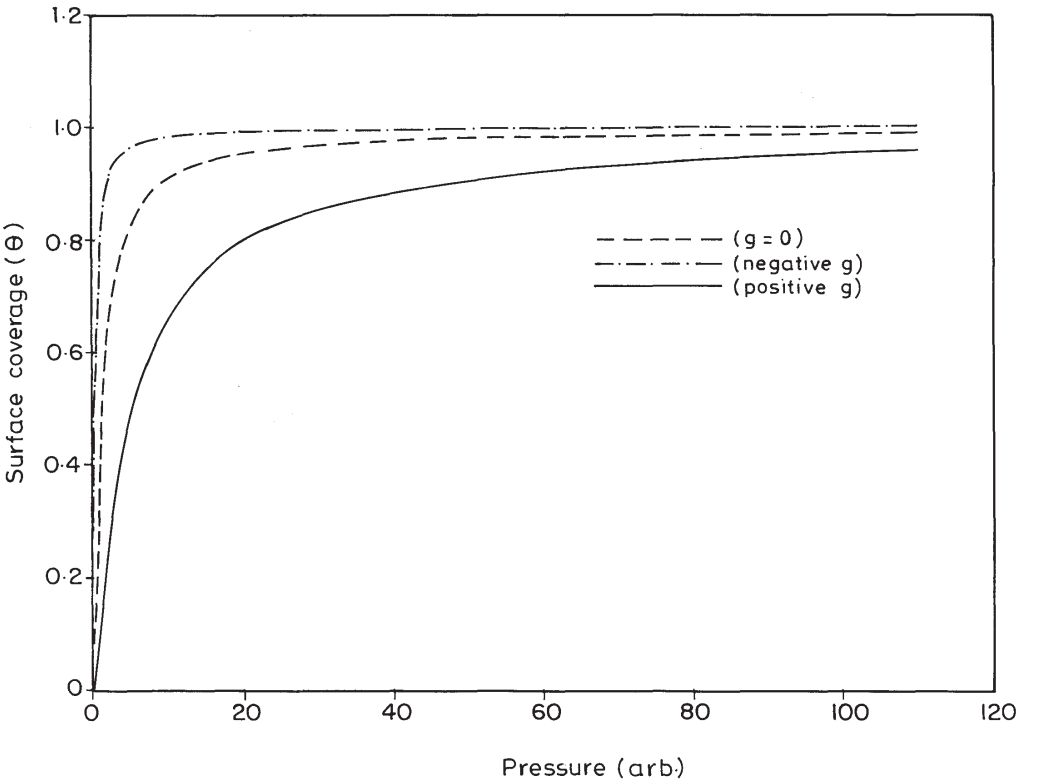

Figure 2.18 Langmuir and Frumkin adsorption isotherms

or

$$\Gamma_i = \frac{RT}{2g} \ln\left(\beta_i a_{i,\infty}\right) \ln\left(\frac{\Gamma_s - \Gamma_i}{\Gamma_i}\right) \tag{2.89}$$

Γ_s being the maximum surface excess. A positive value of g implies attractive interaction and negative g repulsive interaction. When $g = 0$ and putting $\Gamma_i / \Gamma_s = \theta$, the Langmuir isotherm is obtained from (2.88). Additionally, comparison of (2.87) and (2.88, 2.89) shows that the Temkin isotherm is a special case of the Frunkin isotherm when $\Gamma_i / \Gamma_s = 0.5$. Both, the logarithmic Temkin and the Langmuir isotherm can be considered as special cases of the Frumkin isotherm. Typical plots of Langmuir and Frumkin isotherms are given in Fig. 2.18. The fundamentals and derivation of various monolayer and multilayer adsorption isotherms and the factors influencing them are discussed in detail elsewhere.[36]

2.5.7 Electrokinetic phenomena and colloids: the zeta potential

A colloidal system consists of a disperse phase suspended in a dispersion medium, which does not separate with time.[37–39] All combinations of gases, liquids and solids are possible except for a

gas dispersed in a gas. Normally, the term colloid refers to a solid suspended in a liquid, the solid particles having diameters between 10^{-7} cm and 10^{-5} cm. The solid particles are charged, which causes repulsion between the particles and offers temporal stability to the colloidal system. There has been increasing interest in colloids due to their possible use as electrodes for electrolysis, each particle acting as the anode and cathode at a time. Their particular advantage is the large surface area exposed to solution in relation to their solid volume. Since the particles are charged there is an interfacial region, which exhibits many of the properties of the interfacial region of a solid electrode. Therefore, the study of colloids can also lead to a better knowledge of the double layer region, especially for ionic solids and semiconductors.

The electrokinetic phenomenon that results from the movement of a solid phase with surface charge relative to an electrolyte-containing liquid phase is a useful phenomenon in case of colloidal particles. An applied electric field induces movement of ions, or movement of ions induces an electric field, and can be divided into two categories.

1. *Electrophoresis*: Charged solid particles (colloidal particles) moving through the liquid under the influence of an electric field. It is called *sedimentation potential* when it is due to gravitational force.
2. *Electroosmosis*: Liquids moving past charged solid surfaces (or possibly through membranes) under the influence of an electric field. It is called *streaming potential* when it is due to an applied pressure difference . These effects are normally studied in fine capillaries in order to maximize the ratio of the solid surface area to the liquid volume.

The four manifestations of the electrokinetic effect are summarized in Table 2.2

Table 2.2 Electrokinetic phenomena

Mobile phase / Stationary phase	Phenomenon (Force applied)	Property measured
Solid / Liquid	Electrophoresis (Electric field)	Electrophoretic mobility via mass transport measurement, microscope or Doppler effect
Solid / Liquid	Sedimentation potential (Force of gravity)	Potential difference
Liquid / Solid	Electroosmosis (Electric field)	Rate of liquid movement, pressure
Liquid / Solid	Streaming potential (Pressure)	Potential difference

The size of the particles that is calculated from these experiments corresponds to particle dimensions plus the double layer thickness. The shear plane is defined such that inside it the adsorbed species are rigidly held, and outside it there is free movement. The shear plane can therefore be associated roughly with the outer Helmholtz plane. The value of the electrostatic potential at the shear plane with respect to the value in bulk solution is called the electrokinetic or zeta potential, ξ.[40]

In the presence of a large quantity of inert electrolyte, the entire potential drop is confined to within the compact layer and ξ is zero. By application of an appropriate potential at an electrode one can also achieve $\xi = 0$; this value of the potential is called the *isoelectric point*. This is, in fact, not equal to the point of zero charge, as the value of the latter is affected by the presence of specifically adsorbed species.

Electrophoresis

In electrophoresis the solid moves through a liquid phase due to the application of an electric field. The forces acting on the particles are similar to those that act on solvated ions. These are:

- Force of the electric field on the particle
- Frictional forces
- Force due to the action of the electric field on ions of charge opposite to that of the particle within the double layer (relaxation effect)
- Induction forces in the double layer caused by the electric field (electrophoretic retardation)

The electrophoretic mobility, u_e, is calculated by solving the Poisson equation with the appropriate boundary conditions. The final relation is

$$u_e = \frac{2\varepsilon\xi E}{3\eta} f\,(a/x_{DL}) \tag{2.90}$$

where ε is the permitivity, η the absolute viscosity, E the electric field strength, and $f(a/x_{DL})$ a numerical factor, where a is the particle radius and x_{DL} the double layer thickness, that varies according to the type of force described above. This depends on particle size and double layer thickness. For every small particle in dilute solution, the double layer is thick and $f(a/x_{DL}) \rightarrow 1$ (negligible relaxation effect). For large particles in concentrated solutions, where the double layer is thin, $f(a/x_{DL}) \rightarrow 1.5$ (negligible electrophoretic retardation). All other situations lead to intermediate numerical factors. Measurements of electrophoretic mobility, using equation (2.90) with the appropriate numerical factor, lead to values for the zeta potential.

As can be seen, electrophoretic mobility depends on ionic strength of the solution since double layer thickness decreases with increase in concentration of the electrolyte. It also depends on the surface charge of the particles. If this charge varies in colloidal particles of similar dimensions, then electrophoresis provides a basis for their separation. For example, in proteins the surface charge varies with pH according to the protein identity.

Sedimentation potential

Colloidal particles are affected by the force of gravity, either natural or through centrifugation. Sedimentation of the particles often gives rise to an electric field. This occurs because the particles move, whilst leaving some of their ionic atmosphere behind. These potentials are usually difficult to measure and may cause undesirable side-effects in ultracentrifugation. This can be minimized by the addition of large concentrations of inert electrolyte.

Electroosmosis

In electroosmosis, the stationary and mobile phases are exchanged in relation to electrophoresis. Since the measurement of the rate of movement of a liquid through a capillary is difficult, the force that it exerts is measured, that is, the *electroosmotic pressure*, or alternatively, the volume of liquid transported through a capillary in a given time interval. The electroosmotic velocity, v_{eo}, is given as

$$v_{eo} = \frac{\varepsilon \xi E}{\eta} \tag{2.91}$$

which is of the same form as equation (2.90) for electrophoresis, when $f(a/x_{DL}) \rightarrow 1.5$, since the capillary radius is much larger than the double layer thickness.

The volume flow of liquid, V_f, is $v_{eo}A$, where A is the cross-sectional area of the capillary. A current of magnitude $I = AkE$ will flow, in which k is the solution conductivity, and thus the electroosmotic flow, flux per unit electric current at zero pressure difference, is given by

$$\frac{V_f}{I} = \frac{v_{eo}A}{I} = \frac{\varepsilon \xi}{k\eta} \tag{2.92}$$

Streaming potential

If a pressure difference, ΔP, is applied between the extremes of a capillary that exerts a potential difference, this is known as streaming potential:

$$\Delta\phi = \frac{e\xi}{k\eta}\Delta P \tag{2.93}$$

By comparing this with equation (2.92) one can see a close relationship between streaming potential and electroosmotic flow.

Limitations in the calculation of the zeta potential

Quantitative measurements of electrokinetic phenomena permit the calculation of the zeta potential by use of the appropriate equations. However, in the deduction of the equations certain approximations are made, because in the interfacial region physical properties such as concentration, viscosity, conductivity, and dielectric constant differ from the corresponding values in bulk solution, which is not taken into account. Corrections to compensate these approximations have been introduced, as well as consideration of non-spherical particles and particles of dimensions comparable to the diffuse layer thickness.

References

[1] Glasstone S, *Thermodynamics for Chemists*, D.Van Nostrand Company, Inc., Princeton, N.J., 1963.

[2] Kirkwood JA, Oppenheim I, *Chemical Thermodynamics*, McGraw-Hill Book Company, New York, 1961.

[3] Lewis AN, Randall M, *Thermodynamics and the Free Energy of Chemical Substances*, McGraw-Hill Book Company, New York, 1928.

[4] Klotz IM, *Introduction to Chemical Thermodynamics*, WA Benjamin, Inc., New York, 1964.

[5] Guggenheim EA, *Thermodynamics*, 14th ed., Interscience Publishers, Inc., New York, 1957.

[6] Wall FT, *Chemical Thermodynamics*, 2nd ed., WH Freeman and Company, San Francisco, 1965; Ludher WF, *A Different Approach to Thermodynamics*, Reinhold Publishing Corporation, New York, 1967; Benson RS, *Advanced Engineering Thermodynamics*, Pergamon Press, New York, 1967.

[7] Carnot S, *Reflections on Motive Power of Heat* (English translation by RH Thurston), American Society Mechanical Engineers, 1953.

[8] Justi E, Winsel AW, *Kalte-Verbrennung*, Franz Steiner Verlag, Wiesbaden, 1962.

[9] Berger C, Fuel cells, *Advan. Chem. Ser.*, American Chemical Society, Washington, D.C., 47(14), 1965.

[10] Swinkels DAJ, *J. Electrochem. Soc.*, 113:6, 1966.

[11] Lippmann G, *Compt. Rend.*, 76:407, 1873.

[12] Heyrovsky J, *Chem. Listy*, 16:246, 1922.

[13] Von Helmholtz HLF, *Ann. Physik*, 89:211, 1853; 7:337, 1879.

[14] Gouy G, *Compt. Rend.*, 149:654, 1910; Chapman DL, *Phil. Mag.*, 25:475, 1913.

[15] MacInnes D, *Principles of Electrochemistry*, Dover, New York, 7, 1962.

[16] Stern O, *Z. Elektrochem.*, 30:508, 1924.

[17] Grahame DC, *Chem. Rev.*, 41:411, 1947.

[18] Bockris JO'M, Devanathan MA, Muller K, *Proc. R.. Soc.*, A274:55, 1963.

[19] Damaskin BB, Frumkin AN, *Electrochim, Acta*, 19:173, 1974; Damaskin BB, Palm UV, Salve MA, *Elektrokhimiya*, 12:232, 1976; Damaskin BB, *J. Electroanal. Chem.*, 75:359, 1977.

[20] Trasatti S, in Ref. 4, 25–48.

[21] Parsons R, *Chem. Rev.*, 90:813, 1990.

[22] Martynov GA, Salem RR, Electrical double layer at a metal-dilute electrolyte solution interface, *Lecture Notes in Chemistry*, Springer-Verlag, Berlin, 33, 1983.

[23] Goodsman J, *Electrochemistry: Theoretical Foundations*, Wiley-Interscience, New York, 232–239, 1987.

[24] Morrison SR, *Electrochemistry at Semiconductors and Oxidized Metal Electrodes*, Plenum, New York, 1980.

[25] Hamelin A, *Modern Aspects of Electrochemistry*, Conway BE, White RE, Bockris JO'M (eds.), Plenum, New York, 16:1–101, 1985.

[26] Cervino RM, Triaca WE, Arvia AJ, *J. Electroanal. Chem.*, 182:51, 1985.

[27] Uosaki K, Kita H, *Modern Aspects of Electrochemistry*, Conway BE, White RE, Bockris JO'M (eds.), Plenum, New York, 18:1–60, 1986.

[28] Hamnett A, *Comprehensive Chemical Kinetics*, Compton RG (ed.), Elsevier, Amsterdam, 27(2), 1987.

[29] Narayanan R, Viswanathan B, *Chemical and Electrochemical Energy Systems*, University Press India Ltd., 1998.
[30] Schottky W, *Zeit fur Physik*, 113:367, 1939.
[31] Mott NF, *Proc. Roy. Soc.*, A171:27, 1939.
[32] Davydov B, *J. Physics USSR*, 1:169, 1939.
[33] Brattain WH, Garrett CGB, *Phys. Rev.*, 99:376, 1955.
[34] Kalyanasundaram K, *Energy Resources Through Photochemistry and Catalysis*, Graetzel M (ed.), Academic Press, New York, 1983.
[35] Viswanathan B, *Catalysis: Principles and Applications*, Viswanathan B, Sivasanker S., Ramaswamy AV (eds.), Narosa Publishing House, New Delhi, 20:289–301, 2002.
[36] Viswanathan B, *Catalysis: Principles and Applications*, Viswanathan B, Sivasanker S., Ramaswamy AV (eds.), Narosa Publishing House, New Delhi, 1:1–33, 2002.
[37] Hunter RJ, *Foundations of Colloid Science*, Oxford University Press, New York, 1987.
[38] Everett DH, *Basic Principles of Colloid Science*, Royal Society of Chemistry, London, 1988.
[39] Kitahara A, Watanabe A, *Electrical Phenomena at Interfaces*, Dekker, New York, 1984.
[40] Hunter RJ, *Zeta Potential in Colloid Science*, Academic Press, New York, 1981.